MOLECULAR PHYLOGENY OF THE "TEMPERATE HERBACEOUS TRIBES" OF PAPILIONOID LEGUMES: A SUPERTREE APPROACH

MARTIN F. WOJCIECHOWSKI[1,2], MICHAEL J. SANDERSON[1], KELLY P. STEELE[3], AND AARON LISTON[4]

[1]**Section of Evolution and Ecology, University of California, Davis, California 95616 USA**
[2]**Museum of Paleontology and University/Jepson Herbaria, University of California, Berkeley, California 94720 USA**
[3]**Department of Biology, California State University, Hayward, California 94545 USA**
[4]**Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331 USA**

Abstract

Molecular phylogenies provide a framework to discuss relationships in the vast temperate herbaceous radiation of papilionoid legumes comprised of tribes Galegeae, Carmichaelieae, Cicereae, Hedysareae, Trifolieae, Vicieae, as well as among some members of the tropical tribe Millettieae (*Callerya, Wisteria*, and related genera). The taxa form a monophyletic group marked by the loss of the chloroplast DNA inverted repeat. The sister group to this clade includes Loteae, Coronilleae, and Robinieae. These two clades together comprise a clade we refer to as "Hologalegina", the monophyly of which is strongly supported by multiple sources of molecular data. Hologalegina is the sister group to the clade comprising Millettieae and the other predominantly Old World tropical tribes Phaseoleae, Abreae, Desmodieae, Psoraleeae, and Indigofereae, contrasting with the traditional view that the largely temperate, herbaceous tribes that form Hologalegina are derived from within Millettieae.

Introduction

The synthesis presented in this paper portrays the phylogenetic relationships among the major clades within the temperate radiation of primarily herbaceous, papilionoid legumes centered in the Old World around the large, heterogeneous tribe Galegeae (see nomenclature note below). Our goal is to highlight the recent progress that has been made in this area based on results from numerous published (and unpublished) molecular phylogenetic studies. Results from these studies have made it possible to reconstruct a comprehensive phylogeny of this large assemblage at the tribal, generic level, and in some instances, to the species level.

Many of our present day concepts of temperate, herbaceous papilionoids date to Dormer's (1945, 1946) work on vegetative anatomy and morphology. His "epulvinate series" includes tribes Galegeae, Carmichaelieae, Cicereae, Hedysareae, Loteae s.l. (i.e., including Coronilleae), Thermopsideae, Trifolieae, and Vicieae. This series was refined and recircumscribed by Polhill (1981d, 1994) as the "temperate herbaceous group" (Table 1), and excluded Thermopsideae Yakovlev. The temperate herbaceous

TABLE 1. Taxonomic diversity of Polhill's "temperate herbaceous tribes", Robinieae and related genera of tribe Millettieae.

Genus and Tribe	Number of genera/species[1]	Distribution	References[2]
Galegeae (Bronn) T. & G.	20/ ca. 3,150	cosmopolitan	
Carmichaelieae Hutch.	2/25	New Zealand, Lord Howe Island	Heenan (1998a, 1998b)
Cicereae Alefeld	1/40	Mediterranean, Asia	
Hedysareae DC.	7/265	Eurasia, Africa, Himalayas, Mediterranean	
Trifolieae (Bronn) Benth.	7/477	cosmopolitan	
Vicieae (Adans.) DC.	5/298	Eurasia, North America, temperate South America, east Africa	
Millettieae Hutch.			Geesink (1984)
Callerya (Endl.) Geesink	19	Himalaya to southern China, southeast Asia to Australia	Schot (1994)
Wisteria Nuttall (including *Wisteria japonica* Siebold & Zucc.)	9	temperate and subtropical eastern Asia and eastern USA	
Afgekia Craib	3	China, southeast Asia	Sanderson and Wojciechowski (1996)
"IRLC" total (species)	**45/4,286**		
Coronilleae (Adans.) Boiss.	6/54	Europe, Mediterranean, western Asia, tropical Africa, southern South America	
Loteae DC.	4/128	north temperate regions, Mediterranean, Africa, western Asia	
Robinieae (Benth.) Hutch.	12/146	Americas, Caribbean, pantropical (*Sesbania* only)	Lavin and Sousa (1995), Lavin (1995)
Loteae s. l. + Robinieae total	**22/328**		
"Hologalegina" total	**67/4,614**		

[1] Source: Polhill and Raven (1981)
[2] Reference for most recent "treatment"; otherwise, Polhill and Raven (1981).

group was largely distinguished from other predominately temperate tribes like Thermopsideae by the accumulation of the non-protein amino acid canavanine rather than alkaloids in seeds. This group contains many of the economically important and familiar temperate legumes, such as alfalfa, clovers, lentils, peas, vetches, chickpeas, licorice root, and locoweeds, as well as many of the model systems used for studies of nitrogen fixation (Sprent and McKey, 1994), legume biology/genomics (Cook, 1999), symbiotic root nodule development (Schultze and Kondorosi, 1998; Szczyglowski et al., 1998; Bras et al., 2000), and bacterial-plant symbioses/coevolution (Freiberg et al., 1997; Doyle, 1998; Wernegreen and Riley, 1999). Members of this group share a combination of features other than the accumulation of canavanine, notably a predominantly herbaceous habit, epulvinate compound leaves, stipules adnate to the petiole, base chromosome numbers of $n = 7$ or 8, and centers of species diversity primarily in temperate regions of the Old World. Polhill (1981b) also included *Glycyrrhiza* L. in Galegeae, which except for pulvinate leaves, has the other characteristics of this group.

Galegeae and its closest relatives have been considered to form a "natural grouping" derived from the tropical, largely woody tribe Millettieae, originating in temperate regions of Eurasia during the early Tertiary and subsequently diversifying in similar habitat in the New World. The relationship of this "galegoid" alliance to other tribal groups that show "advanced" floral features, such as the predominantly New World tribes Adesmieae (Benth.) Hutch., Amorpheae Boriss., and Aeschynomeneae (Benth.) Hutch., has remained obscure. These tribes, depicted by Polhill (1981d) as derived from Old World Galegeae and/or Millettieae, do not accumulate canavanine and instead share features typical of Dalbergieae Bronn ex DC. and New World tropical Sophoreae Sprengel. Further complicating this picture is Robinieae. Genera in this tribal group have been variously aligned with Galegeae and Millettieae on the basis of canavanine accumulation, but also with Dalbergieae and Aeschynomeneae because of a geographical concentration in the New World and predominantly tropical distribution and woody habit (reviewed in Lavin and Sousa, 1995).

New interpretations of papilionoid tribal relationships began to emerge from studies using molecular data (restriction fragment length polymorphisms or "RFLPs") beginning in the early 1990s (for earlier review, see Doyle, 1995). The first of these to have a major impact (Lavin et al., 1990) investigated the distribution of a rare structural mutation, loss of one copy of the 25-kb "inverted repeat" (IR; containing ribosomal RNA and other genes) in chloroplast DNA (Palmer et al., 1987). This mutation, observed in most temperate and herbaceous legumes surveyed, has been considered an unequivocal marker for a group almost exclusively comprising one of Polhill's lineages, which includes the temperate herbaceous tribes (except Loteae s.l.). Furthermore, the fact that this mutation was shared by the temperate *Wisteria* of Millettieae, but not by members of other legume tribes with temperate distribution (Loteae s.l., Robinieae, Genisteae (Adans.) Benth., and Thermopsideae), provided some of the first evidence suggesting an independent origin of these temperate, largely herbaceous tribes (Lavin et al., 1990). A subsequent study by Liston (1995) confirmed and expanded the results of Lavin et al. (1990), strengthening the hypothesis that loss of the inverted repeat represents a derived feature with a single origin in papilionoids. These studies also suggested that this mutation occurred later in the evolution and diversification of the subfamily, and certainly after the origin of canavanine synthesis.

Since the work of Lavin et al. (1990), a number of phylogenetic analyses of chloroplast and nuclear DNA sequence variation have revealed further support for the monophyly of the IR-lacking clade ("IRLC") as well as provided some resolution as to the relationships among members of this clade and its close relatives. Studies aimed at resolving higher level relationships within the clade have utilised chloroplast encoded *rpo*C genes (Liston and Wheeler, 1994) and *trn*L intron sequences (Wojciechowski et al., 1999), and nuclear rDNA ITS (Sanderson and Wojciechowski,

1996; Wojciechowski et al., 1999), while studies using the chloroplast *rbc*L gene (Doyle et al., 1997; Käss and Wink, 1997; Doyle et al., 2000), *mat*K gene (Hu et al., 2000), and nuclear phytochrome genes (Lavin et al., 1998) have focused more on relationships among papilionoids in general or the Millettieae-Phaseoleae complex (i.e., Millettieae, Phaseoleae DC., Abreae (Wight & Arn.) Hutch., Desmodieae (Benth.) Hutch., Psoraleeae (Benth.) Rydb.). In addition, a number of studies have focused on the molecular systematics of individual genera or specific tribes, such as *Lathyrus* L. (Asmussen and Liston, 1998), *Medicago* L. (Bena et al., 1998a, 1998b, 1998c; Downie et al., 1998), and Loteae (Allan and Porter, in press). In this paper, we summarise the results from many of these studies, including several higher level analyses (Doyle et al., 1997; Hu et al., 2000; Hu, 2000), as well as our own unpublished studies of sequences from the *mat*K and *rpo*C regions of chloroplast genome, in the form of a supertree analysis to reconstruct the phylogenetic relationships among the temperate herbaceous tribes and their relatives with more tropical distribution, the Millettieae-Phaseoleae complex, Indigofereae (Benth.) Rydb., Loteae, and Robinieae.

Methods

Detailed descriptions of the taxa surveyed, molecular methods employed in data collection (DNA isolation, PCR amplification of chloroplast and nuclear genes/regions, restriction analyses, and DNA sequencing) and phylogenetic analyses we have performed are provided in our publications (e.g., Sanderson and Doyle, 1993; Liston and Wheeler, 1994; Steele and Vilgalys, 1994; Wojciechowski et al., 1999) and those of authors whose studies were included here (see Table 1). A complete list and description of taxa sampled, methods for PCR amplification, sequencing, and phylogenetic analyses of the *mat*K gene and their implications will be provided elsewhere (Steele et al., unpubl.; Wojciechowski et al., unpubl.).

For this paper, we take two approaches to reconstructing phylogenetic relationships in this large group of papilionoids. First, we present preliminary results from parsimony analyses of sequences of the complete chloroplast *mat*K gene, sampling representatives from 33 genera of the temperate herbaceous tribes (total of 41 genera of papilionoids) culled from a larger data set spanning most papilionoid tribes, and use that phylogeny as a reference for a brief discussion of higher level relationships. Second, using those results as a framework, we have employed "supertree" construction methods, a strategy for building larger, progressively more inclusive or composite phylogenies from a number of partially overlapping smaller trees (Sanderson et al., 1998), to provide a more complete hypothesis of relationships within this group. Constructing phylogenetic supertrees in this way resembles "taxonomic congruence" approaches in that the raw data are analysed individually and the resultant trees are combined (assuming taxonomic overlap). This method has the added advantage of permitting the combination of trees derived from data that are heterogeneous or incompatible, i.e., derived from different data types or methods of phylogenetic analyses (Bininda-Emonds and Bryant, 1998).

The method we have used here for supertree construction, "matrix representation with parsimony" or "MRP" (Baum, 1992; Ragan, 1992), is essentially a parsimony analysis of the phylogenetic signal within each data set without the confounding noise of homoplasy *within* each data set (conflicts *between* data sets may still exist). MRP represents the pattern of relationships within a single tree derived from an analysis of a particular data matrix ("source tree") as a series of binary "elements" or pseudo-characters, each describing a node such that all descendants of that node are scored as "1", all other taxa scored as "0". This method converts the topology of each source tree into an equivalent data matrix, which are then combined into a "supermatrix" of unique taxa. Taxa missing from any given source tree are scored as "?" in the supermatrix for those particular elements. Trees are rooted either using a

hypothetical – "all zero" – outgroup, which we have employed here, or a real outgroup taxon that is common to all source trees. More detailed discussion of the theoretical framework, basic methodologies used in the construction of supertrees, as well as the limitations of matrix representation, are presented in Bininda-Emonds and Bryant (1998) and Sanderson et al. (1998). To date, this method has been used to produce a composite phylogeny for all primate species (203 taxa, over 100 studies; Purvis, 1995) and all extant species of the Carnivora (271 taxa, 177 studies; Bininda-Emonds et al., 1999). These two supertree studies have included phylogenies derived from analyses of cytogenetic, molecular, morphological, and behavioral characters, and trees generated using parsimony, maximum likelihood, neighbor-joining, and distance methods. Here we extend this approach to begin constructing a composite tree of this large clade of papilionoid legumes, incorporating the results from a limited number of studies and only those which have utilised molecular data (DNA sequences and RFLPs), for phylogeny reconstruction.

Candidate source trees were identified from the list of molecular studies on "temperate herbaceous group" taxa compiled from the literature that is presented in Table 2, plus several of our own unpublished data sets. Single source trees were obtained from each individual study, chosen with several criteria in mind. Only the most recent, and most taxonomically complete study by a researcher or group which has published more than one paper (or made available unpublished data) using the same data source(s) was used; for example, Bena et al. (1998b) was chosen over Bena et al. (1998a, 1998c); Wojciechowski et al. (1999) instead of Wojciechowski et al. (1993). In addition, source trees were included if and only if they shared at least two taxa in common, the minimum requirement for formal supertree assembly (Sanderson et al., 1998). For this reason, studies such as Gauthier et al. (1997), Hayashi et al. (1998), Lavin et al. (1991), Mayer and Soltis (1994), van Oss et al. (1997), and Yamakazi et al. (1994) were not included in the supertree analysis presented here. One possible way to overcome this requirement so as to incorporate the results from studies such as these in supertree construction would be to substitute a single taxon (e.g., a species) that is represented in a supertree – as a place holder – with a source tree for a larger, more inclusive taxon (e.g., the genus) containing it as a member, if in fact the larger taxon has been shown to be monophyletic. However, the monophyly of these taxa have not, or may have not, been definitively demonstrated.

Initially, a supermatrix was constructed by creating a file in PAUP* version 4.02b (Swofford, 1999) containing one single most parsimonious tree derived from analyses of the data from each of twelve studies. These trees were compiled into a single matrix (NEXUS format) using the SUPER function of the program 'r8s' (Sanderson, 1997) then further amended using the data editor of MacClade 3.07 (Maddison and Maddison, 1992), ultimately incorporating data from a total of 22 studies. A source tree based on the *rbc*L results of Doyle et al. (1997) but containing only those taxa that overlapped with taxa from the temperate herbaceous tribes and selected outgroups present in other source trees was constructed manually and incorporated into the final matrix. Additional single characters (scored as '0's and '1's), such as ones corresponding to data consistent with and/or supporting the monophyly of Neo-Astragalus (Wojciechowski et al., 1999), *Trifolium* (Liston, unpubl.; Steele, unpubl.), the IR-lacking clade (Lavin et al., 1990; Liston, 1995), and *Lonchocarpus* as sister taxon to *Tephrosia* (Hu et al., 2000; Hu, 2000) were also incorporated into the matrix. Specific, well-supported topological constraints, were also included in some analyses (see below).

Maximum parsimony analyses of the final supermatrix (572 taxa, 722 elements) were performed with PAUP* 4.03b using heuristic search strategies only. Initially, searches were conducted using SIMPLE and CLOSEST addition sequences with TBR (tree-bisection-reconnection) branch swapping, MAXTREES set to 1000 or 5000 and MULTREES option in effect (this required 18 to 24 hr of processor time on a Macintosh G4 computer). Searches using RANDOM addition sequences (1000 or

TABLE 2. List of phylogenetic studies on taxa from "temperate herbaceous tribes" using molecular data. Number of taxa refers to number in each data set/tree used, or number pruned from study for use, in supertree construction. †Studies from which source tree or data are used in construction of supertree presented in this paper; *studies from which single character data (presence or absence of inverted repeat in cpDNA) derived. Data are nucleotide sequence data unless otherwise noted. Gene/regions: *matK*, *rbcL*, *rpoC*, *trnL*, intron = chloroplast DNA genes; ETS = nuclear ribosomal DNA external transcribed spacer; ITS = nuclear ribosomal DNA internal transcribed spacers; *PHY* = nuclear phytochrome genes. RFLP = restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism; RAPD = random amplified polymorphic DNA.

Study	Taxonomic Focus	No. of Taxa	Gene/region
Allan and Porter (2000)†	Loteae s. l.	42	ITS
Asmussen and Liston (1998)†	*Lathyrus*	44	cpDNA RFLP
Bena et al. (1998a, 1998b†, 1998c)	*Medicago*	53	ITS, ETS
Downie et al. (1998)†	*Medicago*	62	ITS
Doyle et al. (1996)	Fabaceae		cpDNA RFLP
Doyle et al. (1997)†	Fabaceae	10/84	*rbcL*
Fennel et al. (1998)†	*Vicia*	10	*trnL*
Gauthier et al. (1997)	*Lotus*		cpDNA RFLP
Hayashi et al. (1998)	*Glycyrrhiza*		*rbcL*
Hu et al. (2000)†	Millettieae s. l.	19/62	*matK*
Käss and Wink (1995, 1996, 1997)	Fabaceae, Papilionoideae		*rbcL*, ITS
M. Lavin (unpubl., Montana State Univ.)†	Robinieae	36	ITS
Lavin and Doyle (1991)	*Sphinctospermum*		cpDNA RFLP
Lavin and Marriot (1997)†	*Astragalus*	10	cpDNA RFLP
Lavin et al. (1991)	*Gliricidia*		cpDNA RFLP
Lavin et al. (1998)	Millettieae		*PHY*

TABLE 2 continued

Lavin et al. (1990)[†*]	Papilionoideae	28/95	cpDNA RFLP
Lavin (1995); Lavin and Sousa (1995)	Robinieae		cpDNA RFLP
Liston (1992)	*Astragalus*		*rpoC* RFLP
Liston (1995)[†*]	Papilionoideae	43/61	cpDNA RFLP
Liston and Wheeler (1994)[†]; A. Liston (unpubl.)[†]	Astragalean, IR-Lacking clades	70,60	*rpoC* RFLP
Liston, Steiner and Taylor (unpubl.)[†]	New World *Trifolium*	64	ITS
Mayer and Soltis (1994)	*Lens*		cpDNA RFLP
Sanderson and Doyle (1993)[†]	*Astragalus*	33	cpDNA RFLP
Sanderson and Liston (1995)	Galegeae		cpDNA RFLP, ITS
Sanderson and Wojciechowski (1996)[†]	Temperate Herbaceous Clade	41	ITS
Sharma et al. (1996)	*Lens*		AFLP, RAPD
K. P. Steele (unpubl.)[†]	Trifolieae, Vicieae, Cicereae	40	*matK*, ITS
Steiner et al. (1997)	New World *Trifolium*		ITS, RAPD
van Oss et al. (1997)	*Lens*		cpDNA RFLP
Wagstaff et al. (1999)[†]	Carmichaelieae	39	ITS
Watson et al. (in press)[†]	Old World *Trifolium*	65	cpDNA RFLP, ITS
Wojciechowski and Sanderson (1995)	IRLC		*trnL*
Wojciechowski et al. (1993)	*Astragalus*		ITS
Wojciechowski et al. (1999)[†]	*Astragalus*, Astragalean clade	140,34	ITS, *trnL*
M. Wojciechowski (unpubl.)[†]	Hologalegina	52	*matK*
Yamazaki et al. (1994)	*Glycyrrhiza*		RFLP, RAPD

5000 replications with one tree saved per replicate, MAXTREES set to 1000 or 5000, respectively, MULTREES option not invoked) with no branch-swapping were also performed; branch swapping (TBR) was then conducted on "best only" or "all" trees saved from a particular set of random additions. The resulting set(s) of most parsimonious trees derived from these analyses were summarised using consensus methods. Because of the nature of the data, the usual measures of clade support (i.e., Bremer's decay index or bootstrap proportions) have a different interpretation, or may be inappropriate in the context of estimating the robustness of nodes in the composite tree and so were not determined; values are typically low because of the small number of source trees or because of conflict among them (Bininda-Emonds et al., 1999).

Because MRP generates composite trees solely by combining the topologies of different source trees, there is no inherent consideration of either the overall support for the topology of any given source tree or of any differential support (decay index or bootstrap proportions) for individual nodes on a single source tree (Bininda-Emonds and Bryant, 1998). The use of these same quantitative measures of clade support as a means for incorporating relative support for individual nodes in the MRP analyses (i.e., by weighting matrix elements) requires that two conditions be met: 1) the chosen metric must be used for all source trees, and 2) the values of the chosen metric must provide a comparable measure of the relative support for a given nodes across studies, regardless of the characteristics of the original data and algorithm used to produce the source tree (Bininda-Emonds and Bryant, 1998). Since neither of these conditions were met in the present study, weighted analyses, that is with respect to individual matrix characters, were not conducted. Instead, topological constraints were included as a way of incorporating "differential support" for specific nodes that were consistently identified in a number of the source trees. All groups constrained to be monophyletic in our supertree analyses corresponded to the well-supported nodes (by bootstrap, parsimony jackknife analyses) present in the *mat*K analysis described here and identified in certain previous studies (i.e., IRLC, Robinieae + Loteae, "Hologalegina", see below).

Results

matK evidence

Maximum parsimony analyses employed heuristic searches (with TBR branch swapping) of complete sequences of the *mat*K gene for 52 papilionoid taxa (1620 characters, insertion-deletion positions excluded) identified six most parsimonious trees of length 1748, one of which is shown in Figure 1. The strict consensus tree is well resolved (not shown) and most relationships are well-supported by parsimony bootstrap analysis (Fig. 1) as well as by neighbor joining bootstrap (data not shown). Although generally consistent with the previous *rbc*L phylogeny of Doyle et al. (1997), our sampling for *mat*K from taxa of the temperate herbaceous tribes was much more extensive and the results provide better support overall. Several of the major clades, in particular the IRLC, the Robinieae + Loteae clade, and the clade comprised of these two ("Hologalegina"), were much more strongly supported (all > 95% by bootstrap analyses) by the *mat*K analyses compared to that of the *rbc*L data. These strongly supported clades formed the basis of the topological constraints imposed in the supertree reconstruction described above. Some resolution of clades within the IRLC is suggested by the *mat*K results, specifically the sister group relationship of the Hedysaroid clade plus *Caragana* Fabr. to the Astragalean clade, although this is not strongly supported by the present analysis. Interestingly, the *mat*K data place the pantropical genus *Sesbania* Adanson of Robinieae as the closest sister group to Loteae s.l. (Fig. 1), a relationship only recently suggested (albeit indirectly) by the *rbc*L results of Käss and Wink (1997).
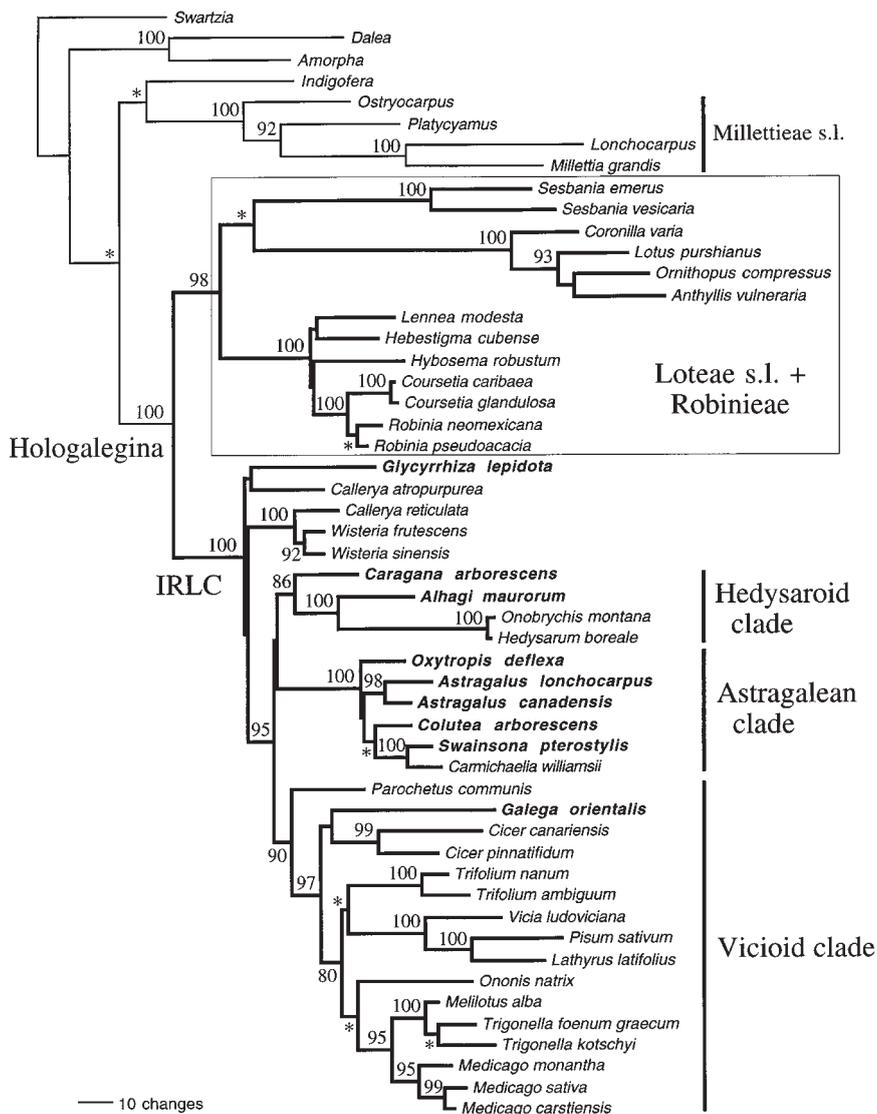
FIG. 1. Hypothesis of phylogenetic relationships within Hologalegina (indicated by thick lines), based on maximum parsimony analyses (heuristic) of complete chloroplast *mat*K gene sequences (1521 characters). Tree is one of six most parsimonious trees of length 1748 steps: CI = 0.604; HI = 0.396; RI = 0.751; RC = 0.454. *Swartzia* is designated as the outgroup. Parsimony bootstrap proportions shown only for those clades found in bootstrap majority-rule tree (100 replicates) and only for clades with values greater than 80% (clades with bootstrap values between 50% and 80% are indicated by asterisks). Members of tribe Galegeae are indicated in bold type. "Clades" are informal taxa recognized within Hologalegina; tribes Loteae s.l. and Robinieae are boxed.

Supertree

Heuristic searches of the final supermatrix using SIMPLE and CLOSEST addition sequences consistently produced sets of most parsimonious trees of 847-854 steps (shortest uncovered using any search strategy), depending on the topological constraints imposed. Searches using RANDOM additions, with no branch swapping invoked, typically converged on sets of most parsimonious trees that were 100–125 steps longer; when these resultant trees were retained in memory and then swapped (TBR) to completion, the length of most parsimonious trees uncovered never equaled that of the shortest trees identified in searches employing SIMPLE addition sequences. To convey a reasonable estimate of the phylogenetic information that is present in sets of most parsimonious trees uncovered during heuristic searches, we present a majority-rule consensus rather than a strict consensus of our supertree results (Fig. 2). This tree represents the 90% majority-rule consensus of 5000 equally most parsimonious trees of length 849 steps (MAXTREES set to 5000) generated from heuristic analysis of the supermatrix containing 572 taxa, with all members of Hologalegina topologically constrained, but otherwise unresolved, to form a single clade during searches. This tree, which contains 518 of 571 possible internal nodes – making it some 91% resolved compared to a strictly bifurcating tree, contains representatives of all nine tribes and 54 of the 67 genera from Hologalegina (Robinieae + Loteae s.l., 80 taxa; IRLC, 481 taxa; Table 1), in addition to ten outgroup taxa from the Millettieae–Phaseoleae complex, Indigofereae, Amorpheae, and Swartzieae DC. For comparison, the strict consensus of this same set of 5000 most parsimonious trees contained 321 of 571 possible internal nodes resolved (56%). The degree of taxonomic overlap among the phylogenetic studies used here in supertree construction, can be depicted by a tree-graph, such as that shown in Fig. 3, in which each node represents a tree from one of the studies with internodes connecting two trees that share two or more taxa.

The sheer size of the full composite tree, complete with names of all taxa, that is represented by Fig. 2 precludes its reproduction as a single figure in this volume; thus we provide an electronic copy of the same composite tree that is available for viewing on the first author's website, at the URL "**http://loco.ucdavis.edu/wojo/HGsupertree.html**". A hard copy of the full-size composite tree showing all taxon names can be obtained from the first author.

The inclusion and/or placement of several taxa in the composite tree are problematic and deserve comment. In the Doyle et al. (1997) study of *rbc*L sequences from 84 legume taxa, *Bolusanthus* Harms (Sophoreae) was sister to Loteae s.l., and this taxon was included as such in the supertree analysis here. However, this relationship is not substantiated by phylogenetic analyses of *mat*K sequences (Hu et al., 2000; Wojciechowski et al., unpubl.), *trn*L intron (Pennington et al., 2000), or nuclear *PHY* gene sequences (Lavin et al., 1998), which consistently find *Bolusanthus* nested within other Sophoreae (*Ormosia* Jackson, *Acosmium* Schott), close to *Poecilanthe* Benth. (Millettieae) and Brongniartieae (Benth.) Hutch., among basal papilionoid groups. *Coronilla varia* L., *Colutea arborescens* L., and *Lotus purshianus* (Benth.) Clements are examples of taxa we will term here *incertae sedis cladus*; that is, in the context of their position(s) in the supertree. For these taxa, their placements near the base of clades including in part other species of the same genus does not necessarily imply paraphyly of the genus and is (most likely) the consequence of matrix representation. MRP, like other consensus methods, places taxa that are not well resolved or well represented in source trees (in these cases taxa that were present in only a single source tree) at the base of the least inclusive clade common to or containing those taxa and their sister group(s) found in the original source tree(s).
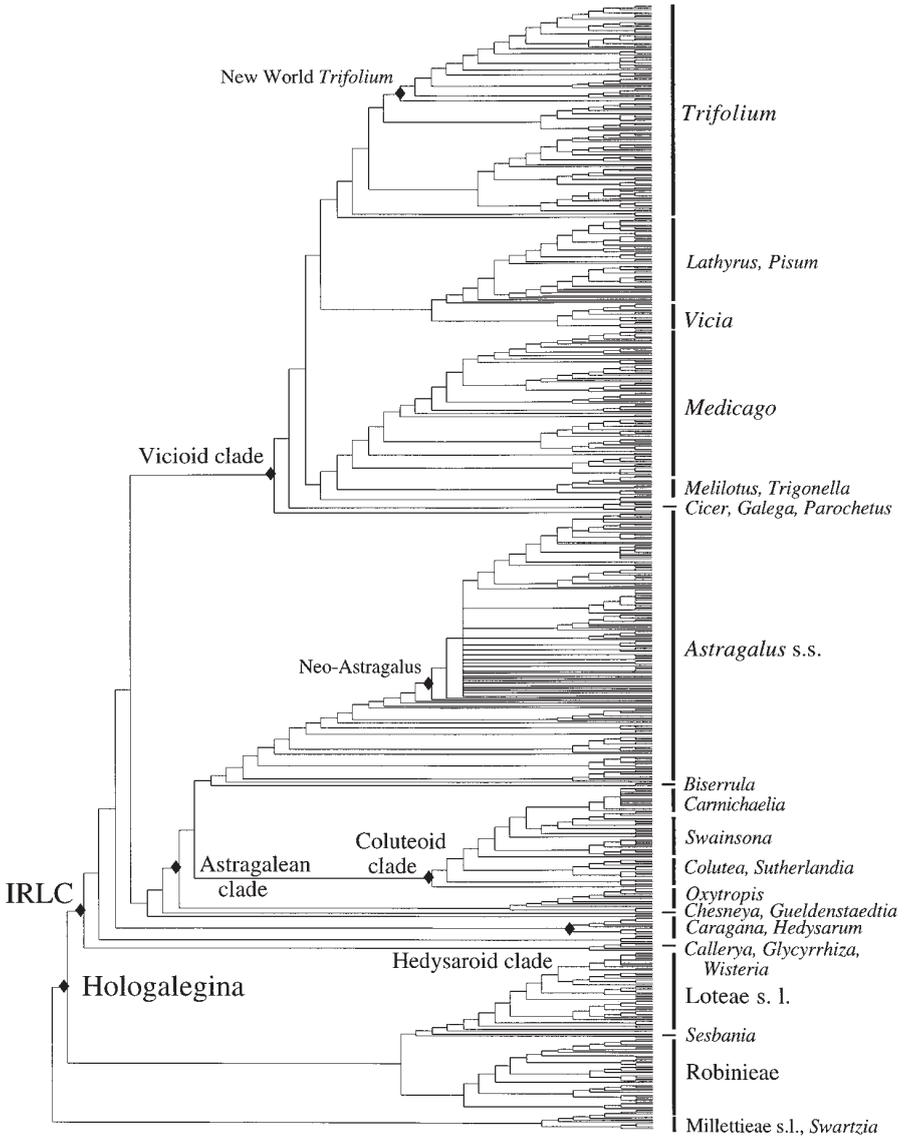
FIG. 2. The composite "supertree" of Hologalegina, based upon maximum parsimony analyses (heuristic search options) of matrix containing 571 legume taxa and 722 'elements', with monophyly of Hologalegina topologically constrained. Hypothetical 'outgroup' for analyses omitted from figure. Tree is 90% majority-rule consensus tree of 5000 most parsimonious trees of length 849 steps: CI = 0.850; HI = 0.150; RI = 0.973; RC = 0.828. Basal nodes for major clades described in text marked by diamonds (♦).

Discussion

In this report we present a molecular phylogeny of the group comprised of Polhill's (1981d, 1994) temperate herbaceous tribes plus Robinieae, a clade generally resolved, but not always well-supported, by several recent higher level studies (Doyle et al., 1997; Käss and Wink, 1997; Lavin et al., 1998; Hu et al., 2000). Given that the clade we designate here as Hologalegina was found supported in a number of previous studies, including the *mat*K analyses presented here, it was topologically constrained to be monophyletic in the supertree analyses. It is important to emphasise that the composite tree (Fig. 2) is merely a most parsimonious synthesis resulting from an analysis of a number of disparate sources of molecular data; therefore, the original studies (Table 2) should be consulted for any and all direct evidence supporting or refuting particular relationships presented on this tree. We highlight some of the more interesting or unexpected results to emerge from this analysis.

Paraphyly of Galegeae

The currently proposed circumscription of Galegeae (Polhill, 1994) represents a minor revision of the last major treatment by Polhill (1981b), by inclusion of a few newly recognised genera (*Spongiocarpella* Yakovl. et Ulzij. and *Neodielsia* Harms not sampled here; *Astragalus* L. segregate *Astracantha* Podlech sampled here) and recognition of a separate subtribe containing *Alhagi* Gagnebin. Attempts to circumscribe the members of this morphologically diverse, cosmopolitan tribe have been hampered by the view that Galegeae form the transition between the pulvinate and epulvinate condition, or the transition between the woody and herbaceous habit. The molecular data obtained in the last few years, however, are unequivocal in demonstrating the paraphyly of Galegeae. Indeed, the least inclusive clade that contains Galegeae sensu Polhill comprises all members of that tribe plus all members of the tribes Carmichaelieae, Cicereae, Hedysareae, Trifolieae, and Vicieae, as well as the genera *Callerya* (*Millettia* segregate, sensu Geesink, 1984; Schot, 1994), *Wisteria* (including *Wisteria japonica* Siebold & Zucc., a species that is generally placed in *Millettia* but considered part of *Callerya* by Geesink (1984)), and most likely *Afgekia* (Hu, 2000; see below) of the tropical tribe Millettieae. This clade includes some 45 genera and more than 4,000 species, roughly one-third of all papilionoids. Although there exists some uncertainty in branching order at its base (i.e., among species of *Callerya*, *Glycyrrhiza*, and *Wisteria*), this clade corresponds precisely to the monophyletic group marked by the loss of one copy of the inverted repeat in chloroplast DNA (the IRLC). Note that the IRLC is equivalent to what we have previously referred to as the "temperate herbaceous clade", or THC (Sanderson and Wojciechowski, 1996), plus the genera *Callerya*, *Wisteria*, and relatives such as *Afgekia*. Traditional members of Galegeae occupy basally-branching positions in each of the three major clades within the IRLC, the Astragalean (where genera of Galegeae dominate), Hedysaroid, and Vicioid clades, as well as in the IRLC itself. This affirms the long-held view of Galegeae as the foundation from which the epulvinate series arose (Polhill, 1981d; p. 202). However, contrary to Polhill's view (1981d; p. 202), Galegeae is nested within Hologalegina, which is the sister group to the Millettieae–Phaseoleae complex plus Indigofereae, and not a "temperate offshoot" of Millettieae.

IRLC

At the time of Lavin et al.'s (1990) survey of the distribution of the inverted repeat mutation in legume chloroplast DNAs, Polhill's hypothesis of the higher-level systematics of the Papilionoideae (Polhill, 1981d; Fig. 3), was the most comprehensive to date. The results from this work reinforced the belief that Millettieae held a pivotal position in the radiation of a large portion of papilionoids, and thus the phylogenetic
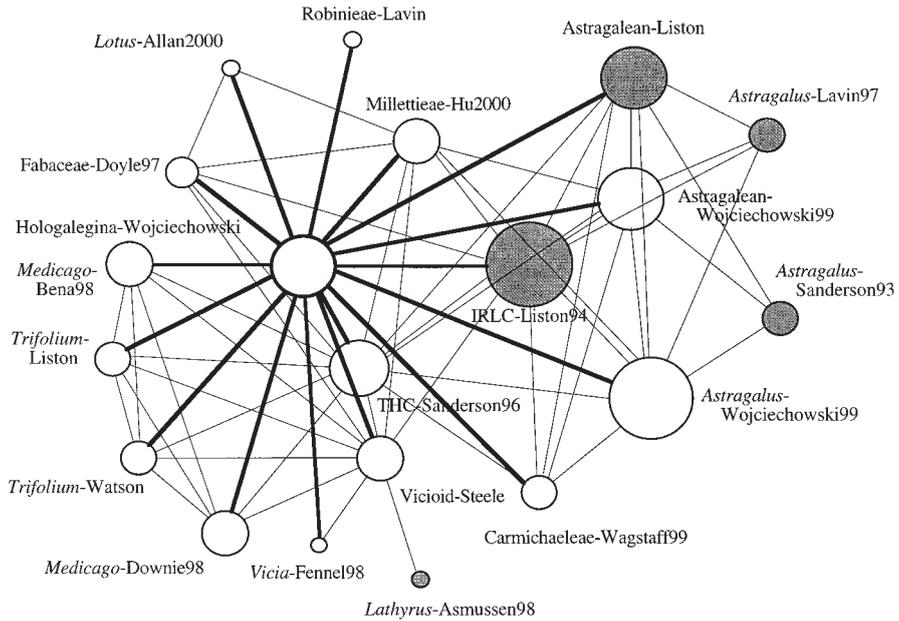
FIG. 3. 'Tree-graph' (sensu Sanderson et al., 1998) of the phylogenies in Hologalegina supertree (Fig. 2). Each node represents a phylogenetic tree taken from 20 of the 22 studies (single character datasets of Lavin et al. (1990) and Liston (1995) not included in tree-graph) used in supertree construction (Table 2; taxon-author and year abbreviated), with internodes connecting trees that share taxa in common. The size of each node is proportional to total number of taxa that study shares with all other studies. Shaded nodes signify RFLP-based studies, clear nodes indicate nucleotide sequence-based studies. Studies connecting to *mat*K phylogeny of Hologalegina (Fig. 1; Wojciechowski, unpubl.) indicated by thick lines.

resolution of this tribe was crucial to an understanding of relationships at this level. Liston (1995) demonstrated that other Millettieae, including species of *Wisteria* and the tropical *Callerya* shared this mutation, thereby suggesting that at least part of Millettieae was the sister group to most of the temperate herbaceous legumes. While the existence of the IRLC was not evident from cladistic analyses of non-molecular characters (e.g., see Chappill, 1995), analyses of molecular data, including the chloroplast *rbc*L (Doyle et al., 1997; Käss and Wink, 1997), *mat*K (Hu, 2000; Hu et al., 2000; Wojciechowski, unpubl.), and nuclear phytochrome genes (Lavin et al., 1998), have conclusively shown that *Callerya* and *Wisteria* are more closely related to members of Galegeae (i.e., are part of IRLC) than they are to those of Millettieae s.l. Moreover, other members of Millettieae, considered morphologically related to *Callerya* and *Wisteria*, may also belong to the IRLC: these form a small, predominantly southeast Asian group that includes *Afgekia*, *Antheroporum* Gagnep., *Endosamara* Geesink and *Sarcodum* Lour. (Lavin et al., 1998; B. Shrire, Royal Botanic Gardens, Kew, pers. communication). Work is now in progress to verify loss of the inverted repeat in these taxa (Wojciechowski, unpubl.). Lastly, these same studies have demonstrated unequivocally that members of Robinieae and Loteae s.l. together form the sister group to the IRLC.

In all analyses emphasising broad taxonomic sampling across papilionoids (Sanderson and Wojciechowski, 1996; Doyle et. al., 1997; Lavin et al., 1998; Hu et al., 2000; Wojciechowski et al., unpubl.) the IRLC is consistently resolved and often strongly supported (by bootstrap or parsimony jackknife) as a monophyletic group. Within the IRLC three large clades, containing about 97% of the species diversity, the Astragalean, Hedysaroid, and Vicioid clades, are also well-supported. The relationships among these clades and the remaining genera in Galegeae, *Glycyrrhiza, Caragana, Calophaca* Fisch., *Halimodendron* Fisch. ex DC., *Gueldenstaedtia* Fisch., and *Chesneya* Lindl. ex Endl. are only partly resolved (e.g., Sanderson and Wojciechowski, 1996). In the *mat*K analysis presented here, relationships among these three large clades are poorly resolved (Fig. 1), but suggest that the Astragalean clade is sister to the clade comprising *Caragana* and its allies and the Hedysaroid clade. Together they are sister to the Vicioid clade. However, in the supertree analysis (Fig. 2), *Caragana* and the Hedysaroid clade form the sister group to the Astragalean and Vicioid clades. In both analyses, *Callerya, Glycyrrhiza*, and *Wisteria* are strongly supported as the basal most lineages of the IRLC, sister taxa to the rest of the clade, although relationships among them are not yet clearly resolved.

Astragalean clade

The largest of the three large clades in the IRLC is the "Astragalean" clade. Details of the composition and phylogenetic relationships across the Astragalean clade based on analyses of nrDNA ITS sequence, chloroplast RFLP and *trn*L intron sequence data have been discussed previously (Sanderson and Liston, 1995; Sanderson and Wojciechowski, 1996; Wojciechowski et al., 1999). Briefly, the Astragalean clade comprises more than 3,000 species in four smaller, well-supported clades. These include (1) a clade comprising the vast majority of species in *Astragalus (Astragalus* s.s., Fig. 2); (2) the "Coluteoid clade", comprising all of Galegeae subtribe Coluteinae sensu Polhill (1981b), at least some "outlier" members of *Astragalus*, and the tribe Carmichaelieae; (3) a monophyletic *Oxytropis* DC.; and (4) the species pair *Biserrula pelecinus* L. and *Astragalus epiglottis* L. The species *Astragalus vogelii* (Webb.) Bronm. may represent a fifth clade (not indicated on Fig. 2). The relationships among these five taxa (four clades and *A. vogelii*), however, are largely unresolved in these studies.

The results of *mat*K sequence analysis (Fig. 1) are consistent with these earlier studies although fewer taxa have been sampled. The supertree analysis begins to provide some resolution of taxa within the Astragalean clade (Fig. 2). *Astragalus* s. s., with the Mediterranean species *B. pelecinus* and *A. epiglottis* as sister group, is the sister group to the clade comprising the large, circumboreal genus *Oxytropis* and the Asian–Australasian Coluteoid clade. Contrary to commonly held ideas dating back to the early 19th century (discussed in Wojciechowski et al., 1999), both the *mat*K and supertree results indicate that *Oxytropis* is not the closest sister taxon to *Astragalus*, nor is it nested within it; instead *Oxytropis* is the sister group to most of the rest of the Astragalean clade. The small Asian genera *Chesneya* and *Gueldenstaedtia*, with the enigmatic *Astragalus lusitanicus* Lam., form the sister lineage to the Astragalean clade, relationships first suggested by nrDNA ITS analyses (Sanderson and Wojciechowski, 1996) and now supported by the *rpo*C RFLP data (Liston, unpubl.). Originally described as *Phaca baetica* L., the taxon *A. lusitanicus* also shares some morphological characters with *Chesneya* and *Gueldenstaedtia*, consistent with its placement here (Fig. 2). But, while its inclusion in the genus *Astragalus* no longer appears justified, *Astragalus lusitanicus* is certainly not a member of tribe Sophoreae as proposed by Podlech (1993; as *Erophaca baetica* (L.) Boiss.) since both the *rpo*C results (Liston, unpubl.) and the lack of an inverted repeat in the chloroplast genome (Liston, 1995) clearly place this taxon within the IRLC.

Coluteoid Clade – Carmichaelieae

Nested within the Astragalean clade is a moderately well-supported clade ("Coluteoid") comprising all of Galegeae subtribe Coluteinae and the endemic New Zealand tribe Carmichaelieae (Sanderson and Wojciechowski, 1996). Carmichaelieae (Polhill, 1981a) comprises some 45 species of trees, shrubs and lianas in four extant genera, *Carmichaelia* R. Br., *Chordospartium* Cheesm., *Corallospartium* J. B. Armst., and *Notospartium* Hook., and the extinct *Streblorrhiza* Endl. The tribe has been considered related to Galegeae through the large Australian genus *Swainsona* Salisb. and small New Zealand genus *Clianthus* Sol. ex Lindl. (Polhill, 1981a). Recent phylogenetic studies of morphological and anatomical characters by Heenan (1998a, 1998c) have led to a revised and expanded circumscription of *Carmichaelia* (to include *Chordospartium*, *Corallospartium*, *Notospartium*), concomitant with a reduction in the number of recognised species. Heenan (1998b) also proposed segregation of the sole New Zealand representative of *Swainsona*, *Swainsona novae-zelandiae* Hook. f., to a new monotypic genus *Montigena* Heenan. Subsequent analyses of nrDNA ITS sequences provide support for a monophyletic Australian–New Zealand "Carmichaelinae" clade that now includes all members of *Carmichaelia*, *Clianthus*, *Montigena*, and *Swainsona* (Wagstaff et al., 1999), derived from the predominately northern temperate members of the Coluteoid clade (Sanderson and Wojciechowski, 1996). The results presented here (Fig. 2) are consistent with those of Wagstaff et al. (1999) in showing the New Zealand genera *Carmichaelia*, *Clianthus*, and *Montigena* are nested within the Australian *Swainsona*, but there remains some uncertainty as to placement of the species *Carmichaelia ramosa* and *Carmichaelia carmichaeliae* (Hook. f.) Heenan, and thus the monophyly of *Carmichaelia* as reported by these authors.

Vicioid Clade

A clade containing members of tribes Cicereae, Trifolieae, and Vicieae, with the genus *Galega* L. as sister group, was first suggested by *rpo*C analysis (Liston and Wheeler, 1994) and received further support from nrDNA ITS analysis (Sanderson and Wojciechowski, 1996). The *mat*K and supertree results presented here extend and modify these early observations. The monotypic *Parochetus* Buch.-Ham. ex D. Don whose placement in Trifolieae has been questioned (Small, 1987) is shown in the *mat*K analysis in (Fig. 1) to be sister to all remaining members of the Vicioid clade; thus its continued inclusion in Trifolieae renders the tribe paraphyletic. *Galega* and Cicereae are sister to the remaining Vicioid clade. The genus *Ononis* L., whose placement in Trifolieae also has been questioned (Small, 1987), is sister to a clade corresponding to Trifolieae subtribe Trigonellinae (Schulz) E. Small. Although there is low bootstrap support for this relationship, there is 80% bootstrap support for the inclusion of *Ononis* within a group formed by Trifolieae (minus *Parochetus*) and Vicieae.

The monophyly of *Trifolium* L. is strongly supported in the *mat*K analysis (Fig. 1; Steele, unpubl.) and is apparent in the supertree (Fig. 2), which incorporates nrDNA ITS results from many Old World (Watson et al., in press) and New World (Liston et al., unpubl.) species. The New World species of *Trifolium* are strongly supported as monophyletic and derived from Old World taxa (Steiner et al., 1997; Liston et al., unpubl.), but relationships among the species are not well-resolved. This parallels the situation observed in *Astragalus* (Wojciechowski et al., 1999), which like *Trifolium* has a large number of morphogically diverse species in western North America. The monophyly of New World *Trifolium* has not been previously hypothesised, while the monophyly of the vast majority of New World *Astragalus* (aneuploid "Neo-Astragalus") was first proposed on the basis of extensive cytological data (reviewed in Wojciechowski et al., 1999). The relationship of *Trifolium* to other genera in Trifolieae (*Ononis*, *Medicago* L., etc.) is not certain; *Trifolium* appears sister to the Vicieae, although bootstrap support for that relationship is low. There is also strong support

for the monophyly of subtribe Trigonellinae, which consists of *Medicago*, *Trigonella* L., and *Melilotus* Mill. The subtribe is united by the presence of stipules that are adnate to the petiole, as compared to sheathing stipules in *Trifolium*, or free stipules as in *Parochetus* and Vicieae (Kupicha, 1981; Small, 1987). Note that *Melilotus* is sister to (Fig. 1) or possibly nested within *Trigonella* (Fig. 2). Expanded taxon sampling of Vicieae and Trifolieae for *mat*K and nrDNA ITS sequences (Steele and Wojciechowski, unpubl.) will undoubtedly resolve generic level relationships in this agriculturally important group.

Caragana – Hedysaroid Clade

The Hedysaroid clade was one of three large (but not yet well sampled), strongly supported clades to emerge from analyses of the temperate herbaceous clade based on nrDNA ITS data (Sanderson and Wojciechowski, 1996). This clade reunites the small, south central Asian genus *Alhagi* of Galegeae with the primarily Eurasian tribe Hedysareae, where it has been traditionally allied on the basis of similar jointed, indehiscent fruits (Hutchinson, 1964). In the analyses presented here (Fig. 1, 2), *Caragana* (along with its shrubby Eurasian allies *Calophaca* and *Halimodendron* of Galegeae subtribe Astragalinae; Polhill, 1981b) form a monophyletic group that is sister to the Hedysaroid clade. The clade is centered in Eurasia and to a lesser extent in North America. This relationship is consistent with Polhill's (1981c) view of Hedysareae as forming a single complex dominated by *Hedysarum* L. and *Onobrychis* Mill., plus a few minor segregates, and derived from the "astragaloid part" of Galegeae. The *mat*K results suggest a sister group relationship of this clade to the Astragalean clade, while the supertree analysis (Fig. 2) positions this clade outside both the Astragalean and Vicioid clades. Greater sampling of *Caragana* and its relatives, as well as within *Hedysarum* and *Onobrychis* is needed to clearly resolve relationships within this clade and as well as within Hedysareae itself, which have always been difficult to circumscribe (Polhill, 1981c).

Loteae (s.l., including Coronilleae)

Generic circumscriptions in Loteae and Coronilleae are currently in a state of flux and undergoing considerable revision (Lassen, 1989; Sokoloff, 1998; Kirkbride, 1999; Allan and Porter, in press). Likewise, the higher level relationships of the tribes have remained problematic. Loteae and Coronilleae traditionally have been regarded as derived elements of the temperate tribes characterised by the loss of the chloroplast DNA inverted repeat (especially Galegeae, and to a lesser extent Hedysareae; Polhill, 1981e). In contrast, possession of a similar determinate root nodule morphology and the apparent homology of their inflorescence with the pseudoraceme of Millettieae, Phaseoleae and Psoraleeae suggests a close relationship of Loteae and Coronilleae to these latter tropical tribes (Lavin et al., 1990). An analysis of *rbc*L data (Doyle et al., 1997) was the first study to suggest a close (sister group) relationship of Loteae and Coronilleae to *Robinia* (Robinieae), and *Bolusanthus* (Sophoreae), although this was not strongly supported. While the relationship with Robinieae is now well supported by both *mat*K data (Hu et al., 2000; Wojciechowski, unpubl.) and nrDNA ITS data (Hu, unpubl.), the latter relationship receives no support from other molecular data (Hu et al., 2000; Lavin, unpubl.). That Loteae s.l. is predominantly Mediterranean and African in distribution and is related to Robinieae via an apparent sister group relationship with *Sesbania* (Fig. 1) is consistent with the "Boreotropical hypothesis". This hypothesis posits a northern tropical origin for and explains the phylogenetic patterns of Robinieae (Lavin, 1995; Lavin and Sousa, 1995) and thus Loteae as well. However, further sampling is necessary to identify more clearly relationships within Loteae s.l. (e.g., Allan and Porter, in press), the basal relationships of this clade (i.e, monophyly of *Sesbania* and Loteae s.l.) as well as that of the remaining Robinieae. Regardless, the origin and early diversification of the Loteae–Robinieae clade is

distinct from that of its sister group, the IRLC, which apparently originated in eastern Asia and later migrated to the New World, diversifing almost exclusively in temperate regions.

Robinieae

Circumscriptions of the tribe Robinieae have undergone fundamental change in recent years. As currently circumscribed (Lavin and Sousa, 1995), the tribe includes 12 genera primarily centered in the Neotropics and warm temperate North America; only *Sesbania* is pantropical but with a diversity of continental endemics in North America and Africa. The taxonomy of the genera of Robinieae has been variously anchored to Galegeae since Rydberg's treatments of the 1920s, mostly as elements in subtribe Robiniinae or aligned in different subtribes (Lavin and Sousa, 1995). Polhill and Sousa (1981) completely realigned the circumscription of Robinieae at the tribal level around Robiniinae, in the process emphasising a close relationship of Robinieae to tropical tribes such as Aeschynomeneae and Millettieae centered in the New and Old Worlds, respectively. Following the circumscription of Polhill and Sousa (1981), Lavin and Sousa (1995) presented the first comprehensive taxonomy of the tribe based on phylogenetic relationships of all constituent genera, but they concluded the closest relatives of Robinieae were more likely to be found among Old World genera of tribes Millettieae (*Millettia, Wisteria*) and Galegeae (*Caragana*), rather than with the New World tropical tribes Aeschynomeneae or Dalbergieae. From the *mat*K analyses (Fig. 1; Hu et al., 2000; Wojciechowski, unpubl.), supertree analysis (Fig. 2), and nuclear *PHY* gene results (Lavin et al., 1998), is it now certain that the closest relatives of Robinieae are indeed the temperate herbaceous tribes, specifically Loteae s.l., and that the Loteae–Robinieae clade is sister to the IR-lacking clade. Relationships within the more restricted Robinieae (i.e., excluding *Sesbania*) placing *Lennea* Klotzsch and *Hebestigma* Urban together as sister taxa to a clade containing two groups, one comprised of the genera *Hybosema* Harms, *Poitea* Ventenat, and *Gliricidia* H. B. K., and a second containing *Coursetia* DC., *Peteria* A. Gray, *Sphinctospermum* Rose, *Genistidium* I. M. Johnston, *Olneya* A. Gray, and *Robinia* L., are now supported by chloroplast RFLP (Lavin and Sousa, 1995) and *mat*K sequences (Wojciechowski, unpubl.), nuclear *PHY* gene (Lavin et al., 1998) and nrDNA ITS sequences (Lavin, unpubl.).

Hologalegina

Understanding the phylogenetic relationships among the mainly temperate herbaceous tribes and their tropical relatives is important in elucidating the evolutionary history and higher-level systematics of the Papilionoideae. Until the advent of molecular data, many of these relationships went largely undetected. The monophyly of the clade we refer to as "Hologalegina" is now substantiated by evidence from both chloroplast *mat*K (Hu et al., 2000; Wojciechowski, unpubl.) and *rbc*L (Doyle et al., 1997) sequences, as well as from nuclear rDNA ITS (Hu, 2000) and *PHY* gene (Lavin et al., 1998) sequences; hence, our rationale for constraining its monophyly during the supertree analysis. The clade comprises a very large proportion of the subfamily with temperate distribution, including nine traditionally recognised tribes in their entirety and at least 67 genera. Remarkably, more than 4,600 species, or about $\frac{1}{4}$ of the species diversity of the entire family Fabaceae, are included in this clade. The recognition of this clade resolves long standing uncertainties surrounding the higher level relationships of Loteae (s.l.) and Robinieae, as well as those of the closest tropical relatives to the temperate tribes which have lost the chloroplast DNA inverted repeat. Hologalegina is the sister group to the clade comprising genera of Millettieae, Indigofereae, and other, mainly Old World tropical tribes *rather* than being derived from it as has been traditionally held. This finding also stands in contrast to the views of Geesink (1984) who held that Millettieae was a paraphyletic stem group from which many monophyletic offshoots, including the predominantly temperate tribes centered around Galegeae, arose.

The estimate of Hologalegina phylogeny presented in this paper is intended to be a working hypothesis. Additional data should be incorporated into future analyses of this type to increase taxonomic breadth and extent of representation. This is the first attempt to bring together the vast amount of molecular evidence which has accumulated on the systematics of the "higher" papilionoids and summarise it in a phylogenetic context. The phylogenetic resolution provided by analyses of multiple molecular datasets will certainly guide future taxon sampling for both molecular and non-molecular data in the effort to resolve the remaining uncertain relationships in this large group of legumes. Comprehensive, robust phylogenies of these legumes are essential to a more complete understanding of a variety of important evolutionary issues such as the co-evolution of legumes and their nitrogen-fixing bacteria, as well the origin and diversification of the major groups of this family.

Note. Recently, J. Reveal (1997; pers. communication) has noted several suprageneric names currently in widespread use in Fabaceae (e.g., Polhill, 1994) appear to be invalid as prescribed by the International Code of Botanical Nomenclature (ICBN), two of which have relevance here. According to Reveal, Astragaleae has priority over Galegeae, and Fabeae must be used instead of Vicieae. Considering the potential nomenclatural conservation of the names Galegeae and Vicieae, we prefer to maintain the traditional usage in this paper.

Acknowledgments

References

Allan, G. and Porter, J.M. (In press). 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*.

Asmussen, C.B. and Liston, A. (1998). Chloroplast DNA characters, phylogeny, and classification of *Lathyrus* (Fabaceae). *American Journal of Botany* 85: 387–401.

Baum, B.R. (1992). Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* 41: 3–10.

Bena, G., Jubier, M.F., Olivieri, I., and Lejeune, B. (1998a). Ribosomal external and internal transcribed spacers: Combined use in the phylogenetic analysis of *Medicago* (Leguminosae). *Journal of Molecular Evolution* 46: 299–306.

Bena, G., Lejeune, B., Prosperi, J.-M., and Olivieri, I. (1998b). Molecular phylogenetic approach for studying life-history evolution: the ambiguous example of the genus *Medicago* L. *Proceedings of the Royal Society of London, series B* 265: 1141–1151.

Bena, G., Prosperi, J.-M., Lejeune, B., and Olivieri, I. (1998c). Evolution of annual species of the genus *Medicago*: a molecular phylogenetic approach. *Molecular Phylogenetics and Evolution* 9: 552–559.

Bininda-Emonds, O.R.P. and Bryant, H.N. (1998). Properties of matrix representation with parsimony analyses. *Systematic Biology* 47: 497–508.

Bininda-Emonds, O.R.P., Gittleman, J.L., and Purvis, A. (1999). Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews* 74: 143–175.

Bras, C.P., Jorda, M.A., Wijfjes, A.H.M., Harteveld, M., Stuurman, N., Thomas-Oates, J.E., and Spaink, H.P. (2000). A *Lotus japonicus* nodulation system based on heterologous expression of the fucosyl transferase NodZ and the acetyl transferase NolL in *Rhizobium leguminosarum*. *Molecular Plant-Microbe Interactions* 13: 475–479.

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.

Cook, D.R. (1999). *Medicago truncatula* – a model in the making! *Current Opinion in Plant Biology* 2: 301–304.

Dormer, K.J. (1945). An investigation of the taxonomic value of shoot structure in angiosperms with especial reference to Leguminosae. *Annals of Botany (London)* 9: 141–153.

Dormer, K.J. (1946). Vegetative morphology as a guide to the classification of Papilionate. *New Phytologist* 45: 145–161.

Downie, S.R., Katz-Downie, D.S., Rogers, E.J., Zujweski, 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 *rbc*L 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 *rbc*L 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. (1996). The distribution and phylogenetic significance of a 50-kb chloroplast DNA inversion in the flowering plant family Leguminosae. *Molecular Phylogenetics and Evolution* 5: 429–438.

Fennel, S.R., Powell, W., Wright, F., Ramsay, G. and Waugh, R. (1998). Phylogenetic relationships between *Vicia faba* (Fabaceae) and related species inferred from chloroplast *trn*L sequences. *Plant Systematics and Evolution* 212: 247–259.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A., and Perret, X. (1997). Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* 387: 394–401.

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. (1984). Scala Millettiearum. E. J. Brill/Leiden University Press, Leiden.

Hayashi, H, Hosono, N., Kondo, M., Hiraoka, N., and Ikeshiro, Y. (1998). Phylogenetic relationship of *Glycyrrhiza* plants based on *rbc*L sequences. *Biological and Pharmaceutical Bulletin* 21: 782–783.

Heenan, P.B. (1998a). Phylogenetic analysis of the *Carmichaelia* complex, *Clianthus*, and *Swainsona* (Fabaceae), from Australia and New Zealand. *New Zealand Journal of Botany* 36: 21–40.

Heenan, P.B. (1998b). *Montigena* (Fabaceae), a new genus endemic to New Zealand. *New Zealand Journal of Botany* 36: 41–51.

Heenan, P.B. (1998c). An emended circumscription of *Carmichaelia*, with new combinations, a key, and notes on hybrids. *New Zealand Journal of Botany* 36: 53–63.

Hu, J.-M. (2000). The 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. 299–320. 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 *trn*K/*mat*K sequences, and implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418–430.

Hutchinson, J. (1964). The genera of flowering plants. Oxford University Press, London.

Käss, E. and Wink, M. (1995). Molecular phylogeny of the Papilionoideae (Family Leguminosae): *rbc*L 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 *rbc*L sequences. *Biochemical Systematics and Ecology* 24: 365–378.

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

Kirkbride, J.H. (1999). Lotus systematics and distribution. In P. R. Beuselink (editor). Special publication No. 28, pp. 1–20. American Society of Agronomy and Crop Science Society of America, Madison, WI.

Kupicha, F.K. (1981). Vicieae. In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 377–381. Royal Botanic Gardens, Kew.

Lassen, P. 1989. A new delimitation of the genera *Coronilla*, *Hippocrepis*, and *Securigera* (Fabaceae). *Willldenowia* 19: 49–62.

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. and Doyle, J.J. (1991). Tribal relationships of *Sphinctospermum* (Leguminosae): Integration of traditional and chloroplast DNA data. *Systematic Botany* 16:162–172.

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. and Marriott, H. (1997). *Astragalus molybdenus* s.l. (Leguminosae): higher taxonomic relationships and identity of constituent species. *Systematic Botany* 22: 199–217.

Lavin, M., Mathews, S., and Hughes, C. (1991). Chloroplast DNA variation in *Gliricidia sepium* (Leguminosae): Intraspecific phylogeny and tokogeny. *American Journal of Botany* 78:1576–1585.

Lavin, M. and Sousa S., M. (1995). Phylogenetic systematics and biogeography of the tribe Robinieae. *Systematic Botany Monographs* 45: 1–165.

Liston, A. (1992). Variation in the chloroplast genes *rpo*C1 and *rpo*C2 of the genus *Astragalus* (Fabaceae): evidence from restriction site mapping of a PCR-amplified fragment. *American Journal of Botany* 79: 953–961.

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 *rpo*C1 and *rpo*C2. *Biochemical Systematics and Ecology* 22: 377–388.

Maddison, W.P, and Maddison, D.R. (1992). MacClade. Analysis of phylogeny and character evolution, version 3. Sinauer Associates, Sunderland, MA.

Mayer, M.S. and Soltis, P.S. (1994). Chloroplast DNA phylogeny of *Lens* (Leguminosae) – origin and diversity of the cultivated lentil. *Theoretical and Applied Genetics* 87: 773–781.

Palmer, J.D., Osorio, B., Aldrich, J., and Thompson, W.F. (1987). Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. *Current Genetics* 11: 275–286.

Pennington, T., Lavin, M., Ireland, H., Klitgaard, B., Preston, J., and Hu, J.-M. (2000). Phylogenetic relationships of basal papilionoid legumes based upon sequence of the chloroplast intron *trn*L. In: P. S. Herendeen and A. Bruneau (editors). Advances in legume systematics, part 9, pp. 233–248. Royal Botanic Gardens, Kew.

Podlech, D. (1993). Beiträge zur Kenntnis der Gattung *Astragalus* L. (Leguminosae) III–VI. *Sendtnera* 1: 267–272.

Polhill, R.M. (1981a). Carmichaelieae. In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 364–366. Royal Botanic Gardens, Kew.

Polhill, R.M. (1981b). Galegeae. In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 357–363. Royal Botanic Gardens, Kew.

Polhill, R.M. (1981c). Hedysareae. In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 367–370. Royal Botanic Gardens, Kew.

Polhill, R.M. (1981d). 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. (1981e). Loteae (and Coronilleae). In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 371–375. 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. and Sousa, M. (1981). Robinieae. In: R. M. Polhill and P. H. Raven (editors). Advances in legume systematics, part 1, pp. 283–288. Royal Botanic Gardens, Kew.

Purvis, A. (1995). A composite estimate of primate phylogeny. *Philosophical Transactions of the Royal Society of London, Series B* 348: 405–421.

Ragan, M.A. (1992). Phylogenetic inference based on matrix representation of trees. *Molecular Phylogenetics and Evolution* 1: 53–58.

Reveal, J.L. (1997). Suprageneric names in Fabaceae published prior to 1870. *The Bean Bag* 45.

Sanderson, M.J. and Doyle, J.J. (1993). Phylogenetic relationships in North American *Astragalus* (Fabaceae) based on chloroplast DNA restriction site variation. *Systematic Botany* 18: 395–408.

Sanderson, M.J. and Liston, A. (1995). Molecular phylogenetic systematics of Galegeae, with special reference to *Astragalus*. In: M. D. Crisp and J. J. Doyle (editors). Advances in legume systematics, part 7: Phylogeny, pp. 331–350. Royal Botanic Gardens, Kew.

Sanderson, M.J., Purvis, A., and Henze, C. (1998). Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology and Evolution* 13: 105–109.

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.

Schot, A.M. (1994). A revision of *Callerya* Endl. (including *Padbruggea* and *Whitfordiodendron*) (Papilionoideae: Millettieae). *Blumea* 39: 1–40.

Schultze, M. and Kondorosi, A. (1998). Regulation of symbiotic root nodule development. *Annual Reviews of Genetics* 32: 33–57.

Sharma, S.K., Knox, M.X., and Ellis, T.H.N. (1996). AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theoretical and Applied Genetics* 93: 751–758.

Small, E. (1987). Generic changes in Trifolieae subtribe Trigonellinae. In: C. H. Stirton (editor). Advances in legume systematics, part 3, pp. 169–181. Royal Botanical Gardens, Kew.

Sokoloff, D.D. (1998). On the principles of revision of the tribe Loteae (Papilionaceae). In: Problems of the botanical sciences at the boundary of XX and XXI centuries, vol. 2, pp. 163–164. St. Petersburg (in Russian).

Sprent, J.I. and McKey, D. (editors). (1994). Advances in legume systematics, part 5: The nitrogen factor. 241 pp. Royal Botanical Gardens, Kew.

Steele, K.P. and Vilgalys, R. (1994). Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *mat*K . *Systematic Botany* 19: 126–142.

Steiner, J.J., Robinson, W.A., Liston, A., and Taylor, N.L. (1997). ITS and RAPD phylogenetic hypotheses and the ecological distributions of North American *Trifolium* L. (Fabaceae). *American Journal of Botany* 84 (suppl): 235–236.

Swofford, D.L. (1999). PAUP*, Phylogenetic Analysis Using Parsimony (*and Other Methods); version 4.0. Sinauer Associates, Sunderland, MA.

Szczyglowski, K., Shaw, R.S., Wopereis, J., Copeland, S., Hamburger, D., Kasiborski, B., Dazzo, F.B., and deBruijn, F.J. (1998). Nodule organogenesis and symbiotic mutants of the model legume *Lotus japonicus. Molecular Plant-Microbe Interactions* 11: 684–697.

van Oss, H., Aron, Y., and Ladizinsky, G. (1997). Chloroplast DNA variation and evolution in the genus *Lens* Mill. *Theoretical and Applied Genetics* 94: 452–457.

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.F. (In press). Molecular phylogeny of Old World *Trifolium* (Fabaceae), based on plastid and nuclear markers. *Plant Systematics and Evolution.*

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.

Wojciechowski, M.F. and Sanderson, M.J. (1995). Phylogenetic analysis of DNA sequence variation in the leucine tRNA gene group I intron of the chloroplast genome: an example from the Fabaceae. *American Journal of Botany* 82: s172.

Wojciechowski, M.F., Sanderson, M.J., Baldwin, B.G., and Donoghue, M.J. (1993). Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 80: 711–722.

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 *trn*L intron data. *Systematic Botany* 24: 409–437.

Yamazaki, M., Sato, A., Shimomura, K., Saito, K., and Murakoshi, I. (1994). Genetic relationships among *Glycyrrhiza* plants determined by RAPD and RFLP analyses. *Biological and Pharmaceutical Bulletin* 17: 1529–1531.