

## REVIEWS

# Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades

## The Legume Phylogeny Working Group<sup>1</sup>

This paper was compiled by Anne Bruneau,<sup>2</sup> Jeff J. Doyle,<sup>3</sup> Patrick Herendeen,<sup>4</sup> Colin Hughes,<sup>5</sup> Greg Kenicer,<sup>6</sup> Gwilym Lewis,<sup>7</sup> Barbara Mackinder,<sup>6,7</sup> R. Toby Pennington,<sup>6</sup> Michael J. Sanderson<sup>8</sup> and Martin F. Wojciechowski<sup>9</sup> who were equally responsible and listed here in alphabetical order only,

with contributions from Stephen Boatwright,<sup>10</sup> Gillian Brown,<sup>11</sup> Domingos Cardoso,<sup>12</sup> Michael Crisp,<sup>13</sup> Ashley Egan,<sup>14</sup> Renée H. Fortunato,<sup>15</sup> Julie Hawkins,<sup>16</sup> Tadashi Kajita,<sup>17</sup> Bente Klitgaard,<sup>7</sup> Erik Koenen,<sup>5</sup> Matt Lavin<sup>18</sup>, Melissa Luckow,<sup>3</sup> Brigitte Marazzi,<sup>8</sup> Michelle M. McMahon,<sup>19</sup> Joseph T. Miller,<sup>20</sup> Daniel J. Murphy,<sup>21</sup> Hiroyoshi Ohashi,<sup>22</sup> Luciano P. de Queiroz,<sup>12</sup> Lourdes Rico,<sup>7</sup> Tiina Särkinen,<sup>23</sup> Brian Schrire,<sup>7</sup> Marcelo F. Simon,<sup>24</sup> Elvia R. Souza,<sup>12</sup> Kelly Steele,<sup>25</sup> Benjamin M. Torke,<sup>26</sup> Jan J. Wieringa<sup>27</sup> and Ben-Erik van Wyk<sup>28</sup>

1 Recommended citation: LPWG (2013)

2 *Institut de recherche en biologie végétale and Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke East, Montreal, QC H1X 2B2, Canada*

3 *Department of Plant Biology, 412 Mann Library Building, Cornell University, Ithaca, New York 14853, U.S.A.*

4 *Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, Illinois 60022, U.S.A.*

5 *Institute for Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland*

6 *Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, U.K.*

7 *Herbarium, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, U.K.*

8 *Department of Ecology and Evolutionary Biology - 1041 E. Lowell, Tucson, Arizona 85721-0088, U.S.A.*

9 *School of Life Sciences, Arizona State University, Tempe, Arizona 85287-4501, U.S.A.*

10 *Department of Biodiversity and Conservation Biology, University of the Western Cape, Private Bag X17, Belville 7535, Cape Town, South Africa*

11 *University of Melbourne Herbarium, School of Botany, The University of Melbourne, Victoria 3010, Australia*

12 *Departamento de Ciências Biológicas, Universidade Estadual de Feira de Santana, Av. Transnordestina s/n, Novo Horizonte, 44036-900, Feira de Santana, Bahia, Brazil*

13 *Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia*

14 *Department of Biology, North Carolina Center for Biodiversity, Howell Science Complex, Mailstop 551, East Carolina University Greenville, North Carolina 27858-4353, U.S.A.*

15 *Instituto de Recursos Biológicos, CIRN-INTA, N. Repetto y Los Reseros s/nº, Hurlingham 1686, Prov. de Buenos Aires, Argentina*

16 *Plant Science Laboratories, University of Reading, Whiteknights, Reading, RG6 6AS, U.K.*

17 *Graduate School of Science, Chiba University, Department of Biology, 1–33 Yayoi, Inage, Chiba 263-8522, Japan*

18 *Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, Montana 59717, U.S.A.*

19 *School of Plant Sciences, University of Arizona, Tucson, Arizona 85721-0036, U.S.A.*

20 *Centre for Australian National Biodiversity Research, CSIRO Plant Industry, The Australian National Herbarium, GPO Box 1600, Canberra, ACT 2601, Australia*

21 *Royal Botanic Gardens Melbourne, Private Bag 2000, Birdwood Avenue, South Yarra 3141, Victoria, Australia*

22 *Herbarium, Botanical Garden, Tohoku University, Sendai, 980-0862, Japan*

23 *Natural History Museum, Cromwell Road, London SW7 5BD, U.K.*

24 *Embrapa Recursos Genéticos e Biotecnologia, PqEB, Caixa Postal 02372, 70770-917 Brasília-DF, Brazil*

25 *Department of Applied Sciences and Mathematics, Arizona State University Polytechnic, 6073 S. Backus Mall, Mesa, Arizona 85212, U.S.A.*

26 *Institute of Systematic Botany, The New York Botanical Garden, Bronx, New York 10458, U.S.A.*

27 *Netherlands Centre for Biodiversity Naturalis (section NHN), Herbarium Vadense, Biosystematics Group, Wageningen University, Generaal Foulkesweg 37, 6703 BL Wageningen, Netherlands*

28 *Department of Botany and Plant Biotechnology, University of Johannesburg, South Africa*

Author for correspondence: Toby Pennington, [t.pennington@rbge.org.uk](mailto:t.pennington@rbge.org.uk)

**Abstract** The Leguminosae, the third-largest angiosperm family, has a global distribution and high ecological and economic importance. We examine how the legume systematic research community might join forces to produce a comprehensive phylogenetic estimate for the ca. 751 genera and ca. 19,500 species of legumes and then translate it into a phylogeny-based classification. We review the current state of knowledge of legume phylogeny and highlight where problems lie, for example in taxon sampling and phylogenetic resolution. We review approaches from bioinformatics and next-generation sequencing, which can facilitate the production of better phylogenetic estimates. Finally, we examine how morphology can be incorporated into legume phylogeny to address issues in comparative biology and classification. Our goal is to stimulate the research needed to improve our knowledge of legume phylogeny and evolution; the approaches that we discuss may also be relevant to other species-rich angiosperm clades.

**Keywords** Caesalpinioideae; Fabaceae; Leguminosae; low-copy nuclear genes; Mimosoideae; multiple sequence alignment; Papilionoideae; phylogenetic inference

Received: 2 July 2012; revision received: 4 Dec. 2012; accepted: 30 Jan. 2013

## ■ INTRODUCTION

By whatever criteria are used to measure evolutionary success, the legume family (Leguminosae) is one of the most successful lineages of flowering plants. With ca. 751 genera and ca. 19,500 species (Lewis & al., 2005; updates in this paper), it is the third-largest angiosperm family. It has a global distribution spanning all major biomes and forming ecologically important constituents of temperate, Mediterranean, tropical, arid, seasonally dry, rain forest, and savanna ecosystems (Schrire & al., 2005). The family presents spectacular morphological and life history diversity, from giant rain forest trees and woody lianas, to desert shrubs, ephemeral herbs, herbaceous twining climbers, aquatics and fire-adapted savanna species (Lewis & al., 2005); it shows a significantly higher than average species diversification rate over the last 60 million years than angiosperms as a whole (Magallon & Sanderson, 2001); finally, it harbours the largest genus of flowering plants, *Astragalus* L. (Sanderson & Wojciechowski, 1996), and some of the most rapidly evolving plant clades (Richardson & al., 2001; Hughes & Eastwood, 2006; Scherson & al., 2008).

In addition, the legumes represent one of the most phenomenal examples of manipulation and utilization of a plant family by human cultures worldwide. This has involved the domestication of a set of globally important food crops, such as soybean (*Glycine max* (L.) Merr.), culinary beans (various species of *Phaseolus* L. and *Vicia faba* L.), groundnut (*Arachis hypogaea* L.), lentil (*Lens culinaris* Medik.), chickpea (*Cicer arietinum* L.) and pea (*Pisum sativum* L.), as well as important temperate and tropical forage crops such as alfalfa (*Medicago sativa* L.), clovers (*Trifolium* L.) and leucaena (*Leucaena leucocephala* (Lam.) de Wit). The nitrogen-fixing ability of many legumes provides an important source of biological nitrogen in agriculture and natural ecosystems, benefitting sustainable agricultural productivity, and providing essential ecosystem services. Nitrogen-fixing legume trees (e.g., *Calliandra* Benth., *Acaciella* Britton & Rose, *Gliricidia* Kunth, *Inga* Mill., *Leucaena* Benth.) form fundamental components of tropical agroforestry, forest restoration and soil improvement. While some genera, such as *Acacia* Mill., *Mimosa* L., *Prosopis* L., *Parkinsonia* L., *Pueraria* DC. and *Ulex* L., harbour species that have negative impacts

as prominent invasive weeds (Richardson & Rejmánek, 2011; Richardson & al., 2011), many legume species are ornamentals (e.g., *Amherstia* Wall., *Delonix* Raf., *Erythrina* L., *Laburnum* Fabr., *Lathyrus* L., *Lupinus* L., *Acacia* Mill., *Wisteria* Nutt.) or provide medicines (e.g., *Glycyrrhiza* L.).

The legumes are a good example of a family that was first recognized based on a small set of conspicuous morphological characters (Jussieu, 1789)—the legume itself, and a hard seed coat with a distinctive palisade layer of twisted cell walls in the epidermis and hypodermal cells that are usually hour-glass shaped—that has stood the test of time and of subsequent molecular analysis. The monophyly of the family has never been questioned despite the continuing lack of very strong bootstrap support (91% or less) even in the most recent analyses, a feature that is attributable to short phylogenetic branches among families within the order Fabales (Bello & al., 2009, 2010, 2012). The Leguminosae is usually divided into three subfamilies, but there are those who continue to recognize three separate families (e.g., Steyermark & al., 1998, 1999, 2001; Cullen & al., 2011) despite this view being widely regarded as untenable (e.g., Lewis & Schrire, 2003; Lewis & al., 2005). The subfamilies are typically characterized as being easy to distinguish based on morphological characters, but the Caesalpinioideae is not a monophyletic group (e.g., Doyle & al., 1997; Wojciechowski & al., 2004; Bruneau & al., 2008).

With the legume family being so important ecologically and economically, it should be no surprise that it has been a special focus of taxonomists since the time of Candolle (1825) and Bentham (1865). In a landmark effort to compile the voluminous quantity of legume taxonomic work from Bentham's time onward to the early 1980s, Roger Polhill (Kew), Peter Raven (Missouri) and collaborators organized the first International Legume Conference at the Royal Botanic Gardens, Kew in 1978, and subsequently published the conference proceedings as *Advances in Legume Systematics* (Polhill & Raven, eds., 1981). The original two volumes of *Advances in Legume Systematics* set in motion a seemingly exponential increase in interest in legume biology and systematics that has culminated in numerous publications, including important edited volumes, many of which are in the *Advances in Legume Systematics* series (eleven volumes published from 1981 to 2003).

In the late 1980s and early 1990s, molecular data, increasingly from DNA sequences, began to influence ideas of relationships across Leguminosae. These new data were synthesized by Polhill (1994), who updated the earlier (Polhill & Raven, 1981) classification of the family. The legume systematics team originally led by Polhill at the Royal Botanic Gardens, Kew, later coordinated the synthesis of the next decade of legume molecular systematic studies in *Legumes of the World* (Lewis & al., 2005). This volume provided information on all 727 legume genera recognized at that time. *Legumes of the World* was a landmark in organizing the latest information on the legumes, but the phylogenetic content of the book was implicit rather than explicit. In 2005 a phylogeny with adequate genus-level sampling of the whole family was not available and it was thus not possible to propose a fully revised phylogenetic classification of the family. Because the Leguminosae is the subject of active research across a global network, many new data have been published at the genus and suprageneric levels since 2005. In addition a series of increasingly well-sampled family-wide molecular phylogenies is available (e.g., Doyle & al., 1997; Käss & Wink, 1995, 1996, 1997; Kajita & al., 2001; Wojciechowski & al., 2004; Simon & al., 2009). *Legumes of the World* is being developed as an online Web resource for information about legume genera that can complement, link to and build upon the success of the online International Legume Database & Information Service (ILDIS; <http://www.legumes-online.net>; Bisby, 1993).

In the past few years there has been a growing interest in the legume systematics community to pool expertise and data and to take advantage of new approaches in genetics and bioinformatics. Such collaboration and innovation would facilitate the production of the comprehensive phylogenetic estimate and revised classification that is needed both by legume systematists as well as other consumers of systematics data. This has led to the formation of the Legume Phylogeny Working Group (LPWG),<sup>1</sup> which aims to develop collaborative research towards a comprehensive phylogeny and classification for Leguminosae.

In this paper, authored by the LPWG, we outline the current state of knowledge of legume phylogeny, giving detailed summaries for each of the three traditionally recognized subfamilies. We highlight where particular problems lie, for example in terms of taxon sampling and phylogenetic resolution in such a large phylogenetic tree. We then review new approaches in bioinformatics and from next-generation DNA sequencing that might help to resolve some of these problems. We discuss how morphology can be incorporated into our phylogenies to help to address questions in both legume biology and classification. If we can accomplish what we envision—to improve our understanding of the evolutionary history of legumes and to deliver a phylogeny-based classification—the products will

not only serve a very broad spectrum of researchers, but in addition may serve as an example for other species-rich angiosperm clades.

## ■ BUILDING A HIGH-RESOLUTION MOLECULAR PHYLOGENY OF LEGUMES

Before reviewing the current state of legume phylogeny, it is important to outline our ultimate goal—a high-resolution phylogenetic tree sampling as many species as possible. Though the utility of a phylogenetic framework for understanding evolution is clear, it may be less obvious why the most desirable goal both for legumes and other species-rich angiosperm clades is to build such a high-resolution tree. Many problems are best addressed only with dense sampling of species, including ecological questions about community assembly (e.g., Webb & al., 2002; Pennington & al., 2009) or the origins of biomes (e.g., Crisp & al., 2009). Moreover, studies of diversification patterns are sensitive to taxon sampling biases, and studies of trait evolution benefit from the increased power of large numbers of gains and losses (Aliscioni & al., 2011). It is therefore clear that a well resolved, densely sampled tree will be needed to answer many key questions in the comparative biology of legumes.

## ■ THE CURRENT STATUS OF LEGUME PHYLOGENY

The next sections summarize in detail what is known of the phylogeny of each legume subfamily, emphasizing studies published subsequent to *Legumes of the World* (Lewis & al., 2005).

### Caesalpinioideae

Subfamily Caesalpinioideae is a paraphyletic group at the base of the Leguminosae and from which are derived the monophyletic subfamilies Mimosoideae and Papilionoideae (Fig. 1). Species of this diverse subfamily occur primarily in tropical and subtropical regions of South America, Africa and Southeast Asia as trees, some extremely large, as well as lianas and shrubs. The subfamily includes approximately 2250 species in 171 genera, currently divided into four tribes: Cercideae, Detarieae, Cassieae and Caesalpinieae (Lewis & al., 2005). Of these, only the former two are supported as monophyletic in recent phylogenetic analyses of plastid sequence data, which to date have sampled 166 genera (Herendeen & al., 2003a; Bruneau & al., 2001, 2008; unpub. data from several researchers). The tribal limits, informal generic groupings and the generic limits of certain large genera (e.g., *Bauhinia* L. s.l., *Caesalpinia* L. s.l.) that were proposed by Lewis & al. (2005) are generally well supported and consistently resolved in recent phylogenetic studies, whereas others are tentative arrangements that merit further study with thorough taxon sampling.

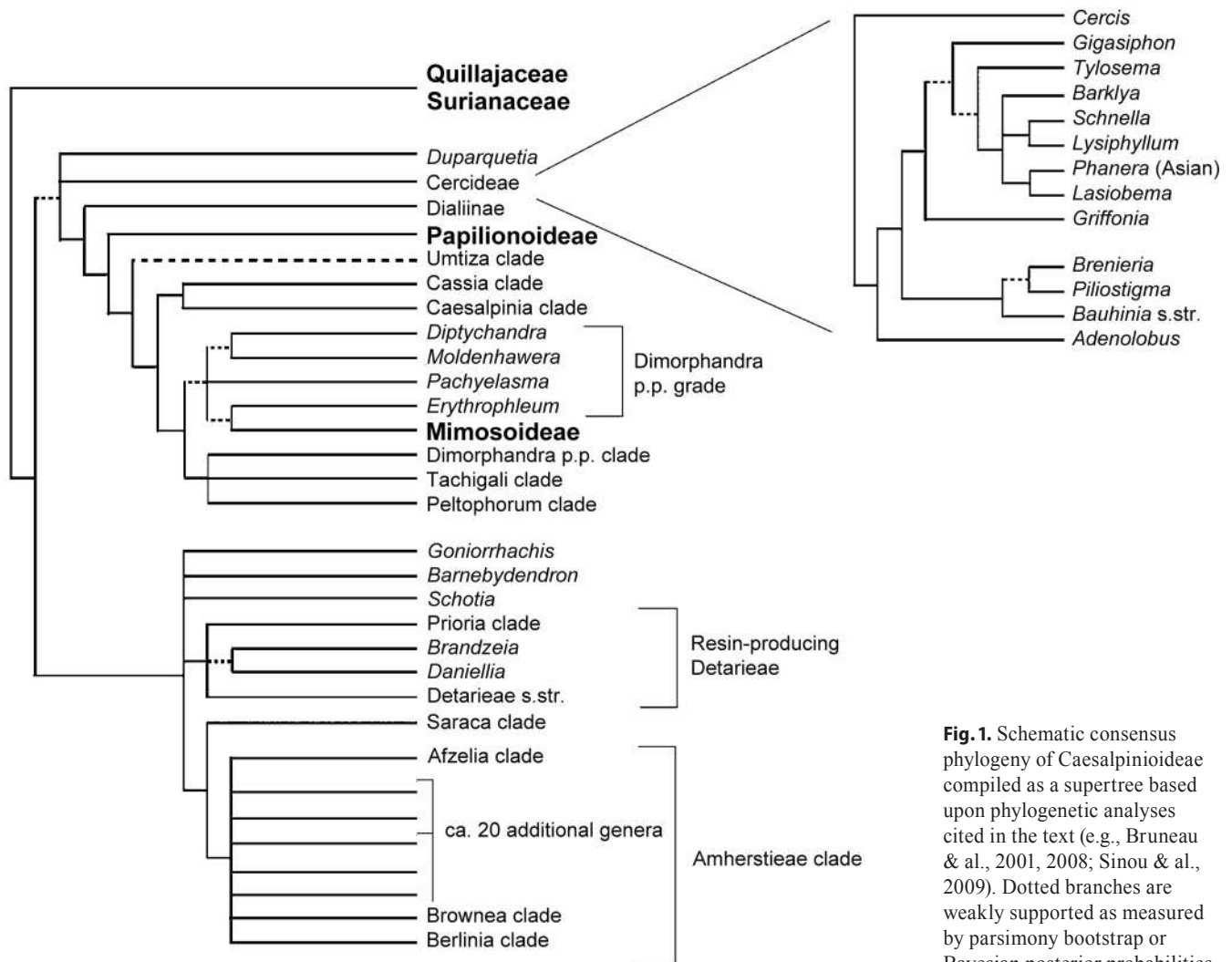
Relationships among the basal nodes of the legumes are not well supported (Fig. 1), with Cercideae, Detarieae and *Duparquetia* Baill. alternatively resolved as the sister group

1 The LPWG was established at a meeting convened in 2010 in Phoenix, Arizona, with A. Bruneau, P.S. Herendeen, C.E. Hughes, M. Lavin, G.P. Lewis, M. Luckow, B. Mackinder, B. Marazzi, M.M. McMahon, R.T. Pennington, M.J. Sanderson, K.P. Steele and M.F. Wojciechowski as its initial members.

to the remaining legumes depending on taxon sampling, locus sequenced and method of phylogenetic analysis (Bruneau & al., 2008; Bello & al., 2009, 2012). However, most recent phylogenetic evaluations of the family consider Cercideae to take that position (e.g., Doyle & al., 2000; Kajita & al., 2001; Bruneau & al., 2001, 2008; Herendeen & al., 2003a; Wojciechowski, 2003; Wojciechowski & al., 2004; Bello & al., 2009). Cercideae and Detarieae are individually strongly supported as monophyletic, but relatively few molecular or morphological characters support their wider relationships (e.g., Herendeen & al., 2003a) and the position of the monospecific and morphologically unique West African genus *Duparquetia* (Banks & al., 2003; Herendeen & al., 2003a; Prenner & Klitgaard, 2008) remains uncertain (Fig. 1).

The monophyly of tribe Cercideae is well supported in all morphological (Chappill, 1995; Herendeen & al., 2003a) and molecular (e.g., Käss & Wink, 1996; Doyle & al., 1997, 2000; Bruneau & al., 2001, 2008) (Fig. 1) phylogenetic studies. Members of this tribe share a number of unique vegetative and floral morphological features that support its monophyly

(e.g., Wunderlin & al., 1981). The most obvious of these features are the characteristic leaves that typically are simple or unifoliolate with a single joined pulvinus; the lamina is entire or often bilobed, though rarely two free leaflets are present. Relationships among genera within the tribe have been more problematic, especially relative to the large and complex pantropical *Bauhinia* s.l. Most recent analyses place *Cercis* L. as sister to the rest of the tribe, followed by *Adenolobus* (Harvey ex Benth. & Hook. f.) Torre & Hillc. as sister to two major clades (Fig. 1). One clade includes *Griffonia* Baill. weakly supported as sister to *Brenieria* Humbert, *Piliostigma* Hochst, and *Bauhinia* s.str., whereas the other clade includes all other segregate genera of *Bauhinia* s.l. (Sinou & al., 2009) (Fig. 1). These analyses support the recent taxonomic treatment by Lewis & Forest (2005) who, based on preliminary molecular analyses and previous taxonomic treatments, recognised the genera *Barklya* F. Muell., *Gigasiphon* Drake, *Lasiobema* Miq., *Lysiphyllum* de Wit., *Phanera* Lour. and *Tylosema* (Schweinf.) Torre & Hillcoat as distinct from *Bauhinia* s.str. In addition, recent analyses suggest that American *Phanera* (= *Schnella*



**Fig. 1.** Schematic consensus phylogeny of Caesalpinioideae compiled as a supertree based upon phylogenetic analyses cited in the text (e.g., Bruneau & al., 2001, 2008; Sinou & al., 2009). Dotted branches are weakly supported as measured by parsimony bootstrap or Bayesian posterior probabilities.



Raddi) might best be considered a genus distinct from the Asian (= *Phanera* s.str.) species (Lewis & Forest, 2005; Sinou & al., 2009, unpub. data; Wunderlin, 2010).

The tribe Detarieae as circumscribed by Mackinder (2005) is strongly supported as monophyletic in all recent phylogenetic analyses. Detarieae comprises mostly African genera, but also includes New World and Asian taxa, with several continental disjunctions between close generic pairs and a few within genera. Morphologically, this tribe is highly diverse and classification systems have been modified numerous times over the past decades to accommodate accumulating information on the group (e.g., Léonard, 1957; Cowan & Polhill, 1981a, b; Breteler, 1995; Wieringa, 1999; Mackinder, 2005). The broad pattern that emerges is weakly supported monophyly of the “resin-producing Detarieae” (sensu Fougère-Danezan & al., 2007), including the “Prioria” and the “Detarieae s.str.” clades (Fig. 1), which are sister to a large “Amherstieae” clade (sensu Bruneau & al., 2001). The positions of the South African *Schotia* Jacq. and of the South American *Goniorrhachis* Taub. and *Barnebydendron* J.H. Kirk. are unresolved relative to one another, but these genera are always sister to the resin-producing Detarieae and Amherstieae clades (Fig. 1).

Morphologically, the resin-producing Detarieae are extremely variable, particularly in their floral morphology, with numerous differences in sepal, petal and stamen number, but most members of this clade produce bicyclic diterpenes, a characteristic unique to this group (Langenheim, 2003; Fougère-Danezan & al., 2007).

Within the Amherstieae clade, few strongly supported subclades are consistently resolved. These are the mainly Asian “Saraca” (*Saraca* L., *Enderstia* Steenis & de Wit, *Lysidice* Hance) and “Afzelia” (*Afzelia* Sm., *Intsia* Thouars, *Brodriquesia* R.S. Cowan) clades (Bruneau & al., 2008), the New World “Brownea” clade (*Brownea* Jacq., *Browneopsis* Huber, *Ecuadendron* D.A. Neill, *Heterostemon* Desf., *Paloue* Aubl., *Elizabetha* Schomburgk ex Benth., *Paloveopsis* R.S. Cowan; Redden & Herendeen, 2006; Redden & al., 2010) and the entirely African “Berlinia” clade (Wieringa & Gervais, 2003; Mackinder & al., 2010; Mackinder & Pennington, 2011). The remaining genera and small generic groups of the Amherstieae clade form a large polytomy.

The “Dialiinae” clade, most recently defined to include subtribes Dialiinae and Labicheinae of the Cassieae as recognised by Irwin & Barneby (1981), together with *Poeppegia* C.S. Presl, is strongly supported as sister to the Papilionoideae plus the clade that includes the Mimosoideae, and most Caesalpinieae and Cassieae lineages (Bruneau & al., 2008) (Fig. 1). The monospecific Neotropical *Poeppegia*, which was placed in its own generic group of the Caesalpinieae by Polhill & Vidal (1981), is sister to this entire clade in most analyses. A number of Dialiinae clade species have determinate inflorescences (primarily simple or compound cymes or dichasia), produce drupes or samaras, and the clade is characterised by a high frequency of floral organ loss, all of which are unusual features for legumes (Tucker, 1998; Herendeen & al., 2003a; Zimmerman & al., 2013; E. Zimmerman, unpub. data). Generic relationships within the Dialiinae clade are, with few exceptions,

not highly resolved, while infrageneric relationships are entirely unknown.

The “Umtiza” clade, recognised by Herendeen & al. (2003b) in their combined morphological and molecular analyses, is only weakly supported as monophyletic in other recent analyses. Despite their disjunct geographical distribution, the seven genera of the Umtiza clade share a number of unique morphological features. For example, the flowers are usually small and greenish, and most genera are dioecious, both of which characters are not often encountered in the Caesalpinioideae (Herendeen & al., 2003b). The temperate genera *Gleditsia* L. and *Gymnocladus* Lam. and the South African *Umtiza* Sim form a strongly supported monophyletic group, defined by a number of morphological features. Similarly, the grouping of the Mediterranean *Ceratonia* L. with the Southeast Asian *Acrocarpus* Wight ex Arn. and Madagascan *Tetrapterocarpon* Humbert, is also supported by several morphological characters. However, the relationship of these two clades with the Caribbean *Arcoa* Urb. is poorly resolved in recent molecular analyses that include the nuclear sucrose synthase gene (Manzanilla & Bruneau, 2012).

Ongoing phylogenetic analyses with nuclear gene sequences, additional plastid loci and increased species-level sampling have allowed us to better resolve relationships among the other Cassieae and Caesalpinieae lineages. For example, recent analyses recover a clade that includes *Cassia* and *Senna* Mill. (Marazzi & al., 2006; Marazzi & Sanderson, 2010) as sister taxa, and sister to a clade that includes *Melanoxyton* Schott, *Recordoxylon* Ducke, *Batesia* Spruce ex Benth. and *Chamaecrista* Moench (Manzanilla & Bruneau, 2012). In the Caesalpinioideae, the ability to nodulate is restricted to a few genera, three of which, *Chamaecrista*, *Melanoxyton* and *Recordoxylon*, are members of this clade (though nodulation records for the latter are ambiguous; Sprent, 2001, 2009). The position of *Vouacapoua* Aubl. is problematic, with conflicting results between studies by Bruneau & al. (2008) and Haston & al. (2003, 2005).

The “Caesalpinia” clade includes all of the *Caesalpinia* s.l. segregate genera (*Caesalpinia* s.str., *Coulteria* Kunth, *Erythrostemon* Klotzsch, *Guilandina* L., *Libidibia* (DC.) Schldl., *Mezoneuron* Desf., *Poincianella* Britton & Rose, *Tara* Molina) recognised by Lewis (2005), as well as *Cordeauxia* Hemsl., *Balsamocarpon* Clos, *Hoffmannseggia* Cav., *Stahlia* Bello, *Pomaria* Cav., *Haematoxylum* L., *Moullava* Adans., *Pterolobium* R. Br. ex Wight & Arn., *Stuhlmannia* Taub., *Lophocarpinia* Burkart, *Stenodrepanum* Harms and *Zuccagnia* Cav. *Pterogyne* Tul. generally is resolved as sister to the Caesalpinia clade genera, but always with low bootstrap support, and certain analyses resolve *Pterogyne* as sister to the Cassia clade. In contrast, *Cordeauxia* and *Stuhlmannia* together are strongly supported as sister to all other remaining Caesalpinia clade genera (Haston & al., 2005; Manzanilla & Bruneau, 2012). The remaining genera are resolved into two distinct clades as indicated by several studies with varying taxon sampling (e.g., Simpson & al., 2003; Bruneau & al., 2008; Nores & al., 2012; R. Fortunato, E. Gagnon, G.P. Lewis, C.E. Hughes, S. Sotuyo, unpub. data). The first clade comprises a well-supported group

with *Balsamocarpon* Clos, *Hoffmannseggia* Cav., *Stenodrepanum*, *Stahliia*, *Libidibia* and *Zuccagnia*, sister to a clade that includes *Erythrostemon*, *Poincianella*, *Pomaria* and possibly *Cenostigma* Tul. The second clade includes *Guilandina*, *Pterolobium*, *Tara*, *Coulteria*, *Mezoneuron*, *Moullava*, *Caesalpinia* s.str., *Lophocarpinia* and *Haematoxylum*. In the combined plastid and nuclear DNA sequence analysis by Manzanilla & Bruneau (2012), the *Caesalpinia* clade is weakly supported as sister to the *Cassia* clade, and together these two clades are sister to all remaining *Caesalpinieae* clades, including the *Mimosoideae*.

As circumscribed by Haston & al. (2003, 2005), the “*Peltophorum*” clade is a well-supported group comprised of eight genera. Generic-level relationships are well resolved based on a diversity of plastid (Haston & al., 2005; Bruneau & al., 2008) (Fig. 1) and nuclear DNA sequences (Manzanilla & Bruneau, 2012). *Bussea* Harms and *Peltophorum* (Vogel) Benth. form a strongly supported monophyletic group, as do *Delonix*, *Lemuropisum* H. Perrier, *Colvillea* Bojer, *Conzattia* Rose and *Parkinsonia* L., with *Schizolobium* Vogel being sister to this latter clade. The genus *Delonix* as presently circumscribed is not supported as monophyletic with the monospecific genera *Colvillea* and *Lemuropisum* nested within it based on phylogenetic studies of plastid data (Simpson & al., 2003; Haston & al., 2005), several nuclear loci (M. Babineau, unpub. data), pollen morphology (Banks & al., 2003) and population genetic studies (Rivers & al., 2011). In both *Peltophorum* subclades, African and Madagascan taxa are grouped with South American genera, and although no unique morphological synapomorphies are apparent for the entire clade, the genera share a combination of features, such as bipinnate leaves, generally yellow petals, and narrow seeds (Haston & al., 2005).

The South American “*Tachigali*” clade sensu Haston & al. (2005), which comprises *Arapatiella* Rizzini & A. Mattos, *Jacqueshuberia* Ducke and *Tachigali* Aubl. (including *Sclerolobium* Vogel), is supported as monophyletic in all recent molecular phylogenetic analyses.

The relationship between the *Peltophorum* clade, the *Tachigali* clade and genera of the *Dimorphandra* group as circumscribed by Polhill & Vidal (1981) and Polhill (1994) is not well resolved (Fig. 1). It also appears that the *Dimorphandra* group is not monophyletic. *Burkea* Hook., *Dimorphandra* Schott, *Mora* Benth., *Stachyothyrsus* Harms, *Dinizia* Ducke (the latter transferred from the *Mimosoideae* following Luckow & al. (2000, 2003)) and possibly *Campsiandra* Benth. form a weakly supported clade (Bruneau & al., 2008). However, four other *Dimorphandra* group genera, *Diptychandra* Tul., *Moldenhawera* Schrad., *Pachyelasma* Harms and *Erythrophleum* Afzel. ex R. Br., are a paraphyletic grade at the base of the *Mimosoideae*, and this entire clade (*Mimosoideae*+*Dimorphandra* p.p. grade) is sister to the group comprising the *Peltophorum*, *Tachigali* and *Dimorphandra* p.p. clades (cf. Manzanilla & Bruneau, 2012). All other caesalpinoid genera known to nodulate occur in one or the other of these *Dimorphandra* groups (i.e., *Dimorphandra*, *Campsiandra*, *Moldenhawera*, *Erythrophleum*) or in the *Tachigali* clade (*Tachigali*, including synonym *Sclerolobium*) (Sprent, 2001, 2009; Doyle, 2011).

## Mimosoideae

Mimosoids have a pantropical distribution and form ecologically abundant elements in all major tropical biomes, including seasonally dry tropical forests and deserts (the succulent biome sensu Schrire & al., 2005), savannas (e.g., Ratter & al., 2003), and rain forests (e.g., Richardson & al., 2001). Mimosoids are also diverse in their life-history strategies, ranging from giant trees to aphyllous shrubs, woody lianas, functionally herbaceous geoxylic suffrutices, and even a few truly herbaceous and aquatic species. Mimosoids are the second-largest legume subfamily, with ca. 3271 species (Lewis & al., 2005), and although *Caesalpinioideae* has many more genera, the large species number in *Mimosoideae* reflects high species-richness in several of its 83 genera, especially *Acacia* s.str. (1000+ species; Murphy & al., 2010; González-Orozco & al., 2011) and *Mimosa* (ca. 540 spp.; Barneby, 1991; Bessega & Fortunato, 2011; Simon & al., 2011).

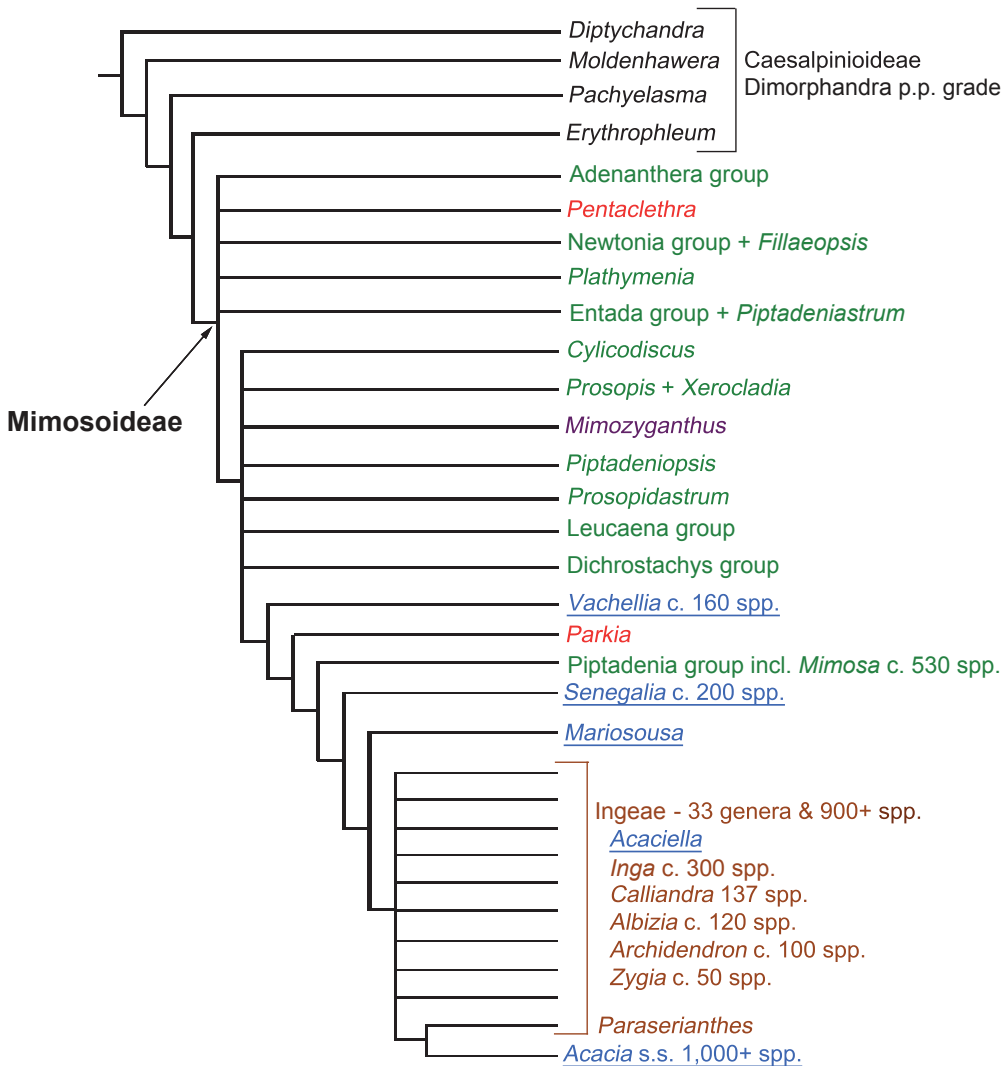
While the subfamily *Mimosoideae* has been consistently supported as monophyletic in all recent molecular phylogenetic analyses, delimitation of the subfamily remains to be satisfactorily resolved (Luckow & al., 2000, 2003; Bruneau & al., 2008). Recent analyses of the *Caesalpinioideae* provided support for a monophyletic mimosoid lineage nested in a grade of four caesalpinoid genera in the *Dimorphandra* group (see above), *Diptychandra*, *Moldenhawera*, *Pachyelasma* and *Erythrophleum*, with *Erythrophleum* as a possible sister group to *Mimosoideae* (Bruneau & al., 2008) (Fig. 2). Several *Dimorphandra* group genera have characteristics similar to taxa in the *Mimosoideae* (e.g., bipinnate leaves, alternate leaflets, small regular flowers; Luckow & al., 2000), and the recent analyses by Manzanilla & Bruneau (2012) suggest *Chidlowia* Hoyle might be grouped within the *Mimosoideae*, but this requires further verification. Bruneau & al. (2008) recognized that the sampling of mimosoids in their study was sparse, and although the relationships of many of the closely related caesalpinoid lineages remained incompletely resolved or lacking support (see above and Bruneau & al., 2008), a more recent and densely sampled legume- and mimosoid-wide phylogeny, based on *matK* sequences for 839 terminals, including 201 mimosoids (Simon & al., 2009) reiterated the difficulty of exactly delimiting *Mimosoideae*.

Early legume-wide molecular phylogenies based on *rbcl* (e.g., Doyle & al., 1997; Kajita & al., 2001) and *matK* (Wojciechowski & al., 2004; Lavin & al., 2005), although supporting the monophyly of the subfamily, shed little light on tribal or generic relationships within the subfamily because of sparse taxon sampling. More densely sampled mimosoid-wide analyses, based on plastid *trnL* and *trnK* intron and *matK* gene sequences, were presented in studies by Luckow & al. (2000, 2003). These revealed the non-monophyly of the traditionally recognized tribes, *Parkieae*, *Mimoseae*, *Acacieae* and *Ingeae* (Fig. 2), demonstrated the problems surrounding delimitation of the subfamily as a whole in relation to some closely related caesalpinoid genera, and documented a general lack of robustly supported resolution across the backbone of the mimosoid tree. Subsequent studies (e.g., Miller & al., 2003; Brown & al., 2008,

2011; Bouchenak-Khelladi & al., 2010; Gómez-Acevedo & al., 2010; Miller & Seigler, 2012) have confirmed these results, most notably the incongruence between these recent phylogenies and the traditional tribal classification (Bentham, 1875; Elias, 1981; Polhill, 1994) (Fig. 2). This traditional classification was based on a handful of conspicuous flower characters (notably number of stamens and fusion/or not of stamens into a staminal tube) that are now shown to be homoplastic. With the demise of the monogeneric tribe Mimosygantheae (Fortunato, 2005) following the discovery that *Mimosyganthus* Burkart is placed along with *Piptadeniopsis* Burkart and *Prosopidastrum* Burkart within the clade that includes the informal Leucaena and Dichrostachys groups (Luckow & al., 2005) (Fig. 2), all five tribes sensu Bentham (1875), Elias (1981) and Polhill (1994) have been demonstrated to be non-monophyletic, and a new tribal classification of the mimosoids remains to be established.

In contrast to the caesalpinioids and papilionoids, there are relatively few large higher-level mimosoid clades that are resolved with robust support across the mimosoid phylogeny.

Amongst members of the former Mimosaeae, a number of monophyletic groups are apparent, albeit some with only moderate support (Luckow & al., 2000, 2003, 2005; Lewis & al., 2005; Simon & al., 2009; Bouchenak-Khelladi & al., 2010), several of these corresponding to a large degree with the informal groups established by Lewis & Elias (1981) and Luckow & al. (2005) (Fig. 2). These include: the Adenanthera group (*Adenanthera* L., *Tetrapleura* Benth., *Amblygonocarpus* Harms, *Calpocalyx* Harms, *Pseudoprosopis* Harms, *Xylia* Benth.), the Newtonia group (*Newtonia* Baill., *Fillaeopsis* Harms), the Entada group (*Entada* Adans., *Elephantorrhiza* Benth., *Piptadeniastrum* Brenan), the informal Dichrostachys group (*Alantsilodendron* Villiers, *Calliandropsis* H.M. Hern. & P. Guinet, *Dichrostachys* (DC.) Wight & Arn. and *Gagnebina* Neck. ex DC.), the Leucaena group (*Desmanthus* Willd., *Kanaloa* Lorence & K.R. Wood, *Leucaena*, *Schleinitzia* Warb. ex Nevling & Niezgod), with the addition of *Prosopidastrum*, *Piptadeniopsis* and *Mimosyganthus* within this clade (Luckow & al., 2005), the Prosopis group (*Neptunia* Lour., *Prosopis*,



**Fig. 2.** Schematic consensus phylogeny of Mimosoideae based on Luckow & al. (2000, 2003, 2005), Simon & al. (2009), Bouchenak-Khelladi & al. (2010), Brown & al. (2011) and Miller & al. (2011). Informal generic groups of Luckow (2005) are indicated where they correspond to clades. The non-monophyly of the traditional tribes is indicated by colours: red, Parkieae; green, Mimosaeae; blue, Acaciae; brown, Ingeae; the monospecific tribe Mimosygantheae is indicated in purple. The five segregate genera currently recognized following the disintegration of the non-monophyletic *Acacia* s.l. are underlined. Species numbers are indicated for genera with more than 50 species.



*Xerocladia* Harv.) and the large clade comprising members of the former tribe Ingeae plus *Acacia* s.str. which accounts for ca. 2000 of the 3200 species of mimosoids (Fig. 2). Lack of resolution and support is particularly stark across this large Ingeae plus *Acacia* s.str. clade (Fig. 2) and currently there are no satisfactory generic groupings for the 34 genera that are placed here (Polhill, 1994; Luckow & al., 2000, 2003; Lewis & al., 2005), apart from the well-supported sister-group relationship between *Paraserianthes* I.C. Nielsen and *Acacia* s.str. (Brown & al., 2011; Miller & al., 2011). The informal Piptadenia group (Lewis & al., 2005), comprising *Adenopodia* C. Presl (as of yet unsampled in published phylogenies), *Anadenanthera* Speg., *Microlobius* C. Presl, *Parapiptadenia* Brenan, *Piptadenia* Benth., *Pityrocarpa* Britton & Rose, *Pseudopiptadenia* Rauschert, *Stryphnodendron* Mart. and the large genus *Mimosa*, has not been resolved as monophyletic with even moderate support so far (Jobson & Luckow, 2007; Simon & al., 2009; Bouchenak-Khelladi & al., 2010). The lack of resolution and support for these clades/groups along with a mostly unresolved backbone for the rest of the mimosoid phylogeny, mean that generic and tribal level relationships within the subfamily remain very poorly known, as is apparent from Figure 2.

Considerable attention has been devoted in recent phylogenetic studies to unravelling the non-monophyly of *Acacia* s.l. A series of analyses have established clear support for at least five independent lineages scattered widely across the mimosoid clade—variously placed with members of the former Mimosaeae (*Vachellia* Wight & Arn.), nested within the former Ingeae (*Acacia* s.str., *Acaciella*) or as successive sister groups to the former Ingeae+*Acacia* s.str. clade (*Senegalia* Raf., *Mariosousa* Seigler & Ebinger) (Miller & Bayer, 2001, 2003; Maslin & al., 2003; Seigler & al., 2006b; Bouchenak-Khelladi & al., 2010; Miller & Seigler, 2012; see also Murphy, 2008 for a review of phylogenetic studies and the classification of *Acacia* s.l.). There is now widespread support for the recognition of *Acacia* s.str. and four segregate genera, although there is evidence to suggest that *Senegalia* is non-monophyletic (Miller & Seigler, 2012), and despite the controversy accompanying consequent nomenclatural changes (e.g., Moore & al., 2010).

Aside from the prominent unravelling of *Acacia* s.l., numerous other generic changes have been made over the last three decades, notably by Nielsen (1981), Rico Arce (1991, 1992, 1999), Lorence & Wood (1994), Polhill (1994), Barneby & Grimes (1996, 1997), Barneby (1998), Jobson & Luckow (2007), and Villiers (2002), as summarized by Brown (2008) for the Ingeae, and synthesized by Lewis & al. (2005). Despite these advances, many cases of non-monophyletic genera revealed in recent phylogenies remain to be fully resolved. There is evidence to suggest that several important genera are, or may be, non-monophyletic, in particular *Piptadenia* (Jobson & Luckow, 2007), *Prosopis* (Catalano & al., 2008); *Albizia* Durazz and *Archidendron* F. Muell. (Brown & al., 2008 and unpub. data; Souza & al., subm.; Kyalangaililwa & al., in press), *Leucochloron* Barneby & J.W. Grimes (Almeida & al., unpub. data) and *Entada* (Luckow, unpub. data). Several others are potentially nested within other genera (e.g., *Elephantorrhiza* within *Entada* p.p. (Luckow & al., 2003), *Marmaroxylon* Killip

within *Zygia* P. Browne, *Guinetia* L. Rico & M. Sousa within *Calliandra* (Brown & al., unpub. data; Souza & al., subm.), while the status of several more (e.g., *Hydrochorea* Barneby & J.W. Grimes, *Cathormion* Hassk., *Hesperalbizia* Barneby & J.W. Grimes) remains uncertain (Rico Arce, 1992, 1999; Brown & al., unpub. data). It is clear that generic delimitation across the mimosoids remains in a state of considerable flux. Establishing a new generic and tribal classification for the mimosoids must wait until these generic delimitation issues are largely resolved.

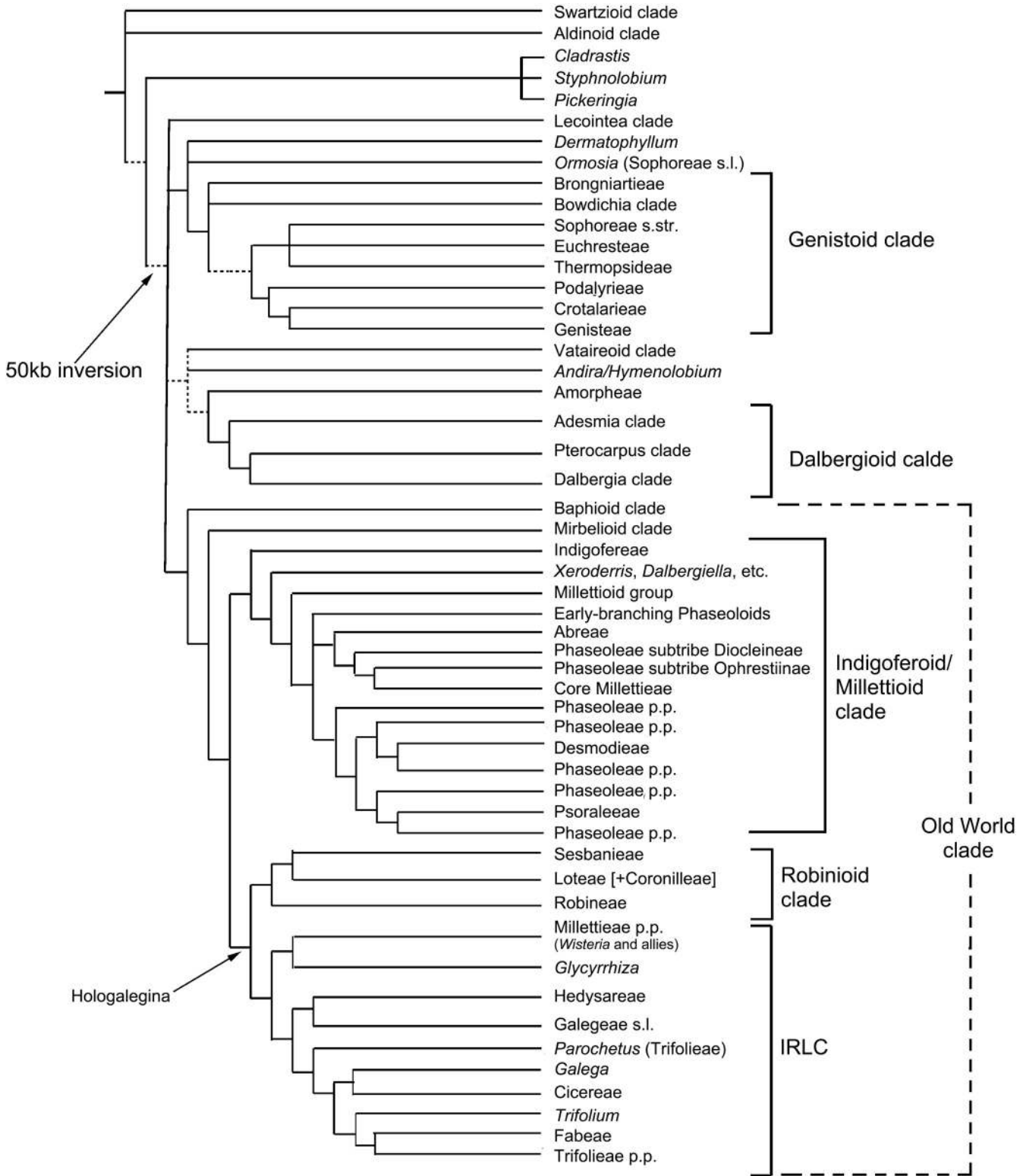
## Papilionoideae

Papilionoideae is the largest legume subfamily with 13,800 species across 28 tribes in 478 genera (Lewis & al., 2005). It includes many species of economic importance, especially the main pulse legume crops such as soybean. This economic importance is partly responsible for the large number of phylogenetic studies, many of which focus on individual genera. Subfamily-wide studies have been relatively rare (e.g., Doyle & al., 1997, Wojciechowski & al., 2004), but there have been an increasing number of studies that have focused at tribal level, or on a number of tribes (e.g., Crisp & al., 2000; Hu & al., 2000; Pennington & al., 2001; Lavin & al., 2001; Crisp & Cook, 2003; Boatwright & al., 2008a, b; Egan & Crandall, 2008).

**Swartzieae, Sophoreae and allies.** — The first branching papilionoid lineages, which have been termed “basal papilionoids” (e.g., Pennington & al., 2001) or “early-branching papilionoids” (e.g., Cardoso & al., 2012a) are largely tropical woody groups, and comprise almost all Swartzieae, many Sophoreae, all Dipterygeae and a few Dalbergieae—ca. 45 genera in total. Floral morphology is diverse in these lineages (Ireland & al., 2000), with some of the swartzioid genera superficially caesalpinoid-like with reduced petal numbers and many stamens. This led to suggestions that some of these genera might not belong in Papilionoideae (e.g., Polhill 1994), but consistent with the presence of isoflavonoids characteristic of Papilionoideae but absent in Caesalpinioideae, all phylogenetic analyses indicate strong support for a monophyletic subfamily that includes them (e.g., Doyle & al., 1997; Wojciechowski & al., 2004).

The basal nodes of the papilionoids are not well resolved (Fig. 3), but some clades are well supported. Studies have consistently recovered a “Swartzioid” clade that includes the species-rich *Swartzia* and several smaller tropical genera (e.g., Ireland & al., 2000; Pennington & al., 2001; Wojciechowski & al., 2004; Torke & Schaal, 2008) (Fig. 3). Among the remaining genera of “basal papilionoids,” Ireland & al. (2000) recovered a moderately supported “Aldinoid” clade that included the Neotropical genera *Aldina* Endl., *Amburana* Schwacke & Taub., *Dussia* Krug. & Urb. ex Taub and *Myrospermum* Jacq., as well as the African genera *Cordyla* Lour. and *Mildbraediodendron* Harms. The exact composition of the Aldinoid clade is uncertain, partly due to phylogenetic studies employing different loci and taxon sampling (e.g., Pennington & al., 2001; Wojciechowski & al., 2004), but these studies suggest





**Fig. 3.** Schematic phylogeny of Papilionoideae compiled as a supertree based upon phylogenetic analyses cited in the text (e.g., Lavin & al., 2001; Pennington & al., 2001; Crisp & Cook, 2003; Wojciechowski & al., 2004; Boatwright & al., 2008a; Egan & Crandall, 2008; Simon & al., 2009). Dotted branches are weakly supported as measured by parsimony bootstrap or Bayesian posterior probabilities.

that it also includes *Angylocalyx* Taub., *Xanthocercis* Baill., *Myroxylon* L. f., *Myrocarpus* Allemão) and all Dipterygeae.

The exact branching pattern among the Swartzioideae and Aldinoideae clades and the remainder of the Papilionoideae is poorly resolved (Fig. 3). Initial phylogenies using the *trnL* intron tentatively suggested that the Swartzioideae clade was sister to all other papilionoid legumes (Ireland & al., 2000; Pennington & al., 2001), but with weak bootstrap support. However, *matK* analyses (Wojciechowski & al., 2004) showed a combination of swartzioideae and aldinoids as sister to all remaining papilionoids, but with weak bootstrap support.

All phylogenies that have sampled widely amongst papilionoids (e.g., Doyle & al., 1997; Pennington & al., 2001; Wojciechowski & al., 2004) resolve a node that separates the Swartzioideae and Aldinoideae lineages plus several other genera (e.g., *Cladrastis* Raf., *Styphnolobium* Schott, *Pickeringia* Nutt.) from a large monophyletic group containing all other papilionoids, though in all cases with weak bootstrap support (Fig. 3). The latter group is marked by an inversion of 50 kb in the Large Single Copy region of the plastid genome that is situated between the *accD* and *trnK* genes. The 50kb inversion was first reported by Palmer & Thompson (1982) and screened using PCR and restriction-site mapping techniques (Doyle & al., 1996). The rapid PCR screening method is hampered by difficulties that may relate to non-specific primer binding (Doyle & al., 1996; Russell, 2004). Despite these difficulties, most genera screened show results consistent with their placement in phylogenies, though results for others were mixed and warrant further study. Although large plastid genome rearrangements are often considered highly stable, and thus informative phylogenetic markers, this is not always the case (e.g., Hoot & Palmer, 1994). Complete sequencing of plastids for papilionoid taxa with and without the 50kb inversion should be a priority to investigate primary homology and to resolve the primer site issues, and is easily feasible using next-generation sequencing techniques. The 50kb inversion node is also problematic in the sense that it may be genuinely cryptic, not apparently supported by a morphological synapomorphy. Interestingly, one feature that may map to this branch is the ability to nodulate (Doyle, 2011) because, with the exception of genera in the Swartzioideae clade, all of which can nodulate, the other Papilionoideae genera placed outside of this 50kb inversion clade do not nodulate (Sprent, 2009).

Included within the “50kb inversion” clade are three major clades and other smaller clades, the relationships amongst which are not clearly resolved (Fig. 3). The small clades include some genera currently assigned to Sophoreae and Dalbergieae, and are: (1) a “Vataireoid” clade (sensu Ireland & al., 2000) comprising *Luetzelburgia* Harms, *Sweetia* Spreng., *Vatairea* Aubl. and *Vataireopsis* Ducke (Mansano & al. 2004); (2) a “Lecointeoid” clade (sensu Ireland & al., 2000) comprising *Exostyles* Schott, *Harleyodendron* R.S. Cowan, *Lecointea* Ducke, *Uribea* Dugand & Romero and *Zollernia* Wied-Neuw. & Nees (Mansano & al., 2004); (3) *Andira* Lam. plus *Hymenolobium* Benth.; (4) *Dermatophyllum* Scheele (syn. *Calia*, see Gandhi & al., 2011). The relationships of these groups to the larger clades of Genistoids, Dalbergioids, and the remaining Papilionoideae remain an open question.

**Clade I: Genistoid.** — The “Genistoid” clade (Fig. 3) as a whole is defined by accumulation of quinolizidine alkaloids and a base chromosome number of  $n = 9$  (Pennington & al., 2001; Kite & Pennington, 2003; Wojciechowski & al., 2004). Both of these traits are also found in *Dermatophyllum* and *Ormosia* Jacks., and these genera are clearly closely related to the genistoids, but are not resolved within the clade in recent analyses (Wojciechowski & al., 2004). Within the clade, the membership of genistoids appears fairly stable (e.g., Pennington & al., 2001; Wojciechowski & al., 2004). Brongniartieae (ca. 100 tropical and subtropical New World species, including the recently discovered monospecific *Tabaroa* L.P. Queiroz & al. (Queiroz & al., 2010) and ca. 45 species in Australia (Ross & Crisp, 2005; Queiroz & al., 2010) is resolved as monophyletic and sister to the remainder of the genistoid groups, though not with high bootstrap support (Edwards & Hawkins, 2007; Cardoso & al., 2012a, b). A “Bowdichia” clade comprising *Bowdichia* Kunth., *Diploptropis* Benth. and the *Acosmium* s.l. segregates *Leptolobium* Vogel and *Guianodendron* Schütz Rodrigues & A.M.G. Azevedo (all formerly included in Sophoreae; Cardoso & al., 2012a, b) appears as sister to the remaining lineages, which comprise the “core genistoids” (sensu Crisp & al., 2000; Genisteeae, Crotalarieae, Euchresteeae, Podalyrieae, Sophoreae s.str., Thermopsidaeae). Tribe Sophoreae clearly requires re-circumscription, though relationships of species of *Sophora* s.l. need careful study to ascertain their correct placement (Heenan & al., 2004; Boatwright & Van Wyk, 2011). *Euchresta* Benn. (Euchresteeae) is thought to be closely allied to the Sophoreae s.str., based on *rbcl* data (Kajita & al., 2001). The tribal delimitations of Podalyrieae, Crotalarieae and Genisteeae, as proposed by Van Wyk & Schutte (1995) and Schutte & Van Wyk (1998a) were supported by molecular phylogenetic studies (Crisp & al., 2000; Van der Bank & al., 2002). Chemosystematic studies provided important clues (Van Wyk, 2003) and may have similar value in other tribes. Examples include  $\alpha$ -pyridone-type quinolizidine alkaloids (transfer of the *Argyrolobium* group from Crotalarieae to Genisteeae); methylated anthocyanins in petals, presence of canavanine in seeds and the absence of alkaloids (exclusion of *Hypocalyptus* from Podalyrieae/Liparieae); carboxylic acid esters of quinolizidine alkaloids (transfer of *Calpurnia*, and more recently *Cadia*, to Podalyrieae) and 3'-hydroxydaidzein as a major seed isoflavone in Podalyrieae and *Cadia*. Recent papers by Boatwright and collaborators have helped to clarify generic relationships in Crotalarieae (Boatwright & al., 2008a; 2009, 2011), resulting in several changes at generic level, and the placement of the morphologically disparate, radial-flowered *Cadia* Forssk. in Podalyrieae (Boatwright & al., 2008b).

**Clade II: Dalbergioid.** — The “Dalbergioid” clade (sensu Lavin & al., 2001) (Fig. 3) was first identified by Doyle & al. (1997) using phylogenetic analysis of *rbcl* sequences, but was studied comprehensively and named by Lavin & al. (2001), using a combination of plastid (*matK*, *trnL*) and nuclear (rDNA ITS) sequences and morphological data to assign 44 genera to the clade. The clade includes predominantly tropical and subtropical species from familiar street trees such as *Dalbergia sissoo* Roxb. ex DC. to herbaceous crop plants, such as the groundnut

or peanut (*Arachis hypogaea*). A monophyletic Amorpheae (100% bootstrap support, see McMahon & Hufford, 2004) is sister to the Dalbergioid clade, which comprises all genera formerly ascribed to the tribes Aeschynomeneae, Adesmieae, the subtribe Bryinae (Desmodieae) and most genera of Dalbergieae. An unexpected addition here is the radially-symmetric-flowered *Acosmium* s.str., long thought to belong to the genistoids (Cardoso & al., 2012a). Additionally, the recently discovered monospecific *Maraniona* C.E. Hughes & al. is a dalbergioid (Hughes & al., 2004). Outliers formerly assigned to Dalbergieae, but not members of the Dalbergioid clade are *Andira*, *Hymenolobium*, *Vataireopsis*, and *Vatairea* (Pennington & al., 2001; Lavin & al., 2001; Wojciechowski & al., 2004). None of these genera are resolved as sister to the Dalbergioid clade, so their relationships remain unclear. The dalbergioids are defined by a unique, “aeschynomenoid” root nodule morphology, which is not shared by *Andira* and *Hymenolobium* (which have indeterminate nodule morphology) or *Vatairea* and *Vataireopsis* (which do not nodulate; Sprent, 2009).

**Clade III: Old World clade.** — The remaining Papilionoideae comprise a large, predominantly Old World clade (Fig. 3). The “Baphioid Clade” is sister to the rest of this clade and is a group of west-central African woody genera (trees, lianas, shrubs) including *Baphiopsis* Benth. ex Bak. (Swartzieae), plus *Baphia* Afzel. ex Lodd. and five other genera of tribe Sophoreae. The “Mirbelioid” clade, comprising Mirbelieae (25 genera, Australasian) and Bossiaeeae (6 genera, Australian) (Wojciechowski & al., 2004) is sister to the remaining taxa in clade III. Hypocalyptae, a monogeneric, southern African tribe of three species (Schutte & Van Wyk, 1998b) was resolved as sister to the Mirbelioid clade by Wojciechowski & al. (2004), though with weak support. Recent revisions of some genera in the Bossiaeeae have resulted in the recircumscription of *Muelleranthus* Hutch., *Ptychosema* Benth. and *Aenictophyton* A.T. Lee with the recognition of the new monotypic *Paragoodia* I. Thomps., and description of several new species (Thompson, 2011a, b, c). The Old World Southern Hemisphere Baphioid and Mirbelioid clades are sister to the remainder of the papilionoids (Pennington & al., 2001; Wojciechowski & al., 2004). The mirbelioids, together with the remaining Papilionoideae, comprise a clade defined by the ability to accumulate canavanine (Wojciechowski & al., 2004). The remaining Papilionoideae are further split into two major subclades: the “Indigoferoid/Millettioid” clade and the “Hologalegina” clade, which is further split into the “Robinoid” clade and “Inverted Repeat-Lacking Clade” (IRLC) (Fig. 3).

• *Indigoferoid/Millettioid.* — Indigoferae is a pantropical, mainly woody tribe, with extensions into the subtropics and the Mediterranean-type Fynbos biome of southern Africa. It is placed as sister to the Millettioid group (Hu & al., 2000; Kajita & al., 2001) and has been well characterised in studies by Schrire and collaborators (see Schrire & al., 2009). The Indigoferae is strongly supported as monophyletic with the novel finding that the monospecific Madagascan *Disynstemon* R. Vig., previously placed in tribe Millettieae, is sister to the tribe (Schrire & al., 2009). With insufficient apomorphies to expand the circumscription of Indigoferae to include

this genus, Schrire & al. maintain it as part of a basal grade, together with *Xeroderris* Roberty, *Dalbergiella* Baker f., *Platy-cyamus* Benth. and *Austrosteenisia* Geesink. This alliance gives rise to the tribes Abreae, Millettieae, Phaseoleae and Indigoferae, although relationships among them are poorly supported. Recent studies based on ITS, *rbcL* and *matK* in the Millettioid group have done much to clarify relationships in this diverse and widespread clade, which includes many of the important legume crops such as soybean and common bean. The Millettioid group comprises a grade of Millettieae nesting two larger clades, one “Phaseoloid” and one of “core Millettieae” and allies (Hu & al., 2000, 2002; Kajita & al., 2001). Both Phaseoleae and Millettieae s.l. are polyphyletic and await a comprehensive revision at tribal level. One of the two larger clades includes the tribe Abreae and representatives of Phaseoleae subtribes Ophrestinae and Diocleinae, with *Rhodopsis* Urb. (Phaseoleae subtribe Erythrinae), resolved as sister to the core Millettieae group (Kajita & al., 2001). The clade with remaining members of Phaseoleae s.l. incorporates tribes Desmodieae and Psoraleae. This paraphyly has been addressed in a series of papers by Delgado-Salinas, Lavin and collaborators (Riley-Hulting & al., 2004; Thulin & al., 2004; Delgado-Salinas & al., 2006, 2011), and by Stefanović & al. (2009) for genera in the polyphyletic Glycininae (Lackey, 1981). Recent analyses of the Millettioid group are revealing non-monophyly of several genera, for example *Lonchocarpus* (Silva & al., 2012), *Vigna* (Delgado-Salinas & al., 2011), *Galactia* P. Browne (Sede & al., 2009), *Dioclea* Kunth (Queiroz & al., unpub.), *Desmodium* Desv. (Kajita & al., unpub.), *Otholobium* (Egan & Crandall, 2008), and *Pueraria* (Egan & al., unpub.). This body of research is clarifying generic limits and moving towards re-delimiting tribes.

• *Hologalegina.* — The Hologalegina clade contains many temperate groups including *Astragalus* L., the most species-rich plant genus with 2300 to 2500 species, as well as many important food and fodder crops (lentils, chickpeas, peas, clovers), and the genetic model legumes *Lotus japonicus* (Regel) K. Larsen and *Medicago truncatula* Gaertn. Kajita & al. (2001) and Wojciechowski & al. (2004) both showed Hologalegina to be split into two major clades, the Robinoid and IRLC (Fig. 3). The Robinoid clade contains three subclades, corresponding largely to tribes Sesbanieae (monogeneric), Loteae (22 genera) and Robinieae (11 genera). Each of these subclades is well supported, although the exact relationship among the three is not well resolved (Lavin & al., 2003; Wojciechowski & al., 2004). *Sesbania* Scop. is placed either as sister to Robinieae (Lavin & al., 2003) or to Loteae (Wojciechowski & al., 2004), and because of this equivocal position, it was treated as a separate tribe in *Legumes of the World* (Lewis & al., 2005).

The IRLC was defined by Wojciechowski & al. (2000) on the basis of the loss of one copy of the inverted repeat in the plastid genome (Lavin & al., 1990; Liston 1995), and like the 50kb inversion clade is cryptic, lacking any obvious morphological synapomorphies. The traditional tribes Galegeae, Cicereae, Fabeae (formerly Viciae; see Greuter & al., 2000), Hedysareae, and Trifolieae, as well as some members of Millettieae comprise the IRLC. *Glycyrrhiza* L. (Galegeae) plus



*Callerya atropurpurea* (Wall.) A. Schott (Millettieae) are resolved as a poorly supported clade that is sister to another group consisting of other Millettieae (*Afgekia* W.G. Craib, *Endosamara* Geesink, *Callerya* Endl., *Wisteria* Nutt.) that together form the sister group to the remaining IRLC (Hu & al., 2000, 2002; Hu & Chang, 2003). The remainder of the IRLC falls into two large clades (Wojciechowski & al., 2004). One of these contains a monophyletic Hedysareae (Ahangarian & al., 2007) and much of Galegeae s.l. (including *Astragalus*), but not *Galega* L., which falls into the second large clade. This second clade consists of *Galega*, and Trifolieae, which is paraphyletic because tribes Cicereae and Fabae are nested within it. Both species of *Parochetus* Ham. ex G. Don (Trifolieae), which are tropical montane herbs, are consistently resolved as sister to the remainder of this clade (Wojciechowski & al., 2000; Steele & Wojciechowski, 2003). The agriculturally important *Melilotus* Mill. is nested within *Trigonella* L. (Steele & Wojciechowski, 2003).

Fabae are consistently resolved as a monophyletic group, although its relationships to Trifolieae remain unclear. *Trifolium* L. may be sister to Fabae rather than to the remainder of Trifolieae (Steele & Wojciechowski, 2003; Wojciechowski & al., 2004). It is apparent that *Lens* Mill. is nested in *Vicia* L., and that *Pisum* L. and *Vavilovia* Fed. are embedded in *Lathyrus* L. (e.g., Steele & Wojciechowski, 2003). These are large tribes of significant economic importance and re-circumscriptions of the genera are imminent.

## ■ SOLUTIONS TO PROBLEMS IN LEGUME PHYLOGENETICS

The reviews of each legume subfamily above make clear that although progress has been relatively swift since the 1990s, we are still a long way from achieving the goal of a phylogeny sampling all the estimated 19,500 legume species. Though a phylogeny sampling virtually all of the ca. 751 currently recognized legume genera may be achieved soon, better species-level sampling and a more robustly supported and well-resolved tree will be needed to answer key questions in the comparative biology of legumes. Such an improved phylogenetic estimate is also required to test the monophyly of individual genera, perhaps the most important remaining issue in legume taxonomy. Given the current incomplete knowledge of legume phylogeny, we need to move beyond the present approaches whereby single researchers or small collaborative groups gather and analyze molecular sequence data from a few loci for a relatively limited taxon sampling. Moreover, how do we mine the publicly available databases (e.g., GenBank) to utilize the data that already exist and are being produced at an accelerating rate? What is the most feasible strategy for achieving a robustly supported, well resolved and densely sampled phylogeny of legumes based on multiple molecular sequences that maximizes the number of species sampled? In the following sections we focus on ideas and challenges in this effort to reconstruct the phylogeny of legumes - ideas and challenges pertinent to other species-rich angiosperm clades.

## Taxon sampling

The first major step towards a phylogeny of legumes fully sampled at species-level is one that samples all ca. 751 accepted genera. Of course to evaluate the monophyly of all these genera would require sampling multiple species per genus, something that has not yet been achieved in many cases. An additional important reason for including multiple species per genus is to mitigate for problems of misidentification (to genus level) or laboratory errors (e.g., cross-sample contamination) when a genus is represented by a single accession (e.g., see Mansano & al., 2004). How to improve species sampling is therefore a key concern.

Even if we set the requirement for generic sampling at the most minimal level, that is we consider a genus to have been sampled if at least a single nucleotide sequence is available in GenBank, then how well sampled are the legumes? To address this question, a list of accepted names of legume genera was compiled based on Lewis & al. (2005) but amended to take account of subsequent taxonomic generic changes. The accepted name list was used to make an audit of legume nucleotide sequence data held in GenBank and this preliminary examination indicated that there are 83 genera (ca. 11%) for which no nucleotide sequence data from any genomic region were publicly available in January 2012 (Table 1). In comparison, only about 8% of the ca. 880 genera of Orchidaceae do not have at least one locus sampled of *rbcL*, *matK* or ITS (M. Chase, pers. comm. 2012).

It is clear that sampling gaps are not evenly distributed across subfamilies. Twelve mimosoid genera remain unsampled in published phylogenies, but sequence data have been generated for ten of these in the last few years (Brown & al., unpub. data; Souza & al., subm.), leaving just two genera, *Aubrevillea* F. Pellegrin (2 spp.) and *Lemurodendron* Villiers & Guinet (1 sp.), that lack any known DNA sequence data. Similarly, most caesalpinoid genera have been sampled, but we still lack sequence data for six tropical, mostly wet-forest genera, *Pseudomacrolobium* Hauman, *Leucostegane* Prain, *Orphanodendron* Barneby & J.W. Grimes, *Uittienia* Steenis and *Androcalymma* Dwyer, all of which are monospecific (except *Leucostegane*, which has two species) and a few of which are possibly now extinct. In contrast, more generic-level sampling gaps remain in Papilionoideae (14.9% of genera vs. 2.4% for Mimosoideae and 4.1% for Caesalpinioideae); however, this is by far the most diverse of the subfamilies.

To focus on these sampling gaps, creating a unified list of holdings in existing DNA banks rich in legume DNAs would be a useful first step (these include the Royal Botanic Garden Edinburgh, Royal Botanic Gardens, Kew, and the Jardim Botânico do Rio de Janeiro). Several other avenues can be pursued, including targeted field collecting, sampling from existing herbarium holdings where these are recent enough and/or suitably preserved to hold sufficiently intact DNA, gathering together unpublished sequence data from a network of collaborators and possibly locating relevant species of known provenance in cultivation in botanic gardens. To facilitate fieldwork, we added information concerning the geographical

**Table 1.** Accepted names and distributions of legume genera for which no sequence data has been deposited in GenBank (January 2012). \* indicates genera for which sequence data have been added to GenBank during the preparation of this paper; \*\* indicates genera no longer recognised (Silva & al., 2012).

Accepted name	Number of species in the genus and geographical distribution
<i>Akschindlium</i>	1 sp.; Indo-China
<i>Amphimas</i> *	3–4 spp.; WC Africa
<i>Androcalymma</i>	1 sp.; endemic to the upper Amazon Basin in Brazil
<i>Antheroporum</i> *	ca. 4 spp.; Asia (SW China and Indo-China)
<i>Aphyllodium</i>	7 spp.; 6 spp. N Australia and Papuasias (New Guinea); 2 spp. Malesia, S China (Hainan), Indo-China, India and Sri Lanka
<i>Arthroclianthus</i>	ca. 30 spp.; endemic to New Caledonia (1 sp. extending to Vanuatu)
<i>Aubrevillea</i>	2 spp.; WC and W Africa
<i>Austrodolichos</i>	1 sp.; N Australia
<i>Baphiastrum</i>	1 sp.; WC Africa (Cameroon, Gabon, Congo [Brazzaville], Congo [Kinshasa] and Central African Republic)
<i>Barbieria</i>	1 sp.; S Mexico, C America, Caribbean and western S America
<i>Bergeronia</i> **	1 sp.; S America (Brazil, Bolivia, Paraguay and Argentina)
<i>Blanchetiodendron</i>	1 sp.; E Brazil (Bahia and Minas Gerais)
<i>Burkilliodendron</i>	1 sp.; Asia (Malaya [Perak])
<i>Camoensia</i> *	2 spp.; WC Africa to Angola
<i>Carrissoa</i>	1 sp.; SW Africa (Angola)
<i>Christia</i>	ca. 10 spp.; India to China, Malesia and Australia, and most diverse in Indo-China (6 spp.) and China (5 spp.)
<i>Chrysoscias</i>	3–4 spp.; South Africa (S parts of W Cape)
<i>Clitoriopsis</i>	1 sp.; Africa (Congo [Kinshasa] and Sudan)
<i>Cochlianthus</i> *	2 spp.; W China (Yunnan, Sichuan) and Himalayas (Nepal)
<i>Codariocalyx</i> *	2 spp.; Sri Lanka, India, Indo-China, Malesia, China and Taiwan
<i>Corethroedendron</i>	4 spp.; C Asia to E Siberia
<i>Cruddasia</i>	ca. 2 spp.; NE Indian subcontinent; Indo-China (Myanmar [Burma], Thailand)
<i>Desmodiastrum</i>	4 spp.; India and Malesia (E Java)
<i>Diphylarium</i>	1 sp.; Indo-China (Laos and Vietnam)
<i>Droogmansia</i>	ca. 5 spp., or often estimated at over 20 spp.; SC to W Zambezi to Sudanian Africa
<i>Dunbaria</i> *	20 spp.; SE Asia (centred in Indo-China to S China, Indian subcontinent, Malesia and Papuasias; 1 sp. to E Asia); 2 spp. extending to N Australia
<i>Dysolobium</i>	4 spp.; SE Asia (E Indian subcontinent, Indo-China, SW China, Malesia)
<i>Eleiotis</i>	2 spp.; India, Indo-China (Myanmar) and Sri Lanka
<i>Eminia</i> *	ca. 4 spp.; Africa (Zambezi region)
<i>Guianodendron</i> *	1 sp.; S America (Guyana and Brazil)
<i>Hanslia</i>	2 spp.; Malesia, Papuasias (centred in New Guinea), Vanuatu and Australia (N Queensland)
<i>Haplormosia</i>	1 sp.; WC and W Africa (Sierra Leone to S Nigeria and E Cameroon to Gabon)
<i>Hegnara</i>	1 sp.; Indo-China (including Myanmar [Burma]) and Malesia
<i>Herpyza</i>	1 sp.; Cuba
<i>Kalappia</i>	1 sp.; endemic around Malili in Sulawesi (previously Celebes)
<i>Lackeya</i>	1–2 spp.; SE U.S.A.
<i>Lemurodendron</i>	1 sp.; NE Madagascar (narrow endemic SSW of Vohemar)
<i>Leptodesmia</i>	3 spp.; Madagascar, 1 sp. also in India
<i>Leucostegane</i>	2 spp.; Malesia (Malay Peninsula and Borneo [Sarawak])
<i>Luzonia</i>	1 sp.; Philippines (Luzon and Leyte)
<i>Macropsychanthus</i>	ca. 2 spp.; Papuasias, Micronesia (possibly Philippines)

Table 1. Continued.

Accepted name	Number of species in the genus and geographical distribution
<i>Macrosamanea</i>	11 spp.; S America, most diverse and numerous in the Amazon basin extending N into the Orinoco valley and the Guianas
<i>Margaritolobium</i> **	1 sp.; S America (Venezuela)
<i>Mecopus</i>	1 sp.; India, Indo-China, S China (Hainan) and Malesia (Java)
<i>Meizotropis</i>	2 spp.; Asia (subcontinental India and W Indo-China)
<i>Melliniella</i>	1 sp.; WC Africa
<i>Micklethwaitia</i>	1 sp.; Zanzibar-Inhambane region of SE Africa (Mozambique)
<i>Monarthrocarpus</i>	1 sp.; E Malesia (Sulawesi, Philippines, Moluccas) and Papuasias (New Guinea)
<i>Monopteryx</i> *	3–4 spp.; northern S America (Colombia, Venezuela, French Guiana and Amazonian Brazil)
<i>Neocolletia</i>	1 sp.; Indo-China (Myanmar [Burma]) and Malesia (Java)
<i>Neorudolphia</i>	1 sp.; Caribbean (Puerto Rico)
<i>Nephrodesmus</i>	6 spp.; endemic to New Caledonia
<i>Ohwia</i>	2 spp.; India, Indo-China, Malesia, China (2 spp.) to Japan
<i>Orphanodendron</i>	1–2 spp.; Colombia
<i>Oryxis</i>	1 sp.; S America (Brazil [Minas Gerais])
<i>Ostryocarpus</i>	1–2 spp.; WC Africa (Guineo-Congolian region)
<i>Ougeinia</i>	1 sp.; India and W Nepal
<i>Painteria</i>	3 spp.; Mexico (Mexican Plateau, with 1 sp. disjunct in lowland Tamaulipas)
<i>Panurea</i> *	2 spp.; S America (Colombia and N Brazil)
<i>Paracalyx</i> *	6 spp.; NE Africa (Ethiopia, Somalia) and Socotra (5 spp.); Indian subcontinent, Indo-China (1 sp.)
<i>Paratephrosia</i>	1 sp.; Australia (W Australia, N Territory, S Australia, Queensland)
<i>Periandra</i>	6 spp.; Brazil (all spp.), Bolivia
<i>Petaladenium</i>	1 sp.; Brazil (Rio Negro)
<i>Pseudoeriosema</i>	ca. 4 spp.; Africa (mostly Zambezi and Sudanian to Somalia-Masai regions)
<i>Pseudomacrolobium</i>	1 sp.; WC Africa (Congo [Kinshasa])
<i>Ptychlobium</i> *	3 spp.; NE Africa and Arabia (Somalia-Masai region); southern Africa (S Zambezi and Kalahari-Highveld regions)
<i>Pycnospora</i> *	1 sp.; Africa (Somalia-Masai and Lake Victoria regions), India, SE and E Asia and Australia
<i>Pyranthus</i>	6 spp.; W, S and C Madagascar
<i>Requienia</i>	3 spp.; W to NE Africa (Sahelian zone); southern Africa (S Zambezi and Kalahari-Highveld regions)
<i>Sakoanala</i>	2 spp.; Madagascar (1 sp. in coastal E Madagascar; the other in NW and W Madagascar)
<i>Sarcodum</i>	ca. 3 spp.; Asia (S China, Indo-China, Malesia, Papuasias)
<i>Sartoria</i>	1 sp.; S Turkey
<i>Serianthes</i>	ca. 18 spp.; Indo-China (Thailand), throughout SE Asia and Pacific (Malesia, Papuasias, Micronesia, Melanesia and W Polynesia; 6 spp. restricted to New Caledonia where the genus shows most variation)
<i>Spirotropis</i> *	2 or 3 spp.; northern S America (Venezuela, Guyana, Surinam and French Guiana)
<i>Spongiocarpella</i>	ca. 7 spp.; SC Asia to W China, most diverse in C Asia
<i>Streblorrhiza</i>	1 sp., now extinct; Phillip Island, near Norfolk Island (between Australia and New Zealand)
<i>Tetragonolobus</i>	6 spp.; Mediterranean to eastern Europe, Caucasus
<i>Thailentadopsis</i>	3 spp.; Sri Lanka (1 sp.) and Indo-China (1 sp., Thailand; 1 sp. S Vietnam)
<i>Tibetia</i> *	4 spp.; mainly Sino-Himalayan region
<i>Trifidacanthus</i>	1 sp.; Malesia (Lombok and Flores, Philippines [Luzon]), Indo-China (S Vietnam) and S China (Hainan)
<i>Uittienia</i>	1 sp.; W Malesia: Sumatra and Borneo (Sabah and Kalimantan)
<i>Uleanthus</i>	1 sp.; S America (Amazonian Brazil)
<i>Viguieranthus</i>	23 spp.; Madagascar



range derived from Lewis & al. (2005) to the list of 83 missing genera (Table 1). As might be predicted, geographic representation is uneven with overall under-representation greater in genera and species from the three main tropical regions. Of those three, generic under-representation is most acute in Southeast Asia, which is clearly a priority region for targeted field work and a survey of existing herbarium collections. The pattern of under-representation at species level is slightly different, being broadly in line with overall species diversity in the three main tropical regions—more missing species belong to genera from tropical America and Southeast Asia than from tropical Africa.

We accept that complete sampling at generic level is unlikely to be attainable. Some genera have already been the subject of extensive searches in their native range and attempts to amplify DNA from the only known herbarium vouchers have been unsuccessful. The task of achieving a more complete generic and species level sampling by new fieldwork is challenging and urgent because some biomes harbouring high levels of legume diversity and endemism, such as tropical dry forests, are severely threatened globally (Pennington & al., 2009). Genera that are highly threatened should be prioritized in collecting efforts before it is too late. Many of the genera that fall into these categories are monospecific including *Androcalymma*, *Herpyza* C. Wright, *Lemurodendron*, and *Luzonia* Elmer. Unfortunately, in the worst cases, genera are endemic in regions where all original vegetation has been lost (e.g., *Streblorrhiza* Endl. on Phillip Island; Heenan, 2001), or are effectively unattainable in war zones. In cases where taxa are extinct, improvements to protocols working with degraded DNA from herbarium specimens offer some hope, especially some techniques using next-generation sequencing technologies (see below).

Many legume genera are small; ca. 500 contain ten or fewer species of which ca. 200 are monospecific (Lewis & al., 2005) and often restricted in their geographic range. Only 41 genera contain 100 or more species. This pattern is exemplified in Mimosoideae, where 23 of 83 genera are monospecific, but there are 1000+ species in *Acacia* s.str. (Murphy & al., 2010; Gonzalez-Orozco & al., 2011), the second-largest legume genus after *Astragalus*, ca. 300 spp. of *Inga* Mill. (Pennington, 1997), ca. 540 spp. of *Mimosa* (Bessegga & Fortunato, 2011; Simon & al., 2011), ca. 200 spp. of *Senegalia* (Seigler & al., 2006a) and ca. 160 spp. of *Vachellia* (Seigler & Ebinger, 2005). We must consider special sampling requirements for larger genera across the whole family, especially those known or thought to be polyphyletic (e.g., *Albizia* s.l., *Cynometra*+*Maniltoa* Scheff., *Macrobium* Schreb.) where a single accession representing each monophyletic subgroup would be a desirable initial baseline. This requires a detailed pre-existing knowledge of the phylogenetic structure of such genera, but where this is not available we must prioritise species-level phylogenetic investigations of known non-monophyletic genera. As an interim measure, published alpha taxonomic revisions of such genera may be used to guide sampling towards an adequate representation of both infrageneric structure and geographical range.

## New approaches to building a high-resolution molecular phylogeny of legumes

In the following section we examine how to exploit publicly available databases (e.g., GenBank) to utilize existing sequence data to produce a robustly supported, well resolved and densely sampled phylogeny of legumes based on multiple molecular sequences that maximizes the number of species sampled.

**Sequence data currently available for phylogenetic inference.**— Although the single locus phylogenies based on *matK*, *trnL* and *rbcL* have revealed much about legume phylogeny, many taxa not sequenced for these genes have been sequenced for other loci. There is a tremendous number of sequences in GenBank that are potentially useful for inferring phylogenies, but many of these data have yet to be exploited, owing to various reasons such as lack of phylogenetic breadth or questionable orthology. A cursory examination of the number of nucleotide sequences in GenBank's core nucleotide database available in July 2011 revealed a total of more than 450,000 sequences from more than 6100 legume species and infraspecific taxa in 640 genera, with 17,443 sequences (3.9%) from the five loci most often used for phylogenetic studies in many plant groups, including legumes (*matK* gene, *rbcL* gene, *trnL* intron, *psbA-trnH* spacer, and the internal transcribed spacers ITS1 and ITS2 of nrDNA). Of the 17,443 sequences, 3631 sequences were from *matK* and *rbcL*, and of these a number either were partial or represented multiple accessions from the same taxon. In addition, the majority of the taxa were represented only by a sequence from one of these loci. This clearly illustrates the current limitations imposed by the nature of the data themselves.

Combining loci is necessary and appropriate to generate the largest well resolved tree possible, but recent theoretical studies highlight a number of challenges. As the number of empirical studies is growing, there is increasing awareness of problematic assembly issues in large-scale “all-in” projects (i.e., data mining all available sequence data for a set of terminal taxa; McMahon & Sanderson, 2006; Hejnal & al., 2009; Smith & al., 2009, 2011a; Peters & al., 2011), and many of these obstacles are likely to impede similar efforts in Leguminosae. Some of these challenges are driven by inconsistencies in the data, such as different circumscriptions of the *matK* “region” by different workers and incomplete taxonomic sampling of the various loci such as noted above. Others are more fundamental and methodological, such as basic choices between alternative inference methods, such as supermatrix, supertree, or gene tree/species inference, or the inherent differences in evolutionary histories ascertained from nuclear vs. plastid regions due to different modes of inheritance. Other less well understood but equally important issues relate to the assembly and alignment of multi-locus datasets, which we focus on here.

**Taxon coverage of loci.**— The PhyLoTA database (Sanderson & al., 2008; <http://phylota.net>) provides a snapshot of the taxonomic distribution of nucleotide sequences in GenBank. PhyLoTA Release 184 includes 644 legume genera, indicating a high taxonomic coverage at the generic level. However, the distribution of this sampling effort is patchy across loci. The

four loci more frequently used for genus coverage are *trnL* (501 genera), *matK* (459), nrDNA ITS (429), and *rbcL* (288), and of course the taxon overlap between loci is only partial. These numbers are approximations because of the vagaries of clustering algorithms and locus length heterogeneity, but they give a sense of depth of coverage. Proportional species-level coverage is much lower, at about 6200 species with at least one sequence in GenBank. An important caveat is that the “species” found in GenBank include formally described species (e.g., *Medicago sativa* L.), as well as informally identified taxa (e.g., “*Astragalus* sp. Sanderson 2509”), for which NCBI assigns a distinct taxon ID, which in effect reduces taxonomic coverage.

**Multiple sequence alignment.** — The recent “mega-tree” of angiosperms by Smith & al. (2011a) included several thousand species of legumes in its multi-locus analysis based on multiple sequence alignments of the best sampled four loci named above, plus *atpB* and *trnK* (the *trnK* intron sequences flanking the *matK* gene, included as a separate locus). This analysis included spacer sequences that are unlikely to have ever been aligned at this phylogenetic scale in legumes (or at least such alignments have never been published). Most legume systematists probably despair when examining a progressively broader taxonomic sample of ITS sequences. For example, this locus was barely alignable across the genus *Vigna* Savi. s.l. (Delgado Salinas & al., 2006) or the much smaller *Leucaena*, where additional issues of many potentially non-functional ITS copy types further complicated alignments (Hughes & al., 2002). The risk in using conventional progressive multiple sequence alignment procedures (Smith & al., 2011b used MAFFT) at this taxonomic level is that they undercount insertion/deletion (indel) events (Loytynoja & Goldman, 2005, 2008), leading to alignments that are too short with respect to their likely evolutionary histories. This may lead to the recovery of clades that do not exist because of errors in assessment of primary homology when columns of nucleotide characters are forced together incorrectly.

To explore this alignment issue, M.J. Sanderson and M.F. Wojciechowski sampled 1000 legume ITS sequences randomly across those available from GenBank and built multiple sequence alignments using the progressive aligner MUSCLE (Edgar, 2004) and PRANK (Loytynoja & Goldman, 2008), which attempts to infer more of the true indels in the sequences’ history. The average length of the input sequences was 640 nucleotides. With default alignment parameters, the length of the PRANK alignment was three times that of the MUSCLE alignment (9243 nucleotide characters vs. 2985). Clearly PRANK is inferring many more indels to achieve an alignment. Comparison of RAxML trees based on these alignments indicated that most differences among topologies were relatively deep in the tree. On the other hand, detailed inspection of the subtree corresponding to previously analyzed narrower datasets within *Hologalegina* (e.g., Wojciechowski & al., 1999) in either MUSCLE or PRANK trees showed more consistency with previous results. This is mirrored in Smith & al.’s (2011a) megatree of angiosperms, where cursory examination of their legume phylogeny suggests that many species-level relationships match known phylogenies, while deeper clades

do not. Thus, more thorough alignment options may be promising in that algorithms for multiple sequence alignment can be scaled up to keep pace with the taxonomic accumulation of nrDNA ITS or other highly informative non-coding sequences, but it is necessary to look carefully at alignment quality when the taxonomic scope is broad. It may also be possible to harness some recent attempts to quantify alignment uncertainty (e.g., the bootstrap procedure of Penn & al., 2010), but this area is not especially well developed, and there are several distinct methods to identify different sources of alignment uncertainty. An alternative solution is to separate datasets like nrDNA ITS into smaller datasets (i.e., corresponding to less inclusive groups), within which alignment quality is preserved, and analyze them as though they were separate “pseudo-loci” (e.g., McMahon & Sanderson, 2006). But this leads to its own challenges, as we discuss next.

**Taxon coverage and missing data.** — The patchiness of data in GenBank almost always leads to multi-locus datasets with significant levels of missing data (Sanderson & al., 2010). Dataset completeness can be increased by limiting the taxa to those with high coverage for loci, or vice versa, and formal algorithms are available to guide this (Sanderson, 2003; Yan & al., 2005), but this strategy almost always reduces the taxon coverage dramatically, which defeats much of the purpose of large-scale phylogenetic studies. Despite generally optimistic appraisals of the effect of missing data on phylogenetic analysis (Kearney, 2002; Wiens, 2003), recent mathematical theory clearly characterizes potential deleterious impacts of partial taxon coverage. In supertree analyses of multiple loci, partial coverage can lead to a multiplicity of equivalent supertrees (Sanderson & al., 2010; Steel & Sanderson, 2010). In supermatrix analyses, partial coverage can induce “terraces” of large numbers of topologically different but equally optimal parsimony or likelihood trees, which greatly multiplies the number of reasonable alternative trees that must be considered by anyone interested in using them (Sanderson & al., 2011). These effects are most pronounced when there are many taxa, few loci and a significant fraction of missing taxa sampled for each locus on average, as for the legume data currently deposited in GenBank. This can also easily occur as a consequence of splitting up a single locus into several more easily aligned pseudo-loci, as described above.

There are several strategies to reduce these problems. One is simply the addition of more loci per taxon. Genomics technologies will no doubt enable more loci to be included and the random effects of missing sequences will become lessened with the power of next-generation sequencing approaches, as seen, for example, in the phylogenomic datasets of Hejnlol & al. (2009), or those emerging from the plant “IKP” project (dePamphilis, Leebens-Mack, pers. comm.; see also Nuclear Genes below). In the meantime, however, there are also some algorithmic approaches. First, the problem of terraces can be eliminated trivially by constructing “decisive” multi-locus datasets (cf. Sanderson & al., 2010). The easiest way to build a large decisive dataset is to select the locus that has the most taxa sampled, and then limit the inclusion of data from additional loci to only those taxa that also have data for the first

locus. Unfortunately, when we initially tried this for legumes, the pattern of taxon overlap between loci made the resulting supermatrices quite sparse. For example, when we selected 1217 *matK* sequences as our first locus, and 660 *rbcL* sequences for the second, we found only 172 *rbcL* sequences that were present for the same taxa as those with *matK*, a clear under-exploitation of available *rbcL* data.

It is possible to improve on this and include more taxa that are unique to one locus. Sanderson & al. (2011) outlined an algorithm for “safe” supermatrix combination that removes the ambiguity associated with terraces in tree space by pruning some (but hopefully not all) taxa unique to one of the loci. The problem addressed by this algorithm has been solved, however, for only two loci, and there are only crude heuristics available to extend it to larger collections of loci. Nonetheless, we think this is a promising avenue to pursue in combining loci to build very large trees for legumes.

At this point, in a new supermatrix analysis developed specifically for this paper, Sanderson and Wojciechowski have assembled datasets of 1276 complete (entire coding sequence) or near complete *matK* nucleotide sequences and 660 complete (from plastid genome sequences) and near complete *rbcL* nucleotide sequences, and over 4000 nrDNA ITS sequences for further analyses. In the case of *matK* and *rbcL*, only those taxa whose nucleotide sequence was accompanied by the amino acid translation were examined and used to identify candidate nucleotide sequences for each locus. Choosing those taxa with amino acid translations in their GenBank records enabled us to identify high-quality DNA sequences and create initial “profile” alignments of *matK* and *rbcL* amino acid sequences using MUSCLE (Edgar, 2004). The amino acid alignments were then used to guide the “reverse alignment” of the corresponding nucleotide sequences using the program tranalign (Rice & al., 2000). The sequences were chosen based on criteria we developed in the course of this project to identify a single complete or, if not complete, the longest sequence (e.g., minimal amount of missing data, ambiguous amino acids/nucleotides) for each unique legume taxon in GenBank (based on “ti”, the unique identification number for the taxon). Throughout this process, the aligned datasets were checked for potential redundant taxa (errors of taxonomic synonymy) and edited for apparent and real errors such as problematic or missing “gi” (unique “Gen-Info Identifier”) and “ti” numbers, and misspellings of taxon names. In several instances, we eliminated all but one of the sequences from multiple varieties or subspecies of a species, but also added a few sequences from taxa that had been missed by our initial screen of GenBank.

Using these individual datasets we have currently constructed a supermatrix of *matK*, *rbcL* and nrDNA ITS sequences from legumes. For *matK* and *rbcL*, we have included only the DNA sequence that corresponds to the protein coding sequence (i.e., no flanking *trnK* intron sequences). For nrDNA ITS, near full length ITS1-5.8S-ITS2 sequences were included, although in a few cases the 5.8S gene was missing. Taxon sampling was adjusted to ensure “decisiveness” of the final matrix (Sanderson & al., 2010) by including only taxa that were sequenced at least for *matK*. Thus, it included 1276 *matK*

sequences, 188 *rbcL* sequences, and 715 ITS sequences, with complete taxon coverage for *matK* and partial taxon coverage for *rbcL* and ITS. As discussed previously, ITS alignments were done with MUSCLE and PRANK and showed considerable differences deeper in the tree, while the *matK* and *rbcL* alignments were done with MUSCLE only and used amino acid translations to guide manual editing of the nucleotide sequence.

Phylogenetic trees were constructed using the GTRCAT model in RAxML v.7.27 (Stamatakis, 2006) without partitioning the model by loci. Optimal trees from a total of 650 runs from different random starting trees were combined in a strict consensus tree to characterize the diversity of local optima found by the program (Fig. 4). We also noted how many distinct optima occurred at the level of smaller clades in the analysis, just to see how the universe of local optima played out within certain clades across the tree (analysis provided by D. Zwickl: Fig. 4). The most compelling message of this analysis is that, despite reducing ambiguity because of partial taxon coverage, there is still a very large solution space associated with this supermatrix, probably only hinted at in the runs we did. That said, we see many parallels in the consensus tree derived from this supermatrix with what is known from previously published studies and described earlier in this paper.

**Low-copy nuclear genes.** — Lack of phylogenetic resolution has proved a persistent problem at various depths in the legume phylogeny. The suggested early and rapid radiation of the family (Lavin & al., 2005) may underlie the lack of resolution at deeper phylogenetic levels amongst major clades along the backbone of the legume phylogeny. For example, in Caesalpinioideae, the relationships amongst the basal nodes are very poorly resolved, as are relationships of caesalpinoid lineages paraphyletic to the Mimosoideae. Exactly the same issue is found in the basal nodes of Papilionoideae, and the uncertainty over the sister group of the Genistoid clade. In more apically nested groups, for example across the large Ingeae+*Acacia* s.str. clade, lack of resolution can perhaps be attributed at least in part to low substitution rates across the mimosoids in *matK* and *rbcL* compared to most other legume lineages (Lavin & al., 2005). Lack of phylogenetic resolution is also acute within species-rich, recently evolved genera such as *Inga*, *Lupinus*, and *Astragalus*. These problems of phylogenetic resolution in legumes at both low and high taxonomic levels may be solved within a few years using data from next-generation sequencing technologies.

The plastid genome has been the primary source of information for reconstructing phylogenies at the genus level and higher since the advent of plant molecular systematics, thanks to its high copy number and single-copy behavior. Although this eliminates problems of paralogy within the plastid genome, the continual incorporation of organellar sequences into the nucleus (e.g., Gaeta & al., 2010) suggests caution in assuming orthology (Arthofer & al., 2011). Absence of historical recombination is one of the most useful phylogenetic properties of the plastid genome, because in theory it means that every nucleotide should track the same historical signal and can thus be combined to produce a single robust phylogeny. This does not, however, eliminate incongruence within the plastid genome



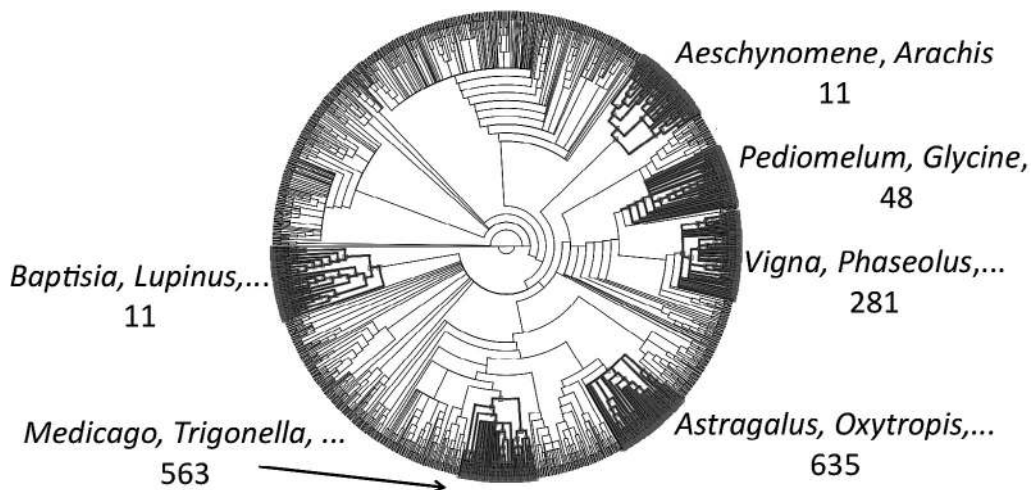
(e.g., Stefanovic & al., 2009). Soon there is likely to be a proliferation of phylogenies using whole plastid genome sequences, as is already happening in grasses (Wu & Ge, 2012). Current work by Koenen & al. is underway to generate a set of whole plastid genome sequences across mimosoids to help to resolve the backbone of that phylogeny and more definitively identify the most variable plastid regions. However, the fact that the entire plastid tracks a single phylogeny is a liability in cases of introgression (“plastid capture”: Rieseberg & Soltis, 1991) and incomplete lineage sorting (aka “deep coalescence”). These problems are often discussed at lower taxonomic levels, but it is not known how long they persist in a phylogeny, and thus to what degree plastid-based studies at the genus or family level may produce precise gene trees that are inaccurate as species trees. This is not a new concern (e.g., Doyle, 1992).

Nuclear genes have long been recognized as a potential complement or alternative to the plastid genome (e.g., Doyle & Doyle, 1999; Sang, 2002). There are tens of thousands of genes in the plant nuclear genome (e.g., 47,845 genes with experimental or database support in *Medicago truncatula*; Young & al., 2011), and a handful of randomly chosen ones can in theory provide numerous independent estimates of organismal phylogeny. The most often used nuclear gene regions are associated with the 18S-5.8S-26S ribosomal gene complex (e.g., especially ITS1 and ITS2), whose combination of high copy number and concerted evolution is both an asset and a liability (e.g., Alvarez & Wendel, 2003; Feliner & Rossello, 2007). ITS has proved very useful in phylogeny estimation of numerous papilionoid and caesalpinoid subclades (e.g., Lavin & al., 2003; Fougère-Danezan & al., 2007) but in some clades, and in mimosoids in particular, paralogy problems have limited its use. In mimosoids this reflects the frequent presence of multiple divergent and/or potentially non-functional pseudogene copy types within individuals in several genera, such as *Leucaena* (Hughes & al., 2002), *Desmanthus* (Hughes & al., 2003) and

sporadically in others, e.g., *Inga* (Richardson & al., 2001) and *Mimosa* (Simon & Hughes, unpub. data). Aside from paralogy issues, in general, the utility of ITS is limited to comparisons among species or closely related genera because alignment is difficult with greater divergence (see discussion above). The 18S and 26S genes, in contrast, are so conservative as to be useful generally only at higher taxonomic levels.

Low-copy nuclear (lcn) genes typically have evolutionarily conserved exons flanking variable introns, making them potentially useful at various taxonomic levels. In addition, divergence rates can vary dramatically among genes even in the same genomic region (e.g., Egan & Doyle, 2010). Such genes are individually single-copy but generally belong to small to large gene families produced by whole-genome and single-gene duplications of varying ages, whose complex birth and death evolution can mimic concerted evolution (Nei & Rooney, 2005). Thus, orthology is a serious issue for lcn genes and must often be confirmed for each taxon in a study, because only orthologous comparisons yield accurate organismal hypotheses. Heterozygosity poses a further problem with lcn genes because it can limit the ability to obtain usable data without isolation of single allelic sequences, for example by cloning or single strand conformation polymorphism (SSCP: e.g., Koopman & Baum, 2010). Additionally, recombination can result in multiple incongruent historical signals being present in a single nuclear sequence (e.g., in resistance genes: Ashfield & al., 2012).

The construction of PCR primers that amplify lcn gene orthologues across a range of taxa can be a daunting task; unlike nuclear ribosomal genes or plastid sequences, primers are rarely “universal”. As the genomics revolution has progressed, the number of available candidate genes has increased dramatically, and orthologous groups of genes have been identified in individual families such as Solanaceae (Wu & al., 2006) and more broadly in land plants (Proost & al., 2009; Duarte & al., 2010). In legumes, similar efforts have been made with some



**Fig. 4.** Strict consensus of 650 distinct locally optimal trees obtained in multiple RAxML searches of the 3-gene 1276-taxon legume dataset. The number of distinct optimal topologies (out of 650) within the six selected clades marked in red is indicated. Abbreviated taxonomic composition of these clades is indicated next to each clade. A Nexus formatted tree file of this tree will be deposited at the Dryad digital data repository (<http://datadryad.org>).

success (e.g., Choi & al., 2004, 2006). It is noteworthy that a very early example of an lcn gene being used to reconstruct a higher-level phylogeny in plants was by Lavin & al. (1998), who used phytochrome sequences to study the millettoid legumes; the recent application of sucrose synthase sequences to caesalpinoid phylogeny (Manzanilla & Bruneau, 2012) also deserves mention.

The economic importance of legumes has meant that genetic resources have been developed for many species, albeit predominantly papilionoids. Whole-genome sequences are now available from the model papilionoid legumes *Lotus japonicus* (Sato & al., 2008) and *Medicago truncatula* (Young & al., 2011), as well as from *Glycine max* (soybean; Schmutz & al., 2010) and *Cajanus cajan* (L.) Huth (Varshney & al., 2011). Collections of expressed sequence tags (ESTs) using Sanger sequencing were developed from many species, from which PCR primers were designed that could be used broadly in the family, even on caesalpinoid and mimosoid species (Choi & al., 2006). The first fruits from the explosion of “next generation” ESTs based on transcriptome sequencing are already available (e.g., chickpea, Hiremath & al., 2011; *Bituminaria bituminosa* (L.) C.H. Stirt., Pazos-Navarro & al., 2011; lentil, Kaur & al., 2011) with many more on their way (e.g., *Inga*, Pennington & al., unpub. data; *Entada*, *Albizia*, and *Microlobius*, Koenen & al. unpub. data; *Acacia*, Miller, unpub. data; numerous phaseoloid legumes, Egan & Doyle, unpub. data; and *Lupinus*, Filatov & al. unpub. data). In addition, 23 legume species spanning the whole family are listed by the 1000 plant transcriptome (1KP) sequencing initiative ([www.onekp.com](http://www.onekp.com)). Large portions of plant genomes are now accessible by methods such as restriction site associated DNA sequencing (RAD-seq; Davey & al., 2011) and genotyping-by-sequencing (GBS; Elshire & al., 2011), and it is only a matter of time before affordable whole genome sequencing is widely accessible. Approaches involving individual genes are facilitated by sequence capture methods that show promise for use with degraded plant material such as herbarium specimens (Grover & al., 2012), and next-generation sequencing can be used to simultaneously sequence many amplicons from multiple samples (e.g., Sakiroglu & al., 2012).

If good quality, high molecular weight nucleic acids can be isolated from legume species, then it is likely that partial to complete sequences of large numbers of lcn gene sequences can be obtained. However, generating massive amounts of data is almost certainly the easiest part of the process of obtaining a phylogeny. The process of constructing a contiguous sequence, particularly from short sequencing reads, with or without a closely related reference genome, can be very difficult given the complex evolutionary patterns of gene families and whole genomes, especially when polyploidy is involved (see Ilut & al., 2012 for a legume example). Even with good (i.e., non-chimeric) sequences the problem of paralogy is not overcome simply by generating a large dataset.

Analysis of multigene, phylogenomic-scale datasets has already received considerable attention in the literature. It is tempting to hope that the mere weight of massive datasets will produce reliable phylogenies, but this is unlikely to be true, and early proclamations that genome-scale sequencing and analysis of resulting supermatrices of concatenated gene sequences

would produce an “end to incongruence” (Gee, 2003) were met with considerable and justified skepticism (e.g., Soltis & al., 2004). Problems include deep coalescence and introgression, which, as noted above, may extend beyond the lower taxonomic levels at which they are generally taken into account. Sampling is another major issue (Soltis & al., 2004), as discussed above. Sanderson and colleagues have explored the problems of incomplete sampling in phylogenomic analyses (see above; Sanderson & McMahon, 2007; Sanderson & al., 2010), and constructed supertrees from several thousand sequences in over 2000 legume species (McMahon & Sanderson, 2006). Some recent examples of phylogenomic studies of angiosperm phylogeny point the way for family-level studies. Burleigh & al. (2011) used gene tree parsimony (Page & Charleston, 1997) to produce a phylogeny from nearly 20,000 gene trees constructed from over half a million sequences from 136 plant taxa. Lee & al. (2011) analyzed a supermatrix from over 20,000 putatively orthologous loci for over 100 taxa. It is noteworthy that both of these analyses produced topologies that were in relatively good agreement with the plastid-dominated APG (2009) phylogeny for angiosperms. The Lee & al. (2011) study is also of interest in that it used genome-wide sampling of genes to formulate functional genomic hypotheses, such as the role of small RNAs in the evolution of monocots.

## ■ INTEGRATING MORPHOLOGY

Progress towards building a comprehensive phylogeny for the legumes has been remarkable, and prospects for the future are excellent. The resulting phylogeny will be an essential resource for addressing a variety of questions. The range of questions becomes much broader if the molecular data are accompanied by data from morphology. In terms of phylogeny reconstruction, morphology can enhance results obtained from DNA sequence data by providing additional phylogenetically informative characters in combined “total evidence” analyses, and morphological data can make possible the inclusion of taxa that are not represented in the DNA sequence dataset. For example, there are a number of taxa that have not been seen in many years and several may be extinct (e.g., the monospecific Amazonian *Androcalymma* in Dialiinae). Although the available herbarium specimens may be inadequate for DNA sequencing, these taxa could be included in the phylogeny if a morphological dataset is analysed simultaneously with the DNA data. Similarly, the phylogenetic position of fossils can be evaluated if there is a morphological data matrix for the relevant extant taxa (e.g., Gandolfo & al., 2011) and recently fossils have become particularly important in analyses of molecular dating and biogeography (Mao & al., 2012; Sauquet & al., 2012). Consideration of morphological data in a phylogenetic context, either through total evidence analyses or optimising morphological characters onto a molecular phylogeny reveals morphological synapomorphies for clades and hence diagnosability in new phylogenetic classifications, as well as many evolutionary insights. For example, bilateral floral symmetry probably evolved multiple times in the legumes, in different ways in different groups within the family. A resolved

phylogeny that incorporates floral development characters can provide insight into the nature of these putative convergences (Pennington & al., 2000; Bello & al., 2012). Finally the addition of morphological data to a DNA sequence dataset can provide greater resolution and clade support in the resulting phylogeny (e.g., Herendeen & al., 2003b; Fougère-Danezan & al., 2010; Redden & al., 2010; Cardoso & al., 2012a), though this is not always the case (e.g., Haston & al., 2003; Russell, 2004).

While the potential benefits of adding morphological data (and other types of non-DNA sequence data) to phylogenetic analyses are clear, there are significant challenges, especially when the effort is scaled up to a group the size of the legume family. The good news is that the legumes are a rich source for morphological characters because the family is extremely diverse biologically, and there have been morphological phylogenetic studies in different clades of the family. Anatomists and morphologists have studied most organ systems and a variety of morphological datasets have been assembled addressing particular organs. The surveys, which provide the raw morphological observations but not necessarily in the form of characters and character states, include: pollen, which shows great diversity especially across caesalpinioideae and mimosoideae (Guinet, 1981; Ferguson & al., 1994; Banks & al., 2003); anther glands (Luckow & Grimes, 1997); extrafloral nectaries (Marazzi & al., 2012, 2013); inflorescence morphology and heterochrony (Grimes, 1999); embryology (Crisp & Cook, 2003); chromosome number (Goldblatt, 1981; Poggio & al., 2008); chemistry (Bisby & al., 1994); wood (Gasson & al., 2003; Evans & al., 2006); floral development (Domenech-Ramirez & Tucker, 1990; Tucker, 2003; Prenner & Klitgaard, 2008); root nodules (Sprent, 2001, 2009); and fruit anatomy and morphology (Lima, 1990; Le Roux & al., 2011), which have featured especially prominently in generic delimitation in mimosoideae, for example in the Piptadenia group and Ingeae (Barneby & Grimes, 1996, 1997). Other morphological studies have focused on clades within the family rather than on particular organ systems. For example, *Dichrostachys* and allies (Luckow, 1995, and in prep.; Luckow & al., 2005), the Pithecellobium group (Grimes, 1995); New World Ingeae (Barneby & Grimes, 1996, 1997; Barneby, 1998), *Desmanthus* (Luckow, 1993); *Parkia* (Luckow & Hopkins, 1995); *Acacia* (Chappill & Maslin, 1995; Grimes, 1995; Rico Arce & Banks, 2000); *Leucaena* (Hughes, 1998); *Prosopis* (Burghardt & Espert, 2007); Caesalpinioideae (Herendeen & al., 2003a); resin-producing Detarieae (Fougère-Danezan & al., 2010), *Aphanocalyx-Bikinia-Tetraberlinia* (Wieringa & Gervais, 2003); Swartzieae and Sophoreae (Herendeen, 1995); the Dalbergioid clade (Lavin & al., 2001); Diocleinae (Queiroz & al., 2003); *Acosmium* s.l. (Rodrigues & Tozzi, 2007).

Ideally, data from these studies can be combined as a starting point in an effort to build a broader morphological dataset that encompasses the entire family. But combining data from independent studies is often challenging, as evidenced by the valiant family-wide morphological cladistic analysis of Chappill (1995), which produced many estimates of relationships incongruent with subsequent molecular phylogenies. Assessing primary homology across a large, morphologically diverse family can be difficult. Different workers define morphological

characters differently. For example, floral symmetry is a complex syndrome involving multiple floral organs and multiple modifications of those organs in different clades within the family. The resulting suite of characters to encode the wide range of floral symmetries encountered in the family would undoubtedly vary from one worker to another. Combining morphological datasets requires careful analysis of all characters to determine which can be merged without modification, and which characters will require re-definition and re-coding in the combined matrix. For this reason legume systematists interested in morphological phylogenetics should communicate with each other as their studies progress to maximize comparability of character definitions and thereby minimize subsequent effort when the studies are combined. Another potential solution to this problem is to combine multiple separate morphological data matrices into a supermatrix. Although this would bring many of the same challenges that apply to molecular supermatrices, it might be an expedient solution for obtaining family-wide morphological data. For example, with different teams working on morphological phylogenetic analyses in the three subfamilies it will undoubtedly be easier to build a morphological supermatrix from the separate character sets and data matrices for Caesalpinioideae, Mimosoideae, and Papilionoideae than it would be to settle on one list of characters and re-code many of the taxa. Operationally, the morphological supermatrix would be structured and function just like a molecular supermatrix, except that the data are based on morphology. Finally, it should be noted that many of the datasets cited above remain incompletely, and in some cases very sparsely, sampled. Considerable effort is required to fill in the gaps, but the benefits provided by a morphological dataset are many. In the evolutionary context of this paper, linking genotype to phenotype is a grand challenge in biology, with implications for understanding homologies of morphological characters and other traits that could complement molecular data in phylogeny reconstruction.

## ■ WHERE ARE WE ALONG THE ROAD TO A FAMILY-WIDE PHYLOGENETIC CLASSIFICATION?

**Higher-level classification.** — It has been known for nearly two decades that the Caesalpinioideae is not monophyletic. As a result, the current three-subfamily scenario in legumes should not persist. At present the legumes are traditionally further divided into 35 tribes (the Mimosoideae, included as a separate tribe in Lewis & al., 2005), was disbanded by Luckow & al. (2005). But the phylogenetic data indicate that many of those 35 tribes are non-monophyletic. As larger tribes would be morphologically highly heterogeneous, if classification is to reflect monophyly, should the 35 tribes now be further divided into over 50? Similarly, if the Caesalpinioideae is to be divided into a number of monophyletic units, how many more subfamilies should be circumscribed—eight, nine, or more? Decisions such as whether to start publishing formal names for strongly supported subclades within the traditional Caesalpinioideae have been a dilemma for legume systematists for the past ten years.



There has been sensible restraint in making formal nomenclatural changes thus far, with the majority view being that while many clades have strong support, others do not and we still need more data for several parts of the legume phylogeny. In other cases, workers are uneasy about formally recognizing clades that have strong molecular support but lack morphological synapomorphies that can be used to diagnose a group. Although better taxon sampling, increased phylogenetic resolution and the integration of morphological data across the whole family are desirable, there is an urgency to produce the revised classification that is so obviously needed. If the legume systematic community is to retain credibility with a wide user-base, then nomenclatural stability is essential (Orthia & al., 2005). To achieve such stability, ideally only well-supported monophyletic, morphologically diagnosable groups should be recognised (Orthia & al., 2005).

Legume nomenclature has used a combination of formal taxonomic names and informal group names since the publication of part one of *Advances in Legume Systematics* (Polhill & Raven, 1981). Informal clade names may help to avoid nomenclatural terrorism whereby those unfamiliar with legume systematics start to publish formal names for potentially transient taxonomic groups. For example, *Deguelia* is placed in Lewis & al. (2005) in a series of nested clades labelled the “non-canavanine group”, the “core Millettieae”, the “MILLETTOID sens. strict.” clade, the “MILLETTOID sens. lat.” clade and the “50kb Inversion clade”. Such informal nomenclatural hierarchy permits communication about generic relationships without requiring formal recircumscription of the clade every time a genus is added or removed, which is important because revising subfamily and tribal classification in a large group such as the Leguminosae cannot be done piecemeal. Although it might be tempting to publish a new subfamily or tribe name when it becomes clear that a particular clade is well supported and needs a name, such an approach will result in ongoing instability and will likely need later revision. As our phylogenies attain greater resolution with adequate support, we will be in an appropriate position to publish formal names for many of these informally named clades, and to propose a revised classification system. Thus, a key issue is to decide when phylogenetic knowledge is solid enough to move forward. Given the current lack of resolution in several areas of the legume tree, and that it would be desirable to add nuclear DNA and morphological characters to our plastid-based phylogenetic trees, informal clade names will continue to be preferable in many cases.

One somewhat contentious issue with respect to nomenclature is what name to apply to the family. The LPWG follows Lewis & Schrire (2003) and Lewis & al. (2005: 1–3) who argued for the continued use of Leguminosae, because the name Fabaceae (the preferred name in APG III) is ambiguous, being used either for the whole family (as an alternative to Leguminosae), or as an alternative to the Papilionaceae when the Papilionoideae is recognised as a separate family. While all the evidence strongly supports the legumes as a monophyletic group, a number of authors and editors (e.g., as in *Flora Malesiana*, *Flora of the Venezuelan Guayana*, *Flora of Australia* and most recently the second edition of the *European Garden Flora*) continue to accept three separate families, leading to

confusion as to what the name Fabaceae encompasses. Recognition of three separate legume families is no longer tenable, and only when this is universally abandoned in favour of the one family position, will it be reasonable to refer to the whole family as the Fabaceae, thus satisfying those who prefer all families to have a standard “-aceae” ending.

**Inter- and intrageneric classification.** — When Bentham & Hooker (1865) published their *Genera Plantarum*, the legume family comprised 399 genera and ca. 6500 species. One hundred and sixteen years later when the first volume of the *Advances in Legume Systematics* series was published (Polhill & Raven, 1981), the number of genera had risen to 650 and species to ca. 18,000. When Polhill (1994), in the *Phytochemical Dictionary of the Leguminosae*, updated Polhill & Raven’s earlier (1981) classification, the number of genera had increased to 671. By the time Lewis & al. (2005) published *Legumes of the World* the number of accepted legume genera totalled 727, and a detailed tally of accepted species gave a total of nearly 19,325. One remaining impediment is that not all the combinations have been published for species that belong to segregate genera, so that a full species list based on the 727 genera in Lewis & al. (2005) is not yet available. Due to inadequate species sampling in phylogenies it is currently difficult to assign species correctly to each segregate genus split from large genera such as *Bauhinia*, *Caesalpinia* and *Acacia*. It is evident that formally publishing new combinations for species in the genus to which they belong is a task that needs to be completed for a number of key genera in the near future.

In the seven years since *Legumes of the World* (Lewis & al., 2005), there has been a continued proliferation of phylogenetic studies at both inter- and intrageneric levels. These are providing more solid, phylogeny-based classifications at lower taxonomic level in many legume groups, and have had the overall effect of continuing to increase the number of genera recognised. Recent phylogenetic studies of genera include *Hoffmannseggia* (Simpson & Ulibarri, 2006); *Phaseolus* (Delgado-Salinas & al., 2006); *Paloue* and related genera (Redden & Herendeen, 2006); *Phyllolobium* Fisch. (Zhang & Podlech, 2006); *Lotus* L. (Degtjareva & al., 2006, 2008); *Piptadenia* (Jobson & Luckow, 2007); *Prosopis* (Burghardt & Espert, 2007; Catalano & al., 2008); *Paraserianthes* I.C. Nielsen (Brown & al., 2011); *Platymiscium* Vogel (Saslis-Lagoudakis & al., 2008); *Mimosa* (Bessega & al., 2008; Simon & al., 2011; Bessega & Fortunato, 2011); *Bauhinia* s.l. (Sinou & al., 2009); *Chamaecrista* (Conceição & al., 2009); *Galactia*, *Camptosema* Hook. & Arn. and allied genera (Sede & al., 2008, 2009); *Centrolobium* Mart. ex Benth. (Pirie & al., 2009); *Lespedeza* Michx. (Nemoto & al., 2010); *Medicago* L. (Steele & al., 2010), *Ononis* L. (Turini & al., 2010); *Adenocarpus* DC. (Cubas & al., 2010); *Podalyria* Willd. (Schutte-Vlok & Van Wyk, 2011); *Lotononis* (DC.) Eckl. & Zeyh. (Boatwright & al., 2011); *Berlinia* Sol. ex Hook. f. (Mackinder & Pennington, 2011); American *Vigna* (Delgado-Salinas & al., 2011); *Pterocarpus* Jacq. (Saslis-Lagoudakis & al., 2011); *Anthyllis* L. (Degtjareva & al., 2012); and *Crotalaria* L. (Le Roux & al., 2011). In addition, there have been recent higher-level studies of generic groups or clades, such as the resin-producing Detarieae (Fougère-Danezan & al., 2007, 2010), *Acacia* and allies

(Brown & al., 2008; Murphy & al., 2010; Miller & al., 2011), *Swarzia* and allies (Torke & Schaal, 2008), phaseoloid legumes (Stefanović & al., 2009), *Lonchocarpus* and allies (Silva & al., 2012); tribes such as Hedysareae (Ahangarian & al., 2007), Crotonarieae (Boatwright & al., 2008a), Podalyrieae (Boatwright & al., 2008b), and Indigoferae (Schrire & al., 2009).

Since 2005, a number of genera have been placed in synonymy, e.g., the two monospecific genera *Ophiocarpus* (Bunge) Ikonn. and *Barnebyella* Podlech have been returned to *Astragalus* (M.F. Wojciechowski, pers. comm.), *Vaughania* S. Moore has been subsumed back into *Indigofera* (Schrire, 2008) and *Pellegriniodendron* (Harms) J. Léonard is now part of *Gilbertiodendron* J. Leonard (Estrella & al., 2012). There has been much published on tribe Loteae: all native New World species formerly placed in *Lotus* are segregated into four genera: *Hosackia* Douglas ex Lindl., *Acmispon* Raf., *Syrmatium* Vogel and *Ottleya* D.D. Sokoloff (e.g., Sokoloff & al., 2007), although Brouillet (2008) only recognises two: *Acmispon* (including *Syrmatium* and *Ottleya*) and *Hosackia*. Degtjareva & al. (2006, 2008) place *Dorycnium* Mill. and *Tetragonolobus* Scop. back in synonymy under *Lotus*, and Degtjareva & al. (2012) show that *Anthyllis* is paraphyletic with respect to *Hymenocarpus* Savi and thus place the latter into synonymy under *Anthyllis*. It is probable that *Securigera* DC. should be placed back into synonymy under *Coronilla* L. (Sokoloff, pers. comm.).

Other generic names have been resurrected from synonymy and are now considered to be accepted genera based on new analyses, e.g., *Acaciella* Britton & Rose (Rico Arce & Bachman, 2006); *Senegalia* (Seigler & al., 2006a), *Pityrocarpa* Britton & Rose (Jobson & Luckow, 2007), *Bionia* Mart. ex Benth. (Queiroz, 2008), *Leptolobium* (Rodrigues & Tozzi, 2008); *Fairchildia* Britton & Rose (Torke & Schaal, 2008); *Calobota* Eckl. & Zeyh. (Boatwright & al., 2009), *Cochlianthus* Trew and *Condylostylis* Piper (Delgado-Salinas & al., 2011), *Euchlora* Eckl. & Zeyh., *Listia* E. Mey. and *Leobordea* Del. (Boatwright & al., 2011), *Ototropis* Nees (Ohashi & Ohashi, 2012a), *Steinbachiella* Harms (Lewis & al., 2012) or taxa have been raised to generic rank from a previously described subdivision of a genus (e.g., *Leptospon* (Benth.) A. Delgado and *Sigmoidotropis* (Piper) A. Delgado (Delgado-Salinas & al., 2011). Some recently recognised segregates have required new generic names, e.g., *Guianodendron* (Rodrigues & Tozzi, 2006), *Mariosousa* (Seigler & al., 2006b), *Wiborgiella* Boatwr. & B.-E. Van Wyk (Boatwright & al., 2009), *Ladeania* A.N. Egan & Reveal (2009), *Ancistrotropis* A. Delgado (Delgado-Salinas & al., 2011), *Ezoloba* B.-E. Van Wyk & Boatwr. (Boatwright & al., 2011), *Helicotropis* A. Delgado (Delgado-Salinas & al., 2011), *Paragoodia* I. Thomps. (Thompson, 2011) and *Verdesmum* H. Ohashi & K. Ohashi (Ohashi & Ohashi, 2012b). In addition, *Heteroflorum* M. Sousa (2005) and *Tabaroa* (Queiroz & al., 2010) are newly discovered genera described from relatively recent field-collected specimens. Lastly, the correct generic name for *Calia* Terán & Berland has been shown to be *Dermatophyllum* (Gandhi & al., 2011). The current estimate is that the Leguminosae comprises 751 genera and ca. 19,500 species, but as might be expected, more new genera are anticipated soon (pers. comm., M. Luckow, B. Mackinder, H. Ohashi, L.P. de Queiroz).

## ■ FUTURE RESEARCH DIRECTIONS: BEYOND HIGHER-LEVEL CLASSIFICATION

The major focus of this paper has been to describe progress towards the development of a new, phylogeny-based, higher-level classification for legumes. We are optimistic that this goal will be achieved soon. To this end, as a community we are collaborating to:

1. complete genus-level taxon sampling, especially in Papilionoideae;
2. increase species-level taxon sampling to refine generic delimitation;
3. deploy next-generation sequencing techniques to gather more DNA sequence data, especially from the nuclear genome;
4. gather more comprehensive morphological datasets across the family;
5. test and optimize supermatrix and/or supertree approaches and procedures to integrate higher-level studies and the growing number of increasingly densely sampled species-level sequence datasets to improve poorly resolved nodes in the legume tree;
6. integrate molecular and morphological datasets.

Looking further forward, research in legume phylogeny must continue to address issues of generic delimitation. Aside from the prominent recent unravelling of *Acacia* s.l. many generic changes have been made over the last three decades in attempts to better reflect monophyly as synthesized by Lewis & al. (2005; and see above). In Mimosoideae, a recent example is *Piptadenia* (Jobson & Luckow, 2007), and there have been numerous generic changes in Ingeae (e.g., Barneby, 1998; Barneby & Grimes, 1996, 1997; Nielsen, 1981; Polhill, 1994; Rico Arce, 1992, 1999; Rico Arce & al., 1999; Villiers 2002). Notable examples in Papilionoideae include the recircumscription of *Sophora* s.l. (Sousa & Rudd, 1993), *Acosmium* s.l. (Rodrigues & Tozzi, 2007; Cardoso & al., 2012a), *Vigna* s.l. (Delgado-Salinas & al., 2011) and *Lonchocarpus* s.l. (Silva & al., 2012). Likewise, in Caesalpinioideae, the splitting of *Cassia* s.l. into *Cassia*, *Senna* and *Chamaecrista* by Irwin & Barneby (1981, 1982), though regarded by some as controversial (e.g., Gentry, 1993) has been amply confirmed by DNA sequence data (e.g., Bruneau & al., 2001; Marazzi & al., 2006; Conceição & al., 2009). Despite these advances, many cases of non-monophyletic genera revealed in recent phylogenies remain to be fully resolved, and this is holding back crucial monographic taxonomic research. Examples include ecologically and economically important genera such as *Bauhinia*, *Caesalpinia*, *Cynometra* and *Macrolobium* in Caesalpinioideae, and the species-rich *Millettia* Wight & Arn. and *Pultenaea* Sm. in Papilionoideae. Resolving these genus level questions will require much denser species-level taxon sampling allied to the identification of phylogenetically more informative plastid and nuclear DNA sequence loci.

Of particular importance are 41 large genera (6% of the Leguminosae) that have over 100 species each and account for a disproportionate percentage of species diversity in the family (12,430 species in total, equivalent to 64% of the species in the

family). Some of these genera have been neglected taxonomically and require more attention and full species-level revision such as *Dalbergia* L. f. (250 spp.), *Aeschynomene* L. (180 spp.), *Millettia* (150 spp.), *Ormosia* (130 spp.) and *Phanera* + *Schnella* (125 spp.). Very encouragingly, however, many large genera are the focus of active species-level phylogenetic studies, in some cases accompanied by taxonomic revisions—e.g., *Inga* (Richardson & al., 2001; Kursar & al., 2009; Dexter & al., 2010; Pennington & al., unpub. data); *Calliandra* (Souza & al., subm.); *Mimosa* (Simon & al., 2011; Bessega & Fortunato, 2011; Simon & al., 2011 and unpub. data); Australian *Acacia* (Murphy & al., 2010; Miller & al., 2011); *Prosopis* (Catalano & al., 2008); *Vachellia* & *Senegalia* (Miller & Bayer, 2003; Bouchenak-Khelladi & al., 2010; Kyalangaililwa & al., in press; Miller & Seigler, 2012); *Swartzia* (Torke & Schaal, 2008; Torke & Mansano, 2009), *Indigofera* (Schrire & al., 2009), *Lupinus* (Hughes & Eastwood, 2006), *Bauhinia* (Sinou & al., 2009).

Our hope in establishing the Legume Phylogeny Working Group is that this group of systematists with a common interest can provide a renewed focus for efforts to achieve completely sampled, robustly supported, well-resolved species-level phylogenies of the large legume genera and can help to couple them with improved higher-level family-wide phylogenetic estimates. If we succeed, the prospects for revisionary taxonomy and comparative biology of legumes are very bright indeed.

## ■ ACKNOWLEDGEMENTS

The meeting convened in 2010 in Phoenix, Arizona where LPWG was established was supported by NSF grant DEB-0829674 to M.J. Sanderson and Michelle McMahon, and a grant from Arizona State University (SoLS RTI) to M.F. Wojciechowski.

## ■ LITERATURE CITED

- Ahangarian, S., Kazempour Osaloo, Sh. & Maassoumi, A.A. 2007. Molecular phylogeny of the tribe Hedysareae with special reference to *Onobrychis* (Fabaceae) as inferred from nrDNA ITS sequences. *Iran. J. Bot.* 13: 64–74.
- Aliscioni, S., Bell, H.L., Besnard, G., Christin, P.-A., Columbus, J.T., Duvall, M.R., Edwards, E.J., Giussani, L., Hasenstab-Lehman, K., Hilu, K.W., Hodkinson, T.R., Ingram, A.L., Kellogg, E.A., Mashayekhi, S., Morrone, O., Osborne, C.P., Salamin, N., Schaefer, H., Spriggs, E., Smith, S.A. & Zuloaga, F. 2011. New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytol.* 193: 304–312.
- Alvarez, I. & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molec. Phylog. Evol.* 29: 417–434. [http://dx.doi.org/10.1016/S1055-7903\(03\)00208-2](http://dx.doi.org/10.1016/S1055-7903(03)00208-2)
- APG III (Bremer, B., Bremer, K., Chase, M.W., Fay, M.F., Reveal, J.L., Soltis, D.E., Soltis, P.S., Stevens, P.F., Anderberg, A.A., Moore, M.J., Olmstead, R.G., Rudall, P.J., Sytsma, K.J., Tank, D.C., Wurdack, K., Xiang, J.Q.-Y. & Zmarzty, S). 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161: 105–121. <http://dx.doi.org/10.1111/j.1095-8339.2009.00996.x>
- Arthofer, W., Schueler, S., Steiner, F.M. & Schlick-Steiner, B.C. 2010. Chloroplast DNA-based studies in molecular ecology may be compromised by nuclear-encoded plastid sequence. *Molec. Ecol.* 19: 3853–3856. <http://dx.doi.org/10.1111/j.1365-294X.2010.04787.x>
- Ashfield, T., Egan, A.N., Pfeil, B.E., Chen, N.W.G., Podicheti, R., Ratnaparkhe, M.B., Ameline-Torregrosa, C., Denny, R., Cannon, S., Doyle, J.J., Geffroy, V., Roe, B.A., Saghai Maroof, M.A., Young, N.D. & Innes, R.W. 2012. Evolution of a complex disease resistance gene cluster in diploid *Phaseolus* and tetraploid *Glycine*. *Pl. Physiol.* 159: 336–354. <http://dx.doi.org/10.1104/pp.112.195040>
- Banks, H., Klitgaard, B.B., Lewis, G.P., Crane, P.R. & Bruneau, A. 2003. Pollen and the systematics of tribes Caesalpinieae and Cassieae (Caesalpinioideae: Leguminosae). Pp. 95–122 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Barneby, R.C. 1991. Sensitivae censitae: A description of the genus *Mimosa* L. (Mimosaceae) in the New World. *Mem. New York Bot. Gard.* 65: 1–835.
- Barneby, R.C. 1998. Silk tree, guanacaste, monkey's earring: A generic system for the synandrous Mimosaceae of the Americas. Part III. *Calliandra*. *Mem. New York Bot. Gard.* 74: 1–223.
- Barneby, R.C. & Grimes, J.W. 1996. Silk tree, guanacaste, monkey's earring: A generic system for the synandrous Mimosaceae of the Americas. Part I. *Abarema*, *Albizia* and allies. *Mem. New York Bot. Gard.* 74: 1–292.
- Barneby, R.C. & Grimes, J.W. 1997. Silk tree, guanacaste, monkey's earring. A generic system for the synandrous Mimosaceae of the Americas. Part II. *Pithecellobium*, *Cojoba* and *Zygia*. *Mem. New York Bot. Gard.* 74: 1–149.
- Bello, M.A., Bruneau, A., Forest, F. & Hawkins, J.A. 2009. Elusive relationships within order Fabales: Phylogenetic analyses using *matK* and *rbcL* sequence data. *Syst. Bot.* 34: 102–114. <http://dx.doi.org/10.1600/036364409787602348>
- Bello, M.A., Hawkins, J.A. & Rudall, P.J. 2010. Floral ontogeny in Polygalaceae and its bearing on the homologies of keeled flowers in Fabales. *Int. J. Pl. Sci.* 171: 482–498. <http://dx.doi.org/10.1086/651945>
- Bello, M.A., Rudall, P.J. & Hawkins, J.A. 2012. Combined phylogenetic analyses reveal interfamilial relationships and patterns of flora evolution in the eudicot order Fabales. *Cladistics* 28: 393–421. <http://dx.doi.org/10.1111/j.1096-0031.2012.00392.x>
- Benthams, G. 1865. Leguminosae. Pp. 434–600 in: Benthams, G. & Hooker, J.D. (eds.), *Genera plantarum*, vol. 1. London: Reeve.
- Benthams, G. 1875. Revision of the suborder Mimoseae. *Trans. Linn. Soc. London* 30: 335–664, pl. 66–70.
- Benthams, G. & Hooker, J.D. 1865. *Genera plantarum*, vol. 1. London: Reeve.
- Bessega, C. & Fortunato, R.H. 2011. Section *Mimadenia*: Its phylogenetic relationships within the genus *Mimosa* (Leguminosae, Mimosoideae) using plastid *trnL-F* sequence data. *Austral. Syst. Bot.* 24: 104–110. <http://dx.doi.org/10.1071/SB10022>
- Bessega, C., Hopp, H.E. & Fortunato, R.H. 2008. Toward a phylogeny of *Mimosa* (Leguminosae: Mimosoideae): A preliminary analysis of southern South American species based on chloroplast DNA sequence. *Ann. Missouri Bot. Gard.* 95: 567–579. <http://dx.doi.org/10.3417/2006012>
- Bisby, F.A. 1993. Species diversity knowledge systems: The ILDIS prototype for legumes. *Ann. New York Acad. Sci.* 700: 159–164. <http://dx.doi.org/10.1111/j.1749-6632.1993.tb26316.x>
- Bisby, F.A., Buckingham, J. & Harborne, J.B. (eds.) 1994. *Phytochemical dictionary of the Leguminosae*. Cambridge: Chapman & Hall.
- Boatwright, J.S. & Van Wyk, B.-E. 2011. The systematic position of *Sophora inhambanensis* (Sophoreae, Fabaceae). *S. African J. Bot.* 77: 249–250. <http://dx.doi.org/10.1016/j.sajb.2010.05.004>
- Boatwright, J.S., Le Roux, M.M., Wink, M., Morozova, T. & Van Wyk, B.E. 2008a. Phylogenetic relationships of tribe Crotalarieae



- (Fabaceae) inferred from DNA sequences and morphology. *Syst. Bot.* 33: 752–761. <http://dx.doi.org/10.1600/036364408786500271>
- Boatwright, J.S., Savolainen, V., Van Wyk, B.-E., Schutte-Vlok, A.L., Forest, F. & Van der Bank, M.** 2008b. Systematic position of the anomalous genus *Cadia* and the phylogeny of the tribe Podalyriaceae (Fabaceae). *Syst. Bot.* 33: 133–147. <http://dx.doi.org/10.1600/036364408783887500>
- Boatwright, J.S., Tilney, P.M. & Van Wyk, B.-E.** 2009. The generic concept of *Lebeckia* (Crotalariaeae, Fabaceae): Reinstatement of the genus *Calobota* and the new genus *Wiborgiella*. *S. African J. Bot.* 75: 546–556. <http://dx.doi.org/10.1016/j.sajb.2009.06.001>
- Boatwright, J.S., Wink, M. & Van Wyk, B.-E.** 2011. The generic concept of *Lotononis* (Crotalariaeae, Fabaceae): Reinstatement of the genera *Euchlora*, *Leobordea* and *Listia* and the new genus *Ezoloba*. *Taxon* 60: 161–177.
- Bouchenak-Khelladi, Y., Maurin, O., Hurter, J. & Van der Bank, M.** 2010. The evolutionary history and biogeography of Mimosoideae (Leguminosae): An emphasis on African acacias. *Molec. Phylog. Evol.* 57: 495–508. <http://dx.doi.org/10.1016/j.ympev.2010.07.019>
- Breteler, F.J.** 1995. The boundary between Amherstieae and Detarieae (Caesalpinioideae). Pp. 53–61 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Brouillet, L.** 2008. The taxonomy of North American Loti (Fabaceae: Loteae): New names in *Acmispon* and *Hosackia*. *J. Bot. Res. Inst. Texas* 2(1): 387–394.
- Brown, G.K.** 2008. Systematics of the tribe Ingeae (Leguminosae-Mimosoideae) over the last 25 years. *Muelleria* 26: 27–42.
- Brown, G.K., Murphy, D.J., Miller, J.T. & Ladiges, P.Y.** 2008. *Acacia* s.s. and its relationship among tropical legumes, tribe Ingeae (Leguminosae: Mimosoideae). *Syst. Bot.* 33: 739–751. <http://dx.doi.org/10.1600/036364408786500136>
- Brown, G.K., Murphy, D.J. & Ladiges, P.Y.** 2011. Relationships of the Australo-Malesian genus *Paraserianthes* (Mimosoideae: Leguminosae) identifies the sister group of *Acacia* sensu stricto and two biogeographical tracks. *Cladistics* 27: 380–390. <http://dx.doi.org/10.1111/j.1096-0031.2011.00349.x>
- Bruneau, A., Forest, F., Herendeen, P.S., Klitgaard, B.B. & Lewis, G.P.** 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Syst. Bot.* 26: 487–514.
- Bruneau, A., Mercure, M., Lewis, G.P. & Herendeen, P.S.** 2008. Phylogenetic patterns and diversification in the caesalpinoid legumes. *Botany* 86: 697–718. <http://dx.doi.org/10.1139/B08-058>
- Burghardt, A.D. & Espert, S.M.** 2007. Phylogeny of *Prosopis* (Leguminosae) as shown by morphological and biochemical evidence. *Austral. Syst. Bot.* 20: 332–339. <http://dx.doi.org/10.1071/SB060403>
- Burleigh, J.G., Bansal, M.S., Eulenstein, O., Hartmann, S., Wehe, A. & Vision, T.J.** 2011. Genome-scale phylogenetics: Inferring the plant tree of life from 18,896 gene trees. *Syst. Biol.* 60: 117–25. <http://dx.doi.org/10.1093/sysbio/syq072>
- Candolle, A.P. de** 1825. *Prodromus systematis naturalis regni vegetabilis*, vol. 2(1). Paris: Treuttel and Würtz.
- Cardoso, D., Lima, H.C. de, Rodrigues, R.S., Queiroz, L.P. de, Pennington, R.T. & Lavin, M.** 2012a. The realignment of *Acosmium* sensu stricto with the dalbergioid clade (Leguminosae, Papilionoideae) reveals a proneness for independent evolution of radial floral symmetry among early branching papilionoid legumes. *Taxon* 61: 1057–1073.
- Cardoso, D., Lima, H.C. de, Rodrigues, R.S., Queiroz, L.P. de, Pennington, R.T. & Lavin, M.** 2012b. The Bowdichia clade of genistoid legumes: Phylogenetic analysis of combined molecular and morphological data and a recircumscription of *Diptotropis*. *Taxon* 61: 1074–1087.
- Catalano, S.T., Vilardi, J.C., Tosto, D. & Saidman, B.O.** 2008. Molecular phylogeny and diversification history of *Prosopis* (Fabaceae: Mimosoideae). *Biol. J. Linn. Soc.* 93: 621–640. <http://dx.doi.org/10.1111/j.1095-8312.2007.00907.x>
- Chappill, J.A.** 1995. Cladistic analysis of the Leguminosae: The development of an explicit phylogenetic hypothesis. Pp. 1–9 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Chappill, J.A. & Maslin, B.R.** 1995. A phylogenetic assessment of tribe Acacieae. Pp. 77–99 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Choi, H., Mun, J., Kim, D., Zhu, H., Baek, J., Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G.B., Young, N.D. & Cook, D.R.** 2004. Estimating genome conservation between crop and model legume species. *Proc. Natl. Acad. Sci. U.S.A.* 101: 15289–15294. <http://dx.doi.org/10.1073/pnas.0402251101>
- Choi, H., Luckow, M.A., Doyle, J. & Cook, D.R.** 2006. Development of nuclear gene-derived molecular markers linked to legume genetic maps. *Molec. Genet. Genomics* 276: 56–70. <http://dx.doi.org/10.1007/s00438-006-0118-8>
- Conceição, A.S., Queiroz, L.P. de, Lewis, G.P., Andrade, M.J.G., Schnadelbach, A.S. & Van den Berg, C.** 2009. Phylogeny of *Chamaecrista* Moench (Leguminosae-Caesalpinioideae) based on nuclear and chloroplast DNA regions. *Taxon* 58: 1168–1180.
- Cowan, R.S. & Polhill, R.M.** 1981a. Detarieae DC. (1825). Pp. 117–134 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Cowan, R.S. & Polhill, R.M.** 1981b. Amherstieae Benth. emend. J. Léon. (1957). Pp. 135–142 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Crisp, M.D. & Cook, L.G.** 2003. Phylogeny and embryo sac evolution in the endemic Australasian papilionoid tribes Mirbelieae and Bossiaeeae. Pp. 253–268 in: Bruneau, A. & Klitgaard, B.B. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Crisp, M.D., Gilmore, S. & Van Wyk, B.-E.** 2000. Molecular phylogeny of the genistoid tribes of papilionoid legumes. Pp. 249–276 in: Herendeen, P.S. & Bruneau, A. (eds.), *Advances in legume systematics*, part 9. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P.** 2009. Phylogenetic biome conservatism on a global scale. *Nature* 458: 754–756. <http://dx.doi.org/10.1038/nature07764>
- Cubas, P., Pardo, C., Tahiri, H. & Castroviejo, S.** 2010. Phylogeny and evolutionary diversification of *Adenocarpus* DC. (Leguminosae). *Taxon* 59: 720–732.
- Cullen, J., Knees, S.G. & Cubey, H.S. (eds.)** 2011. *The European Garden Flora*, vol. 3, *Angiospermae—Dicotyledons: Resedaceae to Cyrillaceae*, 2nd ed. Cambridge: Cambridge University Press.
- Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L.** 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Rev. Genet.* 12: 499–510. <http://dx.doi.org/10.1038/nrg3012>
- Degtjareva, G.V., Kramina, T.E., Sokoloff, D.D., Samigullin, T.H., Valiejo-Roman, C.M. & Antonov, A.S.** 2006. Phylogeny of the genus *Lotus* (Leguminosae, Loteae): Evidence from nrITS sequences and morphology. *Canad. J. Bot.* 84: 813–830. <http://dx.doi.org/10.1139/b06-035>
- Degtjareva, G.V., Kramina, T.E., Sokoloff, D.D., Samigullin, T.H., Sandra, G. & Valiejo-Roman, C.M.** 2008. New data on nrITS phylogeny of *Lotus* (Leguminosae, Loteae). *Wulfenia* 15: 35–49.
- Degtjareva, G.V., Valiejo-Roman, C.M., Samigullin, T.H., Guara-Requena, M. & Sokoloff, D.D.** 2012. Phylogenetics of *Anthyllis* (Leguminosae: Papilionoideae: Loteae): Partial incongruence between nuclear and plastid markers; A long branch problem and

- implications for morphological evolution. *Molec. Phylogen. Evol.* 62: 693–707. <http://dx.doi.org/10.1016/j.ympev.2011.11.010>
- Delgado-Salinas, A., Biber, R. & Lavin, M.** 2006. Phylogeny of the genus *Phaseolus* (Leguminosae): A recent diversification in an ancient landscape. *Syst. Bot.* 31: 779–791. <http://dx.doi.org/10.1600/036364406779695960>
- Delgado-Salinas, A., Thulin, M., Pasquet, R., Weeden, N. & Lavin, M.** 2011. *Vigna* (Leguminosae) sensu lato: The names and identities of the American segregate genera. *Amer. J. Bot.* 98: 1694–1715. <http://dx.doi.org/10.3733/ajb.1100069>
- Dexter, K.G., Pennington, T.D. & Cunningham, C.W.** 2010. Using DNA to assess errors in tropical tree identifications: How often are ecologists wrong and when does it matter? *Ecol. Monogr.* 80: 267–286. <http://dx.doi.org/10.1890/09-0267.1>
- Domenech-Ramírez, J.I. & Tucker, S.C.** 1990. Comparative ontogeny of the perianth in mimosoid legumes. *Amer. J. Bot.* 77: 624–635. <http://dx.doi.org/10.2307/2444809>
- Doyle, J.J.** 1992. Gene trees and species trees: Molecular systematics as one character taxonomy. *Syst. Bot.* 17: 144–163. <http://dx.doi.org/10.2307/2419070>
- Doyle, J.J.** 2011. Phylogenetic perspectives on the origins of nodulation. *Molec. Pl.-Microbe Interact.* 24: 1289–1295. <http://dx.doi.org/10.1094/MPMI-05-11-0114>
- Doyle, J.J. & Doyle, J.L.** 1999. Nuclear protein-coding genes in phylogeny reconstruction and homology assessment: Some examples from Leguminosae. Pp. 229–254 in: Hollingsworth, P., Batsman, R. & Gornall, R. (eds.), *Molecular systematics and plant evolution*. London: Taylor and Francis. <http://dx.doi.org/10.1201/9781439833278.ch12>
- Doyle, J.D., Doyle, J.L., Ballenger, J.A. & Palmer, J.D.** 1996. The distribution and phylogenetic significance of a 50-kb chloroplast DNA inversion in the flowering plant family Leguminosae. *Molec. Phylogen. Evol.* 5: 429–438. <http://dx.doi.org/10.1006/mpev.1996.0038>
- Doyle, J.J., Doyle, J.L., Ballenger, J.A., Dickson, E.E., Kajita, T. & Ohashi, H.** 1997. A phylogeny of the chloroplast gene *rbcl* in the Leguminosae: Taxonomic correlations and insights into the evolution of nodulation. *Amer. J. Bot.* 84: 541–554. <http://dx.doi.org/10.2307/2446030>
- Doyle, J.J., Chappill, J.A., Bailey, C.D. & Kajita, T.** 2000. Towards a comprehensive phylogeny of legumes: Evidence from *rbcl* sequences and non-molecular data. Pp. 1–20 in: Herendeen, P.S. & Bruneau, A. (eds.), *Advances in legume systematics*, part 9. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Duarte, J.M., Wall, P.K., Edger, P.P., Landherr, L.L., Ma, H., Pires, J.C., Leebens-Mack, J. & dePamphilis, C.W.** 2010. Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *B. M. C. Evol. Biol.* 10: 61. <http://dx.doi.org/10.1186/1471-2148-10-61>
- Edgar, R.C.** 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* 32: 1792–1797. <http://dx.doi.org/10.1093/nar/gkh340>
- Edwards, D. & Hawkins, J.A.** 2007. Are Cape floral clades the same age? Contemporaneous origins of two lineages in the genistoids s.l. (Fabaceae). *Molec. Phylogen. Evol.* 45: 952–970. <http://dx.doi.org/10.1016/j.ympev.2007.09.014>
- Egan, A.N. & Crandall, K.A.** 2008. Incorporating gaps as phylogenetic characters across eight DNA regions: Ramifications for North American Psoraleae (Leguminosae). *Molec. Phylogen. Evol.* 46: 532–546. <http://dx.doi.org/10.1016/j.ympev.2007.10.006>
- Egan, A.N. & Doyle, J.** 2010. A comparison of global, gene-specific, and relaxed clock methods in a comparative genomics framework: Dating the polyploid history of soybean (*Glycine max*). *Syst. Biol.* 59: 534–547. <http://dx.doi.org/10.1093/sysbio/syq041>
- Egan, A.N. & Reveal, J.L.** 2009. A new combination in *Pediomelum* and a new genus, *Ladeania*, from Western North America (Fabaceae, Psoraleae). *Novon* 19: 310–314. <http://dx.doi.org/10.3417/2008074>
- Elias, T.S.** 1981. Mimosoideae. Pp. 143–151 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. & Mitchell, S.E.** 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379. <http://dx.doi.org/10.1371/journal.pone.0019379>
- Estrella, M., Devesa, J.A. & Wieringa, J.J.A.** 2012. Morphological re-evaluation of the taxonomic status of the genus *Pellegriniodendron* (Harms) J. Leonard (Leguminosae-Caesalpinioideae-Detarieae) and its inclusion in *Gilbertiodendron* J. Leonard. *S. African J. Bot.* 78: 257–265. <http://dx.doi.org/10.1016/j.sajb.2011.04.006>
- Evans, J.A., Gasson, P.E. & Lewis, G.P.** 2006. Wood anatomy of the Mimosoideae (Leguminosae). *I. A. W. A. J. Suppl.* 5: 1–118.
- Feliner, G.N. & Rossello, J.A.** 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molec. Phylogen. Evol.* 44: 911–919. <http://dx.doi.org/10.1016/j.ympev.2007.01.013>
- Ferguson, I.K., Schrire, B.D. & Shepperson, R.** 1994. Pollen morphology of the tribe Sophoreae and relationships between subfamilies Caesalpinioideae and Papilionoideae (Leguminosae). Pp. 11–32 in: Ferguson, I.K. & Tucker, S.C. (eds.), *Advances in legume systematics*, part 6, *Structural botany*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Fortunato, R.H.** 2005. Mimosoideae. Pp. 184–185 in: Lewis, G.P., Schrire, B.D., Mackinder, B. & Lock, M. (eds.), *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Fougère-Danezan, M., Maumont, S. & Bruneau, A.** 2007. Relationships among resin producing Detarieae s.l. (Leguminosae) as inferred by molecular data. *Syst. Bot.* 32: 748–761. <http://dx.doi.org/10.1600/036364407783390755>
- Fougère-Danezan, M., Herendeen, P.S., Maumont, S. & Bruneau, A.** 2010. Morphological evolution in the variable resin-producing Detarieae (Leguminosae): Do morphological characters retain a phylogenetic signal? *Ann. Bot. (Oxford)* 105: 311–325. <http://dx.doi.org/10.1093/aob/mcp280>
- Gaeta, M.L., Yuyama, P.M.Y., Sartori, D., Fungaro, M.H.P. & Vanzela, A.L.L.** 2010. Occurrence and chromosome distribution of retroelements and NUPT sequences in *Copaifera langsdorffii* Desf. (Caesalpinioideae). *Chromosome Res.* 18: 515–524. <http://dx.doi.org/10.1007/s10577-010-9131-1>
- Gandhi, K.N., Vincent, M.A. & Reveal, J.L.** 2011. *Dermatophyllum*, the correct name for *Calia* (Fabaceae). *Phytoneuron* 57: 1–4.
- Gandolfo, M.A., Hermsen, E.J., Zamalao, M.C., Nixon, K.C., Gonzalez, C.C., Wilf, P., Cuneo, N.R. & Johnson, K.R.** 2011. Oldest known *Eucalyptus* macrofossils are from South America. *PLoS ONE* 6: e21084. <http://dx.doi.org/10.1371/journal.pone.0021084>
- Gasson, P., Trafford, C. & Matthews, B.** 2003. Wood anatomy of Caesalpinioideae. Pp. 63–93 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Gee, H.** 2003. Evolution: Ending incongruence. *Nature* 425: 782. <http://dx.doi.org/10.1038/425782a>
- Gentry, A.H.** 1993. *A field guide to the families and genera of woody plants of northwest South America (Colombia, Ecuador, Peru), with supplementary notes on herbaceous taxa*. Chicago: The University of Chicago Press.
- Goldblatt, P.** 1981. Cytology and the phylogeny of Leguminosae. Pp. 427–463 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 2. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Gómez-Acevedo, S., Rico Arce, L., Delgado-Salinas, A., Magallón, S. & Eguiarte, L.E.** 2010. Neotropical mutualism between *Acacia* and *Pseudomyrmex*: Phylogeny and divergence times. *Molec. Phylogen. Evol.* 56: 393–408. <http://dx.doi.org/10.1016/j.ympev.2010.03.018>



- González-Orozco, C.E., Laffan, S.W. & Miller, J.T.** 2011. Spatial distribution of species richness and endemism of the genus *Acacia* in Australia. *Austral. J. Bot.* 59: 601–609. <http://dx.doi.org/10.1071/BT11112>
- Greuter, W., McNeill, J., Barrie, F.R., Burdet, H.M., Demoulin, V., Filgueiras, T.S., Nicolson, D.H., Silva, P.C., Skog, J.E., Treharne, P., Turland, N.J. & Hawksworth, D.L. (eds.)** 2000. *International Code of Botanical Nomenclature (Saint Louis Code): Adopted by the Sixteenth International Botanical Congress St Louis, Missouri, July–August 1999*. Regnum Vegetabile 138. Königstein: Koeltz Scientific Books.
- Grimes, J.** 1995. Generic relationships of Mimosoideae tribe Ingeae, with emphasis on the New World *Pithecellobium* complex. Pp. 101–122 in: Crisp, M. & Doyle, J.J. (eds.), *Advances in legume systematics, part 7, Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Grimes, J.** 1999. Inflorescence morphology, heterochrony, and phylogeny in the mimosoid tribes Ingeae and Acaciaeae (Leguminosae: Mimosoideae). *Bot. Rev.* 65: 317–347. <http://dx.doi.org/10.1007/BF02857753>
- Grover, C.E., Salmon, A. & Wendel, J.F.** 2012. Targeted sequence capture as a powerful tool for evolutionary analysis. *Amer. J. Bot.* 99: 312–319. <http://dx.doi.org/10.3732/ajb.1100323>
- Guinet, P.** 1981. Mimosoideae: The characters of their pollen grains. Pp. 835–855 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics, part 2*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Haston, E.M., Lewis, G.P. & Hawkins, J.A.** 2003. A phylogenetic investigation of the *Peltophorum* group (Caesalpinieae: Leguminosae). Pp. 149–159 in: Bruneau, A. & Klitgaard, B.B. (eds.), *Advances in legume systematics, part 10, Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Haston, E.M., Lewis, G.P. & Hawkins, J.A.** 2005. A phylogenetic reappraisal of the *Peltophorum* group (Caesalpinieae: Leguminosae) based on the chloroplast *trnL-F*, *rbcL* and *rpS16* sequence data. *Amer. J. Bot.* 92: 1359–1371. <http://dx.doi.org/10.3732/ajb.92.8.1359>
- Heenan, P.B.** 2001. Relationships of *Streblorrhiza* (Fabaceae), an extinct monotypic genus from Phillip Island, South Pacific Ocean. *New Zealand J. Bot.* 39: 9–15. <http://dx.doi.org/10.1080/0028825X.2001.9512714>
- Heenan, P.B., Dawson, M.I. & Wagstaff, S.J.** 2004. The relationship of *Sophora* sect. *Edwardsia* (Fabaceae) to *Sophora tomentosa*, the type species of the genus *Sophora*, observed from DNA sequence data and morphological characters. *Bot. J. Linn. Soc.* 146: 439–446. <http://dx.doi.org/10.1111/j.1095-8339.2004.00348.x>
- Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G.W., Edgecombe, G.D., Martinez, P., Baguna, J., Bailly, X., Jondelius, U., Wiens, M., Muller, W.E.G., Seaver, E., Wheeler, W.C., Martindale, M.Q., Giribet, G. & Dunn, C.W.** 2009. Assessing of bilateral animals with scalable phylogenomic methods. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 276: 4261–4270. <http://dx.doi.org/10.1098/rspb.2009.0896>
- Herendeen, P.S.** 1995. Phylogenetic relationships of the tribe Swartzieae. Pp. 123–132 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics, part 7, Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Herendeen, P.S., Bruneau, A. & Lewis, G.P.** 2003a. Phylogenetic relationships in the caesalpinoid legumes: A preliminary analysis based on morphological and molecular data. Pp. 37–62 in: Bruneau, A. & Klitgaard, B.B. (eds.), *Advances in legume systematics, part 10, Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Herendeen, P.S., Lewis, G.P. & Bruneau, A.** 2003b. Floral morphology in caesalpinoid legumes: Testing the monophyly of the “Umtiza” clade. *Int. J. Pl. Sci.* 164: S393–S407. <http://dx.doi.org/10.1086/376881>
- Hiremath, P.J., Farmer, A., Cannon, S.B., Woodward, J., Kudapa, H., Tuteja, R., Kumar, A., BhanuPrakash, A., Mulaosmanovic, B., Gujaria, N., Krishnamurthy, L., Gaur, P.M., KaviKishor, P.B., Shah, T., Srinivasan, R., Lohse, M., Xiao, Y., Town, C.D., Cook, D.R., May, G.D. & Varshney, R.K.** 2011. Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Pl. Biotechnol. J.* 9: 922–931. <http://dx.doi.org/10.1111/j.1467-7652.2011.00625.x>
- Hoot, S.B. & Palmer, J.D.** 1994. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. *J. Molec. Evol.* 38: 274–281. <http://dx.doi.org/10.1007/BF00176089>
- Hu, J.-M. & Chang, S.-P.** 2003. Two new members of the *Callerya* group (Fabaceae) based on phylogenetic analysis of *rbcL* sequences: *Endosamara racemosa* (Roxb.) Geesink and *Callerya vasta* (Kosterm.) Schot. *Taiwania* 48(2): 118–128.
- Hu, J.-M., Lavin, M., Wojciechowski, M.F. & Sanderson, M.J.** 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trnK/matK* sequences and its implications for evolutionary patterns in Papilionoideae. *Amer. J. Bot.* 87: 418–430. <http://dx.doi.org/10.2307/2656638>
- Hu, J.-M., Lavin, M., Wojciechowski, M.F. & 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. *Syst. Bot.* 27: 722–733
- Hughes, C.E.** 1998. Monograph of *Leucaena* (Leguminosae: Mimosoideae). *Syst. Bot. Monogr.* 55: 1–244. <http://dx.doi.org/10.2307/25027876>
- Hughes, C.E. & Eastwood, R.** 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc. Natl. Acad. Sci. U.S.A.* 103: 10334–10339. <http://dx.doi.org/10.1073/pnas.0601928103>
- Hughes, C.E., Bailey, C.D. & Harris, S.A.** 2002. Divergent and reticulate species relationships in *Leucaena* (Fabaceae) inferred from multiple data sources: Insights into polyploidy origins and nrDNA polymorphism. *Amer. J. Bot.* 89: 1057–1073. <http://dx.doi.org/10.3732/ajb.89.7.1057>
- Hughes, C.E., Bailey, C.D., Krosnick, S. & Luckow, M.A.** 2003. Relationships among genera of the informal *Dichrostachys* and *Leucaena* groups (Mimosoideae) inferred from nuclear ribosomal ITS sequences. Pp. 221–238 in: Bruneau, A. & Klitgaard, B.B. (eds.), *Advances in legume systematics, part 10, Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Hughes, C.E., Lewis, G.P., Daza, Y., A. & Reynel, C.** 2004. *Marañonia*. A new dalbergioid legume genus (Leguminosae: Papilionoideae) from Peru. *Syst. Bot.* 29: 366–374. <http://dx.doi.org/10.1600/036364404774195557>
- Ilut, D.C., Coate, J.E., Luciano, A.K., Owens, T.G., May, G.D., Farmer, A. & Doyle, J.J.** 2012. A comparative transcriptomic study of an allotetraploid and its diploid progenitors illustrates the unique advantages and challenges of RNA-seq in plant species. *Amer. J. Bot.* 99: 383–396. <http://dx.doi.org/10.3732/ajb.1100312>
- Ireland, H., Pennington, R.T. & Preston, J.** 2000. Molecular systematics of the Swartzieae. Pp. 217–231 in: Herendeen, P.S. & Bruneau, A. (eds.), *Advances in legume systematics, part 9*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Irwin, H.S. & Barneby, R.C.** 1981. Cassieae. Pp. 97–106 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics, part 1*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Irwin, H.S. & Barneby, R.C.** 1982. The American Cassiinae. A synoptical revision of Leguminosae tribe Cassieae subtribe Cassiinae in the New World. *Mem. New York Bot. Gard.* 35: 1–918.
- Jobson, R.W. & Luckow, M.** 2007. Phylogenetic study of the genus *Piptadenia* (Mimosoideae: Leguminosae) using plastid *trnL-F* and *trnK/matK* sequence data. *Syst. Bot.* 32: 569–575. <http://dx.doi.org/10.1600/036364407782250544>
- Jussieu, A.L.** 1789. *Genera plantarum*. Paris: apud viduam Herissant et Theophilum Barrois.



- Kajita, T., Ohashi, H., Tateishi, Y., Bailey, C.D. & Doyle, J.J.** 2001. *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. *Syst. Bot.* 26: 515–536.
- Käss, E. & Wink, M.** 1995. Molecular evolution of phylogeny of the Papilionoideae (family Leguminosae): *rbcL* gene sequences versus chemical taxonomy. *Bot. Acta* 108: 149–162.
- Käss, E. & Wink, M.** 1996. Molecular evolution of the Leguminosae: Phylogeny of the three subfamilies based on *rbcL*-sequences. *Biochem. Syst. Ecol.* 24: 365–378. [http://dx.doi.org/10.1016/0305-1978\(96\)00032-4](http://dx.doi.org/10.1016/0305-1978(96)00032-4)
- Käss, E. & Wink, M.** 1997. Phylogenetic relationships in the Papilionoideae (family Leguminosae) based on nucleotide sequences of cpDNA (*rbcL*) and ncDNA (ITS 1 and 2). *Molec. Phylog. Evol.* 8: 65–88. <http://dx.doi.org/10.1006/mpev.1997.0410>
- Kaur, S., Cogan, N.O., Pembleton, L.W., Shinozuka, M., Savin, K.W., Materne, M. & Forster, J.W.** 2011. Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. *B. M. C. Genomics* 12: 265, doi: 10.1186/1471-2164-12-265. <http://dx.doi.org/10.1186/1471-2164-12-265>
- Kearney, M.** 2002. Fragmentary taxa, missing data, and ambiguity: Mistaken assumptions and conclusions. *Syst. Biol.* 51: 369–381. <http://dx.doi.org/10.1080/10635150252899824>
- Kite, G.C. & Pennington, R.T.** 2003. Quinolizidine alkaloid status of *Styphnolobium* and *Cladrastis* (Leguminosae). *Biochem. Syst. Ecol.* 31: 1409–1416. [http://dx.doi.org/10.1016/S0305-1978\(03\)00118-2](http://dx.doi.org/10.1016/S0305-1978(03)00118-2)
- Koopman, M.M. & Baum, D.A.** 2010. Isolating nuclear genes and identifying lineages without monophyly: An example of closely related species from southern Madagascar. *Int. J. Pl. Sci.* 171: 761–771. <http://dx.doi.org/10.1086/654847>
- Kursar, T., Dexter, K.G., Lokvam, J., Pennington, R.T., Richardson, J.E., Weber, M.G., Murakami, E.T., Drake, C., McGregor, R. & Coley, P.D.** 2009. The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proc. Natl. Acad. Sci. U.S.A.* 106: 18073–18078. <http://dx.doi.org/10.1073/pnas.0904786106>
- Kyalangalilwa, B., Boatwright, J.S., Daru, B.H., Maurin, O. & Van der Bank, M.** In press. Phylogenetic position and revised classification of *Acacia* s. l. (Fabaceae: Mimosoideae) in Africa, including new combinations in *Vachellia* and *Senegalia*. *Bot. J. Linn. Soc.*
- Lackey, J.A.** 1981. Phaseoleae. Pp. 301–327 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Langenheim, J.H.** 2003. *Plant resins: Chemistry, evolution, ecology, and ethnobotany*. Portland: Timber Press.
- Lavin, M., Doyle, J.J. & Palmer, J.D.** 1990. Systematic and evolutionary significance of the loss of the large chloroplast DNA inverted repeat in the family Leguminosae. *Evolution* 44: 390–402. <http://dx.doi.org/10.2307/2409416>
- Lavin, M., Eshbaugh, E., Hu, J., Mathews, S. & Sharrock, R.A.** 1998. Monophyletic subgroups of the tribe Millettieae (Leguminosae) as revealed by phytochrome nucleotide sequence data. *Amer. J. Bot.* 85: 412–433. <http://dx.doi.org/10.2307/2446334>
- Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., Lima, H.C. de & Gasson, P.E.** 2001. The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. *Amer. J. Bot.* 88: 503–533. <http://dx.doi.org/10.2307/2657116>
- Lavin, M., Wojciechowski, M.F., Gasson, P., Hughes, C. & Wheeler, E.** 2003. Phylogeny of robinoid legumes (Fabaceae) revisited: *Coursetia* and *Gliricidia* circumscribed, and a biogeographical appraisal of the Caribbean endemics. *Syst. Bot.* 28: 387–409.
- Lavin, M., Herendeen, P.S. & Wojciechowski, M.F.** 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Syst. Biol.* 54: 575–594. <http://dx.doi.org/10.1080/10635150590947131>
- Le Roux, M.M., Van Wyk, B.-E., Boatwright, J.S. & Tilney, P.M.** 2011. The systematic significance of morphological and anatomical variation in the fruits of *Crotalaria* and related genera of the tribe Crotalarieae (Fabaceae). *Bot. J. Linn. Soc.* 165: 84–106. <http://dx.doi.org/10.1111/j.1095-8339.2010.01098.x>
- Lee, E.K., Cibrian-Jaramillo, A., Kolokotronis, S., Katari, M.S., Stamatakis, A., Ott, M., Chiu, J.C., Little, D.P., Stevenson, D.W., McCombie, W.R., Martienssen, R.A., Coruzzi, G. & DeSalle, R.** 2011. A functional phylogenomic view of the seed plants. *PLoS Genet.* 7: e1002411. <http://dx.doi.org/10.1371/journal.pgen.1002411>
- Léonard, J.** 1957. Genera des Cynometreae et des Amherstieae africaines (Leguminosae-Caesalpinioideae). *Mém. Cl. Sci. Acad. Roy. Belgique (8vo)* 30: 1–314.
- Lewis, G.P.** 2005. Caesalpinieae. Pp. 127–161 in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P. & Elias, T.S.** 1981. Mimosae. Pp. 155–168 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P. & Forest, F.** 2005. Cercideae. Pp. 57–67 in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P. & Schrire, B.D.** 2003. Leguminosae or Fabaceae? Pp. 1–3 in: Klitgaard, B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.)** 2005. *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lewis, G.P., Wood, J.R.I. & Lavin, M.** 2012. *Steinbachiella* (Leguminosae: Papilionoideae: Dalbergieae), endemic to Bolivia, is reinstated as an accepted genus. *Kew Bull.* 67: 789–796. <http://dx.doi.org/10.1007/s12225-012-9415-z>
- Lima, H.C. de.** 1990. Tribo Dalbergieae (Leguminosae-Papilionoideae) — Morfologia dos frutos, sementes e plântulas e sua aplicação na sistemática. *Arch. Jard. Bot. Rio de Janeiro* 30: 1–42.
- Liston, A.** 1995. Use of the polymerase chain reaction to survey for the loss of the inverted repeat in the legume chloroplast genome. Pp. 31–40 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Lorence, D.H. & Wood, K.R.** 1994. *Kanaloa*: A new genus of Fabaceae (Mimosoideae) from Hawaii. *Novon* 4: 137–145. <http://dx.doi.org/10.2307/3391582>
- Loytynoja, A. & Goldman, N.** 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. U.S.A.* 102: 10557–10562. <http://dx.doi.org/10.1073/pnas.0409137102>
- Loytynoja, A. & Goldman, N.** 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320: 1632–1635. <http://dx.doi.org/10.1126/science.1158395>
- Luckow, M.** 1993. A monograph of the genus *Desmanthus* (Leguminosae: Mimosoideae). *Syst. Bot. Monogr.* 38: 1–166. <http://dx.doi.org/10.2307/25027822>
- Luckow, M.** 1995. A phylogenetic analysis of the *Dichrostachys* group (Mimosoideae: Mimosae). Pp. 63–75 in: Crisp, M.D. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Luckow, M. & Grimes, J.** 1997. A survey of anther glands in the mimosoid tribes Parkieae and Mimosae. *Amer. J. Bot.* 84: 285–297. <http://dx.doi.org/10.2307/2446002>
- Luckow, M. & Hopkins, H.C.F.** 1995. A cladistic analysis of *Parkia* (Leguminosae: Mimosoideae). *Amer. J. Bot.* 82: 1300–1320. <http://dx.doi.org/10.2307/2446253>
- Luckow, M., White, P.J. & Bruneau, A.** 2000. Relationships among the basal genera of mimosoid legumes. Pp. 165–180 in: Herendeen, P.S. & Bruneau, A. (eds.), *Advances in legume systematics*, part 9. Richmond, U.K.: Royal Botanic Gardens, Kew.

- Luckow, M., Miller, J.T., Murphy, D.J. & Livshultz, T.** 2003. A phylogenetic analysis of the Mimosoideae (Leguminosae) based on chloroplast DNA sequence data. Pp. 197–220 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Luckow, M., Fortunato, R.H., Sede, S. & Livshultz, T.** 2005. The phylogenetic affinities of two mysterious monotypic mimosoids from southern South America. *Syst. Bot.* 30: 585–602. <http://dx.doi.org/10.1600/0363644054782206>
- Mackinder, B.** 2005. Tribe Detarieae. Pp. 68–109 in: Lewis, G., Schrire, B., Mackinder, B. & Lock, M. (eds.), *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Mackinder, B.A. & Pennington, R.T.** 2011. Monograph of *Berlinia* (Leguminosae). *Syst. Bot. Monogr.* 91: 1–117.
- Mackinder, B.A., Wieringa, J.J., Lunenburg, I. & Banks, H.** 2010. Challenges and progress in clarifying the generic limits of *Talbotiella* and *Hymenostegia* (Detarieae: Caesalpinioideae: Leguminosae). Pp. 43–56 in: Van der Burgt, X. (ed.), *Proceedings of the 18th AETFAT Congress*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Magallón, S. & Sanderson, M.J.** 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55: 1762–1780.
- Mansano, V., Bittrich, V., Tozzi, A.M.G. de A. & Souza, A.P. de.** 2004. Composition of the *Lecointea* clade (Leguminosae, Papilionoideae, Swartzieae), a re-evaluation based on combined evidence from morphology and molecular data. *Taxon* 53: 1007–1018. <http://dx.doi.org/10.2307/4135566>
- Manzanilla, V. & Bruneau, A.** 2012. Phylogeny reconstruction in the Caesalpinieae grade (Leguminosae) based on duplicated copies of the sucrose synthase gene and plastid markers. *Molec. Phylog. Evol.* 65: 149–162. <http://dx.doi.org/10.1016/j.ympev.2012.05.035>
- Mao, K., Milne, R.I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R.R. & Renner, S.S.** 2012. The distribution of living Cupressaceae reflects the breakup of Pangea. *Proc. Natl. Acad. Sci. U.S.A.* 109: 7793–7798. <http://dx.doi.org/10.1073/pnas.1114319109>
- Marazzi, B. & Sanderson, M.J.** 2010. Large-scale patterns of diversification in the widespread legume genus *Senna* and the evolutionary role of extrafloral nectaries. *Evolution* 64: 3570–3592. <http://dx.doi.org/10.1111/j.1558-5646.2010.01086.x>
- Marazzi, B., Endress, P.K., Queiroz, L.P. de & Conti, E.** 2006. Phylogenetic relationships within *Senna* (Leguminosae, Cassiinae) based on three chloroplast DNA regions: Patterns in the evolution of floral symmetry and extrafloral nectaries. *Amer. J. Bot.* 93: 288–303. <http://dx.doi.org/10.3732/ajb.93.2.288>
- Marazzi, B., Cécile, A., Simon, M.F., Luckow, M., Delgado Salinas, A. & Sanderson, M.J.** 2012. Locating evolutionary precursors on a phylogenetic tree. *Evolution* 66: 3918–3930. <http://dx.doi.org/10.1111/j.1558-5646.2012.01720.x>
- Marazzi, B., Conti, E., Sanderson, M.J., McMahon, M.M., Bronstein, J.L.** 2013. Diversity and evolution of a trait mediating ant-plant interactions with ants: Insights from extrafloral nectaries in *Senna* (Leguminosae). *Ann. Bot. (Oxford)*. <http://dx.doi.org/10.1093/aob/mcs226>
- Maslin, B.R., Miller, J.T. & Seigler, D.S.** 2003. Overview of the generic status of *Acacia* (Leguminosae: Mimosoideae). *Austral. Syst. Bot.* 16: 1–18. <http://dx.doi.org/10.1071/SB02008>
- McMahon, M.M. & Hufford, L.** 2004. Phylogeny of Amorpheae (Fabaceae: Papilionoideae). *Amer. J. Bot.* 91: 1219–1230. <http://dx.doi.org/10.3732/ajb.91.8.1219>
- McMahon, M.M. & Sanderson, M.J.** 2006. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. *Syst. Biol.* 55: 818–836. <http://dx.doi.org/10.1080/10635150600999150>
- Miller, J.T. & Bayer, R.J.** 2001. Molecular phylogenetics of *Acacia* (Fabaceae: Mimosoideae) based on the chloroplast *matK* coding sequence and flanking *trnK* intron spacer regions. *Amer. J. Bot.* 88: 697–705. <http://dx.doi.org/10.2307/2657071>
- Miller, J.T. & R.J. Bayer.** 2003. Molecular phylogenetics of *Acacia* subgenera *Acacia* and *Aculeiferum* (Fabaceae: Mimosoideae) based on the chloroplast *matK* coding sequence and flanking *trnK* intron spacer regions. *Austral. Syst. Bot.* 16: 27–33. <http://dx.doi.org/10.1071/SB01035>
- Miller, J.T. & Seigler, D.S.** 2012. Evolutionary and taxonomic relationships of *Acacia* s.l. (Leguminosae: Mimosoideae). *Austral. Syst. Bot.* 25: 217–224. <http://dx.doi.org/10.1071/SB11042>
- Miller, J.T., Grimes, J.W., Murphy, D.J., Bayer, R.J. & Ladiges, P.Y.** 2003. A phylogenetic analysis of Acacieae and Ingeae (Mimosoideae: Fabaceae) based on *trnK*, *matK*, *psbA-trnH* and *trnL-trnF* sequence data. *Syst. Bot.* 28: 558–566.
- Miller, J.T., Murphy, D.M., Brown, G.K., Richardson, D.M. & González-Orozco, C.E.** 2011. The evolution and phylogenetic placement of invasive *Acacia* species. *Diversity & Distrib.* 17: 848–860. <http://dx.doi.org/10.1111/j.1472-4642.2011.00780.x>
- Moore, G., Smith, G.F., Figueiredo, E., Demissew, S., Lewis, G., Schrire, B., Rico, L. & Van Wyk, A.E.** 2010. *Acacia*, the 2011 nomenclature section in Melbourne, and beyond. *Taxon* 59: 1188–1195.
- Murphy, D.J.** 2008. A review of the classification of *Acacia* (Leguminosae, Mimosoideae). *Muelleria* 26: 10–26.
- Murphy, D.J., Brown, G.K., Miller, J.T. & Ladiges, P.Y.** 2010. Molecular phylogeny of *Acacia* Mill. (Mimosoideae: Leguminosae): Evidence for major clades and informal classification. *Taxon* 59: 7–19.
- Nei, M. & Rooney, A.P.** 2005. Concerted and birth-and-death evolution of multigene families. *Annual Rev. Genet.* 39: 121–152. <http://dx.doi.org/10.1146/annurev.genet.39.073003.112240>
- Nemoto, T., Yokoyama, J., Fukuda, T., Iokawa, Y. & Ohashi, H.** 2010. Phylogeny of *Lespedeza* (Leguminosae) based on chloroplast *trnL-trnF* sequences. *J. Jap. Bot.* 85: 213–229.
- Nielsen, I.** 1981. Tribe 5. Ingeae. Pp. 173–190 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Nores, M.J., Simpson, B.B., Hick, P., Anton, A.M. & Fortunato, R.H.** 2012. The phylogenetic relationships of four monospecific caesalpinoids (Leguminosae) endemic to southern South America. *Taxon* 61: 790–802.
- Ohashi, H. & Ohashi, K.** 2012a. *Ototropis*, a genus separated from *Desmodium* (Leguminosae). *J. Jap. Bot.* 87: 108–118.
- Ohashi, H. & Ohashi, K.** 2012b. *Verdesmum*, a new genus of Leguminosae: tribe Desmodieae. *J. Jap. Bot.* 87: 299–306
- Orthia, L.A., Cook, L.G. & Crisp, M.D.** 2005. Generic delimitation and phylogenetic uncertainty: An example from a group that has undergone an explosive radiation. *Austral. Syst. Bot.* 18: 41–47. <http://dx.doi.org/10.1071/SB04016>
- Page, R.D.M. & Charleston, M.A.** 1997. From gene to organismal phylogeny: Reconciled trees and the gene tree/species tree problem. *Molec. Phylog. Evol.* 7: 231–240. <http://dx.doi.org/10.1006/mpev.1996.0390>
- Palmer, J.D. & Thompson, W.F.** 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* 29: 537–550. [http://dx.doi.org/10.1016/0092-8674\(82\)90170-2](http://dx.doi.org/10.1016/0092-8674(82)90170-2)
- Pazos-Navarro, M., Dabauza, M., Correal, E., Hanson, K., Teakle, N., Real, D. & Nelson, M.N.** 2011. Next generation DNA sequencing technology delivers valuable genetic markers for the genomic orphan legume species, *Bituminaria bituminosa*. *B. M. C. Genet.* 12: 104. <http://dx.doi.org/10.1186/1471-2156-12-104>
- Penn, O., Privman, E., Landan, G., Graur, D. & Pupko, T.** 2010. An alignment confidence score capturing robustness to guide tree uncertainty. *Molec. Biol. Evol.* 27: 1759–1767. <http://dx.doi.org/10.1093/molbev/msq066>
- Pennington, T.D.** 1997. *The genus Inga: Botany*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Pennington, R.T., Klitgaard, B.B., Ireland, H. & Lavin, M.** 2000. New insights into floral evolution of basal Papilionoideae from



- molecular phylogenies. Pp. 233–248 in: Herendeen, P.S. & Bruneau, A. (eds.), *Advances in legume systematics*, part 9. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Pennington, R.T., Lavin, M., Ireland, H., Klitgaard, B.B., Preston, J. & Hu, J.-M.** 2001. Phylogenetic relationships of basal papilionoid legumes based upon sequences of the chloroplast *trnL* intron. *Syst. Bot.* 26: 537–556.
- Pennington, R.T., Lavin, M. & Oliveira-Filho, A.** 2009. Plant diversity, evolution and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annual Rev. Ecol. Evol. Syst.* 40: 437–457. <http://dx.doi.org/10.1146/annurev.ecolsys.110308.120327>
- Peters, R.S., Meyer, B., Krogmann, L., Borner, J., Meusemann, K., Schutte, K., Niehuis, O. & Misof, B.** 2011. The taming of an impossible child: A standardized all-in approach to the phylogeny of Hymenoptera using public database sequences. *B. M. C. Biol.* 9: 55. <http://dx.doi.org/10.1186/1741-7007-9-55>
- Pirie, M.D., Klitgaard, B.B. & Pennington, R.T.** 2009. Revision and biogeography of *Centrolobium* (Leguminosae, Papilionoideae). *Syst. Bot.* 34: 345–359. <http://dx.doi.org/10.1600/036364409788606262>
- Poggio, L., Espert, S.M. & Fortunato, R.H.** 2008. Citogenética evolutiva en Leguminosae Americanas. *Rodriguésia* 59: 423–433.
- Polhill, R.M.** 1994. Classification of the Leguminosae and complete synopsis of legume genera. Pp. xxxv–lviii in: Bisby, F.A., Buckingham, J. & Harborne, J.B. (eds.), *Phytochemical dictionary of the Leguminosae*. Cambridge: Chapman & Hall.
- Polhill, R.M. & Raven, P.H. (eds.)** 1981. *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Polhill, R.M. & Vidal, J.E.** 1981. Caesalpinieae. Pp. 81–96 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Prenner, G. & Klitgaard, B.B.** 2008. Towards unlocking the deep nodes of Leguminosae: Floral development and morphology of the enigmatic *Duparquetia orchidacea* (Leguminosae, Caesalpinioideae). *Amer. J. Bot.* 95: 1349–1365. <http://dx.doi.org/10.3732/ajb.0800199>
- Proost, S., Van Bel, M., Sterck, L., Billiau, K., Van Parys, T., Van de Peer, Y. & Vandepoel, K.** 2009. PLAZA: A comparative genomics resource to study gene and genome evolution in plants. *Pl. Cell* 21: 3718–3731. <http://dx.doi.org/10.1105/tpc.109.071506>
- Queiroz, L.P. de.** 2008. Re-establishment, synopsis and new combinations in the genus *Bionia* Mart. ex Benth. (Leguminosae: Papilionoideae). *Neodiversity* 3: 13–18.
- Queiroz, L.P. de, Fortunato, R.H. & Giuliatti, A.M.** 2003. Phylogeny of the subtribe Diocleinae (Leguminosae: Papilionoideae) based on morphological characters. Pp. 105–128 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Queiroz, L.P. de, Lewis, G.P. & Wojciechowski, M.F.** 2010. *Tabaroa*, a new genus of Leguminosae tribe Brongniartieae from Brazil. *Kew Bull.* 65: 189–203. <http://dx.doi.org/10.1007/s12225-010-9202-7>
- Ratter, J.A., Bridgewater, S. & Ribeiro, J.F.** 2003. Analysis of the floristic composition of the Brazilian cerrado vegetation III: Comparison of the woody vegetation of 376 areas. *Edinburgh J. Bot.* 60: 57–109. <http://dx.doi.org/10.1017/S0960428603000064>
- Redden, K.M. & Herendeen, P.S.** 2006. Morphology and phylogenetic analysis of *Paloue* and related genera in the *Brownea* clade (Caesalpinioideae: Leguminosae). *Int. J. Pl. Sci.* 167: 1229–1246. <http://dx.doi.org/10.1086/508065>
- Redden, K.M., Herendeen, P.S., Wurdack, K.J. & Bruneau, A.** 2010. Phylogenetic relationships of the Northeastern South American *Brownea* clade of tribe Detarieae (Leguminosae: Caesalpinioideae) based on morphology and molecular data. *Syst. Bot.* 35: 524–533. <http://dx.doi.org/10.1600/036364410792495863>
- Rice, P., Longden, I. & Bleasby, A.** 2000. EMBOSS: The European molecular biology open software suite. *Trends Genet.* 16: 276–277. [http://dx.doi.org/10.1016/S0168-9525\(00\)02024-2](http://dx.doi.org/10.1016/S0168-9525(00)02024-2)
- Richardson, D.M. & Rejmánek, M.** 2011. Trees and shrubs as invasive species—a global review. *Diversity & Distrib.* 17: 788–809. <http://dx.doi.org/10.1111/j.1472-4642.2011.00782.x>
- Richardson, D.M., Carruthers, J., Hui, C., Impson, F.A.C., Miller, J.T., Robertson, M.P., Rouget, M., Le Roux, J.J. & Wilson, J.R.U.** 2011. Human-mediated introductions of Australian *Acacia* species—a global experiment in biogeography. *Diversity & Distrib.* 17: 771–787. <http://dx.doi.org/10.1111/j.1472-4642.2011.00824.x>
- Richardson, J.E., Pennington, R.T., Pennington, T.D. & Hollingsworth, P.H.** 2001. Rapid diversification of a species-rich group of Neotropical rain forest trees. *Science* 293: 2242–2245. <http://dx.doi.org/10.1126/science.1061421>
- Rico Arce, M. de L.** 1991. New species combinations and synonyms for *Zygia*, *Marmaroxylon* and *Pithecellobium* (Leguminosae, Mimosoideae, Ingeae). *Kew Bull.* 46: 493–521. <http://dx.doi.org/10.2307/4110539>
- Rico Arce, M. de L.** 1992. New chromosome counts in Neotropical *Albizia*, *Havardia* and *Pithecellobium* and a new combination for *Albizia* (Leguminosae - Mimosoideae - Ingeae). *Bot. J. Linn. Soc.* 108: 269–274. <http://dx.doi.org/10.1111/j.1095-8339.1992.tb00243.x>
- Rico Arce, M. de L.** 1999. New combinations in Mimosaceae. *Novon* 9: 554–556. <http://dx.doi.org/10.2307/3392164>
- Rico Arce, M. de L. & Bachman, S.** 2006. A taxonomic revision of *Acaciella* (Leguminosae, Mimosoideae). *Anales Jard. Bot. Madrid* 63: 189–244.
- Rico Arce, M. de L. & Banks, H.** 2001. A preliminary survey of pollen and other morphological characters in neotropical *Acacia* subgenus *Aculeiferum* (Leguminosae: Mimosoideae). *Bot. J. Linn. Soc.* 135: 263–270.
- Rico Arce, M. de L., Mario Sousa, S. & Sara Fuentes, S.** 1999. *Guinetia*: A new genus in the tribe Ingeae (Leguminosae: Mimosoideae) from Mexico. *Kew Bull.* 54: 975–981. <http://dx.doi.org/10.2307/4111177>
- Rieseberg, L.H. & Soltis, D.E.** 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Pl.* 5: 65–84.
- Riley-Hulting, E.T., Delgado-Salinas, A. & Lavin, M.** 2004. Phylogenetic systematics of *Strophostyles* (Fabaceae): A North American temperate genus within a Neotropical diversification. *Syst. Bot.* 29: 627–653. <http://dx.doi.org/10.1600/0363644041744464>
- Rivers, M.C., Brummitt, N.A., Nic Lughadha, E. & Meagher, T.R.** 2011. Genetic variation in *Delonix* s.l. (Leguminosae) in Madagascar revealed by AFLPs: Fragmentation, conservation status and taxonomy. *Conservation Genet.* 12: 1333–1344. <http://dx.doi.org/10.1007/s10592-011-0234-9>
- Rodrigues, R.S. & Tozzi, A.M.G. de A.** 2006. *Guianodendron*, a new genus of Leguminosae (Papilionoideae) from South America. *Novon* 16: 129–132. [http://dx.doi.org/10.3417/1055-3177\(2006\)16\[129:GANGOL\]2.0.CO;2](http://dx.doi.org/10.3417/1055-3177(2006)16[129:GANGOL]2.0.CO;2)
- Rodrigues, R.S. & Tozzi, A.M.G. de A.** 2007. Morphological analysis and re-examination of the taxonomic circumscription of *Acosmium* (Leguminosae, Papilionoideae, Sophoreae). *Taxon* 56: 439–452.
- Rodrigues, R.S. & Tozzi, A.M.G. de A.** 2008. Reinstatement of the name *Leptolobium* Vogel (Leguminosae, Papilionoideae, Sophoreae). *Taxon* 57: 980–984.
- Ross, J.H. & Crisp, M.D.** 2005. Tribe Brongniartieae. Pp. 252–259 in: Lewis, G.P., Schrire, B.D. & Mackinder, B. (eds.), *Legumes of the world*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Russell, A.** 2004. *The distribution and taxonomic significance of a 50kb inversion in the Papilionoideae (Leguminosae)*. M.Sc. thesis, University of Edinburgh, Edinburgh, U.K.
- Sakiroglu, M., Sherman-Broyles, S., Story, A., Doyle, J.J., Moore, K.J. & Brummer, E.C.** 2012. Patterns of linkage disequilibrium and association mapping in diploid alfalfa (*M. sativa* L.). *Theor. Appl. Genet.* 125: 577–59. <http://dx.doi.org/10.1007/s00122-012-1854-2>



- Sanderson, M.J. 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19: 301–302. <http://dx.doi.org/10.1093/bioinformatics/19.2.301>
- Sanderson, M.J. & McMahon, M.M. 2007. Inferring angiosperm phylogeny from EST data with widespread gene duplication. *B. M. C. Evol. Biol.* 7: S3. <http://dx.doi.org/10.1186/1471-2148-7-S1-S3>
- Sanderson, M.J. & Wojciechowski, M.F. 1996. Diversification rates in a temperate legume clade: Are there “so many species” of *Astragalus* (Fabaceae)? *Amer. J. Bot.* 83: 1488–1502. <http://dx.doi.org/10.2307/2446103>
- Sanderson, M.J., Boss, D., Chen, D., Cranston, K.A. & Wehe, A. 2008. The PhyLoTA browser: Processing GenBank for molecular phylogenetics research. *Syst. Biol.* 57: 335–346. <http://dx.doi.org/10.1080/10635150802158688>
- Sanderson, M.J., McMahon, M.M. & Steel, M. 2010. Phylogenomics with incomplete taxon coverage: The limits to inference. *B. M. C. Evol. Biol.* 10: 155. <http://dx.doi.org/10.1186/1471-2148-10-155>
- Sanderson, M.J., McMahon, M.M. & Steel, M. 2011. Terraces in phylogenetic tree space. *Science* 333: 448–450. <http://dx.doi.org/10.1126/science.1206357>
- Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenies. *Crit. Rev. Biochem. Molec. Biol.* 37: 121–147. <http://dx.doi.org/10.1080/10409230290771474>
- Saslis-Lagoudakis, C.H., Chase, M.W., Robinson, D.N., Russell, J. & Klitgård, B.B. 2008. Phylogenetics of neotropical *Platymiscium* (Leguminosae: Dalbergiaceae): Systematics, divergence times, and biogeography inferred from nuclear ribosomal and plastid DNA sequence data. *Amer. J. Bot.* 95: 1270–1286. <http://dx.doi.org/10.3732/ajb.0800101>
- Saslis-Lagoudakis, C.H., Klitgård, B.B., Forest, F., Francis, L., Savolainen, V., Williamson, E.M. & Hawkins, J.A. 2011. The Use of phylogeny to interpret cross-cultural patterns in plant use and guide medicinal plant discovery: An example from *Pterocarpus* (Leguminosae). *PLoS ONE* 6(7): e22275. <http://dx.doi.org/10.1371/journal.pone.0022275>
- Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., Sasamoto, S., Watanabe, A., Ono, A., Kawashima, K., Fujishiro, T., Katoh, M., Kohara, M., Kishida, Y., Minami, C., Nakayama, S., Nakazaki, N., Shimizu, Y., Shinpo, S., Takahashi, C., Wada, T., Yamada, M., Ohmido, N., Hayashi, M., Fukui, K., Baba, T., Nakamichi, T., Mori, H. & Tabata, S. 2008. Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 15: 227–239. <http://dx.doi.org/10.1093/dnares/dsn008>
- Sauquet, H., Ho, S.Y.W., Gandolfo, M.A., Jordan, G.J., Wilf, P., Cantrill, D.J., Bayly, M.J., Bromham, L., Brown, G.K., Carpenter, R.J., Lee, D.M., Murphy, D.J., Sniderman, J.M.K. & Udovicic, F. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: The case of *Nothofagus* (Fagales). *Syst. Biol.* 61: 289–313. <http://dx.doi.org/10.1093/sysbio/syr116>
- Scherson, R.A., Vidal, R. & Sanderson, M.J. 2008. Phylogeny, biogeography, and rates of diversification of the New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. *Amer. J. Bot.* 95: 1030–1039. <http://dx.doi.org/10.3732/ajb.0800017>
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. & Jackson, S.A. 2010. Genome sequence of the paleopolyploid soybean. *Nature* 463: 178–183. <http://dx.doi.org/10.1038/nature08670>
- Schrire, B.D. 2008. The Madagascar genus *Vaughania* is reduced to synonymy under *Indigofera* (Leguminosae-Papilionoideae-Indigoferae). *Kew Bull.* 63: 477–479. <http://dx.doi.org/10.1007/s12225-008-9061-7>
- Schrire, B.D., Lavin, M. & Lewis, G.P. 2005. Global distribution patterns of the Leguminosae: Insights from recent phylogenies. *Biol. Skr.* 55: 375–422.
- Schrire, B.D., Lavin, M., Barker, N.P. & Forest, F. 2009. Phylogeny of the tribe Indigoferae (Leguminosae-Papilionoideae): Geographically structured more in succulent-rich and temperate settings than in grass-rich environments. *Amer. J. Bot.* 96: 816–852. <http://dx.doi.org/10.3732/ajb.0800185>
- Schutte, A.L. & Van Wyk, B.-E. 1998a. Evolutionary relationships in the Podalyriaceae and Lipariaceae based on morphological, cytological and chemical evidence. *Pl. Syst. Evol.* 209: 1–31. <http://dx.doi.org/10.1007/BF00991521>
- Schutte, A.L. & Van Wyk, B.-E. 1998b. The tribal position of *Hypocalyptus* Thunb. (Fabaceae). *Novon* 8: 178–182. <http://dx.doi.org/10.2307/3391993>
- Schutte-Vlok, A.L. & Van Wyk, B.-E. 2011. A taxonomic revision of *Podalyria* (Fabaceae). *Syst. Bot.* 36: 631–660. <http://dx.doi.org/10.1600/036364411X583628>
- Sede, S.M., Tosto, D., Gottlieb, A.M., Poggio, L. & Fortunato, R.H. 2008. Genetic relationships in the *Galactia-Camptosema-Collaea* complex (Leguminosae) inferred from AFLP markers. *Pl. Syst. Evol.* 276: 261–270. <http://dx.doi.org/10.1007/s00606-008-0100-5>
- Sede, S.M., Tosto, D., Talia, P., Luckow, M., Poggio, L. & Fortunato, R.H. 2009. Phylogenetic relationships among southern South American species of *Camptosema*, *Galactia* and *Collaea* (Diocleinae: Papilionoideae: Leguminosae) on the basis of molecular and morphological data. *Austral. J. Bot.* 57: 76–86. <http://dx.doi.org/10.1071/BT08091>
- Seigler, D.S. & Ebinger, J.E. 2005. New combinations in the genus *Vachellia* (Fabaceae: Mimosoideae) from the New World. *Phytologia* 87: 139–178.
- Seigler, D.S., Ebinger, J.E. & Miller, J.T. 2006a. The genus *Senegalia* (Fabaceae: Mimosoideae) from the New World. *Phytologia* 88: 38–94.
- Seigler, D.S., Ebinger, J.E. & Miller, J.T. 2006b. *Mariosousa*, a new segregate genus from *Acacia* s.l. (Fabaceae, Mimosoideae) from Central and North America. *Novon* 16: 413–420. [http://dx.doi.org/10.3417/1055-3177\(2006\)16\[413:MANSGF\]2.0.CO;2](http://dx.doi.org/10.3417/1055-3177(2006)16[413:MANSGF]2.0.CO;2)
- Silva, M.J. da, Queiroz, L.P. de, Tozzi, A.M.G.A., Lewis, G.P. & Sousa, A.P. de. 2012. Phylogeny and biogeography of *Lonchocarpus* sensu lato and its allies in the tribe Millettieae (Leguminosae, Papilionoideae). *Taxon* 61: 93–108.
- Simon, M.F., Grether, R., Queiroz, L.P. de, Skema, C., Pennington, R.T. & Hughes, C.E. 2009. Recent assembly of the Cerrado, a Neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proc. Natl. Acad. Sci. U.S.A.* 106: 20359–20364. <http://dx.doi.org/10.1073/pnas.0903410106>
- Simon, M.F., Grether, R., Queiroz, L.P. de, Särkinen, T.E., Dutra, V.F. & Hughes, C.E. 2011. The evolutionary history of *Mimosa* (Leguminosae): Toward a phylogeny of the sensitive plants. *Amer. J. Bot.* 98: 1201–1221. <http://dx.doi.org/10.3732/ajb.1000520>
- Simpson, B.B. & Ulbarri, E.A. 2006. A synopsis of the genus *Hoffmannseggia* (Leguminosae). *Lundellia* 9: 7–33.
- Simpson, B.B., Larkin, L.L. & Weeks, A. 2003. Progress towards resolving relationships of the *Caesalpinia* group (Caesalpinieae: Caesalpinioideae: Leguminosae). Pp. 123–148 in: Klitgaard, B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Sinou, C., Forest, F., Lewis, G.P. & Bruneau, A. 2009. The genus *Bauhinia* s.l. (Leguminosae): A phylogeny based on the plastid *trnL-F* region. *Botany* 87: 947–960. <http://dx.doi.org/10.1139/B09-065>

- Smith, S.A., Beaulieu, J.M. & Donoghue, M.J. 2009. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *B. M. C. Evol. Biol.* 9: 37. <http://dx.doi.org/10.1186/1471-2148-9-37>
- Smith, S.A., Beaulieu, J.M., Stamatakis, A. & Donoghue, M.J. 2011a. Understanding angiosperm diversification using small and large phylogenetic trees. *Amer. J. Bot.* 98: 404–414. <http://dx.doi.org/10.3732/ajb.1000481>
- Smith, S.A., Wilson, N.G., Goetz, F.E., Feehery, C., Andrade, S.C.S., Rouse, G.W., Giribet, G. & Dunn, C.W. 2011b. Resolving the evolutionary relationships of mollusks with phylogenomic tools. *Nature* 480: 364–367. <http://dx.doi.org/10.1038/nature10526>
- Sokoloff, D.D., Degtjareva, G.V., Endress, P.K., Remizowa, M.V., Samigullin, T.H. & Valejo-Roman, C.M. 2007. Inflorescence and early flower development in Loteae (Leguminosae) in a phylogenetic and taxonomic context. *Int. J. Pl. Sci.* 168: 801–833. <http://dx.doi.org/10.1086/518272>
- Soltis, D.E., Albert, V.A., Savolainen, V., Hilu, K., Qiu, Y., Chase, M.W., Farris, J.S., Stefanovic, S., Rice, D.W., Palmer, J.D. & Soltis, P.S. 2004. Genome-scale data, angiosperm relationships, and ‘ending incongruence’: A cautionary tale in phylogenetics. *Trends Pl. Sci.* 9: 477–483. <http://dx.doi.org/10.1016/j.tplants.2004.08.008>
- Sousa, M. 2005. *Heteroflorum*: Un nuevo género del grupo *Peltophorum* (Leguminosae: Caesalpinioideae: Caesalpinieae), endémico para México. *Novon* 15: 213–218.
- Sousa, M. & Rudd, V.E. 1993. Revision del genero *Styphnolobium* (Leguminosae: Papilionoideae: Sophoreae). *Ann. Missouri Bot. Gard.* 80: 270–283. <http://dx.doi.org/10.2307/2399827>
- Sprent, J.I. 2001. *Nodulation in legumes*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Sprent, J.I. 2009. *Legume nodulation: A global perspective*. West Sussex: Wiley-Blackwell. <http://dx.doi.org/10.1002/9781444316384>
- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690. <http://dx.doi.org/10.1093/bioinformatics/btl446>
- Steel, M. & Sanderson, M.J. 2010. Characterizing phylogenetically decisive taxon coverage. *Appl. Math. Lett.* 23: 82–86. <http://dx.doi.org/10.1016/j.aml.2009.08.009>
- Steele, K.P. & Wojciechowski, M.F. 2003. Phylogenetic analyses of tribes Trifolieae and Viciae, based on sequences of the plastid gene *matK* (Papilionoideae: Leguminosae). Pp. 355–370 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Steele, K.P., Ickert-Bond, S., Zarré, S. & Wojciechowski, M.F. 2010. Phylogeny and character evolution in *Medicago* (Leguminosae): evidence from analyses of plastid *trnK/matK* and nuclear *Ga3ox1* sequences. *Amer. J. Bot.* 97: 1142–1155.
- Stefanović, S., Pfeil, B.E., Palmer, J.D. & Doyle, J.J. 2009. Relationships among phaseoloid legumes based on sequences from eight chloroplast regions. *Syst. Bot.* 34: 115–128. <http://dx.doi.org/10.1600/036364409787602221>
- Steyermark, J.A., Berry, P.E., Yatskievych, K. & Holst, B.K. (eds.) 1998. *Flora of the Venezuelan Guayana*, vol. 4. St. Louis: Missouri Botanical Garden Press.
- Steyermark, J.A., Berry, P.E., Yatskievych, K. & Holst, B.K. (eds.) 1999. *Flora of the Venezuelan Guayana*, vol. 5. St. Louis: Missouri Botanical Garden Press.
- Steyermark, J.A., Berry, P.E., Yatskievych, K. & Holst, B.K. (eds.) 2001. *Flora of the Venezuelan Guayana*, vol. 6. St. Louis: Missouri Botanical Garden Press.
- Thompson, I.R. 2010a. A revision of *Plagiocarpus* (Brongniartieae: Fabaceae). *Muelleria* 28: 40–52.
- Thompson, I.R. 2010b. A revision of the leafless species of *Templetonia* (Brongniartieae: Fabaceae). *Muelleria* 28: 53–65.
- Thompson, I.R. 2010c. A revision of *Cristonia* (Brongniartieae: Fabaceae). *Muelleria* 28: 66–73.
- Thompson, I.R. 2011. A revision of *Muelleranthus*, *Ptychosema* and *Aenictophyton* (Fabaceae: Bossiaeeae). *Muelleria* 29: 173–189.
- Thulin, M., Lavin, M., Pasquet, R. & Delgado-Salinas, A. 2004. Phylogeny and biogeography of *Wajira* (Leguminosae): A monophyletic segregate of *Vigna* centered in the Horn of Africa region. *Syst. Bot.* 29: 903–920. <http://dx.doi.org/10.1600/0363644042451035>
- Torke, B.M. & Mansano, V.F. 2009. A phylogenetically based sectional classification of *Swartzia* (Leguminosae-Papilionoideae). *Taxon* 58: 913–924.
- Torke, B.M. & Schaal, B.A. 2008. Molecular phylogenetics of the species-rich neotropical genus *Swartzia* (Leguminosae, Papilionoideae) and related genera of the swartzioid clade. *Amer. J. Bot.* 95: 215–228. <http://dx.doi.org/10.3732/ajb.95.2.215>
- Tucker, S.C. 1998. Floral ontogeny in legume genera *Petalostylis*, *Labichea*, and *Dialium* (Caesalpinioideae: Cassieae), a series in floral reduction. *Amer. J. Bot.* 85: 184–208. <http://dx.doi.org/10.2307/2446307>
- Tucker, S.C. 2003. Floral development in legumes. *Pl. Physiol.* 131: 911–926. <http://dx.doi.org/10.1104/pp.102.017459>
- Turini, F.G., Bräuchler, C. & Heubl, G. 2010. Phylogenetic relationships and evolution of morphological characters in *Ononis* L. (Fabaceae). *Taxon* 59: 1077–1090.
- Van der Bank, M.L., Chase, M.W., Van Wyk, B.-E., Fay, M.F., Van der Bank, F.H., Reeves, G. & Hulme, A. 2002. Systematics of the tribe Podalyrieae (Fabaceae) based on DNA, morphological and chemical data. *Bot. J. Linn. Soc.* 139: 159–170. <http://dx.doi.org/10.1046/j.1095-8339.2002.00051.x>
- Van Wyk, B.-E. 2003. The value of chemosystematics in clarifying relationships in the genistoid tribes of papilionoid legumes. *Biochem. Syst. Ecol.* 31: 875–884. [http://dx.doi.org/10.1016/S0305-1978\(03\)00083-8](http://dx.doi.org/10.1016/S0305-1978(03)00083-8)
- Van Wyk, B.-E. & Schutte, A.L. 1995. Phylogenetic relationships in the tribes Podalyrieae, Liparieae and Crotalarieae. Pp. 283–308 in: Crisp, M. & Doyle, J.J. (eds.), *Advances in legume systematics*, part 7, *Phylogeny*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Varshney, R.K., Chen, W., Li, Y., Bharti, A.K., Saxena, R.K., Schlueter, J.A., Donoghue, M.T.A., Azam, S., Fan, G., Whaley, A.M., Farmer, A.D., Sheridan, J., Iwata, A., Tuteja, R., Penmetsa, R.V., Wu, W., Upadhyaya, H.D., Yang, S., Shah, T., Saxena, K.B., Michael, T., McCombie, W.R., Yang, B., Zhang, G., Yang, H., Wang, J., Spillane, C., Cook, D.R., May, G.D., Xu, X. & Jackson, S.A. 2011. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30: 83–89. <http://dx.doi.org/10.1038/nbt.2022>
- Villiers, J.-F. 2002. *Vigueranthus*. Pp. 271–285 in: Du Puy, D.J., Labat, J.-N., Rabevohitra, R., Villiers, J.F., Bosser, J. & Moat, J. (eds.), *The Leguminosae of Madagascar*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Webb, C.O., Ackerly, D.D., McPeck, M. & Donoghue, M.J. 2002. Phylogenies and community ecology. *Annual Rev. Ecol. Syst.* 33: 475–505. <http://dx.doi.org/10.1146/annurev.ecolsys.33.010802.150448>
- Wiens, J.J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* 52: 528–538. <http://dx.doi.org/10.1080/10635150390218330>
- Wieringa, J.J. 1999. *Monopetalanthus* exit: A systematic study of *Aphanocalyx*, *Bikinia*, *Icuria*, *Michelsonia* and *Tetraberlinia* (Leguminosae, Caesalpinioideae). *Agric. Uni. Wageningen Pap.* 99: 1–320.
- Wieringa, J.J. & Gervais, G.Y.F. 2003. Phylogenetic analyses of combined morphological and molecular data sets on the *Aphanocalyx-Bikinia-Tetraberlinia* group (Leguminosae, Caesalpinioideae, Detarieae s.l.). Pp. 181–196 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.

- Wojciechowski, M.F.** 2003. Reconstructing the phylogeny of legumes (Leguminosae): An early 21st century perspective. Pp. 5–35 in: Klitgaard, B.B. & Bruneau, A. (eds.), *Advances in legume systematics*, part 10, *Higher level systematics*. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Wojciechowski, M.F., Sanderson, M.J. & 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. *Syst. Bot.* 24: 409–437.  
<http://dx.doi.org/10.2307/2419698>
- Wojciechowski, M.F., Sanderson, M., Steele, K.P. & Liston, A.** 2000. Molecular phylogeny of the “Temperate Herbaceous Tribes” of Papilionoid legumes: A supertree approach. Pp. 277–298 in: Herendeen, P. & Bruneau, A. (eds.), *Advances in legume systematics*, part 9. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Wojciechowski, M.F., Lavin, M. & Sanderson, M.J.** 2004. A phylogeny of the legumes (Leguminosae) based on analysis of the plastid *matK* gene sequences resolves many well-supported subclades within the family. *Amer. J. Bot.* 91: 1846–1862.  
<http://dx.doi.org/10.3732/ajb.91.11.1846>
- Wu, F., Mueller, L.A., Crouzillat, D., Petiard, V. & Tanksley, S.D.** 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the euasterid plant clade. *Genetics* 174: 1407–1420.  
<http://dx.doi.org/10.1534/genetics.106.062455>
- Wu, Z. & Ge, S.** 2012. The phylogeny of the BEP clade in grasses revisited: Evidence from the whole-genome sequences of chloroplasts. *Molec. Phylogen. Evol.* 62: 573–578.  
<http://dx.doi.org/10.1016/j.ympev.2011.10.019>
- Wunderlin, R.P.** 2010. New combinations in *Schnella* (Fabaceae: Caesalpinioideae: Cercideae). *Phytoneuron* 49: 1–5.
- Wunderlin, R., Larsen, K. & Larsen, S.S.** 1981. Cercideae. Pp. 107–116 in: Polhill, R.M. & Raven, P.H. (eds.), *Advances in legume systematics*, part 1. Richmond, U.K.: Royal Botanic Gardens, Kew.
- Yan, C.H., Burleigh, J.G. & Eulenstein, O.** 2005. Identifying optimal incomplete phylogenetic data sets from sequence databases. *Molec. Phylogen. Evol.* 35: 528–535.  
<http://dx.doi.org/10.1016/j.ympev.2005.02.008>
- Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., Mayer, K.F.X., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D.R., Meyers, B.C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A.K., Murray, J.D., Naoumkina, M.A., Rosen, B., Silverstein, K.A.T., Tang, H., Rombauts, S., Zhao, P.X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Berges, H., Bidwell, S., Bisseling, T., Choise, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J.J., Duzde, A., Farmer, A.D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., Gonzalez, A.J., Green, P.J., Hallab, A., Hartog, M., Hua, A., Humphray, S.J., Jeong, D., Jing, Y., Joecker, A., Kenton, S.M., Kim, D., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S.L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E.L., Mun, J., Najar, F.Z., Nicholson, C., Noirot, C., O’Bleness, M., Paule, C.R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E.F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A.J., Sherrier, D.J., Shi, R., Sims, S., Singer, S.R., Sinharoy, S., Sterck, L., Violette, A., Wang, B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D.D., White, J.D., Wiley, G.B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Denarie, J., Dixon, R.A., May, G.D., Schwartz, D.C., Rogers, J., Quetier, F., Town, C.D. & Roe, B.A.** 2011. The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480: 520–524.  
<http://dx.doi.org/10.1038/nature10625>
- Zhang, M.-L. & Podlech, D.** 2006. Revision of the genus *Phyllolobium* Fisch. (Leguminosae-Papilionoideae). *Feddes Repert.* 117: 41–64.  
<http://dx.doi.org/10.1002/fedr.200511090>
- Zimmerman, E., Prenner, G. & Bruneau, A.** 2013. Floral morphology of *Apuleia leiocarpa* (Caesalpinioideae: Cassieae), an unusual andromonoecious legume. *Int. J. Pl. Sci.* 174: 154–160.  
<http://dx.doi.org/10.1086/668789>