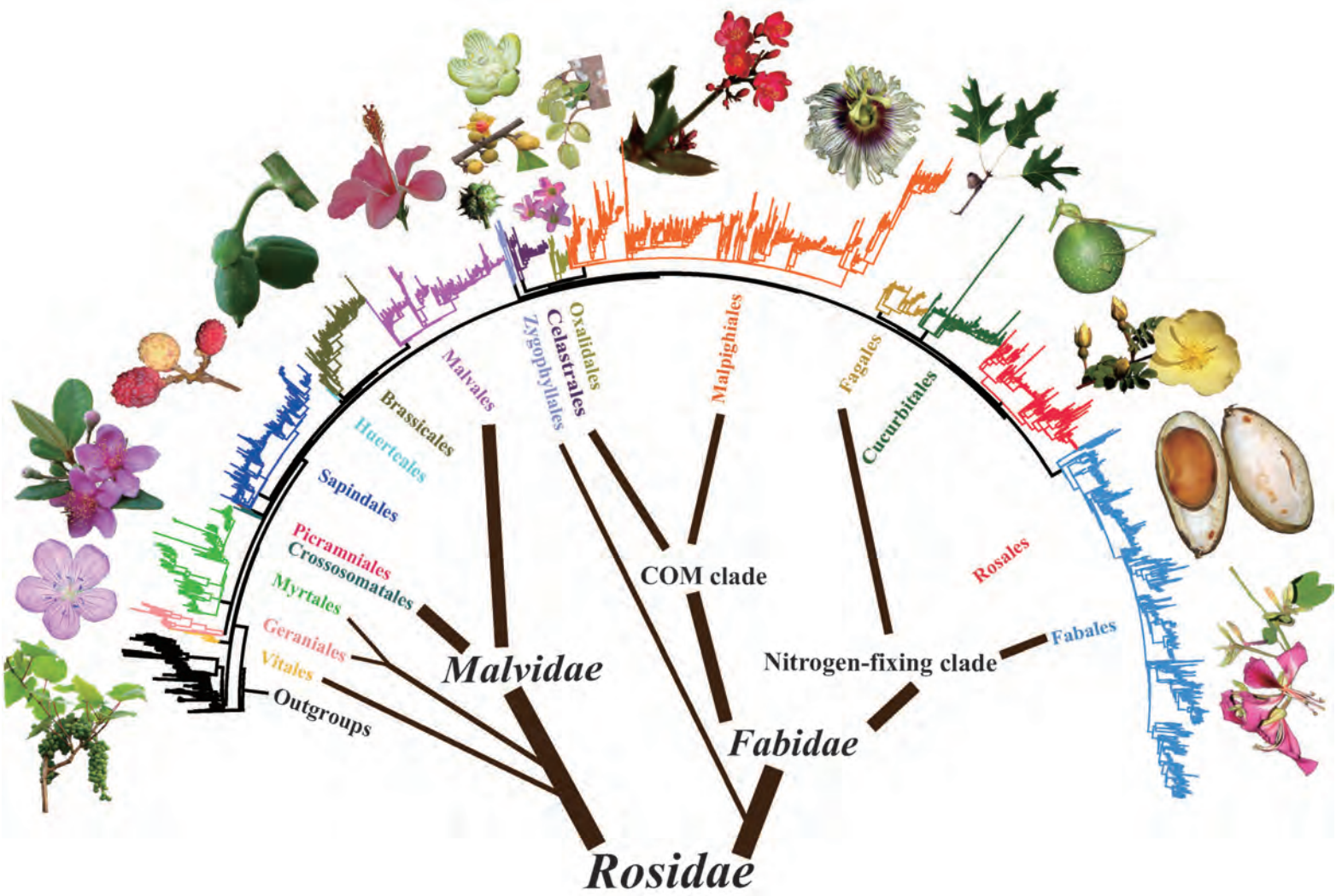
JSE
Journal of Systematics and Evolution

Special Issue: The Tree of Life: China Project



Volume 54 Number 4 July 2016

J S E
Journal of Systematics and Evolution 植物分类学报

ISSN 1674-4918 (Print)
ISSN 1759-6831 (Online)

JEB.JSE.March 16

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12215

Editorial

The Tree of Life: China project

The knowledge of evolutionary relationships is fundamental to all disciplines of biology, yielding novel and profound insights across plant sciences, from comparative genomics, molecular evolution, and plant development, to the study of adaptation, speciation, community assembly, and ecosystem functioning (Forest et al., 2007; Donoghue, 2008; Gehrke & Linder, 2011). Phylogeny (the Tree of Life, TOL) has become the foundation of evolutionary biology. It is accurate to say "Evolutionary biology makes much more sense in the light of phylogeny", as a corollary to Dobzhansky's (1973) famous statement "Nothing in biology makes sense except in the light of evolution."

China harbors 31 362 species, 3328 genera, and 312 families of vascular plants (Wu et al., 1994–2013) and has the richest flora of the Northern Hemisphere (Wu et al., 2003). A well-resolved phylogeny of vascular plants of China has many potential uses in various areas of biology—ecology, conservation genetics, and agriculture—as well as stimulates new research at the interface of evolutionary ecology, phylogenetics, and biogeography, thus clarifying processes that shaped patterns of distribution and diversity of such a rich flora of the Northern Hemisphere (Qian & Ricklefs, 2000; Wang et al., 2009; López-Pujol et al., 2011). Understanding the phylogeny of vascular plants and phylogenetic diversity at this scale will help elucidate fundamental processes underlying plant/animal associations and the assembly of entire ecosystems, and help manage the impact of global challenges to biodiversity and the maintenance of natural resources to humankind.

In June 2007, an international symposium on the TOL was held in Beijing, China. *Journal of Systematics and Evolution* (JSE) organized and published the symposium special issue: Patterns of Evolution and the Tree of Life (JSE vol. 46, no. 3, 2008). Since then, the Chinese botanical community has continued to make contributions to TOL studies. The present special issue aims to present recent progress in reconstructing TOL of the vascular plant genera in China, including the assembly of DNA materials, establishment of co-operation, data generation, tree reconstruction, on how to use the China TOL as a framework to further examine the origin and evolution of major clades in vascular plants, and the floristic relationship between China and other regions of the world as all vascular plants share a common ancestor (Wen et al., 2010; Xiang et al., 2015).

This special issue consists of 11 papers all related to the "giant" phylogeny of the Chinese vascular plants. Chen et al. (2016) sampled 6098 species representing 3114 genera of vascular plants and five genera of bryophytes as out-groups to reconstruct the TOL of the Chinese vascular plants at the generic level. To facilitate further application of such a large-scale phylogeny to other biology fields, the SoTree software was introduced to enable the efficient generation of the phylogenetic trees by providing sub-datasets with interested species lists for studies concerning the origin, ecology, and biogeography of the local flora in China. Using DNA sequences of three plastid genes, Liu (2016) presents a phylogenetic analysis of 259 genera of pteridophytes, which provided evidence documenting the impact of Ren-Chang Ching's integrative classification of pteridophytes. Ten out of 11 orders in Ching's system are consistent with the modern DNA-based phylogeny, whereas four new orders were introduced to avoid paraphyletic orders in the leptosporangiate ferns. Wang et al. (2016) integrated *Leefructus mirus*—one of the earliest eudicot macrofossils—in an exhaustive morphological dataset of extant Ranunculales to improve our understanding of the diversification of this lineage in eudicots. As a result of the integration of this fossil, the authors recovered that basal eudicots experienced an accelerated diversification during the onset of the angiosperm radiation in the Early Cretaceous. Du et al. (2016) sampled 139 genera (in 43 families) representing most families of the aquatic plants worldwide. Their results suggested that aquatic habitats were colonized at least three times during the early radiation of angiosperms, namely by Nymphaeales, Ceratophyllales, and the monocots.

Three of the papers address the phylogeny of angiosperms at the ordinal level or above. Special attention is given to the rosids (*Rosidae*) because the clade contributes not only one-quarter of the extant diversity of angiosperms, including considerable economically important crops and most dominant forest trees, but is also recognized as a major contributor to the angiosperm diversity of China. Using a supermatrix approach, Sun et al. (2016) resolved the phylogeny of *Rosidae* world-wide with a dense sampling scheme (four genes, a total of 9300 taxa representing 2775 genera, 138 families, and 17 orders). They discovered several novel relationships and recognized two families and 467 genera as non-monophyletic. As part of the rosids, the N-fixing clade is one of the largest clades of the angiosperms, containing over 1300 genera, approximately 30 000 species, which are important components of extant temperate and tropical forest. Li et al. (2016a) constructed the most comprehensive and robust global tree of the N-fixing clade to date with a supermatrix to compare with the local tree from the TOL of the Chinese vascular plants. Topologies of the global tree and the local tree are generally congruent and most of the internal supports are greatly improved with dense sampling. Yang et al. (2016) used eight chloroplast markers and one mitochondrial gene, and assembled a matrix of 11 951 characters of 649 genera, covering ca. 54% of the genera of Gentianales, to reconstruct the phylogeny of Gentianales. Topologies of the global Gentianales tree and the Chinese Gentianales tree are largely

congruent. The Gentianales and each family within the order are strongly supported as monophyletic. Relationships among some deep nodes are newly resolved.

This special issue also includes the phylogenetic analyses of four families, of which Asteraceae and Orchidaceae are the first and fourth largest families of the Chinese flora, respectively. Fu et al. (2016) used three plastid markers (*rbcL*, *ndhF*, and *matK*) to reconstruct the phylogeny of 506 genera, approximately one-third of all the genera of the Asteraceae, with a total of 200 Chinese Asteraceae genera included in the analysis. The results are largely congruent with those of earlier studies. A systematic arrangement of all the genera of the Chinese Asteraceae was presented, in which 255 genera (48 introduced), 22 tribes, and 7 subfamilies were recognized. Orchidaceae have ca. 200 genera in China. Li et al. (2016b) investigated the molecular phylogenetic relationships of the higher-level Chinese orchids with 175 genera sampled. The subfamilies, tribes, and subtribes *sensu* Genera Orchidacearum are supported as monophyletic, except that the paraphyletic Diseae, Calyposeae, Vandeae, and Eriinae, and the relationships of Epidendroideae are weakly supported. Five faster-evolving genes (*rbcL*, *matK*, *psaB*, *ycf1*, and Xdh) were used to further reconstruct the phylogenies of the perplexing Epidendroideae. Li & Wen (2016) sampled 96 accessions representing all 20 genera and 50 species of Chinese Araliaceae and 45 closely related taxa to assess the evolutionary relationships of Araliaceae and their biogeographic diversification in China. Their results supported that the Chinese members of Araliaceae were scattered within the Asian Palmate group and the *Aralia–Panax* group with *Osmoxylon* at the base of core Araliaceae. The Chinese Araliaceae have originated in Asia and the distribution pattern of the phylogenetic diversity of Araliaceae corresponds with its taxonomic diversity across the entire region. Cai & Ma (2016) present a case study of phylogeny at the generic and specific levels with nuclear genes, using Brassicaceae taxa as examples. They used three protein-coding nuclear genes, *MLH1*, *SMC2*, and *MCM5*, with up to 10 200 base pairs (in both exons and introns) to reconstruct a phylogeny with multiple species in each of five genera within Brassicaceae for a total of 65 taxa. The combined data revealed high resolution at various phylogenetic depths and their results provided a robust species-level phylogeny for a number of Brassicaceae members and supported an optimistic perspective on the phylogenetic utility of conserved nuclear data for relatively recent clades.

The TOL China project has been carried out as a long-term collaboration among several institutions since March 2009. The State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences (Beijing, China) organized the project and has closely collaborated with the following institutions: Fairylake Botanical Garden (Shenzhen), Computer Network Information Center, Chinese Academy of Sciences (Beijing), Wuhan Botanical Garden, Chinese Academy of Sciences (Wuhan), the Orchid Conservation and Research Center of Shenzhen (Shenzhen), and University of Florida, Gainesville (USA) in the last 7 years. This project was financially supported by grants from the National Natural Science Foundation of China (Grant Nos. 31590822, 31270268, and 31270269), the National Key Basic Research Program of China (Grant No. 2014CB954100), the Chinese Academy of Sciences International Institution Development Program (Grant No. SAJC201315), the Shenzhen Science and Technology Innovation Council Funding (Grant No. KQC 201105310009A), the Chinese Academy of Sciences External Cooperation Program of BIC (Grant No. GJHZ201321), and the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists (Grant No. 2011T1S24).

The project has arguably achieved now its main goal, a phylogeny comprising almost all genera of vascular plants native in China. However, this is only the first step towards an exhaustive understanding of the phylogeny of Chinese plants. In the future, the trees need to be expanded to cover not only other land plant lineages, such as liverworts and mosses, but especially to be expanded to include at least one representative of each plant species occurring in China. These efforts will allow the merging of efforts of the TOL project with those of the DNA barcoding of land plants. The importance of this next step forward was well illustrated by the unique insights provided by the studies published in this special issue of JSE.

Finally, we want to briefly discuss the plurality and singularity of the family names in Latin as there are two different writing ways that coexist. According to Stearn's Botanical Latin (1992), families are female plural nouns and they should be treated as plural nouns in English. However, it is common to find that writers consider a family name in Latin as singular and use the plural noun in combination with a singular verb. This common usage may be consistent in the context that families are considered to correspond to monophyla in contrast to their grammatical status; families are therefore individuals instead of classes in the context of ontology (Minelli, 2012). We prefer using the family names as plural nouns and have consistently used them this way in this special issue.

Zhi-Duan Chen, *An-Ming Lu*
State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
Shou-Zhou Zhang
Shenzhen Key Laboratory of Southern Subtropical Plant Diversity, Fairylake Botanical Garden, Shenzhen and Chinese Academy of Sciences, Shenzhen 518004, China
Qing-Feng Wang
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China
Zhong-Jian Liu
The Orchid Conservation and Research Center of Shenzhen, Shenzhen 518114, China
De-Zhu Li
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
Hong Ma
College of Life Sciences, Fudan University, Shanghai 200433, China
Jian-Hua Li
Biology Department, Hope College, Holland, MI 49423, USA
Douglas E. Soltis
Department of Botany, University of Florida, Gainesville, FL 32611-7800, USA

Pamela S. Soltis
Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
Jun Wen
Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012 USA
China Phylogeny Consortium†
†Members of the China Phylogeny Consortium are listed in the Appendix.

References

Cai LM, Ma H. 2016. Using nuclear genes to reconstruct angiosperm phylogeny at the species level: A case study with Brassicaceae species. *Journal of Systematics and Evolution* 54: 438–452.

Chen ZD, Yang T, Lin L, Lu LM, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, Cao ZY, Liu HM, Wang XM, Wang W, Zhang JB, Meng Z, Cao W, Li JH, Wu SD, Zhao HL, Liu ZJ, Du ZY, Wang QF, Guo J, Tan XX, Su JX, Zhang LJ, Yang LL, Liao YY, Li MH, Zhang GQ, Chung SW, Zhang J, Xiang KL, Li RQ, Soltis DE, Soltis PS, Zhou SL, Ran JH, Wang XQ, Jin XH, Chen YS, Gao TG, Li JH, Zhang SZ, Lu AM, China Phylogeny Consortium. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54: 277–306.

Dobzhansky TH. 1973. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* 35: 125–129.

Donoghue MJ. 2008. A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences USA* 105: 11549–11555.

Du ZY, Wang QF, China Phylogeny Consortium. 2016. Phylogenetic tree of vascular plants reveals the origins of aquatic angiosperms. *Journal of Systematics and Evolution* 54: 342–348.

Forest F, Grenyer R, Rouget M, Davies TJ, Cowling RM, Faith DP, Balmford A, Manning JC, Proches S, der Bank M, Reeves G, Hedderson TAJ, Savolainen V. 2007. Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature* 445: 757–760.

Fu ZX, Jiao BH, Nie B, Zhang GJ, Gao TG, China Phylogeny Consortium. 2016. A comprehensive generic-level phylogeny of the sunflower family: Implications for the systematics of Chinese Asteraceae. *Journal of Systematics and Evolution* 54: 416–437.

Gehrke B, Linder HP. 2011. Time, space and ecology: Why some clades have more species than others. *Journal of Biogeography* 38: 1948–1962.

Li HL, Wang W, Li RQ, Zhang JB, Sun M, Naeem R, Su JX, Xiang XG, Mortimer PE, Li DZ, Hyde KD, Xu JC, Soltis DE, Soltis PS, Li JH, Zhang SZ, Wu H, Chen ZD, Lu AM. 2016a. Global versus Chinese perspectives on the phylogeny of the N-fixing clade. *Journal of Systematics and Evolution* 54: 392–399.

Li MH, Zhang GQ, Lan SR, Liu ZJ, China Phylogeny Consortium. 2016b. A molecular phylogeny of Chinese orchids. *Journal of Systematics and Evolution* 54: 349–362.

Liu HM. 2016. Embracing the pteridophyte classification of Ren-Chang Ching using a generic phylogeny of Chinese ferns and lycophytes. *Journal of Systematics and Evolution* 54: 307–335.

Li R, Wen J. 2016. Phylogeny and diversification of Chinese Araliaceae based on nuclear and plastid DNA sequence data. *Journal of Systematics and Evolution* 54: 453–467.

López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. 2011. Centres of plant endemism in China: Places for survival or for speciation? *Journal of Biogeography* 38: 1267–1280.

Minelli A. 2012. Biological systematics: The state of the art. Springer: Springer Science & Business Media.

Qian H, Ricklefs RE. 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature* 407: 180–182.

Stearn WT. 1992. Botanical Latin (fourth version). Newton Abbot: David & Charles. 101–102.

Sun M, Naeem R, Su JX, Cao ZY, Burleigh JG, Soltis PS, Soltis DE, Chen ZD. 2016. Phylogeny of the *Rosidae*: A dense taxon sampling analysis. *Journal of Systematics and Evolution* 54: 363–391.

Wang W, Dilcher DL, Sun G, Wang HS, Chen ZD. 2016. Accelerated evolution of early angiosperms: Evidence from ranunculalean phylogeny by integrating living and fossil data. *Journal of Systematics and Evolution* 54: 336–341.

Wang ZH, Brown JH, Tang ZY, Fang JY. 2009. Temperature dependence, spatial scale, and tree species diversity in eastern Asia and North America. *Proceedings of the National Academy of Sciences USA* 106: 13388–13392.

Wen J, Ickert-Bond SM, Nie ZL, Li R. 2010. Timing and modes of evolution of eastern Asian–North American biogeographic disjunctions in seed plants. In: Long M, Gu H, Zhou Z eds. *Darwin's heritage today: Proceedings of the Darwin 200 Beijing International Conference*. Beijing: Higher Education Press. 252–269.

Wu ZY, Lu AM, Tang YC, Chen ZD, Li D-Z. 2003. *The families and genera of angiosperms in China: A comprehensive analysis*. Beijing: Science Press.

Wu ZY, Raven P, Hong DY eds. 1994–2013. Flora of China. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.

Xiang JY, Wen J, Peng H. 2015. Evolution of the eastern Asian–North American biogeographic disjunctions in ferns and lycophytes. *Journal of Systematics and Evolution* 53: 2–32.

Yang LL, Li HL, Wei L, Yang T, Kuang DY, Li MH, Liao YY, Chen ZD, Wu H, Zhang SZ. 2016. A supermatrix approach provides a comprehensive genus-level phylogeny for Gentianales. *Journal of Systematics and Evolution* 54: 400–415.

Appendix

Members of the China Phylogeny Consortium

Zhi-Duan CHEN, An-Ming LU, Hong-Zhi KONG, Xiao-Quan WANG, Yin-Zheng WANG, Shi-Liang ZHOU
Institute of Botany, Chinese Academy of Sciences, Beijing, China
Shou-Zhou ZHANG, Xiao-Ming WANG
Fairylake Botanical Garden, Shenzhen, China
Zhong-Jian LIU
The Orchid Conservation and Research Center of Shenzhen, Shenzhen, China
Qing-Feng WANG
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
Jian-Hui LI
Computer Network Information Center, Chinese Academy of Sciences, Beijing, China
De-Zhu LI, Ting-Shuang YI
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Hong MA
Fudan University, Shanghai, China
Douglas E. SOLTIS, Pamela S. SOLTIS
University of Florida, Gainesville, USA
Jian-Hua LI
Hope College, Holland, USA
Cheng-Xin FU
Zhejiang University, Hangzhou, China
Qi-Xin LIU
Nanjing Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12219

Research Article

Tree of life for the genera of Chinese vascular plants

Zhi-Duan Chen[1*], Tuo Yang[1,12], Li Lin[1,12], Li-Min Lu[1], Hong-Lei Li[1,2], Miao Sun[1], Bing Liu[1], Min Chen[1,12], Yan-Ting Niu[1,12], Jian-Fei Ye[1,12], Zhi-Yong Cao[1], Hong-Mei Liu[2], Xiao-Ming Wang[2], Wei Wang[1], Jing-Bo Zhang[1,12], Zhen Meng[3], Wei Cao[3], Jian-Hui Li[3], Sheng-Dan Wu[4], Hui-Ling Zhao[4], Zhong-Jian Liu[5], Zhi-Yuan Du[6], Qing-Feng Wang[6], Jing Guo[4], Xin-Xin Tan[4], Jun-Xia Su[4], Lin-Jing Zhang[4], Lei-Lei Yang[2], Yi-Ying Liao[2], Ming-He Li[5], Guo-Qiang Zhang[5], Shih-Wen Chung[7], Jian Zhang[1], Kun-Li Xiang[1,12], Rui-Qi Li[1], Douglas E. Soltis[8,9,10], Pamela S. Soltis[9,10], Shi-Liang Zhou[1], Jin-Hua Ran[1], Xiao-Quan Wang[1], Xiao-Hua Jin[1], You-Sheng Chen[1], Tian-Gang Gao[1], Jian-Hua Li[11], Shou-Zhou Zhang[2*], An-Ming Lu[1*], and China Phylogeny Consortium[#]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[2]Shenzhen Key Laboratory of Southern Subtropical Plant Diversity, Fairylake Botanical Garden, Shenzhen & Chinese Academy of Sciences, Shenzhen 518004, China
[3]Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China
[4]College of Life Sciences, Shanxi Normal University, Linfen 041004, China
[5]The Orchid Conservation & Research Center of Shenzhen, Shenzhen 518114, China
[6]Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China
[7]Department of Botany, Taiwan Forestry Research Institute, 20 Nan-Hai Road, Taipei 100, China
[8]Department of Biology, University of Florida, Gainesville, FL 32611, USA
[9]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
[10]Genetics Institute, University of Florida, Gainesville, FL 32610, USA
[11]Biology Department, Hope College, Holland, MI 49423, USA
[12]University of Chinese Academy of Sciences, Beijing 100049, China
[#]Members of the China Phylogeny Consortium are listed in the Appendix.
*Authors for correspondence. Z.-D. Chen. E-mail: zhiduan@ibcas.ac.cn. Tel.: +86-010-62836434. Fax: +86-010-62836095. A.-M. Lu. E-mail: anmin@ibcas.ac.cn. Tel.: +86-010-62836448. Fax: +86-010-62836095. S.-Z. Zhang E-mail: szzhang@126.com. Tel.: +86-755-25732477. Fax: +86-755-25734474
Received 22 May 2016; Accepted 23 June 2016; Article first published online 22 July 2016

Abstract We reconstructed a phylogenetic tree of Chinese vascular plants (Tracheophyta) using sequences of the chloroplast genes *atpB*, *matK*, *ndhF*, and *rbcL* and mitochondrial *matR*. We produced a matrix comprising 6098 species and including 13 695 DNA sequences, of which 1803 were newly generated. Our taxonomic sampling spanned 3114 genera representing 323 families of Chinese vascular plants, covering more than 93% of all genera known from China. The comprehensive large phylogeny supports most relationships among and within families recognized by recent molecular phylogenetic studies for lycophytes, ferns (monilophytes), gymnosperms, and angiosperms. For angiosperms, most families in Angiosperm Phylogeny Group IV are supported as monophyletic, except for a paraphyletic Dipterocarpaceae and Santalaceae. The infrafamilial relationships of several large families and monophyly of some large genera are well supported by our dense taxonomic sampling. Our results showed that two species of *Eberhardtia* are sister to a clade formed by all other taxa of Sapotaceae, except *Sarcosperma*. We have made our phylogeny of Chinese vascular plants publically available for the creation of subtrees via SoTree (http://www.darwintree.cn/flora/index.shtml), an automated phylogeny assembly tool for ecologists.

Key words: China, DarwinTree, megaphylogeny, regional tree of life, SoTree, vascular plants.

China is ranked among the top six megadiverse countries of the world (Huang et al., 2013) and has 31 362 species of vascular plants belonging to 3328 genera and 312 families (Wu et al., 1994–2013). Considered as "a plantsman's paradise" and "the mother of gardens", China attracted many botanists in the world to collect specimens and germplasm during the last half of the 19th century and early 20th century. For example, French collector Jean Marie Delavay (1834–1895) collected 200 000 plant specimens, mostly in Yunnan, from 1867 to 1895 (Hu & Watson, 2013). As a major biodiversity hotspot, China continues to attract scientists from around the world to study taxonomy, biogeography, propagation, ecology, and phylogeny of vascular plants (Hong & Blackmore, 2013).

Botanical studies in China date back to at least the Western Zhou Dynasty (ca. 3000 years ago), during which time plant names were written on the carapaces of tortoises. More than 130 species of plants were recorded in *Shi Jing* (*The Book of Poems*) around 6th century BC (Wang, 1994). By the time of

the Ming Dynasty, the Chinese possessed a deep practical knowledge and profound understanding of plants, which is reflected in numerous ancient texts, such as *Shen Nong Ben Cao Jing* (i.e., *Shen Nong Herbal*, 100–200 AD), *Qi Min Yao Shu* (553–544 AD), and *Ben Cao Gang Mu* (1596) (Hu & Watson, 2013; Peng, 2013). Botanical studies of vascular plants in China since the 1920s have contributed greatly to the understanding of plant phylogeny and evolution (Hong et al., 2008; Hu & Watson, 2013; Peng, 2013; Zhang & Li, 2013) and have been compiled in important works, such as the *Florae Reipublicae Popularis Sinicae* (FRPS; 80 volumes and 126 books) and *Flora of China* (FOC; 26 volumes). These two monumental floras were completed in 2004 and 2013, respectively.

Additionally, many individual groups of Chinese plants have been carefully studied and revised since the 1920s. For example, Ching (1940, 1978a, 1978b) revolutionized the classification of ferns, narrowing and redefining the Polypodiaceae family, and his circumscription is supported by molecular phylogenetic studies (see Schneider et al., 2004; Smith et al., 2009). Moreover, Cheng & Fu (1978) proposed a system of classification for Chinese gymnosperms in FRPS Volume 7. Hu (1950) published a polyphyletic classification system for angiosperms that emphasized the characters of pollen (monocolpate *vs.* tricolpate) and stamens (centrifugal *vs.* centripetal). Wu et al. (1998, 2002, 2003) proposed a system for angiosperms involving eight classes and suggested that the angiosperms may have evolved according to a polyphyletic-polychronic-polytopic pattern.

Molecular phylogenetic studies conducted over the past 30 years have greatly improved our understanding of the relationships of vascular plants (Qiu et al., 2006; Fiz-Palacios et al., 2011; Soltis & Soltis, 2013; Ruhfel et al., 2014), lycophytes, ferns (Smith et al., 2009), and seed plants (Chase et al., 1993; APG, 1998; Qiu et al., 1999, 2010; APG II, 2003; APG III, 2009; Soltis et al., 2011; Zhang et al., 2012; Zeng et al. 2014; Lu et al., 2015; APG IV, 2016), with revised classifications for several clades. In China, molecular phylogenetic studies have been expanded to include gymnosperms (Wang et al., 1997, 2000), Campanulaceae (Ge et al., 1997; Hong & Wang, 2015), Fagales (Chen et al., 1998, 1999), Gesneriaceae (Wang & Li, 1998), and Poaceae (Ge et al., 1999), as well as numerous taxa since 2000, and to support other fields such as biogeography, phylogeography, and DNA barcoding (Li et al., 2011a,b; Qiu et al., 2011; Zhang & Li, 2013; Dong et al., 2014; Liu et al., 2014; Wang & Ran, 2014; Wen et al., 2014; Meng et al., 2015a; Xiang et al., 2015).

Ecologists have more commonly applied phylogenetics to the study of community dynamics by using detailed information (e.g., Webb et al., 2002; Cavender-Bares et al., 2009). Such work has been made possible in part by extremely large phylogenies with huge numbers of species (megaphylogenies). Considerable progress has been made in producing these large phylogenies, especially by utilizing sequence databases, such as GenBank, and by using cutting-edge computational algorithms and hardware (Ciccarelli et al., 2006; McMahon & Sanderson, 2006; Sanderson et al., 2008; Smith et al., 2009; Fiz-Palacios et al., 2011; Izquierdo-Carrasco et al., 2011). Nevertheless, community ecologists often require more species in phylogenies to test macroevolutionary hypotheses than are available in most molecular phylogenies (Pearse et al., 2013; Hennequin et al., 2014).

A regional tree of life is a pivotal platform for ecological and biogeographic studies, including studies of phylogenetic diversity (PD), relationships among species within communities, and patterns in geographic distributions of biodiversity, as well as phylogenetic correlation and spatial arrangements of ecological traits (Lu et al., 2014; Qian & Zhang, 2016). In this paper, we describe the results of our efforts to reconstruct a large regional tree of Chinese vascular plants at the generic level. Our data are publicly available via our online molecular data analysis website (DarwinTree: www.darwintree.cn; Meng et al., 2010, 2015b) and software (SoTree), which allows registered users to generate subtrees from our phylogeny by providing a taxonomic list. We hope that our effort will help support future hypothesis-driven ecological, biogeographic, and related studies on Chinese flora.

Material and Methods

Taxon and gene sampling

To facilitate taxonomic sampling, we used the accepted names of plant families and genera of Chinese vascular plants according to FOC (Wu et al., 1994–2013; Table S1). Additionally, we have used and applied names of species and infraspecific taxa according to a list from Dr. Hai-Ning Qin (PE, Institute of Botany, Chinese Academy of Sciences, Beijing) that included a total of 38 115 names of seed plants. For ferns and lycophytes, Zhang et al.'s (2013) concept was adopted as the latest classification system of Chinese "pteridophytes." We used the statistical tools on DarwinTree (http://www.darwintree.cn/) to evaluate the optimal DNA markers to apply to our phylogenetic reconstruction and selected five markers for seed plants, including four chloroplast genes—*atpB*, *matK*, *ndhF*, and *rbcL*—and one mitochondrial gene, *matR;* plastid *atpA*, *atpB*, and *rbcL* were selected for lycophytes and ferns.

We used the purification procedure on DarwinTree (Meng et al., 2010, 2015b) to select and download all the sequences available for the selected markers. In cases where there was more than one sequence available for a species, we selected the longest sequence. The DarwinTree output comprised a summary table including the sequences, binomial names, and GenBank Accession Numbers. To maximize the taxon coverage and minimize missing data, we divided all of the genera of Chinese angiosperms and gymnosperms into three categories based on the number of species in Chinese flora: (i) monotypic (1173 genera), (ii) 2–30 species (1736 genera), and (iii) more than 30 species (267 genera). For the first category, we sampled one species. For the second category, two species were sampled, and priority was given to the species with more DNA sequences available in GenBank. If the number of DNA sequences was equal, priority was given to those distributed in China. For the third category, we sampled at least 10% of the diversity of each genus, taking into consideration infrageneric circumscriptions (subgenus and section), and priority was given to the species distributed in China and with more DNA sequences available in GenBank. After obtaining sequences from GenBank via DarwinTree, we found that there were 781 genera of Chinese vascular plants remaining with sequences unavailable. Thus, we collected samples to represent these genera from the field (for 513 genera) or from herbarium specimens (400 specimens at PE).

DNA extraction, PCR, and sequencing

For the samples collected in the field, we extracted total genomic DNAs from silica gel-dried leaves using a Plant

Genomic DNA Kit (Biomed Co., LTD, Beijing, China) or following a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle & Doyle, 1987). The standard PCR was used to amplify target regions. Primers for PCR are listed in Table S2. We purified the amplified products using the GFX^{TM} PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA). We performed cycle sequencing using BigDye 3.1 reagents, and sequences were generated on an ABI 3730 automated sequencer (Applied Biosystems, ABI, Beijing, China). We assembled and edited the resulting sequences using Geneious 6.1.2 (Drummond et al., 2011). Our methods for obtaining DNA and sequences from herbarium materials were similar to those used by Xu et al. (2015). In particular, we extracted total genomic DNA from herbarium materials following a modified CTAB protocol (Li et al., 2013). We carried out DNA repair PCR in a 50 μL volume containing 40 μL DNA, 5 μL 10× *Taq* buffer, 5 μL deoxynucleotide (dNTP) and 2 U *Taq* under the following conditions: initial denaturation at 94 °C for 4 min, 20 cycles of amplification at 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min, followed by a final extension at 72 °C for 10 min. We ligated the primers with five bases at the 5' end to tag them. Each PCR reaction of 10 μL contained 5.9 μL ddH_2O, 1 μL 10× *Taq* buffer, 1 μL dNTP, 0.5 μL 5 μmol/L forward tagged primer, 0.5 μL 5 μmol/L reverse tagged primer, 0.25 U *Taq* (0.1 μL), and 1 μL repaired DNA under the following conditions: initial denaturation at 94 °C for 3 min, 35 cycles of amplification at 94 °C for 30 s, 52 °C for 30 s, and 72 °C for 30 s, followed by a final extension at 72 °C for 10 min. The PCR products with particular tagged primers were pooled together and sequenced on a Roche 454 sequencer using a standard GS FLX Titanium Sequencing Kit XL+ (454 Life Sciences, Branford, CT, USA). Initially, we classified the sequences based on the different tags and primers. In order to confirm the authenticity of the sequences, we performed operational taxonomic unit cluster analysis in Usearch (Edgar, 2004), and removed the sequences with less than 10× coverage. We assembled and corrected the remaining sequences using Sequencher 5.0 (Gene Codes Co., Ann Arbor, MI, USA). In order to evaluate the reliability of the sequences, all sequences were blasted in GenBank and compared with sequences of the same family. All newly generated sequences were further evaluated by reconstructing a UPGMA tree using PAUP* 4.0 (Swofford, 2002).

Sequence alignment and data set construction

We applied a procedure in Geneious 6.1.2 to conduct sequence alignments. For the relatively fast-evolving markers *matK* and *ndhF*, we further improved the alignments through the following steps: (i) constructed a rough tree of the automatically aligned matrix using the maximum likelihood (ML) criterion; (ii) re-sorted the sequences in the matrix according to their phylogenetic placement; and (iii) further adjusted the alignment in Geneious 6.1.2 so that closely related species possessed similar alignment patterns. We repeated this procedure in Geneious until the alignment could not be further improved. We used all aligned sequences to construct a supermatrix, and gaps were treated as missing data.

Phylogenetic analyses

We performed a preliminary phylogenetic analysis on our supermatrix using ML as implemented in RAxML 8.0.24 (Stamatakis, 2014) on the CIPRES Science Gateway Web server (Miller et al., 2010) using 1000 bootstrap replicates. We applied the GTRCAT model for nucleotide substitution instead of the GTRGAMMA model to reduce run time. We treated each gene as a data partition and used independent model parameters on each. We used five species, *Anthoceros formosae* (hornworts), *Pellia endiviifolia*, *Aneura mirabilis* (liverworts), *Syntrichia ruralis*, and *Physcomitrella patens* (mosses) as out-groups in the analysis.

We checked our preliminary phylogeny by comparing it with previous studies to identify species with probable inappropriate systematic positions. We checked interfamilial relationships according to the Angiosperm Phylogeny website (http://www.mobot.org/MOBOT/research/APweb/) (Stevens, 2001), and infrafamilial and infrageneric relationships according to the most recent, well-supported molecular phylogenetic results on specific groups. For species with unexpected positions, we determined whether any of the sequences were chimeric or incorrectly identified using the BLAST tool in GenBank. We eliminated these sequences from the supermatrix and, when possible, replaced them with data from field or herbarium materials or from GenBank. We performed preliminary phylogenetic analyses and error checking as above until we found no additional errors in the matrix or in the tree.

We generated our final phylogeny under the ML criterion using the revised supermatrix. The final analysis comprised 1000 bootstrap replicates performed using the GTRGAMMA model of nucleotide substitution. Our final tree can be accessed in newick format and viewed in FigTree 1.4.2 (Rambaut, 2009) on our DarwinTree website.

Construction of subtrees

In order to better accommodate the use of our megaphylogeny for ecological studies, we developed the SoTree software for making subtrees from the complete phylogenetic tree of vascular plants. The complete tree is in newick format, which easily facilitates the generation of subtrees. Newick format uses a comma to represent bifurcations between two clades (including tips) (Olsen, 1990). Any clade, or set, has a branch length, except for the outermost set (e.g., an out-group) in the tree. Thus, we applied a novel parsing algorithm, called SoTree, across the complete tree in newick format to determine the branch lengths for each node from the parent node. Here, we provide an example of our algorithm using a simplified complete tree:

$$\big((a{:}n_1, b{:}n_2){:}n_3, (c{:}n_4, d{:}n_5){:}n_6\big); \quad (1)$$

$$\big((a{:}n_1, b{:}n_2){:}n_3, (c{:}n_4, d{:}n_5){:}n_6\big){:}0; \quad (2)$$

In the example, our algorithm first converts set (1) into set (2), with branch length initialized to "0". In the second step, the algorithm splits set (2) according to the middle-most comma and records the current split times and branch lengths. These two steps are applied recursively until all nodes have branch lengths {a:n1,b:n2,c:n4,d:n5}. Meanwhile, the split chain of "a" and "b" is "[s_1,s_2]" and the split chain of "c" and "d" is "[s_1,s_3]". All of the parsing details are presented in Table S3.

The parsed data represent a binary tree data structure from which it is possible to generate subtrees with branch lengths from a user-specified taxonomic list. Thus, to generate and

output a tree, a user may supply an input file comprising a list of input nodes (genus and species names) that occur in the complete tree. For example, input nodes "a,b,c":

- To find a common node for all nodes and record, split chain as follows:
  a:[s_1,s_2];b: [s_1,s_2]; c:[s_1,s_3];
- To classify according to the common node and split chain as follows: "((a,b), c)";
- To adjust the branch lengths according to the split chain of subclades and the branch lengths of parent nodes (see Table S3), as follows:
  "((a:n_1,b:n_2):n_3,c:n_4+n_6):0";
- To perform recursive computing as in the above steps, for any level of integration tree in practice.

Input nodes are "a, b, c"; split numbers are "s_1, s_2, s_3"; and branch lengths are "n_1, n_2, n_3, n_4, n_5, n_6".

Results

Data set

Our supermatrix sampled 6098 species representing 3114 genera (ca. 93% of vascular plant genera in China) and 323 families of Chinese vascular plants with five genera of bryophytes as out-groups. The total samples included 2909 genera of angiosperms, 42 of gymnosperms, 5 of lycophytes, and 158 of ferns (Tables S4, S5). Of these sequences, 1803 were newly generated in this study: 349 *matK*, 569 *rbcL*, 358 *atpB*, 181 *ndhF*, and 346 *matR* (Table S6). For herbarium specimens, 73 sequences (26 for *matK* and 47 for *rbcL*) were newly generated, representing 47 species/genera.

Phylogenetic analysis

Our phylogeny of the Chinese vascular plants was generally well resolved. The 50% ML majority rule bootstrap (BS) consensus tree is shown in Fig. S1. The relationships of orders and families are summarized in Figs. 1 and 2A–2J.

Lycophytes and Ferns

In our analysis, both lycophytes and ferns were recovered as well-resolved clades with the majority of families supported by BS values of 100%. Marattiaceae are sister to the rest of the fern clade, followed by Ophioglossaceae + Psilotaceae and Equisetaceae + leptosporangiate ferns. However, the branching order among Marattiaceae, Ophioglossaceae + Psilotaceae, Equisetaceae, and leptosporangiate ferns is not well supported. The species-rich families of ferns, such as Pteridaceae, Thelypteridaceae, Dryopteridaceae, and Polypodiaceae, were strongly supported. Within the leptosporangiate fern clade, we found some sister relationships, such as Gleicheniaceae sister to Dipteridaceae, Schizaeaceae to Lygodiaceae, and Salviniaceae to Marsileaceae, with moderate to high BS values (Fig. 2A).

Lycophytes were recovered as sister to all seed plants. Within lycophytes, the tree resolved relationships as Lycopodiaceae sister to Isoetaceae + Selaginellaceae (Fig. 2B).

Gymnosperms

Our phylogeny does not support gymnosperms as monophyletic. Gnetales, represented by *Gnetum* and *Ephedra*, were strongly supported as sister to a clade, including all other gymnosperms plus angiosperms (98% BS). Gymnosperms, except Gnetales, formed a well-supported clade, with Cycadaceae as sister to the remainder, followed by Ginkgoaceae, Pinaceae, Podocarpaceae, Araucariaceae, Sciadopityaceae, Cupressaceae, Cephalotaxaceae, and Taxaceae, which were each supported with 100% BS, except Taxaceae (95% BS) (Fig. 2B).

Angiosperms

Within angiosperms, Nymphaeales, including Nymphaeaceae and Cabombaceae, were sister to all other angiosperms, followed by Austrobaileyales, including Schisandraceae (*Schisandra*, *Kadsura*) and Illiciaceae (*Illicium*). Many early diverging clades that correspond to orders and families received strong support of 95%–100% BS, including Piperales comprising Piperaceae, Saururaceae, and Aristolochiaceae; Magnoliales including Annonaceae, Magnoliaceae, and Myristicaceae; and Laurales composed of Lauraceae, Hernandiaceae, and Calycanthaceae. We resolved Chloranthaceae as sister to monocots, but the support value was not high (52% BS) (Fig. 2B).

The monophyly of monocots was strongly supported, with 97% BS. Within monocots, most families were well supported as monophyletic, and the orders and interordinal relationships were also strongly supported. Acoraceae, Alismatales, and Petrosaviaceae were successive sisters to all other monocots with strong support (97% BS, respectively). Within Alismatales, all families are primarily aquatic and were well supported, including Alismataceae, Hydrocharitaceae, Butomaceae, Aponogetonaceae, Scheuchzeriaceae, Juncaginaceae, Cymodoceaceae, Ruppiaceae, Posidoniaceae, Zosteraceae, Potamogetonaceae, Tofieldiaceae, and Araceae. The sister-group relationship between Dioscoreales and Pandanales was weakly supported (BS $<$ 50%).

Dioscoreales were strongly supported as monophyletic (84% BS), as were all of its families (Dioscoreaceae, Burmanniaceae, and Nartheciaceae). Pandanales included the well-supported families Pandanaceae and Stemonaceae, and one monotypic family, Acanthochlamydaceae. Liliales included four strongly supported families with 94% BS, namely Liliaceae, Smilacaceae, Melanthiaceae, and Colchicaceae. Orchidaceae are very diverse in China and were resolved as monophyletic with strong support (100% BS). The other families of Asparagales are Hypoxidaceae (100% BS), Ixioliriaceae, Iridaceae (100% BS), Asphodelaceae (100% BS), and two large families, Amaryllidaceae (88% BS) and Asparagaceae (63% BS) (Fig. 2C).

Arecales, which until recently comprised only Arecaceae but now also include Dasypogonaceae (APG IV, 2016), were strongly supported (100% BS). Zingiberales (98% BS) include the large Zingiberaceae (100% BS), Costaceae, Cannaceae, Marantaceae, Musaceae, and Lowiaceae. Commelinales include Commelinaceae (100% BS), Pontederiaceae (100% BS), and Philydraceae. In Poales, Typhaceae were sister to the rest of the clade, with Bromeliaceae as the second-diverging clade. Cyperaceae (100% BS) were sister to Juncaceae. Poaceae formed a clade with Flagellariaceae, Restionaceae, Xyridaceae, and Eriocaulaceae; the relationships were moderately supported (86% BS). Poaceae were supported as monophyletic with 97% BS. Within Poaceae,

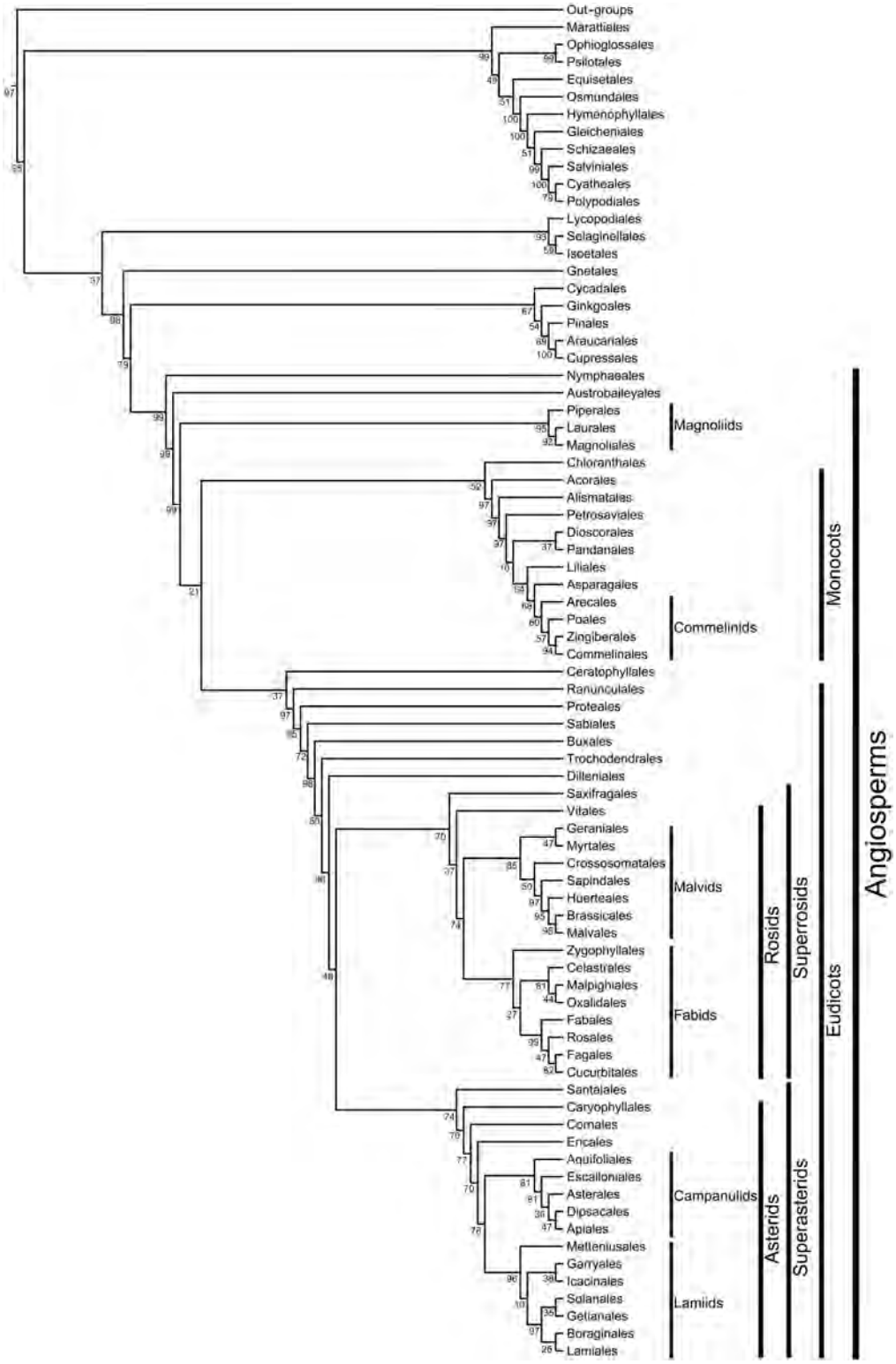


Fig. 1. The ordinal relationships of Chinese vascular plants.

Leptaspis cochleata was sister to the rest of the family (97% BS). The infrafamilial relationships in Cyperaceae and Poaceae were well resolved (Fig. 2D).

Ceratophyllaceae were resolved as sister to eudicots with 37% BS. Ranunculales, including Papaveraceae (100% BS), Eupteleaceae, Lardizabalaceae (100% BS), Circaeasteraceae (100% BS), Menispermaceae (100% BS), Berberidaceae (100% BS), and Ranunculaceae (100% BS), are sister to the rest of the eudicots. Other early-diverging eudicots include Proteales (100% BS; Nelumbonaceae, Platanaceae, and Proteaceae),

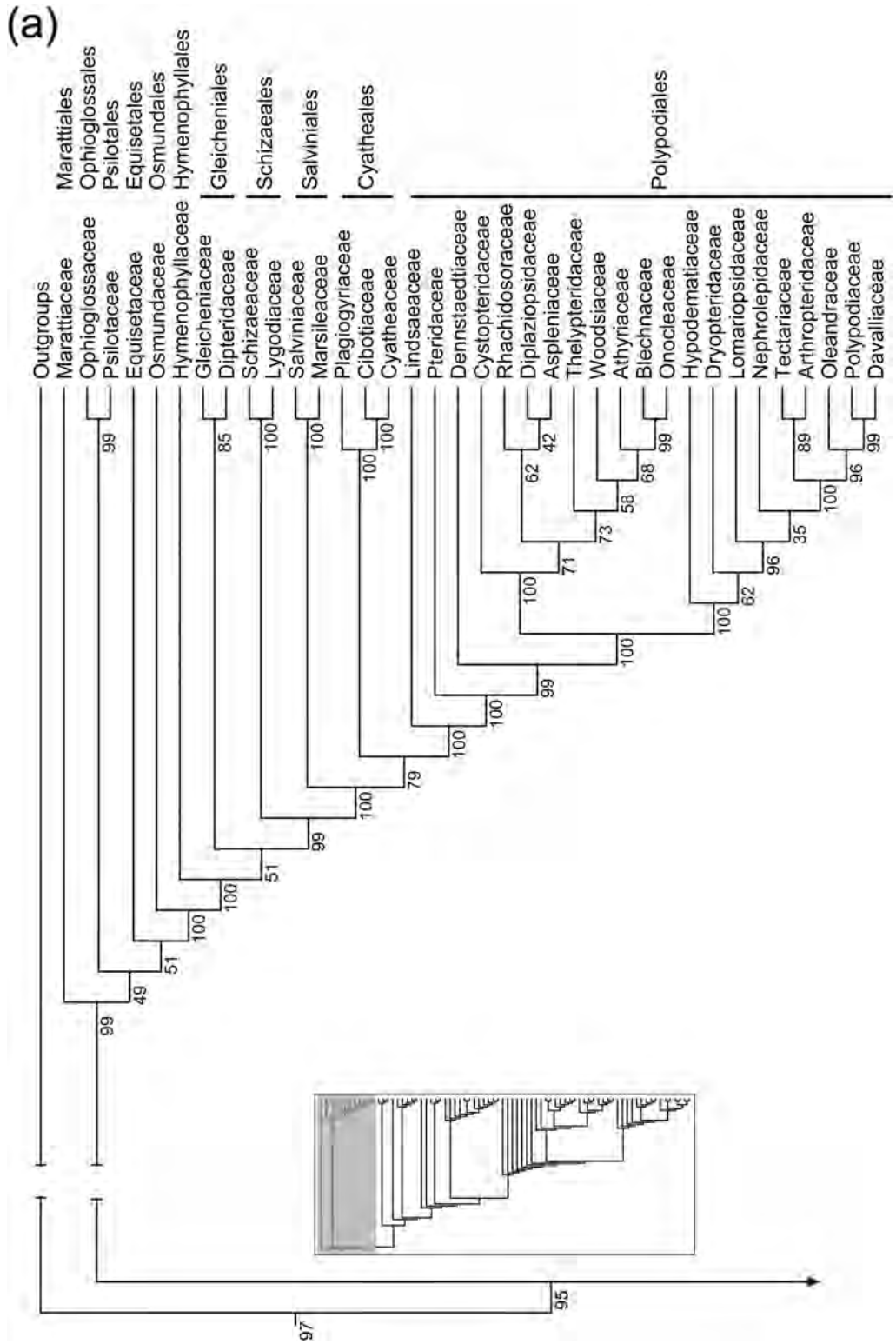


Fig. 2. Phylogenetic relationships among families of major groups of Chinese vascular plants (**A**, Ferns; **B**, Lycophytes, gymnosperms, and early diverging clades of angiosperms; **C**, Chloranthales and monocots, except Commelinids; **D**, Commelinids of monocots, Ceratophyllales, and basal eudicots; **E**, Dilleniales and part of Superrosids; **F**, Part of Superrosids; **G**, Part of Fabids and Santalales; **H**, Caryophyllales, Cornales, and Ericales of Asterids; **I**, Campanulids of Asterids; **J**, Lamiids of Asterids).

Continued

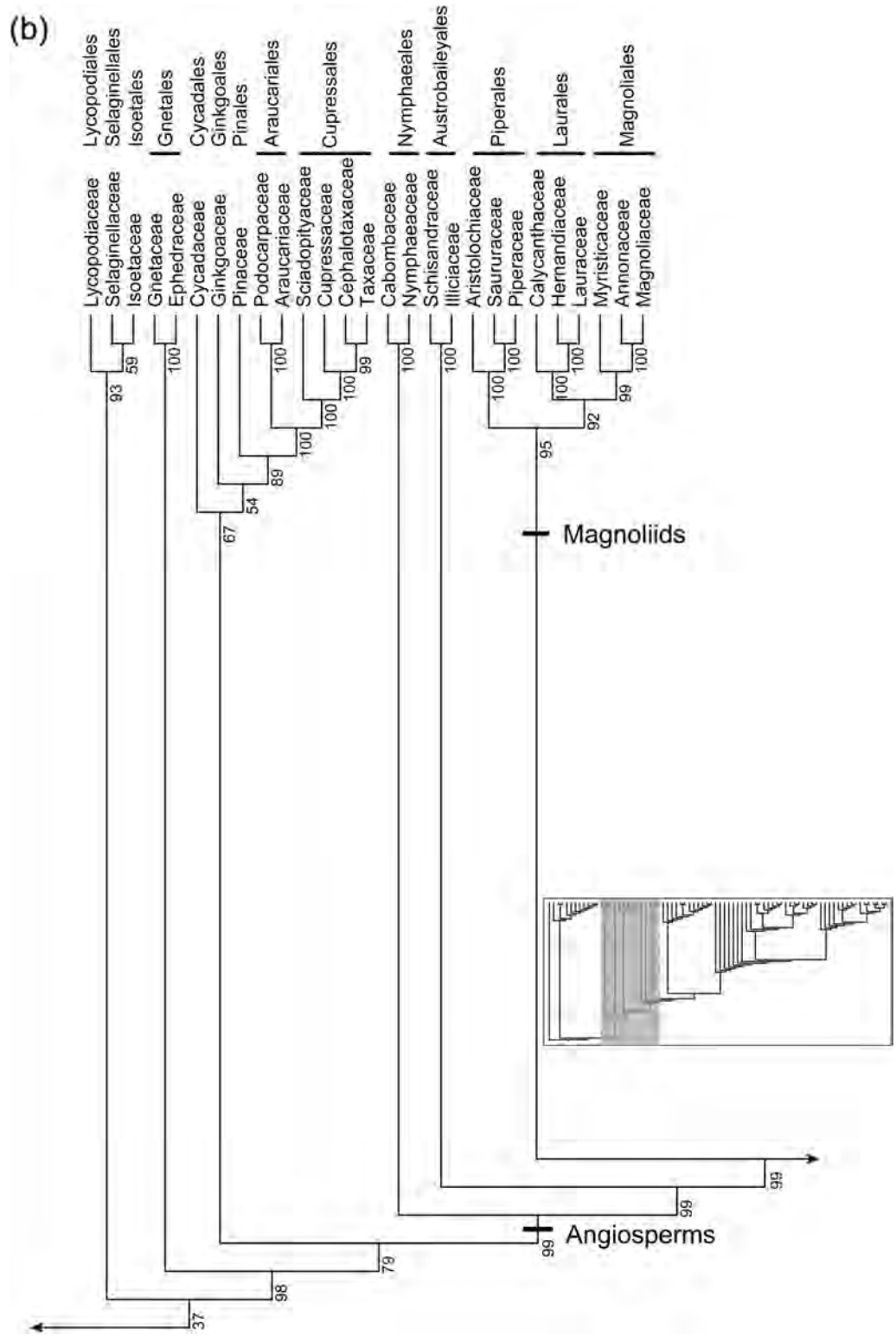


Fig. 2. Continued

Sabiales (*Sabia*, *Meliosma*), Buxales, and Trochodendrales (Fig. 2D). Core eudicots include Dilleniales, superrosids (Saxifragales, rosids), and superasterids (Santalales, Caryophyllales, and asterids). Dilleniales were sister to all other core eudicots (Fig. 2E).

The monophyly of Saxifragales was well supported (89% BS). Within the order, Hamamelidaceae (82% BS), Daphniphyllaceae, Altingiaceae (100% BS), Cercidiphyllaceae, and Paeoniaceae form a weakly supported subclade 1 (60% BS). The other two strongly supported subclades are the core Saxifragales

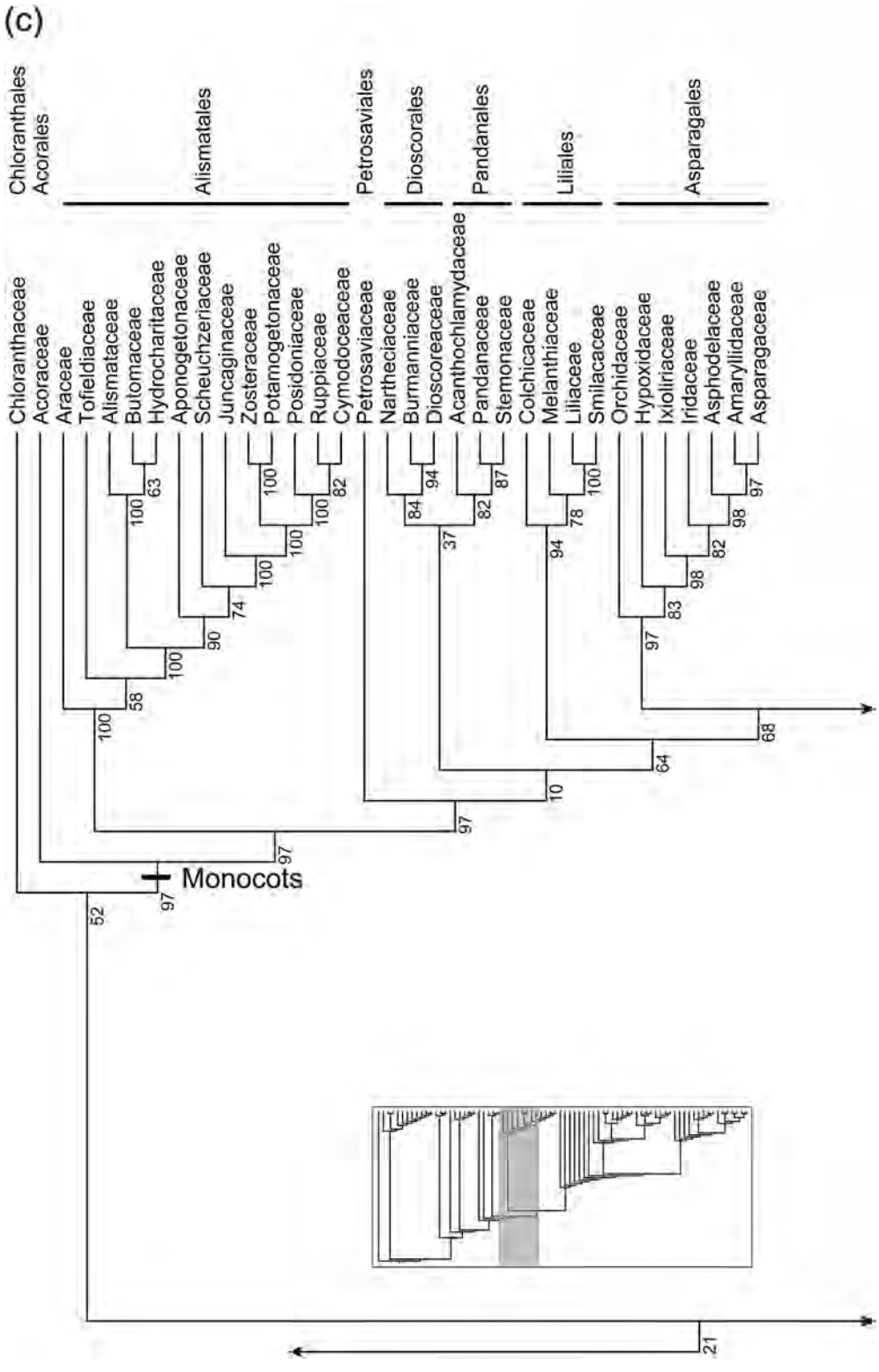


Fig. 2. Continued

(99% BS), comprising Crassulaceae (99% BS), Penthoraceae, and Haloragaceae (100% BS), and subclade 2 including Iteaceae, Grossulariaceae (100% BS), and Saxifragaceae (99% BS). Together, they are sister to the parasitic family Cynomoriaceae, but this relationship is not well supported (44% BS) (Fig. 2E).

Vitales were sister to all other rosids, although this relationship was not well supported (37% BS) (Figs. 1, S1). After Vitales, our phylogeny indicated two major clades of rosids: fabids (77% BS) and malvids (85% BS), which are comparable with *Fabidae* (or eurosids I) and *Malvidae* (or eurosids II),

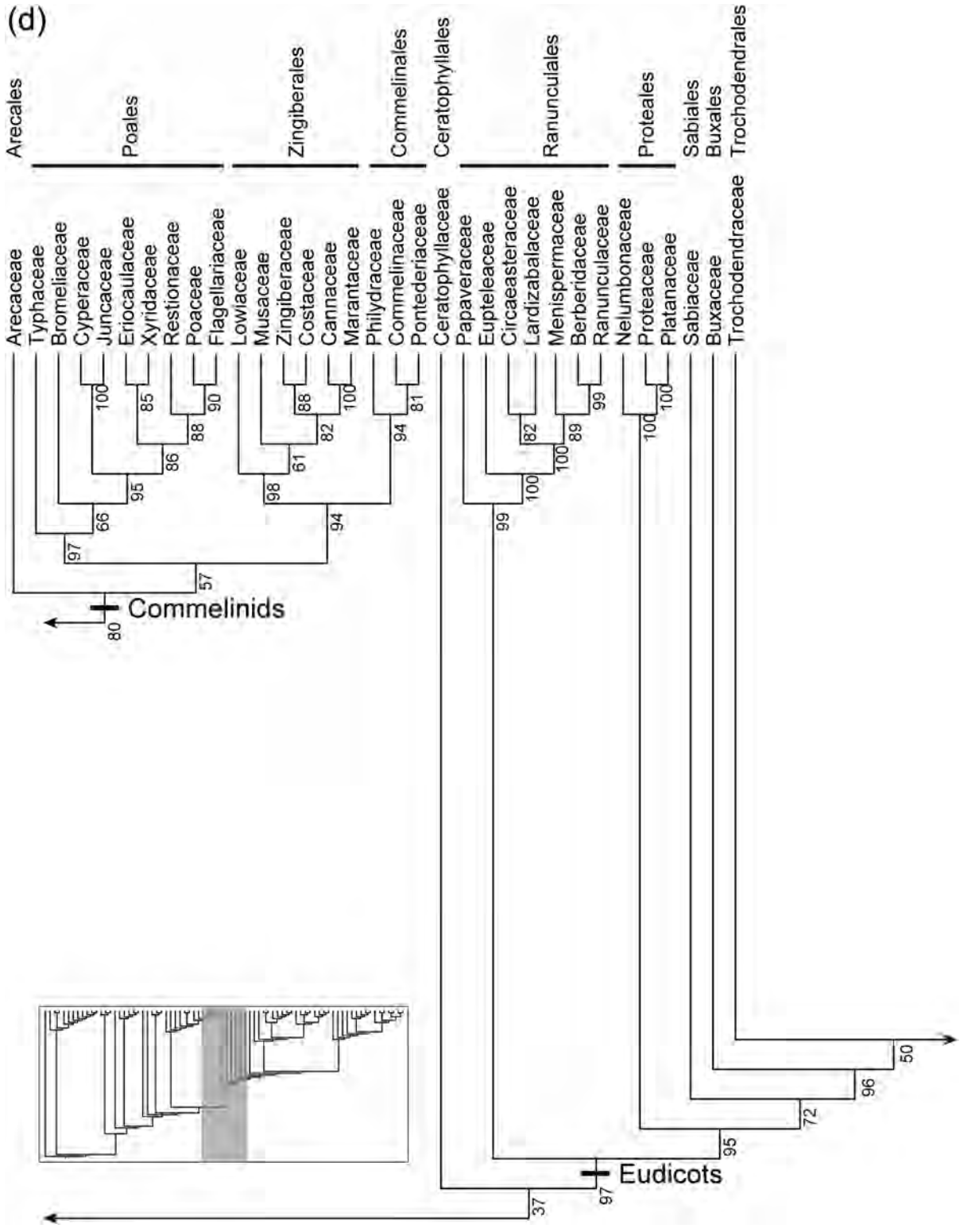


Fig. 2. Continued

respectively, in previous studies (e.g., Soltis et al., 2007, 2011; Wang et al., 2009). In malvids, there are seven orders: Geraniales (Geraniaceae 100% BS); Myrtales (including Lythraceae 94% BS; Onagraceae 99% BS; and Combretaceae, Myrtaceae, Melastomataceae 100% BS; Crypteroniaceae); Crossosomatales (Stachyuraceae, Staphyleaceae 100% BS); Sapindales (100% BS) (Anacardiaceae 85% BS; Meliaceae 92% BS; Simaroubaceae, Rutaceae 99% BS; Nitrariaceae, Burseraceae, Sapindaceae 100% BS; Biebersteiniaceae); Huerteales (100% BS) (Tapisciaceae; Dipentodontaceae 100% BS, including

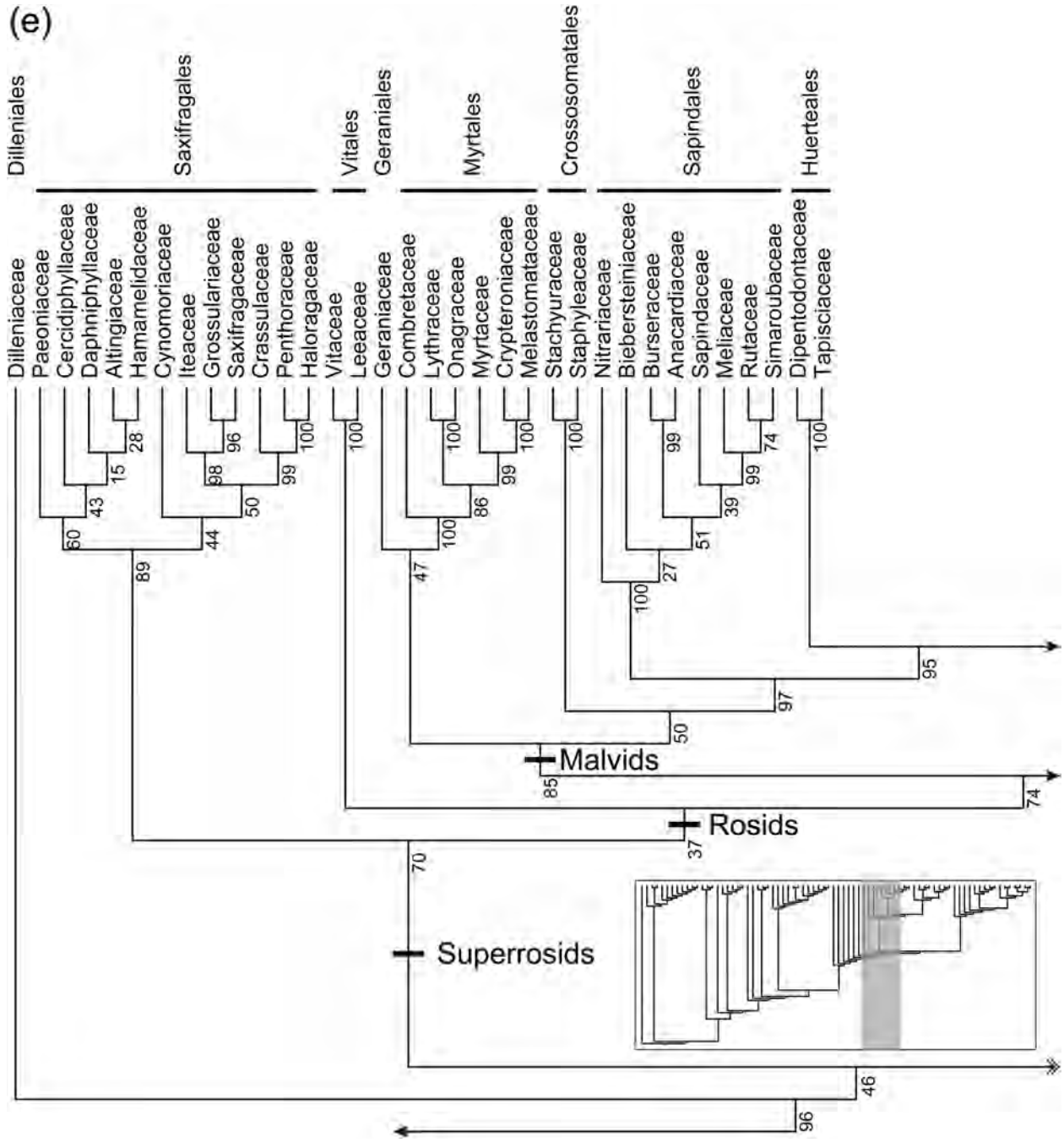


Fig. 2. Continued

Dipentodon and *Perrottetia*); Malvales (99% BS) (Bixaceae; Thymelaeaceae 99% BS; Malvaceae 100% BS; paraphyletic Dipterocarpaceae, and Cistaceae); and Brassicales (97% BS), with several small early-diverging families (Tropaeolaceae, Bretschneideraceae, Moringaceae, Caricaceae, Salvadoraceae) and other families (Capparaceae 86% BS; Brassicaceae 97% BS; Cleomaceae, Resedaceae, Stixaceae 100% BS, and monotypic Borthwickiaceae). All of the orders are well supported (Figs. 2E, 2F). Within fabids (77% BS), three subclades were recovered, one of which is the nitrogen-fixing clade (Soltis et al., 1995, 2005), which contains Fabales (Fabaceae 99% BS; Polygalaceae 100% BS; Surianaceae); Rosales (Rhamnaceae 89% BS; Urticaceae 96% BS; Cannabaceae 98% BS; Rosaceae 99% BS; Elaeagnaceae, Ulmaceae, Moraceae 100% BS); Cucurbitales (Cucurbitaceae, Coriariaceae, Begoniaceae 100% BS; Tetramelaceae); and Fagales (Fagaceae, Juglandaceae, Myricaceae, Casuarinaceae, Betulaceae 100% BS; Rhoipteleaceae). The second subclade includes Celastrales, Oxalidales, and Malpighiales (COM clade; 81% BS). Inside the COM clade, Celastrales (100% BS) (Celastraceae 99% BS; Parnassiaceae 100% BS) were weakly supported as sister to Oxalidales and Malpighiales.

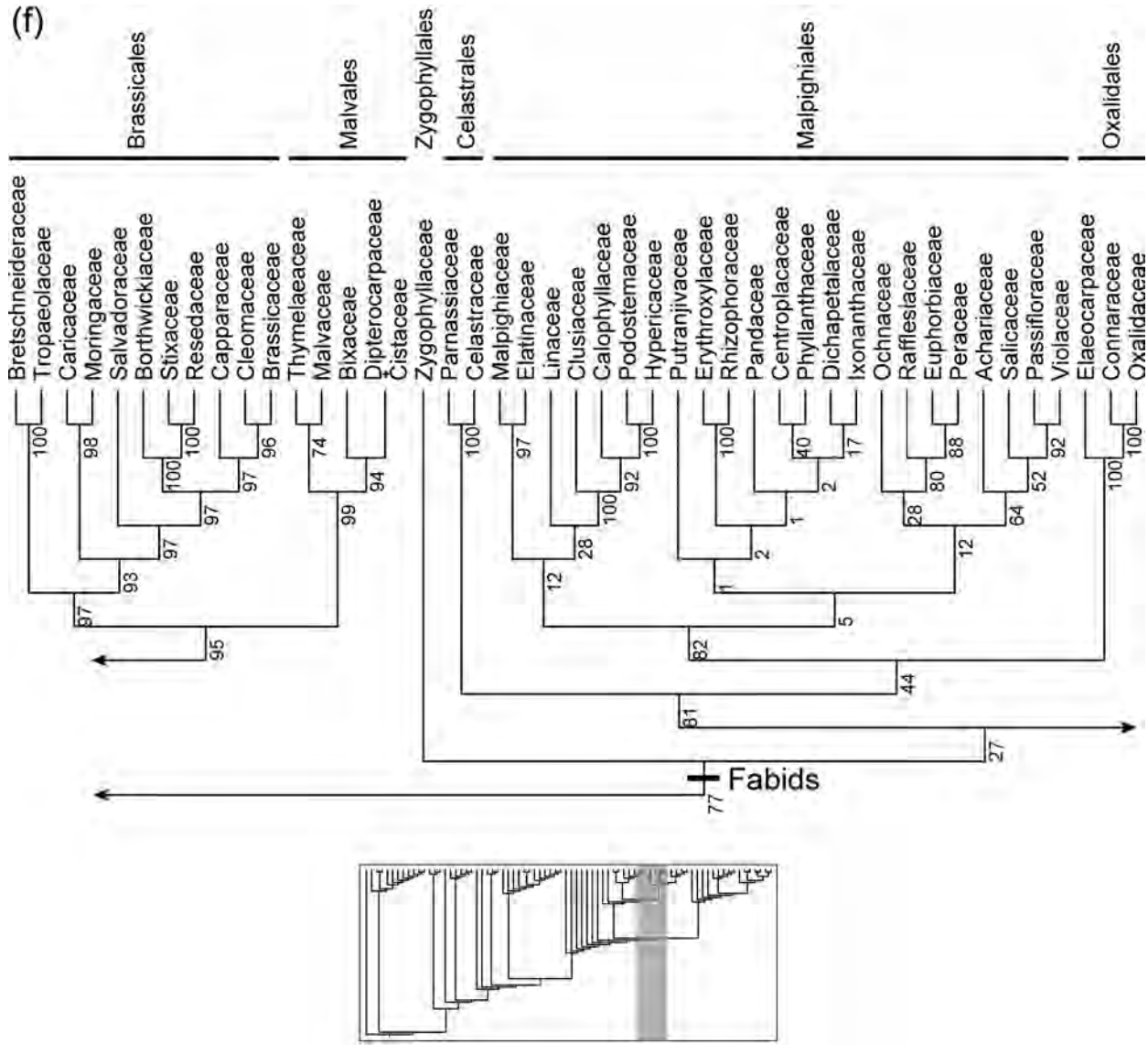


Fig. 2. Continued

Oxalidales (100% BS) included Elaeocarpaceae (100% BS), Oxalidaceae (100% BS), and Connaraceae (100% BS). Malpighiales (82% BS) comprised a large order including the families: Malpighiaceae (98% BS); Elatinaceae (100% BS); Clusiaceae (100% BS); Calophyllaceae (96% BS); Podostemaceae (100% BS); Hypericaceae (100% BS); Phyllanthaceae (99% BS); Centroplacaceae and Putranjivaceae (100% BS); Achariaceae (66% BS); Violaceae (100% BS); Passifloraceae (100% BS); Salicaceae (98% BS, including Flacourtiaceae); Rhizophoraceae (100% BS); Erythroxylaceae (100% BS); Pandaceae and Linaceae (100% BS); Ochnaceae (90% BS); Ixonanthaceae and Dichapetalaceae (82% BS); and Rafflesiaceae, Peraceae, and Euphorbiaceae (85% BS). The third subclade includes Zygophyllales (100% BS) as sister to a clade combining the nitrogen-fixing and COM subclades, but with low support (27% BS). All orders and most of the interordinal relationships in fabids are also well supported (Figs. 2F, 2G).

Our results showed strong support for the monophyletic Santalales and Caryophyllales (97% BS and 100% BS, respectively). Together, Santalales, Caryophyllales, and asterids form a clade (74% BS); within this clade, Caryophyllales are sister to asterids (70% BS). In Santalales, Santalaceae are paraphyletic. In Caryophyllales, the morphologically defined large families, such as Amaranthaceae (97% BS, including Chenopodiaceae), Caryophyllaceae (97% BS), and Polygonaceae (100% BS), all received strong support in our phylogeny. We also recovered a well-supported clade (BS 100%), consisting of Cactaceae (100% BS); Portulacaceae, Talinaceae, and Basellaceae (100% BS); Molluginaceae (100% BS); Aizoaceae (80% BS); and Gisekiaceae, Phytolaccaceae, Petiveriaceae, and Nyctaginaceae (98% BS), and this clade was sister to a clade of Caryophyllaceae and Amaranthaceae. Droseraceae, Nepenthaceae, and Ancistrocladaceae form a clade (87% BS); Frankeniaceae were sister to Tamaricaceae, while Plumbaginaceae were a strongly supported monophyletic family (100% BS) and sister to Polygonaceae (Fig. 2H).

The asterids formed a monophyletic group with 77% BS support. Within asterids, Cornales (100% BS) (Hydrangeaceae,

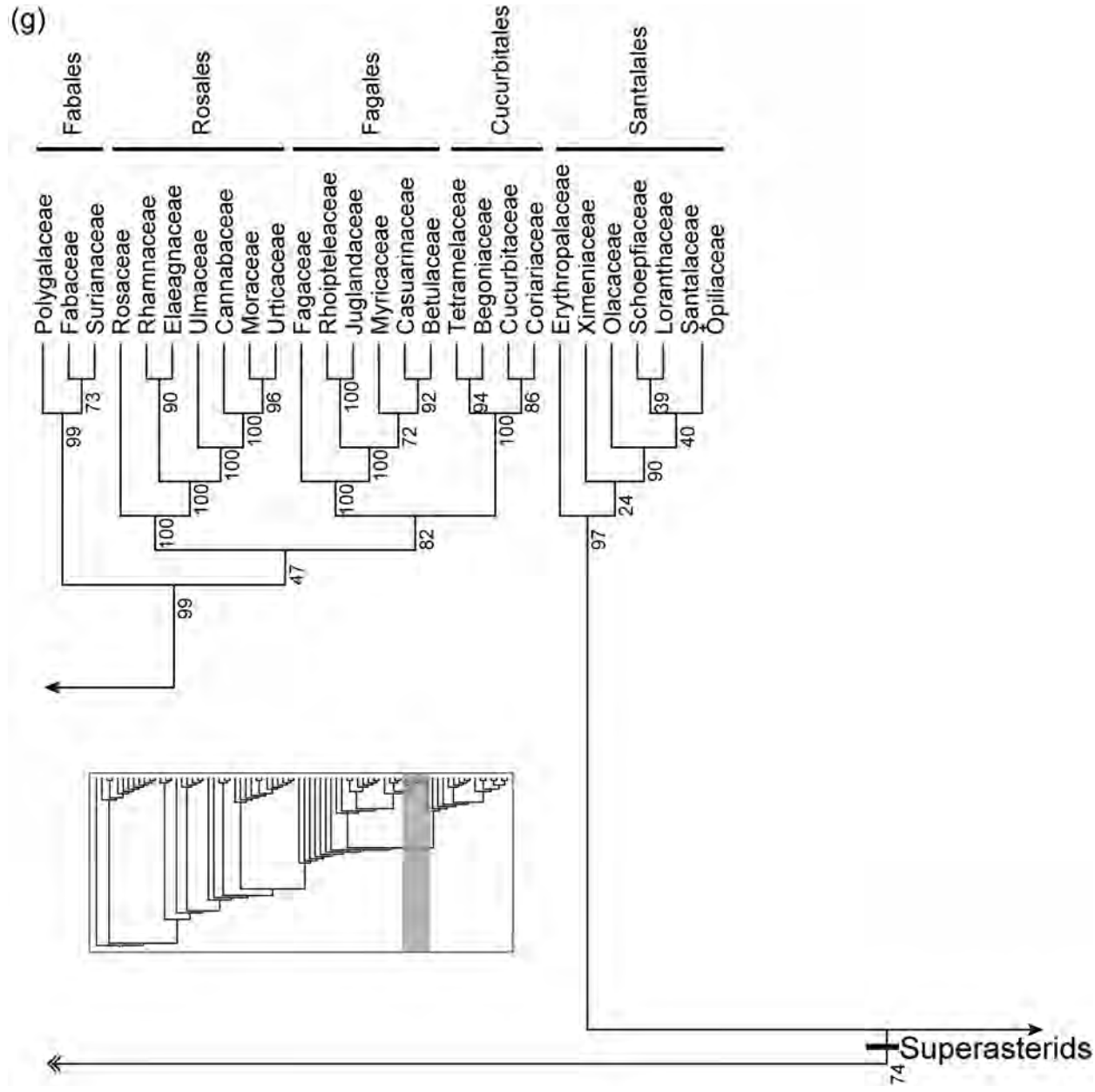


Fig. 2. Continued

Nyssaceae, Cornaceae 100% BS) were sister to all other asterids, which include Ericales (99% BS) (Pentaphylacaceae 66% BS; Styracaceae, Sapotaceae, Actinidiaceae 98% BS; Theaceae, Ericaceae 99% BS; Primulaceae including Mysinaceae, Ebenaceae, Balsaminaceae, Symplocaceae, Diapensiaceae 100% BS; Clethraceae, Lecythidaceae, Polemoniaceae, Sladeniaceae) and a weakly supported clade of euasterids (*sensu* Soltis et al., 2000) (Fig. 2H). Ericales were sister to euasterids with moderate support (70% BS). There are two major clades of euasterids. One was well supported (81% BS) and comprised Aquifoliales (Cardiopteridaceae 98% BS; Aquifoliaceae 100% BS, as sister to the rest of the order; Helwingiaceae, Stemonuraceae), Escalloniales (represented by *Polyosma alangiacea*, Escalloniaceae), Apiales (Apiaceae 53% BS; Araliaceae 69% BS; Pittosporaceae 100% BS; Torricelliaceae), Dipsacales (100% BS) (Adoxaceae, Caprifoliaceae 100% BS) and Asterales (99% BS) (Campanulaceae 100% BS; Asteraceae, Menyanthaceae 99% BS; Goodeniaceae 100% BS; Pentaphragmataceae, Stylidiaceae) (Fig. 2I). The other clade, the lamiids (*sensu* Cantino et al., 2007), was also well supported (96% BS). Metteniusaceae are sister to all other lamiids comprising Garryales (monotypic Eucommiaceae; Garryaceae represented by *Aucuba*) and Icacinaceae; Boraginales (Cordiaceae, 94% BS; Heliotropiaceae, Ehretiaceae, Boraginaceae 100% BS); Solanales (Solanaceae, Convolvulaceae 100% BS; Sphenocleaceae, Hydroleaceae); Gentianales (100% BS) (Rubiaceae, Loganiaceae, Gentianaceae, Apocynaceae 100% BS; Gelsemiaceae); and Lamiales (98% BS) (Plantaginaceae 58% BS; Verbenaceae 94% BS; Gesneriaceae 98% BS; Acanthaceae, Bignoniaceae, Orobanchaceae,

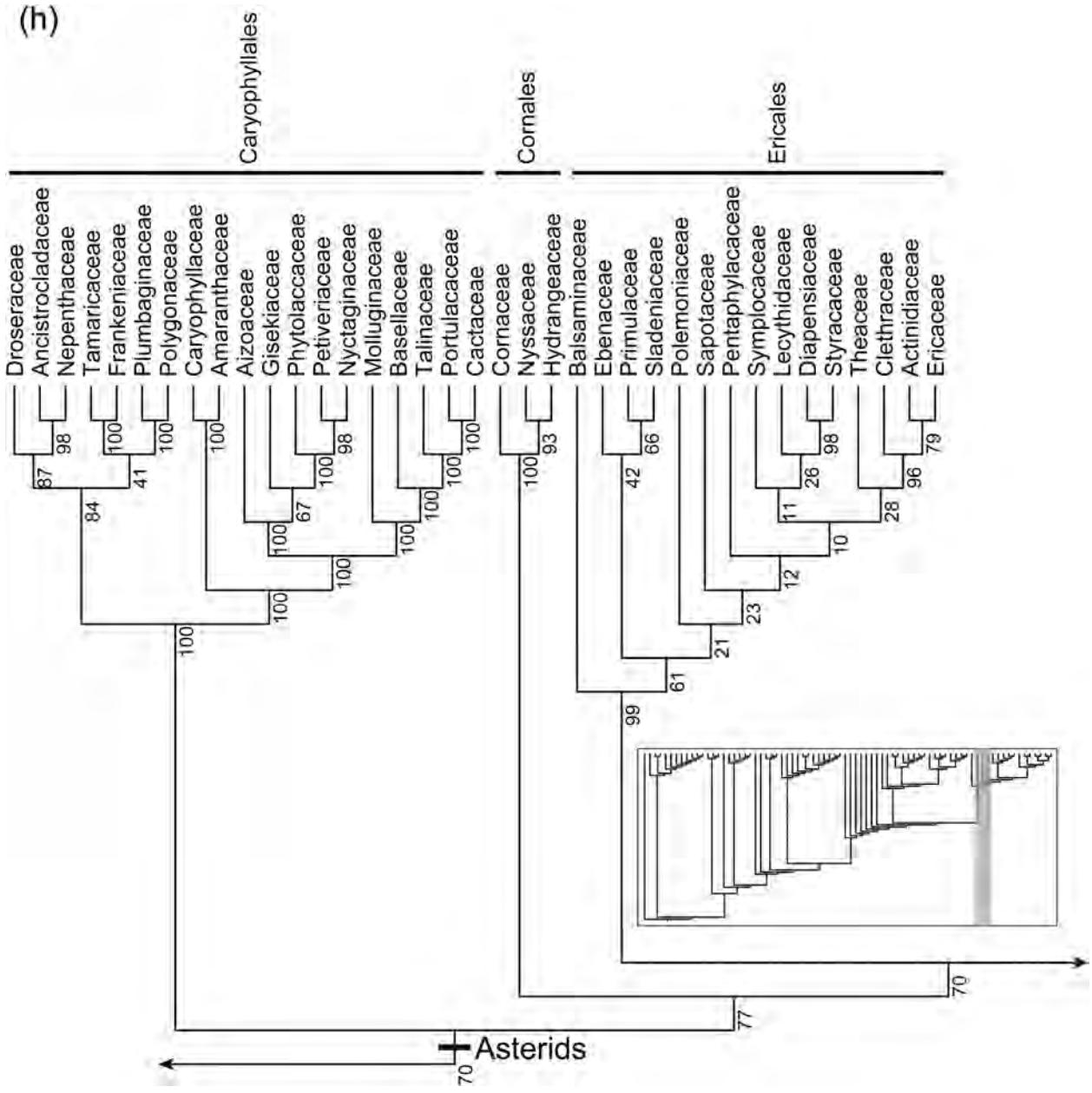


Fig. 2. Continued

Phrymaceae, Lamiaceae 99% BS; Oleaceae, Carlemanniaceae, Scrophulariaceae, Linderniaceae, Lentibulariaceae, Mazaceae 100% BS; Paulowniaceae, Pedaliaceae, Martyniaceae) (Fig. 2J).

Tree converting (SoTree)

Our newly developed software, SoTree, was successful in generating subtrees from our phylogeny of Chinese vascular plants. We developed the software to read and compare an input file of species names with the taxonomic sampling in our phylogeny. Our software automatically performs taxonomic reconciliation using FOC and returns a subtree comprising only species from the input file or closely related species, which have the same systematic relationships according to the previously published phylogeny (Figs. 3, S2, S3).

SoTree returns subtrees in newick format, and this may meet the needs of many in the target user group. However, we have also implemented options for users to select among input and output formats and to manipulate the returned subtree. SoTree also generates and returns a log file. We have implemented SoTree on the DarwinTree website (http://www.darwintree.cn/flora/index.shtml).

Discussion

Our genus-level phylogenetic tree of Chinese vascular plants showed that most traditionally accepted orders and families were monophyletic and strongly supported. Additionally, the phylogeny indicated strong support for some interordinal and

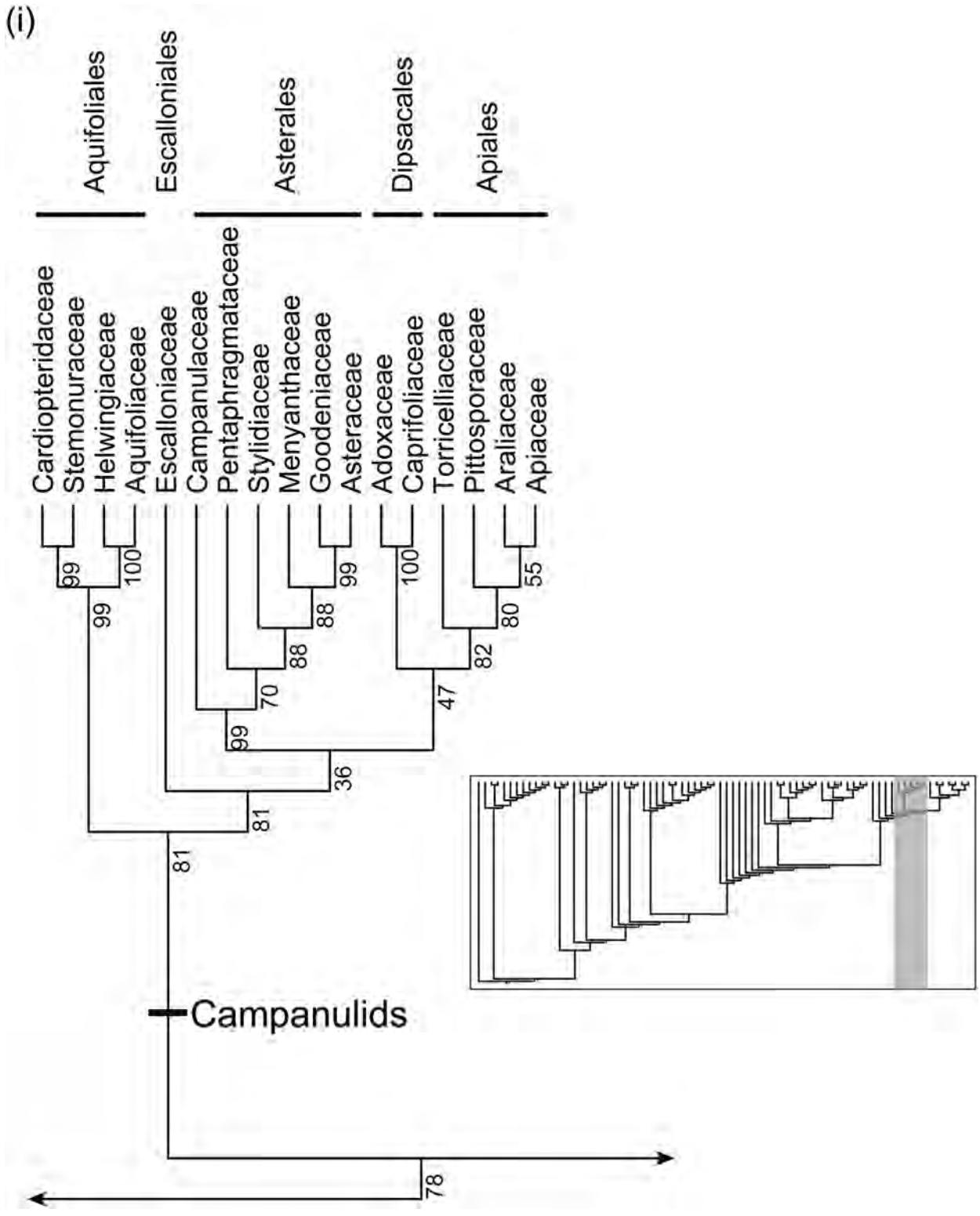


Fig. 2. Continued

higher-level relationships. Within lycophytes, ferns, and gymnosperms, almost all the families and lineages above families are fully resolved and largely consistent with published molecular phylogenies (Qiu et al., 2006; Chaw et al., 2000; Pryer et al., 2001; Fiz-Palacios et al., 2011; Ruhfel et al., 2014). Lycophytes, as shown in the current study, are monophyletic and sister to seed plants. Although this relationship conflicts with the current understanding of tracheophyte relationships (e.g., Pryer et al., 2001), Ruhfel et al. (2014) showed recently that it was an artifact resulting

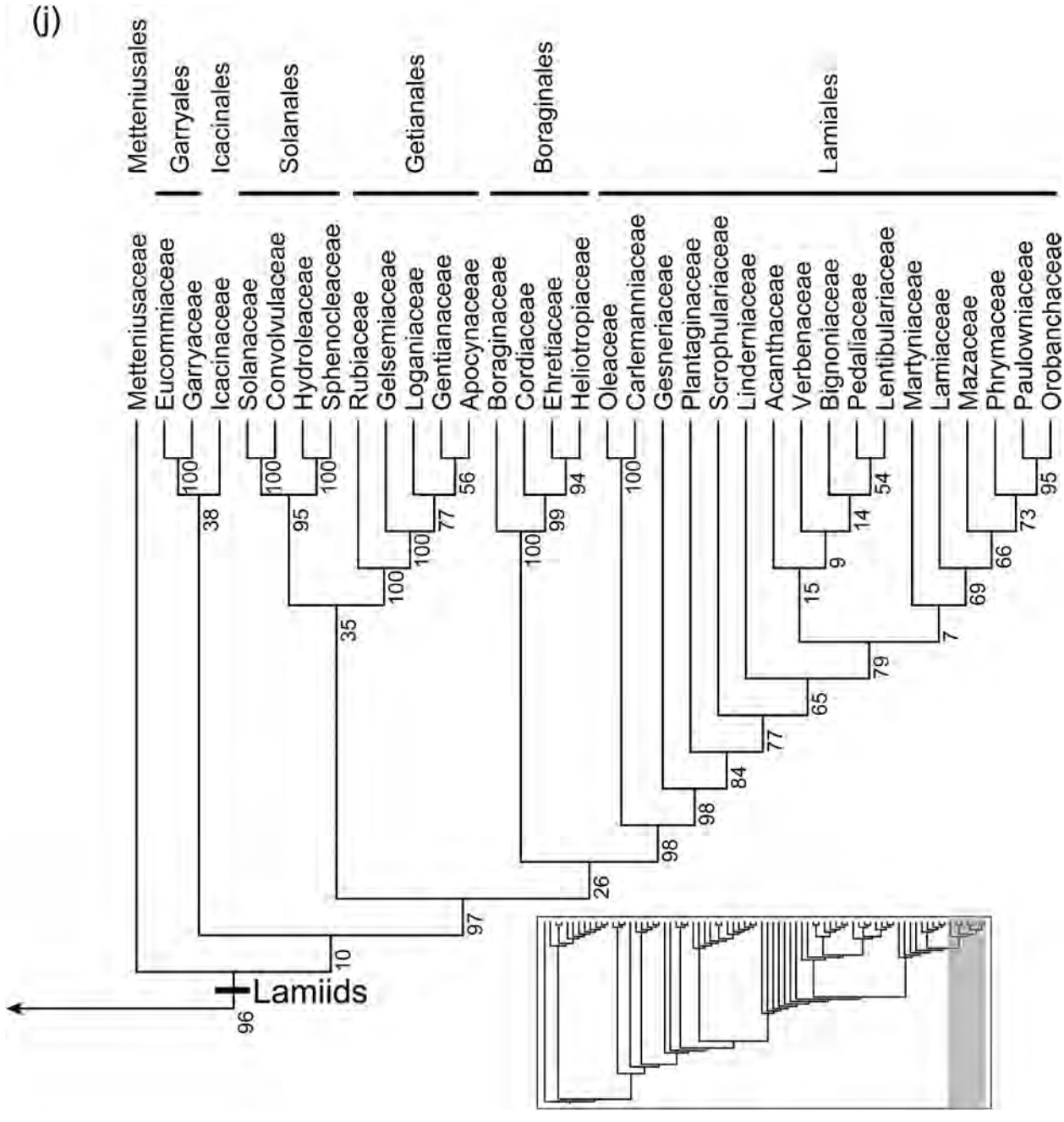


Fig. 2. Continued

from high guanine-cytosine (GC) content. Among ferns, the backbone relationships of the early-diverging lineages are still not fully resolved, even with much denser sampling (Liu, 2016). For example, the position of Equisetaceae in this study was consistent to some extent with previous studies that used similar genes (Pryer et al., 2001; Liu, 2016), but recent studies based on whole plastid genomes recovered this family as sister to a clade comprising Ophioglossaceae and Psilotaceae (e.g., Gao et al., 2013; Lu et al., 2015).

Our results showed that Gnetales were sister to all other seed plants. Prior studies using fast-evolving chloroplast genes (Magallón & Sanderson, 2002; Rydin et al., 2002) showed the same relationships that we recovered with more slowly evolving genes that were more easily aligned for the group (i.e., *atpB*, *rbcL*, and *matK*). In contrast, other previous studies using single or multiple genes from the chloroplast, mitochondrial, and nuclear genomes have supported Gnetales as sister to Pinaceae (Goremykin et al., 1996; Winter et al., 1999; Bowe et al., 2000; Chaw et al., 2000; Frohlich & Parker, 2000; Gugerli et al., 2001; Burleigh & Mathews, 2004; Qiu et al., 2007) or to conifers (Chaw et al., 1997; Burleigh & Mathews, 2004; Ran et al., 2010; also see review by Ickert-Bond & Renner, 2016). Recently, protein-coding sequence data from the plastid genome for 78 genes from 360 green plant taxa recovered

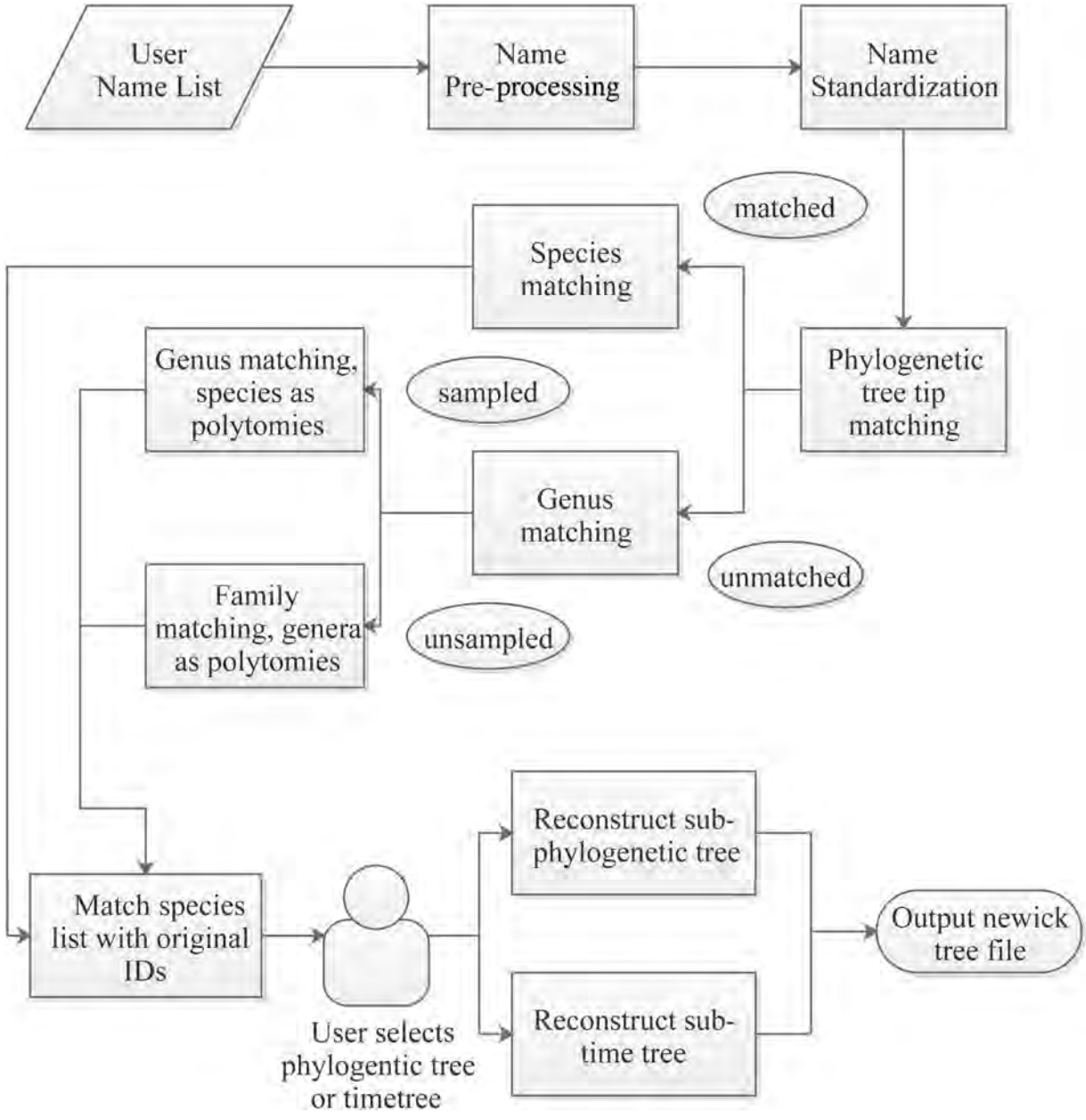


Fig. 3. Flowchart in SoTree to obtain the subtree by providing an input file of species names from users.

Gnetales as sister to non-Pinaceae conifers, placing Gnetales as sister to a clade of Araucariales + Cupressales, and the Gne-Arau-Cup clade is then sister to Pinales (Ruhfel et al., 2014). Our phylogeny showed that within gymnosperms, all families were monophyletic, and Sciadopityaceae (monotypic: *Sciadopitys verticillata* endemic to Japan) are sister to the Taxaceae-Cephalotaxaceae-Cupressaceae clade.

For angiosperms, our results are generally in agreement with previous, large-scale analyses (Qiu et al., 1999; Soltis et al., 2000; Ruhfel et al., 2014), especially the phylogenetic framework of the updated APG classification system (APG IV, 2016). Nymphaeales and Austrobaileyales were the earliest-diverging lineages of extant angiosperms, followed by Piperales, Magnoliales, Laurales, Chloranthaceae, monocots, Ceratophyllaceae, and eudicots. Within eudicots, Ranunculales were sister to all other eudicots, followed by other early-diverging eudicot groups, Proteales, Sabiaceae, Trochodendraceae, Buxaceae, and core eudicots. The clades of core eudicots, such as rosids, asterids, Saxifragales, Santalales, and Caryophyllales, were all well resolved and congruent with results from previous studies (Chase et al., 1993; Olmstead et al., 1992, 2000; Olmstead & Sweere, 1994; Soltis et al., 1999, 2000, 2011; Hilu et al., 2003; Zhu et al., 2007; Burleigh et al., 2009). The regional tree of vascular plants in China recovered families that were consistent with those in the APG classification system and also resolved their infrafamilial relationships. However, our study indicates that some traditionally recognized genera are not monophyletic and need further investigation.

Concept and name of families used in this study: A comparison between APG and FOC

Our sampling scheme followed the taxonomic treatment in FOC. However, our phylogeny is more consistent with APG IV (2016). All of the families recognized by APG were monophyletic in our tree, except Dipterocarpaceae and Santalaceae. Icacinaceae are no longer polyphyletic after *Pittosporopsis* and *Apodytes* were moved to form Metteniusaceae by Stull et al. (2015) during the course of our study. Therefore, we follow the APG system throughout this portion of the discussion (Table 1).

According to FOC, there are 312 families of vascular plants in China and 262 of them are angiosperms; however, APG IV (2016) lists 266 families of angiosperms. In comparison with FOC, some families were phylogenetically nested as close allies in our study. For example, Lemnaceae were merged into Araceae; Zannichelliaceae into Potamogetonaceae; Taccaceae into Dioscoreaceae; Centrolepidaceae to Restionaceae; Tetracentraceae into Trochodendraceae; Leeaceae into Vitaceae; part of Ulmaceae into Cannabaceae; part of Flacourtiaceae into Salicaceae and the rest into Achariaceae; Trapaceae into Lythraceae; Peganaceae into Nitrariaceae; Aceraceae and Hippocastanaceae into Sapindaceae; Cneoraceae into Rutaceae; Bombacaceae, Tiliaceae, and Sterculiaceae into Malvaceae; Viscaceae into Santalaceae; Alangiaceae into Cornaceae; Mastixiaceae into Nyssaceae; Myrsinaceae into Primulaceae; Asclepiadaceae into Apocynaceae; Diervillaceae, Dipsacaceae, Morinaceae, Linnaeaceae, and Valerianaceae into Caprifoliaceae; and *Hydrocotyle* of Apiaceae into Araliaceae. On the other hand, some large families were paraphyletic and are better represented by narrower concepts, such as Acanthochlamydaceae and Amaryllidaceae; Circaeasteraceae (including *Circaeaster* and *Kingdonia*) and Ranunculaceae; Centroplacaceae and Celastraceae; Altingiaceae (including *Altingia*, *Liquidambar* and *Semiliquidambar*) and Hamamelidaceae; Talinaceae and Portulacaceae; Calophyllaceae, Hypericaceae, and Clusiaceae; Peraceae, Phyllanthaceae, Putranjivaceae, and Euphorbiaceae; and Ximeniaceae, Erythropalaceae, Schoepfiaceae, and Olacaceae. Note, however, that the latter narrowly circumscribed families in Santalales have not been recognized by APG IV (2016), pending further study.

The concepts of some families are considerably different in APG IV compared to FOC. For example, Liliaceae *sensu* FOC are morphologically diverse and related to Amaryllidaceae and Asparagaceae. In contrast, APG IV separates Tofieldiaceae, Petrosaviaceae, Nartheciaceae, Melanthiaceae, Colchicaceae, Smilacaceae, Asphodelaceae, and Liliaceae s.s. from Amaryllidaceae and Asparagaceae, and the latter relationship is consistent with our results. The circumscription of Saxifagaceae *sensu* FOC is completely different from that in this study, which supports dividing the family into Iteaceae, Grossulariaceae, Saxifragaceae s.s., Penthoraceae, Parnassiaceae, and Hydrangeaceae. Although the first four of these six families are all part of Saxifragales, Parnassiaceae are part of the rosid order Celastrales, and Hydrangeaceae are in Cornales in the asterid clade. Stixaceae and Borthwickiaceae were separated from Capparaceae recently (Doweld & Reveal, 2008; Su et al., 2012). We maintain the family statuses here, although they were included in an expanded Resedaceae by APG IV (2016). Our results are consistent with the recent splitting of Boraginaceae s.l. into seven families (Weigend et al., 2013; Refulio-Rodríguez & Olmstead, 2014) within an order of Boraginales, although APG IV still recognizes the single large Boraginaceae. Of the seven proposed smaller families, four are native to China: Boraginaceae s.s., Heliotropiaceae, Cordiaceae, and Ehretiaceae.

The most remarkable difference between our results and the FOC classification concerns the Lamiales. In particular, the large Scrophulariaceae of Lamiales in FOC possessed genera that were divided among Phrymaceae, Mazaceae, and Orobanchaceae in our study. Our results also support separating Paulowniaceae and Linderniaceae, representing four genera, from Scrophulariaceae, as in APG IV. Callitrichaceae and Hippuridaceae are included in Plantaginaceae, while Myoporaceae and part of Loganiaceae have been combined with Scrophulariaceae s.s. Our study also showed differences with FOC in the relationships of Verbenaceae of Lamiales. Specifically, our results showed that some genera of Verbenaceae were nested within Lamiaceae, and that *Avicennia* was nested within Acanthaceae.

We suggest retaining monotypic or oligotypic families endemic to or mainly distributed in East Asia, if they possess key innovations or highly distinctive diagnostic characters. Such families include Acanthochlamydaceae, Bretschneideraceae, Illiciaceae, Leeaceae, Parnassiaceae, Rhoipteleaceae, and Toricelliaceae. Recent studies supported merging Bretschneideraceae (*Bretschneidera*) and Acanthochlamydaceae (*Acanthochlamys*) with their sisters, Akaniaceae of Australia (Rodman & Karol, 1996; Bayer & Appel, 2003; APG IV, 2016) and Velloziaceae of South America and Africa (north to the Arabian Peninsula), respectively (Behnke et al., 2000; Mello-Silva, 2005; Mello-Silva et al., 2011). However, we propose maintaining the family status of Bretschneideraceae and Acanthochlamydaceae, because they are clearly morphologically and geographically distinguished from their sisters. Morphologically, Bretschneideraceae have perigynous, zygomorphic flowers that are different from hypogynous, actinomorphic flowers in Akaniaceae (Cronquist, 1981; Bayer & Appel, 2003). Acanthochlamydaceae species are dwarf, caespitose perennial herbs with two pairs of veined leaves, differing from the shrubby or arborescent Velloziaceae, which have multiple pairs of veined leaves (Kubitzki, 1998). We also propose continuing to recognize those seemingly paraphyletic families, such as Dipterocarpaceae and Santalaceae (see discussion of both families in APG IV, 2016), until further molecular and morphological studies can be conducted to resolve their relationships.

Intrafamilial relationships and monophyly of genera

Our dense taxonomic sampling within large families and genera, especially those distributed in China, enabled us to explore intrafamilial relationships and the monophyly of genera. For example, we recovered relationships within the species-rich fern families that were mostly consistent with recently published phylogenetic studies of global scope, such as for Dryopteridaceae (Liu et al., 2007), Polypodiaceae (e.g. Schneider et al. 2004), Pteridaceae (Schuettpelz et al., 2007), and Thelypteridaceae (He & Zhang, 2012). In gymnosperms, our results showed that Cupressaceae contained five major subclades, each with 100% BS support, and that Pinaceae,

Table 1 Comparison of circumscription and the number of families in this study (TOL-gCVP) with that in FOC and APG IV

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
3	Cabombaceae	Cabombaceae	Cabombaceae	
4	Nymphaeaceae	Nymphaeaceae	Nymphaeaceae	
7	Schisandraceae	Schisandraceae & Illiciaceae	Schisandraceae & Illiciaceae	
10	Saururaceae	Saururaceae	Saururaceae	
11	Piperaceae	Piperaceae	Piperaceae	
12	Aristolochiaceae	Aristolochiaceae	Aristolochiaceae	
13	Myristicaceae	Myristicaceae	Myristicaceae	
14	Magnoliaceae	Magnoliaceae	Magnoliaceae	
18	Annonaceae	Annonaceae	Annonaceae	
19	Calycanthaceae	Calycanthaceae	Calycanthaceae	
23	Hernandiaceae	Hernandiaceae	Hernandiaceae	
25	Lauraceae	Lauraceae	Lauraceae	
26	Chloranthaceae	Chloranthaceae	Chloranthaceae	
27	Acoraceae	Acoraceae	Acoraceae	
28	Araceae	Araceae & Lemnaceae	Araceae	Including all genera of Lemnaceae *sensu* FOC
29	Tofieldiaceae	Liliaceae	Tofieldiaceae	Including *Tofieldia* (Liliaceae *sensu* FOC)
30	Alismataceae	Alismataceae	Alismataceae	
31	Butomaceae	Butomaceae	Butomaceae	
32	Hydrocharitaceae	Hydrocharitaceae	Hydrocharitaceae	
33	Scheuchzeriaceae	Scheuchzeriaceae	Scheuchzeriaceae	
34	Aponogetonaceae	Aponogetonaceae	Aponogetonaceae	
35	Juncaginaceae	Juncaginaceae	Juncaginaceae	
37	Zosteraceae	Zosteraceae	Zosteraceae	
38	Potamogetonaceae	Potamogetonaceae & Zannichelliaceae	Potamogetonaceae	Including *Zannichellia* (Zannichelliaceae *sensu* FOC)
39	Posidoniaceae	Posidoniaceae	Posidoniaceae	
40	Ruppiaceae	Ruppiaceae	Ruppiaceae	
41	Cymodoceaceae	Cymodoceaceae	Cymodoceaceae	
42	Petrosaviaceae	Liliaceae	Petrosaviaceae	Including *Petrosavia* (Liliaceae *sensu* FOC)
43	Nartheciaceae	Liliaceae	Nartheciaceae	Including *Aletris* (Liliaceae *sensu* FOC)
44	Burmanniaceae	Burmanniaceae	Burmanniaceae	
45	Dioscoreaceae	Dioscoreaceae & Taccaceae	Dioscoreaceae	Including *Schizocapsa* and *Tacca* *(Taccaceae sensu FOC)*
46	Triuridaceae	Triuridaceae	Triuridaceae	Unsampled
47	Velloziaceae	Amaryllidaceae	Acanthochlamydaceae	Including *Acanthochlamys* (Amaryllidaceae *sensu* FOC)
48	Stemonaceae	Stemonaceae	Stemonaceae	
50	Pandanaceae	Pandanaceae	Pandanaceae	
52	Corsiaceae	Corsiaceae	Corsiaceae	Unsampled
53	Melanthiaceae	Liliaceae	Melanthiaceae	Including *Chionographis*, *Heloniopsis*, *Paris*, *Trillium*, *Veratrum*, *Ypsilandra*, and *Zigadenus* (Liliaceae *sensu* FOC)
56	Colchicaceae	Liliaceae	Colchicaceae	Including *Disporum*, *Gloriosa*, and *Iphigenia* (Liliaceae sensu FOC)
59	Smilacaceae	Liliaceae	Smilacaceae	Including *Heterosmilax* and *Smilax* (Liliaceae *sensu* FOC)
60	Liliaceae	Liliaceae	Liliaceae	
61	Orchidaceae	Orchidaceae	Orchidaceae	
66	Hypoxidaceae	Amaryllidaceae	Hypoxidaceae	Including *Curculigo* and *Hypoxis* (Amaryllidaceae *sensu* FOC)
68	Ixioliriaceae	Amaryllidaceae	Ixioliriaceae	Including *Ixiolirion* (Amaryllidaceae sensu *FOC*)

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
70	Iridaceae	Iridaceae	Iridaceae	
72	Asphodelaceae	Liliaceae	Asphodelaceae	Including *Aloe*, *Dianella*, *Eremurus*, and *Hemerocallis* (Liliaceae *sensu* FOC)
73	Amaryllidaceae	Amaryllidaceae & Liliaceae	Amaryllidaceae	Including *Allium* and *Milula* (Liliaceae *sensu* FOC)
74	Asparagaceae	Asparagaceae, Amaryllidaceae & Liliaceae	Asparagaceae	Including *Agave* (Amaryllidaceae *sensu* FOC), *Anemarrhena*, *Asparagus*, *Aspidistra*, *Barnardia*, *Campylandra*, *Chlorophytum*, *Convallaria*, *Cordyline*, *Disporopsis*, *Diuranthera*, *Dracaena*, *Heteropolygonatum*, *Hosta*, *Liriope*, *Maianthemum*, *Ophiopogon*, *Peliosanthes*, *Polygonatum*, *Reineckea*, *Rohdea*, *Speirantha*, *Theropogon*, *Thysanotus*, and *Tupistra* (Liliaceae *sensu* FOC)
76	Arecaceae/Palmae	Arecaceae	Arecaceae	
78	Commelinaceae	Commelinaceae	Commelinaceae	
79	Philydraceae	Philydraceae	Philydraceae	
80	Pontederiaceae	Pontederiaceae	Pontederiaceae	
83	Lowiaceae	Lowiaceae	Lowiaceae	
85	Musaceae	Musaceae	Musaceae	
86	Cannaceae	Cannaceae	Cannaceae	
87	Marantaceae	Marantaceae	Marantaceae	
88	Costaceae	Costaceae	Costaceae	
89	Zingiberaceae	Zingiberaceae	Zingiberaceae	
90	Typhaceae	Typhaceae	Typhaceae	
91	Bromeliaceae	Bromeliaceae	Bromeliaceae	
93	Xyridaceae	Xyridaceae	Xyridaceae	
94	Eriocaulaceae	Eriocaulaceae	Eriocaulaceae	
97	Juncaceae	Juncaceae	Juncaceae	
98	Cyperaceae	Cyperaceae	Cyperaceae	
99	Restionaceae	Centrolepidaceae & Restionaceae	Restionaceae	
100	Flagellariaceae	Flagellariaceae	Flagellariaceae	
103	Poaceae/Gramineae	Poaceae	Poaceae	
104	Ceratophyllaceae	Ceratophyllaceae	Ceratophyllaceae	
105	Eupteleaceae	Eupteleaceae	Eupteleaceae	
106	Papaveraceae	Papaveraceae	Papaveraceae	
107	Circaeasteraceae	Circaeasteraceae & Ranunculaceae	Circaeasteraceae	Including *Kingdonia* (Ranunculaceae *sensu* FOC)
108	Lardizabalaceae	Lardizabalaceae	Lardizabalaceae	
109	Menispermaceae	Menispermaceae	Menispermaceae	
110	Berberidaceae	Berberidaceae	Berberidaceae	
111	Ranunculaceae	Ranunculaceae	Ranunculaceae	
112	Sabiaceae	Sabiaceae	Sabiaceae	
113	Nelumbonaceae	Nelumbonaceae	Nelumbonaceae	
114	Platanaceae	Platanaceae	Platanaceae	
115	Proteaceae	Proteaceae	Proteaceae	
116	Trochodendraceae	Trochodendraceae & Tetracentraceae	Trochodendraceae	Including *Tetracentron* (Tetracentraceae *sensu* FOC)
117	Buxaceae	Buxaceae	Buxaceae	
120	Dilleniaceae	Dilleniaceae	Dilleniaceae	
122	Paeoniaceae	Paeoniaceae	Paeoniaceae	

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
123	Altingiaceae	Hamamelidaceae	Altingiaceae	Including *Altingia*, *Liquidambar*, and *Semiliquidambar* (Hamamelidaceae *sensu* FOC)
124	Hamamelidaceae	Hamamelidaceae	Hamamelidaceae	
125	Cercidiphyllaceae	Cercidiphyllaceae	Cercidiphyllaceae	
126	Daphniphyllaceae	Daphniphyllaceae	Daphniphyllaceae	
127	Iteaceae	Saxifragaceae	Iteaceae	Including *Itea* (Saxifragaceae *sensu* FOC)
128	Grossulariaceae	Saxifragaceae	Grossulariaceae	Including *Ribes* (Saxifragaceae *sensu* FOC)
129	Saxifragaceae	Saxifragaceae	Saxifragaceae	
130	Crassulaceae	Crassulaceae	Crassulaceae	
133	Penthoraceae	Saxifragaceae	Penthoraceae	Including *Penthorum* (Saxifragaceae *sensu* FOC)
134	Haloragaceae	Haloragaceae	Haloragaceae	
135	Cynomoriaceae	Cynomoriaceae	Cynomoriaceae	
136	Vitaceae	Vitaceae & Leeaceae	Vitaceae & Leeaceae	
138	Zygophyllaceae	Zygophyllaceae	Zygophyllaceae	
140	Fabaceae/Leguminosae	Fabaceae	Fabaceae	
141	Surianaceae	Surianaceae	Surianaceae	
142	Polygalaceae	Polygalaceae	Polygalaceae	
143	Rosaceae	Rosaceae	Rosaceae	
146	Elaeagnaceae	Elaeagnaceae	Elaeagnaceae	
147	Rhamnaceae	Rhamnaceae	Rhamnaceae	
148	Ulmaceae	Ulmaceae	Ulmaceae	
149	Cannabaceae	Cannabaceae & Ulmaceae	Cannabaceae	Including *Aphananthe*, *Celtis*, *Gironniera*, *Pteroceltis*, and *Trema* (Ulmaceae *sensu* FOC)
150	Moraceae	Moraceae	Moraceae	
151	Urticaceae	Urticaceae	Urticaceae	
153	Fagaceae	Fagaceae	Fagaceae	
154	Myricaceae	Myricaceae	Myricaceae	
155	Juglandaceae	Juglandaceae & Rhoipteleaceae	Juglandaceae & Rhoipteleaceae	
156	Casuarinaceae	Casuarinaceae	Casuarinaceae	
158	Betulaceae	Betulaceae	Betulaceae	
162	Coriariaceae	Coriariaceae	Coriariaceae	
163	Cucurbitaceae	Cucurbitaceae	Cucurbitaceae	
164	Tetramelaceae	Tetramelaceae	Tetramelaceae	
166	Begoniaceae	Begoniaceae	Begoniaceae	
168	Celastraceae	Celastraceae, Saxifragaceae & Plagiopteraceae	Celastraceae & Parnassiaceae	Including *Parnassia* (Saxifragaceae *sensu* FOC) and *Plagiopteron* (Plagiopteraceae *sensu* FOC)
170	Connaraceae	Connaraceae	Connaraceae	
171	Oxalidaceae	Oxalidaceae	Oxalidaceae	
173	Elaeocarpaceae	Elaeocarpaceae	Elaeocarpaceae	
176	Pandaceae	Pandaceae	Pandaceae	
179	Rhizophoraceae	Rhizophoraceae	Rhizophoraceae	
180	Erythroxylaceae	Erythroxylaceae	Erythroxylaceae	
181	Ochnaceae	Ochnaceae	Ochnaceae	
183	Clusiaceae/Guttiferae	Clusiaceae	Clusiaceae	Only *Garcinia*
184	Calophyllaceae	Clusiaceae	Calophyllaceae	Including *Calophyllum*, *Mammea*, and *Mesua* (Clusiaceae *sensu* FOC)
185	Podostemaceae	Podostemaceae	Podostemaceae	
186	Hypericaceae	Clusiaceae	Hypericaceae	Including *Cratoxylum*, *Lianthus*, *Hypericum*, and *Triadenum* (Clusiaceae *sensu* FOC)

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
189	Putranjivaceae	Euphorbiaceae	Putranjivaceae	Including *Drypetes* and *Putranjiva* (Euphorbiaceae *sensu* FOC)
190	Centroplacaceae	Celastraceae	Centroplacaceae	Including *Bhesa* (Celastraceae *sensu* FOC)
191	Elatinaceae	Elatinaceae	Elatinaceae	
192	Malpighiaceae	Malpighiaceae	Malpighiaceae	
195	Dichapetalaceae	Dichapetalaceae	Dichapetalaceae	
199	Achariaceae	Flacourtiaceae	Achariaceae	Including *Gynocardia* and *Hydnocarpus* (Flacourtiaceae *sensu* FOC)
200	Violaceae	Violaceae	Violaceae	
202	Passifloraceae	Passifloraceae	Passifloraceae	
204	Salicaceae	Salicaceae & Flacourtiaceae	Salicaceae	Including *Bennettiodendron*, *Carrierea*, *Casearia*, *Flacourtia*, *Homalium*, *Idesia*, *Itoa*, *Poliothyrsis*, *Scolopia*, and *Xylosma* (Flacourtiaceae *sensu* FOC)
205	Peraceae	Euphorbiaceae	Peraceae	Including *Chaetocarpus* (Euphorbiaceae *sensu* FOC)
206	Rafflesiaceae	Rafflesiaceae	Rafflesiaceae	
207	Euphorbiaceae	Euphorbiaceae	Euphorbiaceae	
208	Linaceae	Linaceae	Linaceae	
209	Ixonanthaceae	Erythroxylaceae	Ixonanthaceae	Including *Ixonanthes* (Erythroxylaceae *sensu* FOC)
211	Phyllanthaceae	Euphorbiaceae	Phyllanthaceae	Including *Actephila*, *Antidesma*, *Aporosa*, *Baccaurea*, *Bischofia*, *Breynia*, *Bridelia*, *Cleistanthus*, *Flueggea*, *Glochidion*, *Leptopus*, *Margaritaria*, *Phyllanthodendron*, *Phyllanthus*, *Richeriella*, and *Sauropus* (Euphorbiaceae *sensu* FOC)
212	Geraniaceae	Geraniaceae	Geraniaceae	
214	Combretaceae	Combretaceae	Combretaceae	
215	Lythraceae	Lythraceae & Trapaceae	Lythraceae	Including *Trapa* (Trapaceae *sensu* FOC)
216	Onagraceae	Onagraceae	Onagraceae	
218	Myrtaceae	Myrtaceae	Myrtaceae	
219	Melastomataceae	Melastomataceae	Melastomataceae	
220	Crypteroniaceae	Crypteroniaceae	Crypteroniaceae	
226	Staphyleaceae	Staphyleaceae	Staphyleaceae	
228	Stachyuraceae	Stachyuraceae	Stachyuraceae	
233	Tapisciaceae	Tapisciaceae	Tapisciaceae	
234	Dipentodontaceae	Dipentodontaceae	Dipentodontaceae	
235	Biebersteiniaceae	Biebersteiniaceae	Biebersteiniaceae	
236	Nitrariaceae	Nitrariaceae & Peganaceae	Nitrariaceae	Including *Peganum* (Peganaceae *sensu* FOC)
238	Burseraceae	Burseraceae	Burseraceae	
239	Anacardiaceae	Anacardiaceae	Anacardiaceae	
240	Sapindaceae	Sapindaceae, Aceraceae & Hippocastanaceae	Sapindaceae	Including *Acer* and *Dipteronia* (Aceraceae *sensu* FOC) and *Aesculus* and *Handeliodendron* (Hippocastanaceae *sensu* FOC)
241	Rutaceae	Rutaceae & Cneoraceae	Rutaceae	Including *Harrisonia* (Cneoraceae *sensu* FOC)
242	Simaroubaceae	Simaroubaceae	Simaroubaceae	
243	Meliaceae	Meliaceae	Meliaceae	
247	Malvaceae	Malvaceae, Bombacaceae, Sterculiaceae & Tiliaceae	Malvaceae	Including all genera of Bombacaceae, Sterculiaceae, and Tiliaceae *sensu* FOC
249	Thymelaeaceae	Thymelaeaceae	Thymelaeaceae	

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
250	Bixaceae	Bixaceae	Bixaceae	
251	Cistaceae	Cistaceae	Cistaceae	
253	Dipterocarpaceae	Dipterocarpaceae	Dipterocarpaceae	
254	Akaniaceae	Bretschneideraceae	Bretschneideraceae	
255	Tropaeolaceae	Tropaeolaceae	Tropaeolaceae	
256	Moringaceae	Moringaceae	Moringaceae	
257	Caricaceae	Caricaceae	Caricaceae	
262	Salvadoraceae	Salvadoraceae	Salvadoraceae	
267	Resedaceae	Resedaceae	Resedaceae, Borthwickiaceae & Stixaceae	
268	Capparaceae	Capparaceae & Cleomaceae	Capparaceae	
269	Cleomaceae	Cleomaceae	Cleomaceae	
270	Brassicaceae/Cruciferae	Brassicaceae	Brassicaceae	
273	Olacaceae	Olacaceae	Olacaceae, Erythropalaceae & Ximeniaceae	
274	Opiliaceae	Opiliaceae	Opiliaceae	
275	Balanophoraceae	Balanophoraceae	Balanophoraceae	Unsampled
276	Santalaceae	Santalaceae & Viscaceae	Santalaceae	Including *Arceuthobium, Korthalsella*, and *Viscum* (Viscaceae *sensu* FOC)
278	Schoepfiaceae	Olacaceae	Schoepfiaceae	Including *Schoepfia* (Olacaceae *sensu* FOC)
279	Loranthaceae	Loranthaceae	Loranthaceae	
280	Frankeniaceae	Frankeniaceae	Frankeniaceae	
281	Tamaricaceae	Tamaricaceae	Tamaricaceae	
282	Plumbaginaceae	Plumbaginaceae	Plumbaginaceae	
283	Polygonaceae	Polygonaceae	Polygonaceae	
284	Droseraceae	Droseraceae	Droseraceae	
285	Nepenthaceae	Nepenthaceae	Nepenthaceae	
288	Ancistrocladaceae	Ancistrocladaceae	Ancistrocladaceae	
295	Caryophyllaceae	Caryophyllaceae	Caryophyllaceae	
297	Amaranthaceae	Amaranthaceae & Chenopodiaceae	Amaranthaceae	Including all genera of Chenopodiaceae *sensu* FOC
303	Gisekiaceae	Molluginaceae	Gisekiaceae	Including *Gisekia* (Molluginaceae *sensu* FOC)
304	Aizoaceae	Aizoaceae	Aizoaceae	
305	Phytolaccaceae	Phytolaccaceae	Phytolaccaceae	
306	Petiveriaceae	Phytolaccaceae	Petiveriaceae	Including *Rivina* (Phytolaccaceae *sensu* FOC)
308	Nyctaginaceae	Nyctaginaceae	Nyctaginaceae	
309	Molluginaceae	Molluginaceae	Molluginaceae	
312	Basellaceae	Basellaceae	Basellaceae	
314	Talinaceae	Portulacaceae	Talinaceae	Including *Talinum* (Talinaceae *sensu* FOC)
315	Portulacaceae	Portulacaceae	Portulacaceae	
317	Cactaceae	Cactaceae	Cactaceae	
318	Nyssaceae	Mastixiaceae & Nyssaceae	Nyssaceae	Including all genera of Mastixiaceae *sensu* FOC
320	Hydrangeaceae	Saxifragaceae	Hydrangeaceae	Including *Cardiandra, Decumaria, Deinanthe, Deutzia, Dichroa, Hydrangea, Kirengeshoma, Philadelphus, Pileostegia, Platycrater*, and *Schizophragma* (Saxifragaceae *sensu* FOC)
324	Cornaceae	Cornaceae, & Alangiaceae	Cornaceae	Including *Alangium* (Alangiaceae *sensu* FOC)

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
325	Balsaminaceae	Balsaminaceae	Balsaminaceae	
329	Polemoniaceae	Polemoniaceae	Polemoniaceae	
330	Lecythidaceae	Lecythidaceae	Lecythidaceae	
331	Sladeniaceae	Sladeniaceae	Sladeniaceae	
332	Pentaphylacaceae	Pentaphylacaceae & Theaceae	Pentaphylacaceae	Including *Adinandra*, *Anneslea*, *Cleyera*, *Eurya*, *Euryodendron*, and *Ternstroemia* (Theaceae *sensu* FOC)
333	Sapotaceae	Sapotaceae	Sapotaceae	
334	Ebenaceae	Ebenaceae	Ebenaceae	
335	Primulaceae	Primulaceae & Myrsinaceae	Primulaceae	Including all genera of Myrsinaceae *sensu* FOC
336	Theaceae	Theaceae	Theaceae	
337	Symplocaceae	Symplocaceae	Symplocaceae	
338	Diapensiaceae	Diapensiaceae	Diapensiaceae	
339	Styracaceae	Styracaceae	Styracaceae	
342	Actinidiaceae	Actinidiaceae	Actinidiaceae	
343	Clethraceae	Clethraceae	Clethraceae	
345	Ericaceae	Ericaceae	Ericaceae	
346	Mitrastemonaceae	Rafflesiaceae	Mitrastemonaceae	Including *Mitrastemon* (Rafflesiaceae *sensu* FOC); unsampled
348	Icacinaceae	Icacinaceae	Icacinaceae	
349	Metteniusaceae	Icacinaceae	Metteniusaceae	Including *Apodytes*, *Pittosporopsis*, and *Platea* (Icacinaceae *sensu* FOC)
350	Eucommiaceae	Eucommiaceae	Eucommiaceae	
351	Garryaceae	Aucubaceae	Garryaceae	Including *Aucuba* (Aucubaceae *sensu* FOC)
352	Rubiaceae	Rubiaceae	Rubiaceae	
353	Gentianaceae	Gentianaceae & Loganiaceae	Gentianaceae	Including *Fagraea* (Loganiaceae *sensu* FOC)
354	Loganiaceae	Loganiaceae	Loganiaceae	
355	Gelsemiaceae	Loganiaceae	Gelsemiaceae	Including *Gelsemium* (Loganiaceae *sensu* FOC)
356	Apocynaceae	Apocynaceae & Asclepiadaceae	Apocynaceae	Including all genera of Asclepiadaceae *sensu* FOC
357	Boraginaceae	Boraginaceae	Boraginaceae, Cordiaceae, Ehretiaceae & Heliotropiaceae	
359	Convolvulaceae	Convolvulaceae	Convolvulaceae	
360	Solanaceae	Solanaceae	Solanaceae	
362	Sphenocleaceae	Sphenocleaceae	Sphenocleaceae	
363	Hydroleaceae	Hydrophyllaceae	Hydroleaceae	
365	Carlemanniaceae	Carlemanniaceae	Carlemanniaceae	
366	Oleaceae	Oleaceae	Oleaceae	
369	Gesneriaceae	Gesneriaceae & Scrophulariaceae	Gesneriaceae	Including *Cyrtandromoea* (Scrophulariaceae *sensu* FOC); unsampled
370	Plantaginaceae	Plantaginaceae, Callitrichaceae, Hippuridaceae, Pedaliaceae & Scrophulariaceae	Plantaginaceae	Including *Callitriche* (Callitrichaceae *sensu* FOC), *Hippuris* (Hippuridaceae *sensu* FOC), *Trapella* (Pedaliaceae *sensu* FOC), *Adenosma*, *Bacopa*, *Deinostema*, *Digitalis*, *Dopatrium*, *Ellisiophyllum*, *Gratiola*, *Hemiphragma*, *Lagotis*, *Limnophila*, *Linaria*, *Neopicrorhiza*, *Pseudolysimachion*, *Scoparia*, *Scrofella*, *Stemodia*, *Veronica*, and *Veronicastrum* (Scrophulariaceae *sensu* FOC)

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
371	Scrophulariaceae	Scrophulariaceae, Loganiaceae & Myoporaceae	Scrophulariaceae	Including *Buddleja* (Loganiaceae *sensu* FOC) and *Pentacoelium* (Myoporaceae *sensu* FOC)
373	Linderniaceae	Scrophulariaceae	Linderniaceae	Including *Legazpia*, *Lindernia*, *Picria*, and *Torenia* (Scrophulariaceae *sensu* FOC)
375	Martyniaceae	Martyniaceae	Martyniaceae	
376	Pedaliaceae	Pedaliaceae	Pedaliaceae	
377	Acanthaceae	Acanthaceae & Verbenaceae	Acanthaceae	Including *Avicennia* (Verbenaceae *sensu* FOC)
378	Bignoniaceae	Bignoniaceae	Bignoniaceae	
379	Lentibulariaceae	Lentibulariaceae	Lentibulariaceae	
382	Verbenaceae	Verbenaceae	Verbenaceae	
383	Lamiaceae/Labiatae	Lamiaceae & Verbenaceae	Lamiaceae	Including *Callicarpa*, *Caryopteris*, *Clerodendrum*, *Congea*, *Garrettia*, *Gmelina*, *Hymenopyramis*, *Premna*, *Schnabelia*, *Sphenodesme*, *Tectona*, *Tsoongia*, and *Vitex* (Verbenaceae *sensu* FOC)
384	Mazaceae	Scrophulariaceae	Mazaceae	Including *Dodartia*, *Lancea*, and *Mazus* (Scrophulariaceae *sensu* FOC)
385	Phrymaceae	Phrymaceae & Scrophulariaceae	Phrymaceae	Including *Microcarpaea* and *Mimulus* (Scrophulariaceae *sensu* FOC)
386	Paulowniaceae	Scrophulariaceae	Paulowniaceae	Including *Paulownia* and *Wightia* (Scrophulariaceae *sensu* FOC)
387	Orobanchaceae	Orobanchaceae & Scrophulariaceae	Orobanchaceae	Including *Alectra*, *Brandisia*, *Buchnera*, *Castilleja*, *Centranthera*, *Cymbaria*, *Euphrasia*, *Leptorhabdos*, *Lindenbergia*, *Melampyrum*, *Monochasma*, *Odontites*, *Omphalotrix*, *Pedicularis*, *Petitmenginia*, *Phtheirospermum*, *Pseudobartsia*, *Pterygiella*, *Rehmannia*, *Rhinanthus*, *Siphonostegia*, *Sopubia*, *Striga*, *Triaenophora*, *Triphysaria*, and *Xizangia* (Scrophulariaceae *sensu* FOC)
388	Stemonuraceae	Icacinaceae	Stemonuraceae	Including *Gomphandra* (Icacinaceae *sensu* FOC)
389	Cardiopteridaceae	Cardiopteridaceae & Icacinaceae	Cardiopteridaceae	Including *Gonocaryum* (Icacinaceae *sensu* FOC)
391	Helwingiaceae	Helwingiaceae	Helwingiaceae	
392	Aquifoliaceae	Aquifoliaceae	Aquifoliaceae	
394	Campanulaceae	Campanulaceae	Campanulaceae	
395	Pentaphragmataceae	Pentaphragmataceae	Pentaphragmataceae	
396	Stylidiaceae	Stylidiaceae	Stylidiaceae	
400	Menyanthaceae	Menyanthaceae	Menyanthaceae	
401	Goodeniaceae	Goodeniaceae	Goodeniaceae	
403	Asteraceae/Compositae	Asteraceae	Asteraceae	
404	Escalloniaceae	Saxifragaceae	Escalloniaceae	Including *Polyosma* (Saxifragaceae *sensu* FOC)
408	Adoxaceae	Adoxaceae	Adoxaceae	
409	Caprifoliaceae	Caprifoliaceae, Diervillaceae, Dipsacaceae, Morinaceae, Linnaeaceae & Valerianaceae	Caprifoliaceae	Including all genera of Diervillaceae, Dipsacaceae, Morinaceae, Linnaeaceae, and Valerianaceae *sensu* FOC
411	Torricelliaceae	Toricelliaceae	Torricelliaceae	

Continued

Table 1 Continued

Fam. No.	APG IV	FOC	TOL-gCVP	Difference between TOL-gCVP and FOC
413	Pittosporaceae	Pittosporaceae	Pittosporaceae	
414	Araliaceae	Araliaceae & Apiaceae	Araliaceae	Including *Hydrocotyle* (Apiaceae *sensu* FOC)
416	Apiaceae/Umbelliferae	Apiaceae	Apiaceae	

APG, Angiosperm Phylogeny Group; FOC, *Flora of China*.

which have 10 genera in China, were divided into two main subclades, although one of them received weak support. In angiosperms, intrafamilial relationships are poorly resolved in several families, such as in Gesneriaceae, Apiaceae, Araliaceae, Asteraceae, Brassicaceae, and Orchidaceaee. The subclades and the relationship among them in most families of monocots, rosids, Santalales, Saxifragales, and Caryophyllales are strongly supported.

Our study showed that the following genera are monophyletic, based on the sampling used here: *Actinidia* (Actinidiaceae); *Viburnum* (Adoxaceae); *Allium* (Amaryllidaceae); *Bupleurum* (Apiaceae); *Ilex* (Aquifoliaceae); *Hydrocotyle* (Araliaceae, placed in Apiaceae in FOC); *Aristolochia* (Aristolochiaceae); *Polygonatum* and *Ophiopogon* (Asparagaceae); *Ainsliaea*, *Parasenecio, Taraxacum*, and *Leontopodium* (Asteraceae); *Impatiens* (Balsaminaceae); *Begonia* (Begoniaceae); *Berberis* and *Epimedium* (Berberidaceae); *Betula* (Betulaceae); *Draba* (Brassicaceae); *Adenophora* (Campanulaceae); *Capparis* (Capparaceae); *Lonicera* (Caprifoliaceae); *Polycarpaea* (Caryophyllaceae); *Cornus* (Cornaceae); *Rhodiola* (Crassulaceae); *Hemsleya* (Cucurbitaceae); *Fimbristylis* (Cyperaceae); *Dioscorea* (Dioscoreaceae); *Diospyros* (Ebenaceae); *Elaeocarpus* (Elaeocarpaceae); *Gaultheria* (Ericaceae); *Eriocaulon* (Eriocaulaceae); *Mallotus* (Euphorbiaceae); *Astragalus*, *Oxytropis*, *Indigofera*, *Lespedeza*, *Crotalaria*, and *Bauhinia* (Fabaceae); *Lithocarpus* and *Castanopsis* (Fagaceae); *Gentiana* (Gentianaceae); *Geranium* (Geraniaceae); *Ribes* (Grossulariaceae); *Distylium* (Hamamelidaceae); *Deutzia* and *Philadelphus* (Hydrangeaceae); *Juncus* (Juncaceae); *Scutellaria*, *Salvia*, *Teucrium*, and *Callicarpa* (Lamiaceae); *Machilus* (Lauraceae); *Stephania* (Menispermaceae); *Ficus* (Moraceae); *Syzygium* (Myrtaceae); *Cymbidium* and *Dendrobium* (Orchidaceae); *Pedicularis* (Orobanchaceae); *Corydalis* (Papaveraceae); *Parnassia* (Parnassiaceae, placed in Saxifragaceae in FOC); *Eurya* (Pentaphylacaceae); *Pittosporum* (Pittosporaceae); *Puccinellia* and *Poa* (Poaceae); *Rheum* (Polygonaceae); *Maesa* and *Ardisia* (Primulaceae); *Thalictrum* (Ranunculaceae); *Rhamnus* (Rhamnaceae); *Cotoneaster*, *Spiraea*, and *Rosa* (Rosaceae); *Galium* (Rubiaceae); *Haplophyllum* (Rutaceae); *Populus* (Salicaceae); *Acer* (Sapindaceae); *Chrysosplenium* (Saxifragaceae); *Styrax* (Styracaceae); *Symplocos* (Symplocaceae); *Camellia* (Theaceae); *Pilea* (Urticaceae); *Viola* (Violaceae); and *Ampelopsis* (Vitaceae).

The following genera are not monophyletic: *Bassia* and *Salsola* (Amaranthaceae); *Pleurospermum* (Apiaceae); *Hoya* (Apocynaceae); *Schefflera* (Araliaceae); *Saussurea, Himalaiella, Carpesium*, *Cirsium*, *Erigeron, Vernonia, Ligularia, Senecio, Aster*, and *Artemisia* (Asteraceae); *Mahonia* (Berberidaceae); *Eritrichium* (Boraginaceae); *Codonopsis* (Campanulaceae); *Silene, Arenaria*, and *Stellaria* (Caryophyllaceae); *Euonymus* (Celastraceae); *Ipomoea* (Convolvulaceae); *Sedum* (Crassulaceae); *Carex* and *Cyperus* (Cyperaceae); *Rhododendron* and *Vaccinium* (Ericaceae); *Euphorbia* and *Vernicia* (Euphorbiaceae); *Caragana* (Fabaceae); *Swertia* (Gentianaceae); *Boea* and *Primulina* (Gesneriaceae); *Hydrangea* (Hydrangeaceae); *Hypericum* (Hypericaceae); *Stachys*, *Clerodendrum, Isodon*, and *Micromeria* (Lamiaceae); *Lindera* and *Litsea* (Lauraceae); *Lilium* (Liliaceae); *Dysoxylum* (Meliaceae); *Cocculus* (Menispermaceae); *Adinandra* (Pentaphylacaceae); *Veronica* (Plantaginaceae); *Polygonum* (Polygonaceae); *Yushania, Elymus*, and *Festuca* (Poaceae); *Lysimachia*, *Androsace*, and *Primula* (Primulaceae); *Aglaia, Delphinium, Aconitum, Ranunculus*, *Clematis*, and *Anemone* (Ranunculaceae); *Sorbus, Rubus*, and *Potentilla* (Rosaceae); *Hedyotis, Ophiorrhiza*, and *Wendlandia* (Rubiaceae); *Zanthoxylum* (Rutaceae); *Mitella* and *Saxifraga* (Saxifragaceae); *Daphne* (Thymelaeaceae); and *Cayratia* (Vitaceae).

Conclusions and Prospects

We have provided the first large-scale phylogeny for the genera of Chinese vascular plants. Our analyses of five genes for 6098 species representing 3114 genera of vascular plants and five genera of bryophytes as out-groups resulted in a well-resolved phylogeny. The topology of our regional tree of life for Chinese vascular plants is largely consistent with the circumscriptions of orders and families (APG IV, 2016), as well as with previously reconstructed relationships among these groups (Chase et al., 1993; Soltis et al., 1997, 2000; Savolainen et al., 2000; Hilu et al., 2003; Qiu et al., 2010; Ruhfel et al., 2014). Many efforts have been made to improve phylogenetic accuracy within angiosperms using relatively few representatives of orders or families and many genes (e.g., Soltis et al., 2011, 17 genes, 640 taxa of angiosperms). These analyses have often obtained strong support for the relationships of major clades at familial or ordinal levels and serve as excellent frameworks for testing relationships in future studies with denser taxonomic sampling.

Our genus-level tree of Chinese vascular plants showed some new relationships at familial levels and within families. For example, we found 76% BS support for *Helianthemum scopulicola* of Cistaceae nested within Dipterocarpaceae, which may not be monophyletic (see also discussion in APG IV, 2016). Our results showed that two species of *Eberhardtia* are sister to a clade formed by all other taxa of Sapotaceae, except *Sarcosperma*, which is very helpful for understanding the phylogeny in this family. We also found new subclades within several large families, including

Asteraceae, Orchidaceae, and Apiaceae, and we showed support for the monophyly of some large genera in China with our dense taxonomic sampling: *Pedicularis, Gentiana, Astragalus, Begonia, Pilea, Ficus, Rosa*, and *Berberis*.

In addition to resolving phylogenetic relationships, the regional tree has many other practical applications. In particular, because the tree has sufficiently dense taxonomic sampling, we expect that ecologists can use the tree to study the origin and evolution of Chinese flora, the geographic patterns and conservation of biodiversity, and the structure and dynamics of community (Heath et al., 2008; Lu et al., 2014). Moreover, we expect that the SoTree tool will increase the accessibility of our phylogeny for non-taxonomists engaged in ecological and other studies.

Phylogenetic accuracy and dense sampling at species levels are critical for ecological or other studies that use a phylogenetic framework. Therefore, future studies should extend and improve the tree of life for Chinese vascular plants in the following ways: (i) by extending taxon sampling within the Chinese flora; (ii) by vetting the quality of data, especially the identity of specimens and sequences; and (iii) by improving methods of building large phylogenies, including advances in sequence alignment, computational capacity, and tree building methods (Lu et al., 2014).

Acknowledgements

We wish to thank Tao DENG, Xiao-Yu DONG, Bao-Cai HAN, Shu-Guang JIAN, Peng-Wei LI, Zhen-Yu LI, Qi-Xin LIU, Yan LIU, Yu SONG, Qiang WANG, Yin-Zheng WANG, Lai WEI, Yu-Dong WU, Yong YANG, Sheng-Xiang YU, Chun-Hong ZHANG, Ming-Li ZHANG, Qiang ZHANG, Ren-Bin ZHU, Xiang-Yun ZHU for their support and for providing materials. We thank Herbarium, Institute of Botany, Chinese Academy of Sciences (PE) for approval to obtain pieces of leaves from specimens for DNA extraction and sequencing. This research was supported by the National Natural Science Foundation of China (Grant No. 31590822), the National Key Basic Research Program of China (Grant No. 2014CB954100), the Chinese Academy of Sciences International Institution Development Program (Grant No. SAJC201315), Shenzhen Science and Technology Innovation Council Funding (Grant No. KQC 201105310009A), the National Natural Science Foundation of China (Grant Nos. 31270268 and 31270269), Science and Technology Basic Work (Grant No. 2013FY112100), the Chinese Academy of Sciences Externa1 Cooperation Program of BIC (Grant No. GJHZ201321), the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists (Grant No. 2011T1S24), the US National Science Foundation for Open tree (Grant No. DEB-1208428) and iDigBio (Grant No. EF-1115210).

References

Angiosperm Phylogeny Group (APG). 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.

Angiosperm Phylogeny Group II (APG II). 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants. *Botanical Journal of the Linnean Society* 141: 399–436.

Angiosperm Phylogeny Group III (APG III). 2009. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161: 105–121.

Angiosperm Phylogeny Group IV (APG IV). 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.

Bayer C, Appel O. 2003. Akaniaceae. In: Kubitzki K, Bayer C eds. *The families and genera of vascular plants*. V. Berlin: Springer-Verlag. 21–24.

Behnke HD, Treutlein J, Wink M, Kramer K, Schneider C, Kao PC. 2000. Systematics and evolution of Velloziaceae, with special reference to sieve-element plastids and *rbc*L sequence data. *Botanical Journal of the Linnean Society* 134: 93–129.

Bowe LM, Coat G, dePamphilis CW. 2000. Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences USA* 97: 4092–4097.

Burleigh JG, Hilu KW, Soltis DE. 2009. Inferring phylogenies with incomplete data sets. A 5-gene, 567-taxon analysis of angiosperms. *BMC Evolutionary Biology* 9: 61.

Burleigh JG, Mathews S. 2004. Phylogenetic signal in nucleotide data from seed plants: Implications for resolving the seed plant tree of life. *American Journal of Botany* 91: 1599–1613.

Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis PS, Soltis DE, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56: 822–846.

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW. 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters* 12: 693–715.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu YL, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim KJ, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang QY, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Jr. Learn GH, Graham SW, Barrett SCH, Dayanandan S, Albert VA. 1993. Phylogenetics of seed plants: An analysis of nucleotide-sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–580.

Chaw SW, Parkinson CL, Cheng YC, Vincent TM, Parmer JD. 2000. Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences USA* 97: 4086–4091.

Chaw SM, Zharkikh A, Sung HM, Lau TC, Li WH. 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nuclear 18S rRNA sequences. *Molecular Biology and Evolution* 14: 56–68.

Chen ZD, Manchester SR, Sun HY. 1999. Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology, and paleobotany. *American Journal of Botany* 86: 1168–1181.

Chen ZD, Wang XQ, Sun HY, Han Y, Lu AM. 1998. The systematic position of the Rhoipteleaceae: Evidence from nucleotide sequences of *rbcL* gene. *Acta Phytotaxonomica Sinica* 36: 1–7.

Cheng WC, Fu LK eds. 1978. *Gymnospermae in Flora Reipublicae Popularis Sinicae Tomus 7*. Beijing: Science Press.

Ching RC. 1940. On natural classification of the family Polypodiaceae. *Sunyatsenia* 5: 201–268.

Ching RC. 1978a. The Chinese fern families and genera: Systematic arrangement and historical origin. *Acta Phytotaxonomica Sinica* 16: 1–19.

Ching RC. 1978b. The Chinese fern families and genera: Systematic arrangement and historical origin (cont.). *Acta Phytotaxonomica Sinica* 16: 16–37.

Ciccarelli FD, Doerks T, von-Mering C, Creevey CJ, Snel B, Bork P. 2006. Towards automatic reconstruction of a highly resolved Tree of Life. *Science* 311: 1283–1287.

Cronquist A. 1981. *An integrated system of classification of flowering plants*. New York: Columbia University Press.

Dong W, Cheng T, Li C, Xu C, Long P, Chen C, Zhou S. 2014. Discriminating plants using the DNA barcode *rbcL*: An appraisal based on a large data set. *Molecular Ecology Resources* 14: 336–343.

Doweld A, Reveal JL. 2008. New suprageneric names for vascular plants. *Phytologia* 90: 416–417.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A. 2011. Geneious v5.5.6 [online]. Available from www.geneious.com [Accessed 1 December 2015].

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Fiz-Palacios M, Schneider H, Heinrichs J, Savolainen V. 2011. Diversification of land plants: Insights from a family-level phylogentic analysis. *BMC Evolutionary Biology* 11: 341.

Frohlich MW, Parker DS. 2000. The mostly male theory of flower evolutionary origins: From genes to fossils. *Systematic Botany* 25: 155–170.

Gao L, Wang B, Wang ZW, Zhou Y, Su YJ, Wang T. 2013. Plastome sequences of *Lygodium japonicum* and *Marsilea crenata* reveal the genome organization transformation from basal ferns to core leptosporangiates. *Genome Biology and Evolution* 5: 1403–1407.

Ge S, Sang T, Lu BR, Hong DY. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences USA* 96: 14400–14405.

Ge S, Schaal BA, Hong DY. 1997. A reevaluation of the status of *A. Lobophylla* based on ITS sequence, with reference to the utility of ITS sequence in *Adenophora*. *Acta Phytotaxonomica Sinica* 35: 385–395.

Goremykin V, Bobrova V, Pahnke J, Troitsky A, Antonov A, Martin W. 1996. Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data do not support Gnetalean affinities of angiosperms. *Molecular Biology and Evolution* 13: 383–396.

Gugerli F, Sperisen C, Buchler U, Brunner L, Brodbeck S, Palmer JD, Qiu YL. 2001. The evolutionary split of Pinaceae from other conifers: Evidence from an intron loss and a multigene phylogeny. *Molecular Phylogenetics and Evolution* 21: 167–175.

He LJ, Zhang XC. 2012. Exploring generic delimitation within the fern family Thelypteridaceae. *Molecular Phylogenetics and Evolution* 65: 757–764.

Heath TA, Hedtke SM, Hillis DM. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution* 46: 239–257.

Hennequin G, Vogels TP, Gerstner W. 2014. Optimal control of transient dynamics in balanced networks supports generation of complex. *Neuron* 82: 1394–1406.

Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R, Sauquet H, Neinhuis C, Slotta TAB, Rohwer JG, Campbell CS, Chatrou LW. 2003. Angiosperm phylogeny based on *matK* sequence information. *American Journal of Botany* 90: 1758–1776.

Hong DY, Blackmore S eds. 2013. *Plants of China: A companion to the Flora of China*. Beijing: Science Press.

Hong DY, Chen ZD, Qiu YL, Donoghue MJ. 2008. Tracing patterns of evolution through the Tree of Life: Introduction. *Journal of Systematics and Evolution* 46: 237–238.

Hong DY, Wang Q. 2015. A new taxonomic system of the Campanulaceae s.s. *Journal of Systematics and Evolution* 53: 203–209.

Hu HH. 1950. A new classification of angiosperms based on multiple origins. *Science of China* 1: 243–253.

Hu QM, Watson MF. 2013. Plant exploration in China. In: Hong DY, Blackmore S eds, *Plants of China: A companion to the Flora of China*. Beijing: Science Press. 212–236.

Huang HW, Oldfield S, Qian H. 2013. Global significance of plant diversity in China. In: Hong DY, Blackmore S eds. *Plants of China: A companion to the Flora of China*. Beijing: Science Press. 7–34.

Ickert-Bond SM, Renner SS. 2016. The Gnetales: Recent insights on their morphology, reproductive biology, chromosome numbers, biogeography, and divergence times. *Journal of Systematics and Evolution* 54: 1–16.

Izquierdo-Carrasco F, Smith SA, Stamatakis A. 2011. Algorithms, data structures, and numerics for likelihood-based phylogenetic inference of huge trees. *BMC Bioinformatics* 12: 470.

Kubitzki K ed. 1998. *The families and genera of vascular plants III*. Berlin: Springer-Verlag.

Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB, Fu CX, Zeng CX, Yan HF, Zhu YJ, Sun YS, Chen SY, Zhao L, Wang K, Yang T, Duan YW. 2011a. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences USA* 108: 19641–19646.

Li DZ, Liu JQ, Chen ZD, Wang H, Ge XJ, Zhou SL, Gao LM, Fu CX, Chen SL. 2011b. Plant DNA barcoding in China. *Journal of Systematics and Evolution* 49: 165–168.

Li JL, Wang S, Yu J, Zhou SL. 2013. A modified CTAB protocol for plant DNA extraction. *Chinese Bulletin of Botany* 48: 72–78.

Liu HM. 2016. Embracing the pteridophyte classification of Ren-Chang Ching using a generic phylogeny of Chinese ferns and lycophytes. *Journal of Systematics and Evolution* 54: 307–335.

Liu HM, Zhang XC, Wang W, Qiu YL, Chen ZD. 2007. Molecular phylogeny of the fern family Dryopteridaceae inferred from chloroplast *rbcL* and *atpB* genes. *International Journal of Plant Sciences* 689: 1311–1323.

Liu JQ, Duan YW, Hao G, Ge XJ, Sun H. 2014. Evolutionary history and underlying adaptation of alpine plants on the Qinghai–Tibet Plateau. *Journal of Systematics and Evolution* 52: 241–249.

Lu JM, Zhang N, Du XY, Wen J, Li DZ. 2015. Chloroplast phylogenomics resolves key relationships in ferns. *Journal of Systematics and Evolution* 53: 448–457.

Lu LM, Sun M, Zhang JB, Li HL, Lin L, Yang T, Chen M, Chen ZD. 2014. Tree of life and its applications. *Biodiversity Science* 22: 3–20.

Magallón S, Sanderson MJ. 2002. Relationships among seed plants inferred from highly conserved genes: Sorting conflicting phylogenetic signals among ancient lineages. *American Journal of Botany* 89: 1991–2006.

McMahon MM, Sanderson MJ. 2006. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. *Systematic Biology* 55: 818–836.

Mello-Silva R. 2005. Morphological analysis, phylogenies and classification in Velloziaceae. *Botanical Journal of the Linnean Society* 148: 157–173.

Mello-Silva R, Santos DYAC, Salatino MLF, Motta LB, Cattai MB, Sasaki D, Lovo J, Pita PB, Rocini C, Rodrigues CDN, Zarrei M, Chase MW. 2011. Five vicarious genera from Gondwana: The Velloziaceae as shown by molecules and morphology. *Annals of Botany* 108: 87–102.

Meng HH, Gao XY, Huang JF, Zhang ML. 2015a. Plant phylogeography in arid Northwest China: Retrospectives and perspectives. *Journal of Systematics and Evolution* 53: 33–46.

Meng Z, Chen ZD, Liu HM, He X, Lin XG, Zhang SZ, Li Y, Hu LL, Zhou YC. 2010. Platform construction of land plant phylogeny analysis. *Computer Engineering* 20: 10–20.

Meng Z, Dong H, Li JH, Chen ZD, Zhou YC, Wang XZ, Zhang SZ. 2015b. Darwintree: A molecular data analysis and application environment for phylogenetic study. *Data Science Journal* 14: 10.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*. 14 Nov 2010, New Orleans, LA. 1–8.

Olmstead RG, Kim KJ, Jansen RK, Wagstaff SJ. 2000. The phylogeny of the Asteridae sensu lato based on chloroplast *ndhF* gene sequences. *Molecular Phylogenetics and Evolution* 16: 96–112.

Olmstead RG, Michaels HJ, Scott KM, Palmer JD. 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. *Annals of the Missouri Botanical Garden* 79: 249–265.

Olmstead RG, Sweere JA. 1994. Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.

Olsen G. 1990. "Newick's 8:45" Tree Format Standard [online]. Available at http://evolution.genetics.washington.edu/phylip/newick_doc.html [Accessed 30 May 2016]

Pearse WD, Purvis A, Paradis E. 2013. phyloGenerator: An automated phylogeny generation tool for ecologists. *Methods in Ecology and Evolution* 4: 692–698.

Peng H. 2013. Development of Chinese botany. In: Hong DY, Blackmore S eds, *Plants of China: A companion to the Flora of China*. Beijing: Science Press. 205–211.

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409: 618–622.

Qian H, Zhang J. 2016. Are phylogenies derived from family-level supertrees robust for studies on macroecological patterns along environmental gradients? *Journal of Systematics and Evolution* 54: 29–36.

Qiu YL, Lee JB, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen ZD, Savolainen V, Chase MW. 1999. The earliest angiosperms: Evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402: 404–407.

Qiu YL, Li LB, Wang B, Chen ZD, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, Estabrook GF, Hendry TA, Taylor DW, Testa CM, Ambros M, Crandall-Stotler B, Duff RJ, Stech M, Frey W, Quandt D, Davis CC. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proceedings of the National Academy of Sciences USA* 103: 15511–15516.

Qiu YL, Li LB, Wang B, Chen ZD, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J, Estabrook GF, Hendry TA, Taylor DW, Testa CM, Ambros M, Crandall-Stotler B, Duff RJ, Stech M, Frey W, Quandt D, Davis CC. 2007. The deepest divergences in land plants inferred from phylogenomic evidence. *International Journal of Plant Sciences* 168: 691–708.

Qiu YL, Li LB, Wang B, Xue JY, Hendry TA, Li RQ, Brown JW, Liu Y, Hudson YH, Chen ZD. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* 48: 391–425.

Qiu YX, Fu CX, Comes HP. 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* 59: 225–244.

Rambaut A. 2009. FigTree 1.3.1 [online]. Available at http://tree.bio.ed.ac.uk/software/figtree [Accessed 12 May 2014]

Ran JH, Wang PP, Zhao HJ, Wang XQ. 2010. A test of seven candidate barcode regions from the plastome in *Picea* (Pinaceae). *Journal of Integrative Plant Biology* 52: 1109–1126.

Refulio-Rodríguez NF, Olmstead RG. 2014. Phylogeny of Lamiidae. *American Journal of Botany* 101: 298–299.

Rodman JE, Karol KG. 1996. Molecules, morphology, and Dahlgren's expanded order Capparales. *Systematic Botany* 21: 289–307.

Ruhfel BR, Gitzendanner MA, Soltis PS, Soltis DE, Burleigh JG. 2014. From algae to angiosperms – inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evolutionary Biology* 14: 385–399.

Rydin C, Kallersjo M, Friis EM. 2002. Seed plant relationships and the systematic position of Gnetales based on nuclear and chloroplast DNA: Conflicting data, rooting problems, and the monophyly of conifers. *International Journal of Plant Sciences* 163: 197–214.

Sanderson MJ, Boss D, Chen D, Cranston KA, Wehe A. 2008. The PhyLoTA browser: Processing GenBank for molecular phylogenetics research. *Systematic Biology* 57: 335–346.

Savolainen V, Fay MF, Albach DC, Backlund A, van der Bank M, Cameron KM, Johnson SA, Lledó MD, Pintaud JC, Powell M, Sheahan MC, Soltis DE, Soltis PS, Weston P, Whitten WM, Wurdack KJ, Chase MW. 2000. Phylogeny of the eudicots: A nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bulletin* 55: 257–309.

Schneider H, Smith AR, Cranfil R, Hildebran TJ, Haufle CH, Ranke, TA. 2004. Unraveling the phylogeny of the polygrammoid ferns (Polypodiaceae and Grammitidaceae): Exploring aspects of the diversification of epiphytic plants. *Molecular Phylogenetics and Evolution* 31: 1041–1063.

Schuettpelz E, Schneider H, Huiet L, Windham MD, Pryer KM. 2007. A molecular phylogeny of the fern family Pteridaceae: Assessing overall relationships and the affinities of previously unsampled genera. *Molecular Phylogenetics and Evolution* 44: 1172–1185.

Smith SA, Beaulieu JM, Donoghue MJ. 2009. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology* 9: 1–12.

Soltis DE, Gitzendanner MA, Soltis PS. 2007. A 567-taxon data set for angiosperms: The challenges posed by Bayesian analyses of large data sets. *International Journal of Plant Sciences* 168: 137–157.

Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu YL, Hilu KW, Davis CC, Sanderson MJ, Soltis PS. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis DE, Soltis PS, Chase MW, 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WJ, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris KJ. 2000. Angiosperm

phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

Soltis DE, Soltis PS, Endress PK, Chase MW. 2005. *Phylogeny and evolution of angiosperms*. Sunderland, MA: Sinauer Associates.

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences USA* 92: 2647–2651.

Soltis DE, Soltis PS, Nickrent DL, Johnson LA, Hahn WJ, Hoot SB, Sweere JA, Kuzoff RK, Kron KA, Chase MW, Swensen SM, Zimmer EA, Chaw SM, Gillespie LJ, Kress WJ, Sytsma KJ. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 84: 1–49.

Soltis PS, Soltis DE. 2013. Angiosperm phylogeny: A framework for studies of genome evolution. *Plant Genome Diversity* 2: 1–11.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatic* 30: 1312–1313.

Stevens PF. 2001. Angiosperm phylogeny website. Version 12: July 2012 [online]. Available at www.mobot.org/MOBOT/research/APweb [Accessed 30 June 2016]

Stull GW, Duno de Stefano R, Soltis DE, Soltis PS. 2015. Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *American Journal of Botany* 102: 1794–1813.

Su JX, Wang W, Zhang LB, Chen ZD. 2012. Phylogenetic placement of two enigmatic genera, *Borthwickia* and *Stixis*, based on molecular and pollen data, and the description of a new family of Brassicales, Borthwickiaceae. *Taxon* 61: 601–611.

Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.

Wang HC, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE. 2009. Rosids radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences USA* 106: 3853–3858.

Wang XQ, Han Y, Deng ZR, Hong DY. 1997. Phylogeny of the Pinaceae evidenced by molecular biology. *Acta Phytotaxonomica Sinica* 35: 97–106.

Wang XQ, Li ZY. 1998. The application of sequence analysis of rDNA fragment to the systematic study of the subfamily Cyrtandroideae (Gesneriaceae). *Acta Phytotaxonomica Sinica* 36: 97–105.

Wang XQ, Ran JH. 2014. Evolution and biogeography of gymnosperms. *Molecular Phylogenetics and Evolution* 75: 24–40.

Wang XQ, Tank DC, Sang T. 2000. Phylogeny and divergence times in Pinaceae: Evidence from three genomes. *Molecular Biology and Evolution* 17: 773–781.

Wang ZC. 1994. Development of botany in ancient China. In Wang ZR ed. *History of Chinese Botany*. Beijing: Science Press. 3–8.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

Weigend M, Luebert F, Gottschling M, Couvreur TLP, Hilger HH, Miller JS. 2013. From capsules to nutlets—phylogenetic relationships in the Boraginales. *Cladistics* 30: 508–518.

Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H. 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Frontiers in Genetics* 5: 4.

Winter KU, Becker A, Munster T, Kim JT, Saedler H, Theissen G. 1999. MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. *Proceedings of the National Academy of Sciences USA* 96: 7342–7347.

Wu ZY, Lu AM, Tang YC, Chen ZD, Li DZ. 2002. Synopsis of a new "polyphyletic-polychronic-polytopic" system of the angiosperms. *Acta Phytotaxonomica Sinica* 40: 289–322.

Wu ZY, Lu AM, Tang YC, Chen ZD, Li DZ. 2003. *The families and genera of angiosperms in China: A comprehensive analysis*. Beijing: Science Press.

Wu ZY, Raven P, and Hong DY eds. 1994–2013. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press.

Wu ZY, Tang YC, Lu AM, Chen ZD. 1998. On primary subdivision of the Magnoliophyta - towards a new scheme for an eight-class system of classification of the angiosperms. *Acta Phytotaxonomica Sinica* 36: 385–402.

Xiang JY, Wen J, Peng H. 2015. Evolution of the eastern Asian–North American biogeographic disjunctions in ferns and lycophytes. *Journal of Systematics and Evolution* 53: 2–32.

Xu C, Dong WP, Shi S, Cheng T, Li CH, Liu YL, Wu P, Wu HK, Gao P, Zhou SL. 2015. Accelerating plant DNA barcode reference library construction using herbarium specimens: Improved experimental techniques. *Molecular Ecology Resources* 15: 1366–1374.

Zeng L, Zhang Q, Sun R, Kong HZ, Zhang N, Ma H. 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nature Communications* 5: 4956.

Zhang N, Zeng LP, Shan HY, Ma H. 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytologist* 195: 923–937.

Zhang XC, Wei R, Liu HM, He LJ, Wang L, Zhang GM. 2013. Phylogeny and classification of the extant lycophytes and ferns from China. *Chinese Bulletin of Botany* 48: 119–137.

Zhang YX, Li DZ. 2013. History of and recent advances in plant taxonomy. In: Hong DY, Blackmore S eds. *Plants of China: A companion to the* Flora of China. Beijing: Science Press. 256–278.

Zhu XY, Chase MW, Qiu YL, Kong HZ, Dilcher DL, Li JH, Chen ZD. 2007. Mitochondrial *matR* sequences help to resolve deep phylogenetic relationships in rosids. *BMC Evolutionary Biology* 7: 217.

Appendix

Members of the China Phylogeny Consortium

Zhi-Duan CHEN, An-Ming LU, Hong-Zhi KONG, Xiao-Quan WANG, Yin-Zheng WANG, Shi-Liang ZHOU
Institute of Botany, Chinese Academy of Sciences, Beijing, China
Shou-Zhou ZHANG, Xiao-Ming WANG
Fairylake Botanical Garden, Shenzhen, China
Zhong-Jian LIU
The Orchid Conservation & Research Center of Shenzhen, Shenzhen, China
Qing-Feng WANG
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
Jian-Hui LI
Computer Network Information Center, Chinese Academy of Sciences, Beijing, China
De-Zhu LI, Ting-Shuang YI
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Hong MA
Fudan University, Shanghai, China

Douglas E. SOLTIS, Pamela S. SOLTIS
University of Florida, Gainesville, USA
Jian-Hua LI
Hope College, Holland, USA
Cheng-Xin FU
Zhejiang University, Hangzhou, China
Qi-Xin LIU
Nanjing Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12219/suppinfo

Figure S1 The 50% maximum likelihood (ML) majority-rule bootstrap consensus tree of Chinese vascular plants.

Figure S2 SoTree name pre-processing flowchart.

Figure S3 SoTree name standardization flowchart.

Table S1 Accepted names of families and genera of Chinese vascular plants according to *Flora of China* (FOC). *Indicates the unsampled genus and family in this study.

Table S2 Primers for PCR and sequencing.

Table S3 SoTree algorithm that shows the result of each splitting for complete tree (taking set [2] for example). **Note: set (2)**: "((a:n1,b:n2):n3, (c:n4,d:n5):n6):0;" **weight**: 0, n1, n2,n3, n4, n5, n6.

Table S4 Sampling list of species, genera, and families with GenBank accession numbers.

Table S5 Taxon sampling statistics. *Sometimes in ferns and lycophytes, sequences from several species were selected to represent one genus. In that case, the species number of ferns and lycophytes here was equal to the number of genus.

Table S6 Sampling list of newly generated sequences in this study; each sample with China Phylogenetic Group (CPG) series number or other collection number.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12184

Research Article

Embracing the pteridophyte classification of Ren-Chang Ching using a generic phylogeny of Chinese ferns and lycophytes

Hong-Mei Liu*

Shenzhen Key Laboratory of Southern Subtropical Plant Diversity, Fairylake Botanical Garden, Shenzhen and Chinese Academy of Sciences, Shenzhen 518004, China

*Author for correspondence: E-mail: sorolepidium@hotmail.com. Tel.: 86-755-25735737. Fax: 86-755-25704480.

Received 6 July 2015; Accepted 17 September 2015; Article first published online 22 July 2016

Abstract The phylogenetic relationships of pteridophytes occurring in China were reconstructed using DNA sequences of the three plastid genes, *atpA*, *atpB*, and *rbcL*. The sampling comprised all genera of Chinese pteridophytes—including ferns and lycophytes—with the exception of four small genera. The effort to sample all recorded families and genera in a phylogenetic framework enabled the phylogenetic relationships of all Chinese pteridophytes to be addressed for the first time in a single phylogenetic hypothesis. The results provided strong evidence to support the continuing impact of Ren-Chang Ching's integrative classification of pteridophytes. Ten out of 11 orders accepted by Ching were consistent with the phylogeny, whereas four new orders were introduced to avoid paraphyletic taxa in the leptosporangiate ferns. Of the 63 families considered by Ching, 36 families were supported by molecular data, 22 of those had the same or nearly the same circumscription, and the remaining 14 families were supported but substantially revised. Twenty-eight small families were now accepted as synonyms. A consistent pattern was observed at the generic level. Among the 223 genera considered by Ching, 133 genera were recognized by the phylogeny, although some of them were substantially changed in the context of circumscription, and 90 were now accepted as synonyms. Three endemic genera were incorporated here for the first time in DNA-based phylogenetic analyses, namely *Blechnidium*, *Saxiglossum*, and *Sinephropteris*, which were shown to be nested in *Blechnum*, *Pyrrosia*, and *Asplenium* respectively. This paper tentatively accepts 40 families and 151 genera of ferns and lycophytes occurring in China; the importance of phylodiversity of Chinese pteridophytes is also briefly discussed.

Key words: East Asia, history of plant systematics, monophyly, multiple gene phylogeny, natural classification, tree of life.

The Chinese pteridophyte flora, with approximately 2300 species (Zhang, 2012; Fig. 1), contributes a substantial component to the global pteridophyte diversity, with estimates of 10 000 to 12 000 species (Smith et al., 2006; Moran, 2008). Understanding the pteridophyte flora of China is widely recognized as a major challenge because of the large number of species as well as several other factors codetermining the pteridophyte diversity, such as geographic range, topographic disparity, diverse climates, and geological history (Zhang, 2003b; Wang et al., 2012; Chen et al., 2016). In this context, it is crucial to note that species of ferns and lycophytes new to China continue to be discovered. These "new" species include widely distributed species recorded for China for the first time such as *Didymochlaena truncatula* (Tan et al., 2015), Southeast Asian species only now recognized to occur in China, such as *Selaginella decipiens* (Liu & Zhang, 2004), or species new to science such as *Asplenium cornutissimum* (Jiang et al., 2011). The rediscovery of local species such as *Cystopteris chinensis* (Wei & Zhang, 2014) and *Argyrochosma connectens* (Wang et al., 2015) is also a challenge in terms of the management of changes in the pteridophyte diversity of China. These two examples showed the importance of restudying the generic classification of rare Chinese ferns by reducing *Cystoathyrium* to a synonym of *Cystopteris* and documenting the occurrence of the mainly Neotropical genus *Argyrochosma* in China for the first time.

The 2300 species are considered to represent more than 200 genera, of which 42 are restricted to East Asia and six are endemic to China (modified from Zhang, 2003b; Table 1). These endemic genera comprise only a small proportion of the Chinese pteridophyte diversity (Zhang, 2003b) and many of these segregates are now recognized as synonyms or intrageneric taxa of more widespread genera (e.g., Liu et al., 2007b; Wang et al., 2010). However, some of the small endemic genera are confirmed by molecular data, like *Trichoneuron* (Liu et al., 2016). *Trichoneuron* was introduced to science by Ching (1965) but was treated as a synonym of *Lastreopsis* (e.g., Shing et al., 1999; Chu & He, 2000; Dong & Christenhusz, 2013). Multi-gene evidence supported the reinstatement of this East Asian endemic as a separate genus

Fig. 1. Images of ferns and lycophytes occurring in China representing the morphological and ecological diversity. **A,** Epiphytic *Drynaria propinqua* showing leaf dimorphism (Polypodiaceae). **B,** Epiphytic nest-fern *Asplenium nidus* (Aspleniaceae). **C,** Rock fern *Woodsia* (*Cheilanthopsis*) *elongata* (Woodsiaceae). **D,** Xerophytic rock fern *Aleuritopteris* (*Sinopteris*) *albofusca* (Pteridaceae). **E,** Climbing *Arthropteris palisotii* (Arthropteridaceae). **F,** Epiphytic *Elaphoglossum comforme* (Dryopteridaceae). **G,** *Cheiropleuria integrifolia* (Dipteridaceae) representing a basal lineage with dimorphic leaves. **H,** Apomictic *Polystichum* (*Cyrtogonellum*) *fraxinellum* (Dryopteridaceae). **I,** Thelypteroid *Stegnogramma* (*Dictyocline*) *sagittifolia* (Thelypteridaceae) illustrating the occurrence of complex venation patterns. **J,** Rock fern *Asplenium* (*Ceterachopsis*) *paucivenosum* (Aspleniaceae). **K,** Xerophytic *Selaginella pulvinata* (Selaginellaceae). **L,** Mangrove fern *Acrostichum aureum* (Pteridaceae). **M,** Mountain forest fern *Plagiogyria falcata* (Plagiogyriaceae) showing dimorphic leaves. **N,** *Botrychium yunnanense* (Ophioglossaceae), a representative of ferns with plesiomorphic eusporangia placed at the distinct spike. **O,** East Asian endemic *Blechnum* (*Struthiopteris*) *eburneum* (Blechnaceae). Species are labeled according to the taxonomy of *Flora of China* with one exception, *Stegnogramma sagittifolia* (following Zhang, 2012). Corresponding synonyms are given in brackets.

Table 1 Endemic genera recorded from China and East Asia (modified from Zhang, 2003b)

Family	Genus	*rbcL*	*atpA*	*atpB*	Name in GenBank	Others†
Angiopteridaceae	*Archangiopteris* Christ & Giesenh.	1/+	0/+	1/+		1
Aspleniaceae	*Boniniella* Hayata	1/+	0/+	0/+	*Hymenasplenium cardiophyllum*	1
Aspleniaceae	*Sinephropteris* Mickel	0/+	0/+	0/+		0
Athyriaceae	*Cystoathyrium* Ching	1	0	0		0
Athyriaceae	*Dictyodroma* Ching	1/+	0/+	0/+		1
Athyriaceae	*Kuniwatzukia* Pic. Serm.	1/+	0/+	0/+	*Athyrium cuspidatum*	1
Athyriaceae	*Monomelangium* Hayata	1	1/+	1/+	*Diaplazium pullingeri*	0
Athyriaceae	*Neoathyrium* Ching & Z. R. Wang	1	0	0	*Cornopteris crenulatoserrulata*	1
Athyriaceae	*Pseudocystopteris* Ching	1/+	0/+	0/+		
Blechnaceae	*Blechnidium* T. Moore	0/+	0/+	0/+	*Blechnidium melanopus*	0
Blechnaceae	*Chieniopteris* Ching	1/+	0/+	0/+	*Woodwardia harlandii/kempii*	1
Davalliaceae	*Paradavallodes* Ching	1	0	0	*Davallodes multidentata*	1
Dennstaedtiaceae	*Emodiopteris* Ching & S. K. Wu	0	0	0		0
Dennstaedtiaceae	*Cyrtogonellum* Ching	1/+	1/+	1/+		1
Dryopteridaceae	*Cyrtomidictyum* Ching	1/+	1/+	1/+	*Polystichum lepidocaulon*	1
Dryopteridaceae	*Leptorumohra* (H. Ito) H. Ito	1/+	0/+	1/+		1
Dryopteridaceae	*Lithostegia* Ching	1	0	1		1
Dryopteridaceae	*Phanerophlebiopsis* Ching	1/+	0/+	1/+		1
Dryopteridaceae	*Sorolepidium* Christ	1/+	0	1/+		1
Gymnogrammitidaceae	*Gymnogrammitis* Griffith	1/+	0/+	1/+		1
Monachosoraceae	*Ptilopteris* Hance	1/+	0/+	0/+	*Monachosorum maximowiczii*	1
Onocleaceae	*Pentarhizidium* Hayata	1	1	1		1
Ophioglossaceae	*Mankyua* B. Y. Sun, M. H. Kim & C. H. Kim	1	1	1	complete genome	1
Pleurosoriopsidaceae	*Pleurosoriopsis* Fomin	1/+	0/+	0/+		1
Polypodiaceae	*Arthromeris* (T. Moore) J. Sm.	1	1	1		1
Polypodiaceae	*Caobangia* A. R. Smith & X. C. Zhang	1	0/+	0/+		1
Polypodiaceae	*Drymotaenium* Makino	1/+	1/+	1	*Lepisorus miyoshianus*	1
Polypodiaceae	*Lepidomicrosorium* Ching & K. H. Shing	1/+	0/+	1/+	*Neocheiropteris superficialis*	1
Polypodiaceae	*Metapolypodium* Ching	1/+	0/+	0/+		1
Polypodiaceae	*Neocheiropteris* Christ	1/+	1/+	1/+		1
Polypodiaceae	*Platygyria* Ching & S. K. Wu	1/+	0/+	0/+		1
Polypodiaceae	*Polypodiodes* Ching	1/+	0/+	0/+		1
Polypodiaceae	*Saxiglossum* Ching	0/+	0/+	0/+		0
Sinopteridaceae	*Leptolepidium* K. H. Shing & S. K. Wu	1	0	0		1
Sinopteridaceae	*Sinopteris* C. Chr. & Ching	1/+	0/+	0/+		1
Thelypteridaceae	*Craspedosorus* Ching & W. M. Chu	0	0	0		0
Thelypteridaceae	*Cyclogramma* Tagawa	1/+	0/+	0/+		1
Thelypteridaceae	*Dictyocline* T. Moore	1/+	0/+	0/+		1
Thelypteridaceae	*Glaphyropteridopsis* Ching	1	0	0		1
Thelypteridaceae	*Mesopteris* Ching	1/+	0/+	0		1
Thelypteridaceae	*Trichoneuron* Ching	1	1	1		0
Woodsiaceae	*Cheilanthopsis* Hieron.	1/+	1/+	1/+	*Woodsia elongata/indusiosa*	1

Family and genera concepts follow Ching (1978) and are alphabetically arranged. †Sequences other than *atpA*, *atpB*, and *rbcL*. +, New sequences generated in the present study; 0, absent; 1, present.

of the family Dryopteridaceae (Liu et al., 2016). Unfortunately, several of these putative endemic genera have not been included in phylogenetic studies, thus their generic treatment is still considered ambiguous, such as the segregation of *Blechnidium* from *Blechnum*, the incorporation of *Sinephropteris* in *Asplenium*, and *Saxiglossum* in *Pyrrosia* (Table 1).

The remarkable fern diversity inspired one of the most innovative and influential pteridologist of the 20th century, Ren-Chang Ching (1898–1986). His groundbreaking publications not only contributed greatly to the classification of pteridophytes in China but continue to influence scientific activities in China and even to influence the study of fern classification and taxonomy around the globe because of his emphasis on integrative approaches to plant systematics (Kramer, 1995). Ching's studies, especially Ching (1940), challenged the classifications of pteridophytes established in the first half of the 20th century (e.g., Christ, 1897; Diels, 1899/1900; Christensen, 1938) and inspired all classifications of pteridophytes proposed since the mid-20th century (e.g., Copeland, 1947; Holttum, 1947; Pichi-Sermolli, 1977; Tryon & Tryon, 1982; Kramer, 1990c). Arguably, his integrative approach is still one of the main inspirations to the study of

the classification and taxonomy of pteridophytes, including studies incorporating phylogenetic methods. The influence of Ching on Chinese pteridology cannot be overstated. His classifications (Ching, 1940, 1954, 1978) were the basis for nearly all systematic arrangements of the Chinese pteridophytes until the introduction of post-cladistic classifications (Liu et al., 2008; Zhang, 2012; Zhang et al., 2013c). According to Ching's classification in 1954 (Ching, 1954), Chinese pteridophytes were arranged in nine orders, 41 families, and 161 genera. About 20 years later, Ching re-arranged Chinese pteridophytes into 11 orders, 63 families, and 223 genera (Ching, 1978; Table 2). Although Ching (1978) emphasized that his systematic arrangement was constrained by regional and temporal knowledge, his concepts and arrangements of families and genera of Chinese pteridophytes have been widely adopted by Chinese researchers (e.g., Wu & Ching, 1991) and *Flora Reipublicae Popularis Sinicae* (e.g., Ching et al., 1959, 1990; Chu et al., 1999; Shing et al., 1999; Wu, 1999; Wu & Wang, 1999; Lin et al., 2000; Wu et al., 2000; Kung et al., 2001; Zhang & Zhang, 2004). Many of his concepts were also confirmed in recent phylogenetic studies (e.g., Hypodematiaceae; see Christenhusz et al., 2011). At the same time, some of his treatments were not supported by molecular evidence, for example, several small genera of Dryopteridaceae (Li & Lu, 2006; Liu et al., 2007b, 2007c, 2010; Zhang & Zhang, 2012; Zhang et al., 2012). In the latest summary of previous molecular studies on phylogenetic relationships of ferns and lycophytes, Zhang et al. (2013c) recognized 14 orders, 39 families, and about 140 genera for Chinese ferns and lycophytes. However, there is no comprehensive phylogenetic study of Chinese pteridophytes that enables the exploration of Ching's classification in the context of our improved understanding of fern and lycophyte phylogeny (Liu et al., 2008).

This study aims to expose the continuous influence of Ching's concepts to our understanding of the diversity of ferns and lycophytes in China and at the same time address the phylogenetic relationships of all Chinese pteridophyte families and genera by including all of them in a single phylogenetic hypothesis. The chosen strategy, a large-scale phylogeny including all Chinese pteridophytes, was enabled by the ability to obtain phylogenetic hypotheses comprising several hundreds of taxa. Several outstanding examples showed the potential of large-scale phylogenies to address macroevolutionary/-ecological questions, such as the diversification of angiosperms (Smith et al., 2010, 2011; Xi et al., 2012), bryophytes (Feldberg et al., 2014; Laenen et al., 2014), ferns (Schuettpelz & Pryer, 2007, 2009; Lehtonen, 2011; Hennequin et al., 2014), non-flowering seed plants (Nagalingum et al., 2011), and all land plants (Fiz-Palacios et al., 2011). These studies achieved their objective by combining the increased power of bioinformatics tools (e.g., Miller et al., 2010) with the rapid assembly of DNA sequences for a vast range of taxa. The latter was accomplished by the combination of generating new DNA sequences and retrieving data from open source databases like GenBank (Benson et al., 2012). Currently, the major challenge to this kind of study is the incompleteness of DNA sequence data for a large portion of extant species (Lehtonen, 2011). However, this challenge can be addressed by targeted sequencing efforts focusing on taxa that were not incorporated in previous studies. For example, the recent study on the phylodiversity of fern flora of Mascarene Island comprised 211 out of 232 fern species with DNA sequences newly generated for 122 species (Hennequin et al., 2014). To achieve a phylogeny comprising nearly all Chinese pteridophyte species will require considerable effort but is likely achievable within the next couple of years.

Three main objectives are targeted in this study:

1. To expand the sampling of Chinese ferns and lycophytes to obtain a phylogenetic hypothesis including all families and genera recorded in China as recognized in pre-cladistic treatments (e.g., Ching, 1978) as well as recent taxonomic treatments (e.g., Zhang, 2012; Liu et al., 2013; Zhang et al., 2013c; Wu et al., 2013). The study also considers the results of more specialized studies focusing on particular groups (e.g., *Kaulinia*, Kreier et al., 2008; Zhang, 2012; Zhang et al., 2013c) or new records to China (e.g., *Didymochlaena*, Tan et al., 2015).
2. To re-evaluate the influence of Ching's concepts to our understanding of the Chinese pteridophyte diversity, and to update the classification and relationships of Chinese ferns and lycophytes based on molecular phylogeny.
3. To explore the status of segregates occurring exclusively in China and adjacent regions, especially those that were not inferred in previous studies.

Material and Methods

Taxon sampling

In general, relationships at family and higher levels in ferns and lycophytes are considered to be remarkably robust (Lehtonen, 2011; Rothfels et al., 2015), with the exception of some nodes such as the relationships of horsetails (Equisetales, see Pryer et al., 2001; Schneider, 2007; Schneider et al., 2009), the initial radiation of eupolypods II (Rothfels et al., 2012), and the relationships among the lineages closely related to the epiphytic Davalliaceae–Polypodiaceae clade (Liu et al., 2013, 2014). Therefore, the taxon sampling in this study was specifically focused on the generic level and it was not the aim to collect a substantial part of the species diversity of the Chinese pteridophyte flora. The study also aimed generally not to collect representatives of pteridophyte genera that do not occur in China. However, some species belonging to genera not occurring in China were included if they were considered to be important to clarify the generic relationships of Chinese pteridophytes.

According to Ching's classification, there are 223 genera classified in 63 families and 11 orders of Chinese pteridophytes (Ching, 1978). Since then, 24 generic names were added to the checklist of Chinese pteridophytes as either new genera or new records to China (Table 2; Wu, 1979; Ching & Wu, 1980; Ching & Wang, 1982; Ching & Shing, 1983; Ching et al., 1990; Li, 1990; Wu & Ching, 1991; Wang, 1992; Zhang, 1993; Chu & Zhou, 1994; Chu et al., 1999; Shing et al., 1999; Wu, 1999; Wu & Wang, 1999; Lin et al., 2000; Wu et al., 2000; Zhang & Kung, 2000; Kung et al., 2001; Zhang & Liu, 2004; Zhang & Zhang, 2004; Dong & Zhang, 2005; Xu et al., 2008; Shao & Lu, 2009, 2011; Zhang, 2012; Zhang et al., 2013c; Tan et al., 2015; Wang et al., 2015). Additional 21 genera were newly introduced in the recent English version of *Flora of China* (FOC) (see Wu

Table 2 Comparison of family and genera concepts in pre-cladistic and phylogeny-based classifications

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
Acrostichaceae	*Acrostichum*	1	*Acrostichum*	*Acrostichum*	*Acrostichum*	Pteridaceae
Adiantaceae	*Adiantum*	1	*Adiantum*	*Adiantum*	*Adiantum*	Pteridaceae
Angiopteridaceae	*Angiopteris*	1	*Angiopteris*	*Angiopteris*	*Angiopteris*	Marattiaceae
	Archangiopteris	1	*Angiopteris*	*Angiopteris*	*Angiopteris*	Marattiaceae
Antrophyaceae	*Antrophyum*	1	*Antrophyum*	*Antrophyum*	*Antrophyum*	Pteridaceae
Aspleniaceae	*Asplenium*	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Camptosorus	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Ceterach	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Ceterachopsis	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Neottopteris	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Phyllitis	2	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Sinephropteris	1	*Asplenium*	*Asplenium*	*Asplenium*	Aspleniaceae
	Boniniella	1/7	*Hymenasplenium*	*Hymenasplenium*	*Hymenasplenium*	Aspleniaceae
Athyriaceae	*Allantodia*	1	*Diplazium*	*Diplazium*	*Diplazium*	Athyriaceae
	Anisocampium	1	*Anisocampium*	*Anisocampium*	*Anisocampium*	Athyriaceae
	Athyriopsis	1	*Deparia*	*Deparia*	*Deparia*	Athyriaceae
	Athyrium	1	*Athyrium*	*Athyrium*	*Athyrium*	Athyriaceae
	Callipteris	1	*Diplazium*	*Diplazium*	*Diplazium*	Athyriaceae
	Cornopteris	2	*Cornopteris*	*Cornopteris*	*Cornopteris*	Athyriaceae
	Dictyodroma	1	*Deparia*	*Deparia*	*Deparia*	Athyriaceae
	Diplazium	1	*Diplazium*	*Diplazium*	*Diplazium*	Athyriaceae
	Dryoathyrium	1	*Deparia*	*Deparia*	*Deparia*	Athyriaceae
	Kuniwatzukia	1	*Anisocampium*	*Anisocampium*	*Anisocampium*	Athyriaceae
	Lunathyrium	1	*Deparia*	*Deparia*	*Deparia*	Athyriaceae
	Monomelangium	1	*Diplazium*	*Diplazium*	*Diplazium*	Athyriaceae
	Neoathyrium	3	*Cornopteris*	*Cornopteris*	*Cornopteris*	Athyriaceae
	Pseudocystopteris	1	*Athyrium*	*Athyrium*	*Athyrium*	Athyriaceae
	Triblemma	1	*Deparia*	*Deparia*	*Deparia*	Athyriaceae
	Acystopteris	1	*Acystopteris*	*Acystopteris*	*Acystopteris*	Cystopteridaceae
	Cystoathyrium	1	*Cystoathyrium*	*Cystoathyrium*	*Cystopteris*	Cystopteridaceae
	Cystopteris	1	*Cystopteris*	*Cystopteris*	*Cystopteris*	Cystopteridaceae
	Gymnocarpium	1	*Gymnocarpium*	*Gymnocarpium*	*Gymnocarpium*	Cystopteridaceae
	Diplaziopsis	1	*Diplaziopsis*	*Diplaziopsis*	*Diplaziopsis*	Diplaziopsidaceae
	Rhachidosorus	1	*Rhachidosorus*	*Rhachidosorus*	*Rhachidosorus*	Rhachidosoraceae
Azollaceae	*Azolla*	1	*Azolla*	*Azolla*	*Azolla*	Salviniaceae
Blechnaceae	*Blechnidium*	1	*Blechnidium*	*Blechnum*	*Blechnum*	Blechnaceae
	Blechnum	1	*Blechnum*	*Blechnum*	*Blechnum*	Blechnaceae
	Brainea	1	*Brainea*	*Brainea*	*Brainea*	Blechnaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
	Chieniopteris	1	*Chieniopteris*	*Woodwardia*	*Woodwardia*	Blechnaceae
	Diploblechnum	1	*Diploblechnum*	*Blechnum*	*Blechnum*	Blechnaceae
	Woodwardia	1	*Woodwardia*	*Woodwardia*	*Woodwardia*	Blechnaceae
	Struthiopteris	1	*Struthiopteris*	*Blechnum*	*Blechnum*	Blechnaceae
Bolbitidaceae	*Bolbitis*	1	*Bolbitis*	*Bolbitis*	*Bolbitis*	Dryopteridaceae
	Egenolfia	1	*Bolbitis*	*Bolbitis*	*Bolbitis*	Dryopteridaceae
Botrychiaceae	*Botrychium*	1	*Botrychium*	*Botrychium*	*Botrychium*	Ophioglossaceae
	Botryus	1	*Botrychium*	*Botrychium*	*Botrychium*	Ophioglossaceae
	Sceptidium	1	*Botrychium*	*Botrychium*	*Botrychium*	Ophioglossaceae
Cheiropleuriaceae	*Cheiropleuria*	1	*Cheiropleuria*	*Cheiropleuria*	*Cheiropleuria*	Dipteridaceae
Christenseniaceae	*Christensenia*	1	*Christensenia*	*Christensenia*	*Christensenia*	Marattiaceae
Cyatheaceae	*Alsophila*	1	*Alsophila*	*Alsophila*	*Alsophila*	Cyatheaceae
	Gymnosphaera	1	*Alsophila*	*Gymnosphaera*	*Gymnosphaera*	Cyatheaceae
	Sphaeropteris	1	*Sphaeropteris*	*Sphaeropteris*	*Sphaeropteris*	Cyatheaceae
Davalliaceae	*Araiostegia*	1	*Araiostegia*	*Davallia*	*Davallia*	Davalliaceae
	Davallia	1	*Davallia*	*Davallia*	*Davallia*	Davalliaceae
	Davallodes	1	–	*Davallia*	*Davallia*	Davalliaceae
	Humata	1	*Humata*	*Davallia*	*Davallia*	Davalliaceae
	Paradavallodes	1	*Paradavallodes*	*Davallia*	*Davallia*	Davalliaceae
	Leucostegia	1	*Leucostegia*	*Leucostegia*	*Leucostegia*	Hypodematiaceae
Dennstaedtiaceae	*Dennstaedtia*	1	*Dennstaedtia*	*Dennstaedtia*	*Dennstaedtia*	Dennstaedtiaceae
	Emodiopteris	1	*Dennstaedtia*	*Dennstaedtia*	*Emodiopteris*?	Dennstaedtiaceae
	Microlepia	1	*Microlepia*	*Microlepia*	*Microlepia*	Dennstaedtiaceae
Dicksoniaceae	*Cibotium*	1	*Cibotium*	*Cibotium*	*Cibotium*	Cibotiaceae
Didymochlaenaceae	*Didymochlaena*	4	–	–	*Didymochlaena*	Didymochlaenaceae
Dipteridaceae	*Dipteris*	1	*Dipteris*	*Dipteris*	*Dipteris*	Dipteridaceae
Drynariaceae	*Aglaomorpha*	1	*Aglaomorpha*	*Drynaria*	*Drynaria*	Polypodiaceae
	Drynaria	1	*Drynaria*	*Drynaria*	*Drynaria*	Polypodiaceae
	Photinopteris	5/6	*Photinopteris*	*Drynaria*	*Drynaria*	Polypodiaceae
	Pseudodrynaria	1	*Aglaomorpha*	*Drynaria*	*Drynaria*	Polypodiaceae
Dryopteridaceae	*Acrorumohra*	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
	Arachniodes	1	*Arachniodes*	*Arachniodes*	*Arachniodes*	Dryopteridaceae
	Cyrtogonellum	1	*Polystichum*	*Cyrtogonellum*	*Polystichum*	Dryopteridaceae
	Cyrtomidictyum	1	*Polystichum*	*Cyrtomidictyum*	*Polystichum*	Dryopteridaceae
	Cyrtomium	1	*Cyrtomium*	*Cyrtomium*	*Cyrtomium*	Dryopteridaceae
	Dryopteris	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
	Leptorumohra	1	*Arachniodes*	*Arachniodes*	*Arachniodes*	Dryopteridaceae
	Lithostegia	1	*Arachniodes*	*Arachniodes*	*Arachniodes*	Dryopteridaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
	Nothoperanema	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
	Phanerophlebiopsis	1	*Arachniodes*	*Arachniodes*	*Arachniodes*	Dryopteridaceae
	Polystichum	1	*Polystichum*	*Polystichum*	*Polystichum*	Dryopteridaceae
	Sorolepidium	1	*Polystichum*	*Polystichum*	*Polystichum*	Dryopteridaceae
	Cyclopeltis	1	*Cyclopeltis*	*Cyclopeltis*	*Cyclopeltis*	Lomariopsidaceae
Elaphoglossaceae	*Elaphoglossum*	1	*Elaphoglossum*	*Elaphoglossum*	*Elaphoglossum*	Dryopteridaceae
Equisetaceae	*Equisetum*	1	*Equisetum*	*Equisetum*	*Equisetum*	Equisetaceae
	Hippochaete	1	*Equisetum*	*Equisetum*	*Equisetum*	Equisetaceae
Gleicheniaceae	*Dicranopteris*	1	*Dicranopteris*	*Dicranopteris*	*Dicranopteris*	Gleicheniaceae
	Diplopterygium	1	*Diplopterygium*	*Diplopterygium*	*Diplopterygium*	Gleicheniaceae
	Sticherus	1	*Sticherus*	*Sticherus*	*Sticherus*	Gleicheniaceae
Grammitaceae	*Acrosorus*	1	–	*Grammitis*	*Acrosorus* (?)	Polypodiaceae
	Calymmodon	1	*Calymmodon*	*Grammitis*	*Calymmodon*	Polypodiaceae
	Chrysogrammitis	7	*Chrysogrammitis*	*Grammitis*	*Chrysogrammitis*	Polypodiaceae
	Ctenopteris	1	*Prosaptia*	*Grammitis*	*Prosaptia*	Polypodiaceae
	Ctenopterella	7	*Ctenopterella*	*Grammitis*	*Ctenopterella*	Polypodiaceae
	Dasygrammitis	7	*Dasygrammitis*	*Grammitis*	*Dasygrammitis*	Polypodiaceae
	Grammitis	1	–	*Grammitis*	◇	Polypodiaceae
	Micropolypodium	6	*Micropolypodium*	*Grammitis*	*Micropolypodium*	Polypodiaceae
	Oreogrammitis	7	*Oreogrammitis*	*Grammitis*	*Oreogrammitis*	Polypodiaceae
	Prosaptia	1	*Prosaptia*	*Grammitis*	*Prosaptia*	Polypodiaceae
	Radiogrammitis	7	*Radiogrammitis*	*Grammitis*	*Radiogrammitis*	Polypodiaceae
	Scleroglossum	1	*Scleroglossum*	*Grammitis*	*Scleroglossum*	Polypodiaceae
	Themelium	7	*Themelium*	*Grammitis*	*Themelium*	Polypodiaceae
	Tomophyllum	7	*Tomophyllum*	*Grammitis*	*Tomophyllum*	Polypodiaceae
	Xiphopterella	7	*Xiphopterella*	*Grammitis*	*Xiphopterella*	Polypodiaceae
Gymnogrammitidaceae	*Gymnogrammitis*	1	*Gymnogrammitis*	*Selliguea*	*Selliguea*	Polypodiaceae
Helminthostachyaceae	*Helminthostachys*	1	*Helminthostachys*	*Helminthostachys*	*Helminthostachys*	Ophioglossaceae
Hemionitidaceae	*Anogramma*	1	*Anogramma*	*Anogramma*	*Anogramma*	Pteridaceae
	Cerosora	8	–	*Cerosora*	*Cerosora*	Pteridaceae
	Coniogramme	1	*Coniogramme*	*Coniogramme*	*Coniogramme*	Pteridaceae
	Gymnopteris	1/7	*Paragymnopteris*	*Paragymnopteris*	*Paragymnopteris*	Pteridaceae
	Hemionitis	1/7	*Parahemionitis*	*Hemionitis*	*Hemionitis*	Pteridaceae
	Pityrogramma	1	*Pityrogramma*	*Pityrogramma*	*Pityrogramma*	Pteridaceae
Huperziaceae	*Huperzia*	1	*Huperzia*	*Huperzia*	*Huperzia*	Lycopodiaceae
	Phlegmariurus	1	*Phlegmariurus*	*Huperzia*	*Phlegmariurus*	Lycopodiaceae
Hymenophyllaceae	*Abrodictyum*	1	*Abrodictyum*	*Abrodictyum*	*Abrodictyum*	Hymenophyllaceae
	Callistopteris	1	*Callistopteris*	*Callistopteris*	*Callistopteris*	Hymenophyllaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
	Cephalomanes	1	*Cephalomanes*	*Cephalomanes*	*Cephalomanes*	Hymenophyllaceae
	Crepidomanes	1	*Crepidomanes*	*Crepidomanes*	*Crepidomanes*	Hymenophyllaceae
	Crepidopteris	1	*Crepidomanes*	*Crepidomanes*	*Crepidomanes*	Hymenophyllaceae
	Didymoglossum	9	*Didymoglossum*	*Didymoglossum*	*Didymoglossum*	Hymenophyllaceae
	Gonocormus	1	*Crepidomanes*	*Crepidomanes*	*Crepidomanes*	Hymenophyllaceae
	Hymenophyllum	1	*Hymenophyllum*	*Hymenophyllum*	*Hymenophyllum*	Hymenophyllaceae
	Mecodium	1	*Hymenophyllum*	*Hymenophyllum*	*Hymenophyllum*	Hymenophyllaceae
	Meringium	1	*Hymenophyllum*	*Hymenophyllum*	*Hymenophyllum*	Hymenophyllaceae
	Microgonium	1	*Didymoglossum*	*Didymoglossum*	*Didymoglossum*	Hymenophyllaceae
	Nesopteris	1	*Crepidomanes*	*Crepidomanes*	*Crepidomanes*	Hymenophyllaceae
	Pleuromanes	1	*Hymenophyllum*	*Hymenophyllum*	*Hymenophyllum*	Hymenophyllaceae
	Selenodesmium	1	*Abrodictyum*	*Abrodictyum*	*Abrodictyum*	Hymenophyllaceae
	Trichomanes	1/7	*Vandenboschia*	*Vandenboschia*	*Vandenboschia*	Hymenophyllaceae
Hypodematiaceae	*Hypodematium*	1	*Hypodematium*	*Hypodematium*	*Hypodematium*	Hypodematiaceae
Hypolepidaceae	*Hypolepis*	1	*Hypolepis*	*Hypolepis*	*Hypolepis*	Dennstaedtiaceae
Isoetaceae	*Isoetes*	1	*Isoetes*	*Isoetes*	*Isoetes*	Isoetaceae
Lindsaeaceae	*Lindsaea*	1	*Lindsaea*	*Lindsaea*	*Lindsaea*	Lindsaeaceae
	Osmolindsaea	7	*Osmolindsaea*	*Osmolindsaea*	*Osmolindsaea*	Lindsaeaceae
	Tapeinidium	1	*Tapeinidium*	*Tapeinidium*	*Tapeinidium*	Lindsaeaceae
	Schizoloma	10	*Lindsaea*	*Lindsaea*	*Lindsaea*	Lindsaeaceae
	Stenoloma	1/7	*Odontosoria*	*Odontosoria*	*Odontosoria*	Lindsaeaceae
Lomariopsidaceae	*Lomagramma*	1	*Lomagramma*	*Lomagramma*	*Lomagramma*	Dryopteridaceae
	Lomariopsis	1	*Lomariopsis*	*Lomariopsis*	*Lomariopsis*	Lomariopsidaceae
	Teratophyllum	11	*Teratophyllum*	*Teratophyllum*	*Teratophyllum*	Dryopteridaceae
Loxogrammaceae	*Loxogramme*	1	*Loxogramme*	*Loxogramme*	*Loxogramme*	Polypodiaceae
Lycopodiaceae	*Diphasiastrum*	1	*Lycopodium*	*Lycopodium*	*Lycopodium*	Lycopodiaceae
	Lycopodiastrum	1	*Lycopodiastrum*	*Lycopodium*	*Lycopodium*	Lycopodiaceae
	Lycopodiella	1	*Lycopodiella*	*Lycopodiella*	*Lycopodiella*	Lycopodiaceae
	Lycopodium	1	*Lycopodium*	*Lycopodium*	*Lycopodium*	Lycopodiaceae
	Palhinhaea	1	*Lycopodium*	*Lycopodiella*	*Lycopodiella*	Lycopodiaceae
	Pseudolycopodiella	12	*Lycopodiella*	*Lycopodiella*	*Lycopodiella*	Lycopodiaceae
Lygodiaceae	*Lygodium*	1	*Lygodium*	*Lygodium*	*Lygodium*	Lygodiaceae
Marattiaceae	*Marattia*	1/7	*Ptisana*	*Ptisana*	*Ptisana*	Marattiaceae
Marsileaceae	*Marsilea*	1	*Marsilea*	*Marsilea*	*Marsilea*	Marsileaceae
Monachosoraceae	*Monachosorum*	1	*Monachosorum*	*Monachosorum*	*Monachosorum*	Dennstaedtiaceae
	Ptilopteris	1	*Monachosorum*	*Monachosorum*	*Monachosorum*	Dennstaedtiaceae
Nephrolepidaceae	*Arthropteris*	1	*Arthropteris*	*Arthropteris*	*Arthropteris*	Arthropteridaceae
	Nephrolepis	1	*Nephrolepis*	*Nephrolepis*	*Nephrolepis*	Nephrolepidaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
Oleandraceae	*Oleandra*	1	*Oleandra*	*Oleandra*	*Oleandra*	Oleandraceae
Onocleaceae	*Matteuccia*	1	*Matteuccia*	*Onoclea?*	*Matteuccia*	Onocleaceae
	Onoclea	1	*Onoclea*	*Onoclea*	*Onoclea*	Onocleaceae
	Pentarhizidium	13	*Pentarhizidium*	*Onoclea?*	*Pentarhizidium*	Onocleaceae
Ophioglossaceae	*Ophioderma*	1	*Ophioglossum*	*Ophioglossum*	*Ophioglossum*	Ophioglossaceae
	Ophioglossum	1	*Ophioglossum*	*Ophioglossum*	*Ophioglossum*	Ophioglossaceae
Osmundaceae	*Osmunda*	1	*Osmunda*	*Osmunda*	*Osmunda*	Osmundaceae
	Osmundastrum	7	*Osmundastrum*	*Osmundastrum*	*Osmundastrum*	Osmundaceae
Parkeriaceae	*Ceratopteris*	1	*Ceratopteris*	*Ceratopteris*	*Ceratopteris*	Pteridaceae
Peranemaceae	*Acrophorus*	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
	Diacalpe	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
	Peranema	1	*Dryopteris*	*Dryopteris*	*Dryopteris*	Dryopteridaceae
Plagiogyriaceae	*Plagiogyria*	1	*Plagiogyria*	*Plagiogyria*	*Plagiogyria*	Plagiogyriaceae
Platyceriaceae	*Platycerium*	1	*Platycerium*	*Platycerium*	*Platycerium*	Polypodiaceae
Pleurosoriopsidaceae	*Pleurosoriopsis*	1	*Pleurosoriopsis*	*Pleurosoriopsis*	*Pleurosoriopsis*	Polypodiaceae
Polypodiaceae	*Arthromeris*	1	*Arthromeris*	*Arthromeris*	*Arthromeris*	Polypodiaceae
	Belvisia	1	*Lepisorus*	*Lepisorus*	*Lepisorus*	Polypodiaceae
	Caobangia	14	*Caobangia*	*Lemmaphyllum*	*Lemmaphyllum*	Polypodiaceae
	Christiopteris	1	*Christiopteris*	*Drynaria*	*Drynaria*	Polypodiaceae
	Colysis	1	*Leptochilus*	*Leptochilus*	*Leptochilus*	Polypodiaceae
	Crypsinus–	15	*Selliguea*	–	*Selliguea*	Polypodiaceae
	Dendroglossa	1	*Leptochilus*	*Leptochilus*	*Leptochilus*	Polypodiaceae
	Drymoglossum	1	*Pyrrosia*	*Pyrrosia*	*Pyrrosia*	Polypodiaceae
	Drymotaenium	1	*Lepisorus*	*Lepisorus*	*Lepisorus*	Polypodiaceae
	Kaulinia	8	–	*Kaulinia*	*Kaulinia*	Polypodiaceae
	Himalayopteris	16	*Himalayopteris*	–	*Himalayopteris?*	Polypodiaceae
	Lemmaphyllum	1	*Lemmaphyllum*	*Lemmaphyllum*	*Lemmaphyllum*	Polypodiaceae
	Lepidogrammitis	1	*Lemmaphyllum*	*Lemmaphyllum*	*Lemmaphyllum*	Polypodiaceae
	Lepidomicrosorium	17	*Lepidomicrosorium*	*Lepidomicrosorium*	*Lepidomicrosorium*	Polypodiaceae
	Lepisorus	1	*Lepisorus*	*Lepisorus*	*Lepisorus*	Polypodiaceae
	Leptochilus	1	*Leptochilus*	*Leptochilus*	*Leptochilus*	Polypodiaceae
	Metapolypodium	1	*Metapolypodium*	*Goniophlebium*	*Goniophlebium*	Polypodiaceae
	Microsorum	1	*Microsorum*	*Microsorum*	*Microsorum*	Polypodiaceae
	Neocheiropteris	1	*Neocheiropteris*	*Neocheiropteris*	*Neocheiropteris*	Polypodiaceae
	Neolepisorus	1	*Neolepisorus*	*Neolepisorus*	*Neolepisorus*	Polypodiaceae
	Paraleptochilus	1	*Leptochilus*	*Leptochilus*	*Leptochilus*	Polypodiaceae
	Phymatosorus	1	*Phymatosorus*	*Phymatosorus*	*Microsorum*	Polypodiaceae
	Phymatopteris	1	*Selliguea*	*Selliguea*	*Selliguea*	Polypodiaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
	Platygyria	18	*Lepisorus*	*Lepisorus*	*Lepisorus*	Polypodiaceae
	Polypodiastrum	1	*Polypodiastrum*	*Goniophlebium*	*Goniophlebium*	Polypodiaceae
	Polypodiodes	1	*Polypodiodes*	*Goniophlebium*	*Goniophlebium*	Polypodiaceae
	Polypodium	1	*Polypodium*	*Polypodium*	*Polypodium*	Polypodiaceae
	Pyrrosia	1	*Pyrrosia*	*Pyrrosia*	*Pyrrosia*	Polypodiaceae
	Saxiglossum	1	*Pyrrosia*	*Pyrrosia*	*Pyrrosia*	Polypodiaceae
	Schellolepis	1/7	*Goniophlebium*	*Goniophlebium*	*Goniophlebium*	Polypodiaceae
	Selliguea	1	*Selliguea*	*Selliguea*	*Selliguea*	Polypodiaceae
	Tricholepidium	1	*Tricholepidium*	*Tricholepidium*	*Tricholepidium*	Polypodiaceae
Psilotaceae	*Psilotum*	1	*Psilotum*	*Psilotum*	*Psilotum*	Psilotaceae
Pteridaceae	*Histiopteris*	1	*Histiopteris*	*Histiopteris*	*Histiopteris*	Dennstaedtiaceae
	Pteris	1	*Pteris*	*Pteris*	*Pteris*	Pteridaceae
	Paesia	1	*Paesia*	*Paesia*	*Paesia*	Dennstaedtiaceae
	Pteridium	1	*Pteridium*	*Pteridium*	*Pteridium*	Dennstaedtiaceae
Salviniaceae	*Salvinia*	1	*Salvinia*	*Salvinia*	*Salvinia*	Salviniaceae
Schizaeaceae	*Schizaea*	1	*Schizaea*	*Schizaea*	*Schizaea*	Schizaeaceae
Selaginellaceae	*Selaginella*	1	*Selaginella*	*Selaginella*	*Selaginella*	Selaginellaceae
Sinopteridaceae	*Aleuritopteris*	1	*Aleuritopteris*	*Aleuritopteris*	*Aleuritopteris*	Pteridaceae
	Argyrochosma	22	–	–	*Argyrochosma*	Pteridaceae
	Cheilosoria	1/7	*Cheilanthes*	*Cheilanthes*	*Cheilanthes*	Pteridaceae
	Cryptogramma	1	*Cryptogramma*	*Cryptogramma*	*Cryptogramma*	Pteridaceae
	Doryopteris	1	*Doryopteris*	*Doryopteris*	*Doryopteris*	Pteridaceae
	Doryopteris	7	*Calciphilopteris*	*Calciphilopteris*	*Calciphilopteris*	Pteridaceae
	Leptolepidium	19	*Aleuritopteris*	*Aleuritopteris*	*Aleuritopteris*	Pteridaceae
	Notholaena	1/7	*Cheilanthes*	*Cheilanthes*	*Cheilanthes*	Pteridaceae
	Onychium	1	*Onychium*	*Onychium*	*Onychium*	Pteridaceae
	Pellaea	1	*Pellaea*	*Pellaea*	◇	Pteridaceae
	Sinopteris	1	*Aleuritopteris*	*Aleuritopteris*	*Aleuritopteris*	Pteridaceae
Stenochlaenaceae	*Stenochlaena*	1	*Stenochlaena*	*Stenochlaena*	*Stenochlaena*	Blechnaceae
Taenitidaceae	*Taenitis*	1	*Taenitis*	*Taenitis*	*Taenitis*	Pteridaceae
Tectariaceae	*Ataxipteris*	20	*Ctenitis*	*Ctenitis*	*Ctenitis*	Dryopteridaceae
	Ctenitis	1	*Ctenitis*	*Ctenitis*	*Ctenitis*	Dryopteridaceae
	Dryopsis	21	*Dryopteris*	*Dryopsis*	*Dryopteris*	Dryopteridaceae
	Lastreopsis	1	*Lastreopsis*	*Lastreopsis*	*Lastreopsis*	Dryopteridaceae
	Pleocnemia	1	*Pleocnemia*	*Pleocnemia*	*Pleocnemia*	Dryopteridaceae
	Ctenitopsis	1	*Tectaria*	*Tectaria*	*Tectaria*	Tectariaceae
	Hemigramma	1	*Tectaria*	*Tectaria*	*Tectaria*	Tectariaceae
	Pteridrys	1	*Pteridrys*	*Pteridrys*	*Pteridrys*	Tectariaceae

Continued

Table 2 Continued

Pre-cladistic classification				Phylogeny-based classification		
Family	Genus	Ref.	FOC (2013)	Zhang et al. (2013c)	This study	Family
	Quercifilix	1	*Tectaria*	*Tectaria*	*Tectaria*	Tectariaceae
	Tectaria	1	*Tectaria*	*Tectaria*	*Tectaria*	Tectariaceae
Thelypteridaceae	*Ampelopteris*	1	*Ampelopteris*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Amphineuron	1	*Cyclosorus*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Craspedosorus	1	*Craspedosorus*	–	*Craspedosorus*?	Thelypteridaceae
	Cyclogramma	1	*Cyclogramma*	*Cyclogramma*	*Cyclosorus*	Thelypteridaceae
	Cyclosorus	1	*Cyclosorus*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Dictyocline	1	*Dictyocline*	*Stegnogramma*	*Cyclosorus*	Thelypteridaceae
	Glaphyropteridopsis	1	*Glaphyropteridopsis*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Lastrea	1/7	*Oreopteris*	*Oreopteris*	*Oreopteris*	Thelypteridaceae
	Leptogramma	1	*Leptogramma*	*Stegnogramma*	*Cyclosorus*	Thelypteridaceae
	Macrothelypteris	1	*Macrothelypteris*	*Macrothelypteris*	*Macrothelypteris*	Thelypteridaceae
	Mesopteris	1	*Mesopteris*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Metathelypteris	1	*Metathelypteris*	*Parathelypteris*	*Parathelypteris*	Thelypteridaceae
	Parathelypteris	1	*Parathelypteris*	*Parathelypteris*	*Parathelypteris*	Thelypteridaceae
	Phegopteris	1	*Phegopteris*	*Phegopteris*	*Phegopteris*	Thelypteridaceae
	Pronephrium	1	*Pronephrium*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Pseudocyclosorus	1	*Pseudocyclosorus*	*Cyclosorus*	*Cyclosorus*	Thelypteridaceae
	Pseudophegopteris	1	*Pseudophegopteris*	*Pseudophegopteris*	*Pseudophegopteris*	Thelypteridaceae
	Stegnogramma	1	*Stegnogramma*	*Stegnogramma*	*Cyclosorus*	Thelypteridaceae
	Thelypteris	1	*Thelypteris*	*Thelypteris*	*Thelypteris*	Thelypteridaceae
	Trichoneuron	1	*Lastreopsis*	*Lastreopsis*	*Trichoneuron*	Dryopteridaceae
Vittariaceae	*Vittaria*	1/7	*Haplopteris*	*Haplopteris*	*Haplopteris*	Pteridaceae
	Monogramma	1	*Monogramma*	*Monogramma*	◇	Pteridaceae
	Vaginularia	1	*Monogramma*	*Monogramma*	*Vaginularia*	Pteridaceae
Woodsiaceae	*Cheilanthopsis*	1	*Cheilanthopsis*	*Woodsia*	*Woodsia*	Woodsiaceae
	Eriosoriopsis†	2	–	–	–	Woodsiaceae
	Protowoodsia	1	*Protowoodsia*	*Woodsia*	*Woodsia*	Woodsiaceae
	Woodsia	1	*Woodsia*	*Woodsia*	*Woodsia*	Woodsiaceae

The family and genera concepts are based on Ching (1978) and alphabetically arranged. Arthropteridaceae were accepted (Zhang et al., 2013c) instead to be included in Tectariaceae (Xing et al., 2013); *Pleocnemia* was treated as a member of Tectariaceae in *Flora of China* (Xing et al., 2013) but included in Dryopteridaceae in Zhang et al. (2013c), which was confirmed later by Liu et al. (2014). References: 1, Ching (1978); 2, Wu & Ching (1991); 3, Ching & Wang (1982); 4, Tan et al. (2015); 5, Li (1990); 6, Lin et al. (2000); 7, *Flora of China* (2013); 8, Zhang et al. (2013c); 9, Chu & Zhou (1994); 10, Ching et al. (1959); 11, Dong & Zhang (2005); 12, Zhang & Kung (2000); 13, Zhang & Liu (2004); 14, Xu et al. (2008); 15, Shao & Lu (2009); 16, Shao & Lu (2011); 17, Ching & Shing (1983); 18, Ching & Wu (1980); 19, Wu (1979); 20, Zhang (1993); 21, Wang (1992); 22, Wang et al. (2015). †*Eriosoriopsis* was not accepted here based on invalid publication (see detail in Zhang et al., 2013a); ◇, Not occurring in China; (?), need confirmation of occurrence in China;?, treated as ambiguous in this study; –, no treatment.

et al., 2013). As a result, up to 268 generic names are here considered for Chinese pteridophyte genera; of these, 263 were included in the current sampling (Table S1). Four genera were not sampled because of either a lack of material (e.g., *Craspedosorus*) or the discovery of the Chinese occurrence postdated the analysis of this study (*Argyrochosma*). For the convenience of readers, the taxon name of newly generated sequences follows the taxonomy of the most recent comprehensive treatment of Chinese pteridophytes (Ching, 1978), and the original references, and/or the *Flora Reipublicae Popularis Sinicae* (e.g., Ching et al., 1959, 1990; Chu et al., 1999; Shing et al., 1999; Wu, 1999; Wu & Wang, 1999; Lin et al., 2000; Wu et al., 2000; Kung et al., 2001; Zhang & Zhang, 2004). Alternative generic concepts and/or species combinations are provided as treated in the recently published FOC and latest references (Tables 2, S1).

Efforts were made to incorporate the type species of each genus (Table S1) and more than one species per genus were included when the type species was missing or the available specimens contained incomplete DNA sequences. For a large proportion of species included, only *rbcL* sequences were available. Importantly, the study aimed to include at least one species occurring in China for each genus included. Several hundreds of new DNA sequences were generated to achieve this aim (Table S1). To root the ferns and lycophytes phylogeny, four bryophytes and three seed plants were selected as outgroup taxa.

DNA extraction and polymerase chain reaction amplification and sequencing

Total genomic DNA was extracted from silica-dried materials using either a modified CTAB approach (Doyle & Doyle, 1987) or the Plant Genomic DNA Kit (Tiangen Biotech Co., Beijing, China) following the manufacturer's instructions. For each species, three protein-coding plastid genes (*atpA*, *atpB*, and *rbcL*) were separately amplified using the standard polymerase chain reaction following the protocols reported in previous studies (Liu et al., 2013, 2014). Sequence reactions were carried out using an ABI 3730XL genetic analyzer and the associated BigDye chemistry (Applied Biosystems, Foster City, CA, USA). New sequences were assembled and edited using BioEdit version 7.1.11 (Hall, 1999) and were checked using nucleotide blast searches in GenBank (Benson et al., 2012). All new sequences used in the final analyses were deposited in GenBank (Table S1).

Sequence alignment and phylogenetic analyses

The alignments were generated using MUSCLE (Edgar, 2004) followed by manual adjustment in BioEdit (Hall, 1999). The incomplete 5′- and 3′-ends of the *rbcL* and *atpB* alignments were removed, as were the non-coding regions at the 5′-flank of *atpA*. Two phylogenetic analyses were carried out in this study. First, a maximum likelihood analysis of 720 taxa with only *rbcL* sequence data was carried out. The result of this analysis (results not shown) was considered as the framework for the further combined phylogenetic analysis. Second, a combined dataset of 662 taxa was analyzed comprising three genes representing 263 genera recorded in China. Both datasets were analyzed using the maximum likelihood approach as implemented in RAxML (Miller et al., 2010). The model of molecular evolution was determined using jModeltest 3.7 (Posada & Crandall, 1998) and bootstrap values were obtained by carrying out 100 non-parametric bootstrap replicates (Felsenstein, 1985).

Results

A total of 600 DNA sequences were newly generated, including 228 *rbcL* sequences, 202 *atpB* sequences, and 170 *atpA* sequences (given with both the lab number BOPXXXXXX or FPXXXXXX and GenBank accession numbers in Table S1). Among the 268 genera recorded in China, 264 of them now have molecular data accumulating to the generic sampling coverage of 98.5%. The type species were available for 192 out of 263 genera included, corresponding to type species coverage of 73% (Table S1). The recovered phylogenetic hypothesis was consistent with previously published hypotheses as long as some unsupported branches, e.g., Dipteridaceae not sister to the Gleicheniaceae, are ignored (Fig. 2; the original trees are available on request from the author).

New DNA sequence data were generated for 30 of the 42 genera with restricted distribution to China and/or East Asia (Tables 1, S1). Of these, three genera, namely *Blechnidium*, *Saxiglossum*, and *Sinephropteris* were incorporated into phylogenetic analyses for the first time (Tables 1, S1; Figs. 3–5). The four samples of *Blechnidium melanopus* (= *Blechnum melanopus*) formed a monophylum sister to *Blechnum spicant*—the type species of *Struthiopteris*, whereas *Struthiopteris eburnea* (= *Blechnum eburneum*) was found to be sister to *Blechnum orientale* (Fig. 3). The two *Saxiglossum angustissimum* (= *Pyrrosia angustissima*) samples nested within a clade formed by species belonging to the genus *Pyrrosia* (Fig. 4). Specifically, this species formed a clade together with *Pyrrosia subfurfuracea* and *Pyrrosia piloselloides*—a representative of the former segregate *Dryomoglossum* (Fig. 4). The spleenwort *Sinephropteris delavayi* (=*Asplenium delavayi*), the single species of the genus *Sinephropteris*, was found to have close relationships with temperate species of *Asplenium* including *Asplenium trichomanes*, *A. rhizophyllum*, and *A. ruprechtii* (Fig. 5). The latter two species formed a sister clade corresponding to the satellite genus *Camptosorus*.

The three samplings of *Ataxipteris sinii* (= *Ctenitis sinii*) confirmed the placement of the species in the genus *Ctenitis* (Fig. 6). Newly generated sequences of *Pleurosoriopsis makinoi* confirmed the sister relationships of this enigmatic genus to *Polypodium* (Fig. 4), whereas the newly obtained sequences of *Cheilanthes* (*Cheilosoria*) *hancockii* and *Cheilanthes nudiuscula* (= *Notholaena hirsuta*) confirmed the polyphyly of *Cheilanthes* in China (Fig. 7). The exposure of close relationships of *Cheilanthes nudiuscula* and the Brazilian *C. micropteris*—the type species of *Cheilanthes*—indicated the occurrence of the genus *Cheilanthes* in Southeast Asia (Fig. 7). *Cheilanthes hancockii* instead was sister to *C. chusana*. Together, these two species were nested in a clade comprising representatives of the segregates *Aleuritopteris*, *Leptolepidium*, and *Sinopteris*.

Discussion

Large-scale phylogenetic studies rely often on availability and reliability of the DNA sequence data deposited in public

Fig. 2. Phylogenetic hypothesis depicting relationships of ferns and lycophytes based on the maximum likelihood analysis of the combined dataset including *atpA*, *atpB*, and *rbcL* DNA sequence data. The topology is presented as a phylogram with branch lengths corresponding to the estimated number of substitution events. Families and major groups recognized in the recent classification of extant ferns and lycophytes are indicated (Smith et al., 2006; Christenhusz et al., 2011; Zhang et al., 2013c). Branches considered as ambiguous (bootstrap percentages $\leq$90%) are marked in red.

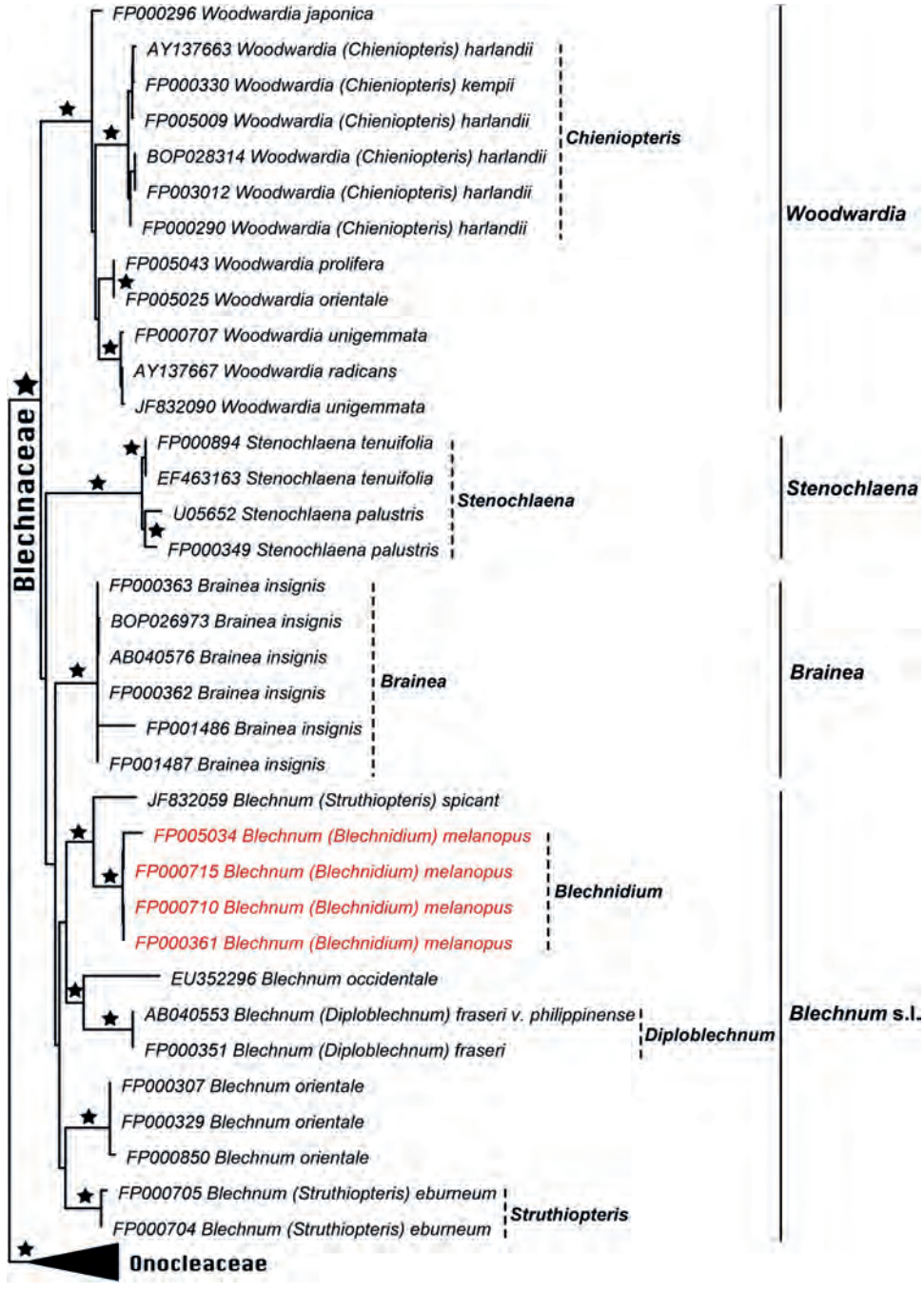


Fig. 3. Phylogram of Blechnaceae. Maximum likelihood bootstrap values are indicated at the branches; stars indicate bootstrap percentages $\geq$90%. Alternative generic concepts are noted at the right side of the figure. Genera are recorded according to the concepts shown in Smith et al. (2006), whereas deviating generic segregates recognized by Ching (1978) are given in parentheses. The newly included genus *Blechnum* (*Blechnidium*) *melanopus* is shown in red. To distinguish specimens, each is given with either a DNA laboratory number (e.g., FPXXXXXX) or the GenBank accession of one of the three genes included per specimen (preferably *rbcL*).

databases such as GenBank to reduce the amount of sequences required to be newly generated. However, this approach is challenged by the incompleteness of the accessible data in respect to taxon sampling (missing taxa) and the DNA regions used. Thus, macro-evolutionary studies often require the identification of major gaps that need to be addressed by generating new DNA sequence data. For example, six (about 2.2%) out of the 268 genera recognized as occurring in China (Table S1) were still lacking DNA sequence data, of which three genera (*Blechnidium*, *Saxiglossum*, and *Sinephropteris*) were added here for the first time (Tables 1, S1). Thus, three genera have still not been sampled, namely *Craspedosorus*,

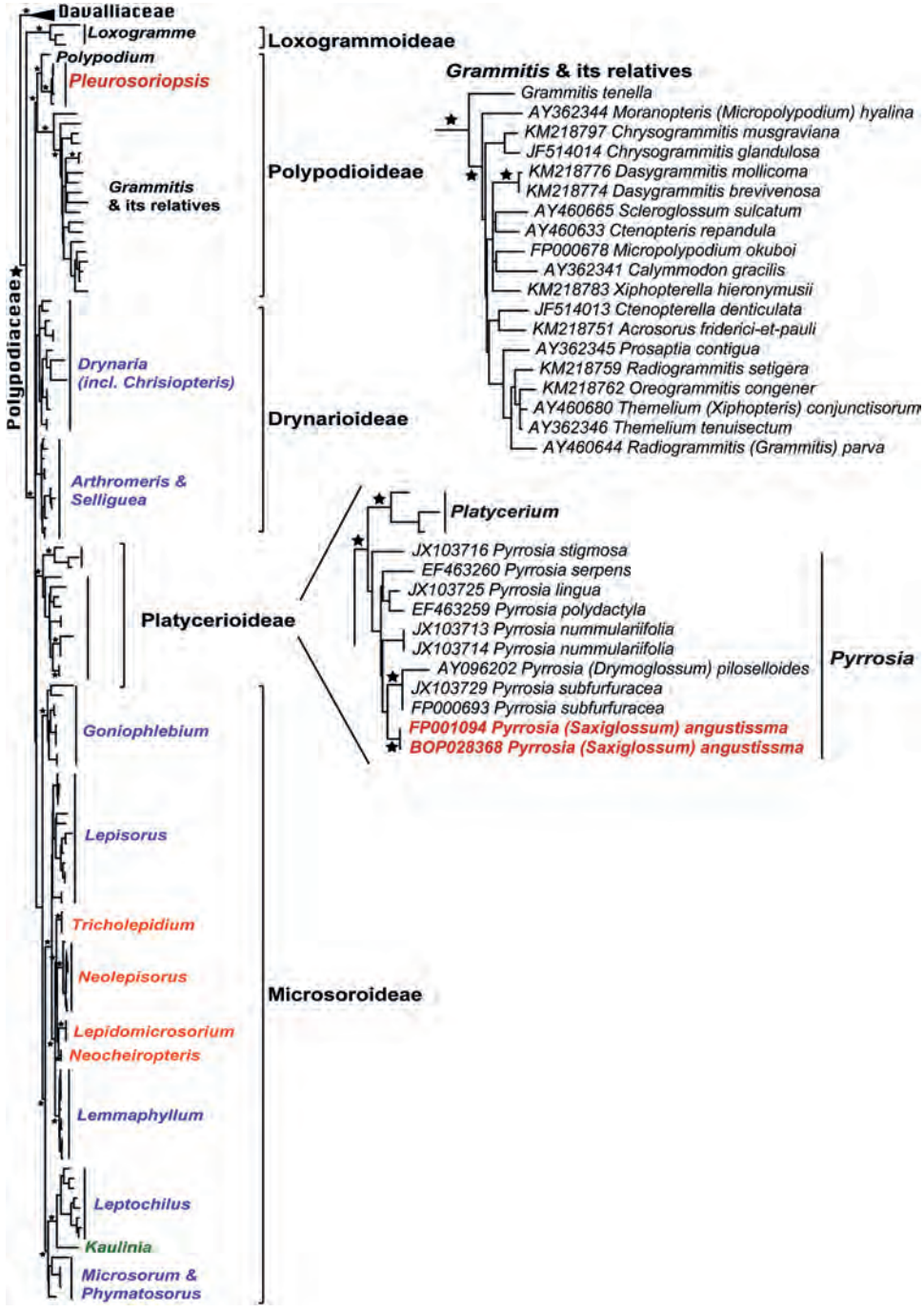


Fig. 4. Summary phylogeny for polygrammoid ferns. The five subfamilies are marked for convenience (Zhang et al., 2013c). The clade of subfamily Platycerioideae is shown on the right side to illustrate the position of the newly sampled *Pyrrosia* (*Saxiglossum*) *angustissma*, marked in red. Similarly, the grammitid clade (= *Grammitis* and its relatives) is shown on the right side to illustrate the newly established generic classification (see Sundue et al., 2014). The monotypic genus *Pleurosoriopsis* is marked in red because sequences of four new specimens confirmed the sister relationship to *Polypodium*. Genera names (with the exception of *Grammitis* and its relatives) follow Zhang et al. (2013c). Genera shown in purple comprise more than one genus recognized by Ching (1978), whereas genera in red were recognized by Ching. The genus *Kaulinia* (green) was only introduced recently (Zhang et al., 2013c).

Emodiopteris, and *Himalayopteris*. The monotypic thelypterioid *Craspedosorus* has only been reported from two isolated locations in northeast Yunnan, and the dennstaedtoid *Emodiopteris* has not been widely accepted since its introduction (see Kramer, 1990e; Yan et al., 2013). The polygrammoid *Himalayopteris* was only recently introduced (Shao & Lu, 2011) but requires confirmation in the context of the phylogeny of Drynarioideae (Fig. 4; see Schneider et al., 2010).

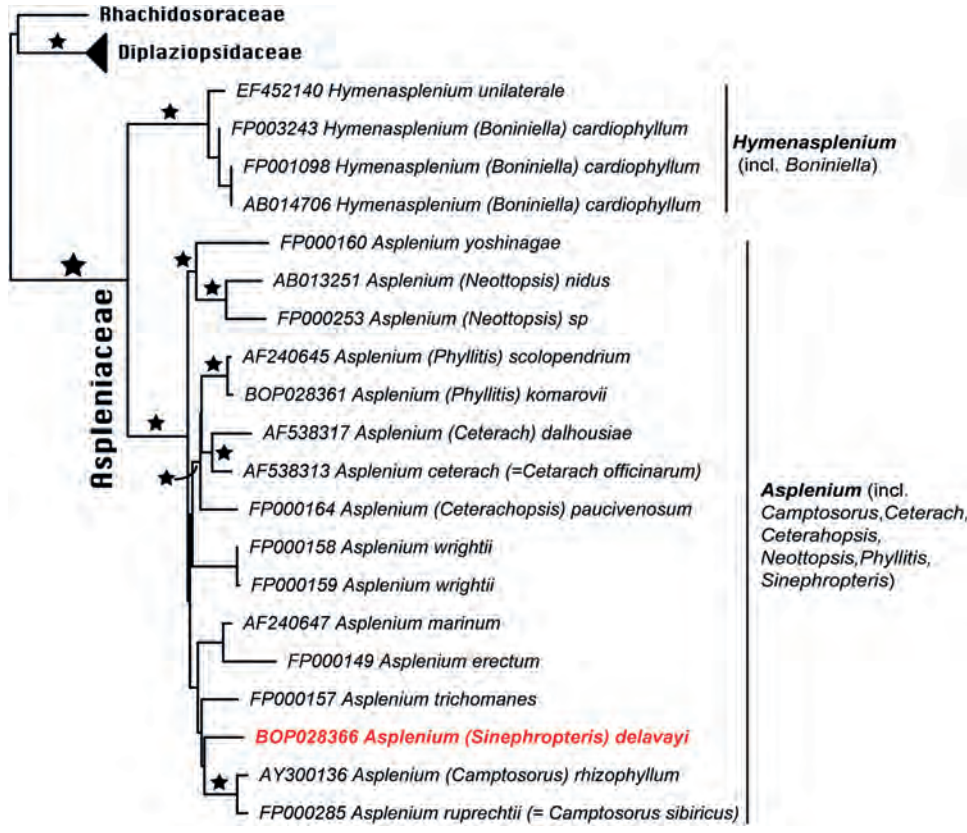


Fig. 5. Phylogram of Aspleniaceae. The position of the newly sampled *Asplenium (Sinephropteris) delavayi* is indicated in red. Generic concept and species treatment follows Schneider et al. (2004a) and Lin & Viane (2013), respectively, corresponding generic segregates recognized by Ching (1978) are given in parentheses.

In addition to the incompleteness of taxonomic and molecular markers, two other issues are considered as major challenges. The first is errors existing in the accessible DNA sequences, caused by a low but existing percentage of erroneously determined nucleotides (Wesche et al., 2004; Shen et al., 2013). Although the error rate of the Sanger sequencing methodology has decreased since its introduction, a low rate of errors is likely maintained as a result of the impact of DNA quality on the sequencing quality, lack of standards in laboratories amplifying and sequencing DNA, or human error. Addressing this issue is rather difficult because most available DNA sequences were not submitted together with records of sequencing quality as required for DNA barcodes (Benson et al., 2012). The identification of plant material is also one of the most serious sources of error, because some sequences have been generated based on misidentified materials or undetected contaminations (human error). This issue can be addressed if several sequences of the same species were obtained from different specimens and preferably in different laboratories. However, conflicts between taxonomy and reconstructed phylogenetic relationships may also be caused by other biological processes such as hybridization (Chang et al., 2013), chloroplast capture (Wang et al., 2012), or other factors (Bauret et al., 2015). In this study, these issues were addressed by applying a strategy that comprised two steps. First, a phylogenetic hypothesis was obtained using multiple specimens per species by integrating newly generated sequences and the existing sequence data accessible in an open source database (GenBank). Then the dataset was reduced by excluding specimens with odd relationships or long branches. The second step was carried out in the context of our current understanding of the taxonomy of the species included.

Overall phylogeny of extant ferns and lycophytes

Given the fact that this study is based on chloroplast genome regions used in the majority of previous studies, it is not a surprise that the recovered phylogenetic relationships (Fig. 2) are highly consistent with previously reported phylogenetic relationships of ferns and lycophytes (Pryer et al., 2001; Schneider et al., 2004b; Schuettpelz & Pryer, 2007; Rai & Graham, 2010; Lehtonen, 2011; Fiz-Palacios et al., 2011). However, this study improved the resolution of several poorly sampled clades and/or uncertainties in which sufficient information was previously missing, including the already mentioned putative segregates.

Classifications consistent with modern phylogeny

The comparison of Ching's classification (Ching, 1978) with the phylogenetic relationships recovered in the current study, as well as post-cladistic classifications (e.g., Smith et al., 2006; Liu et al., 2008; Christenhusz et al., 2011; Zhang et al., 2013c), determined some remarkable patterns, and many higher taxonomic concepts were confirmed. For example,

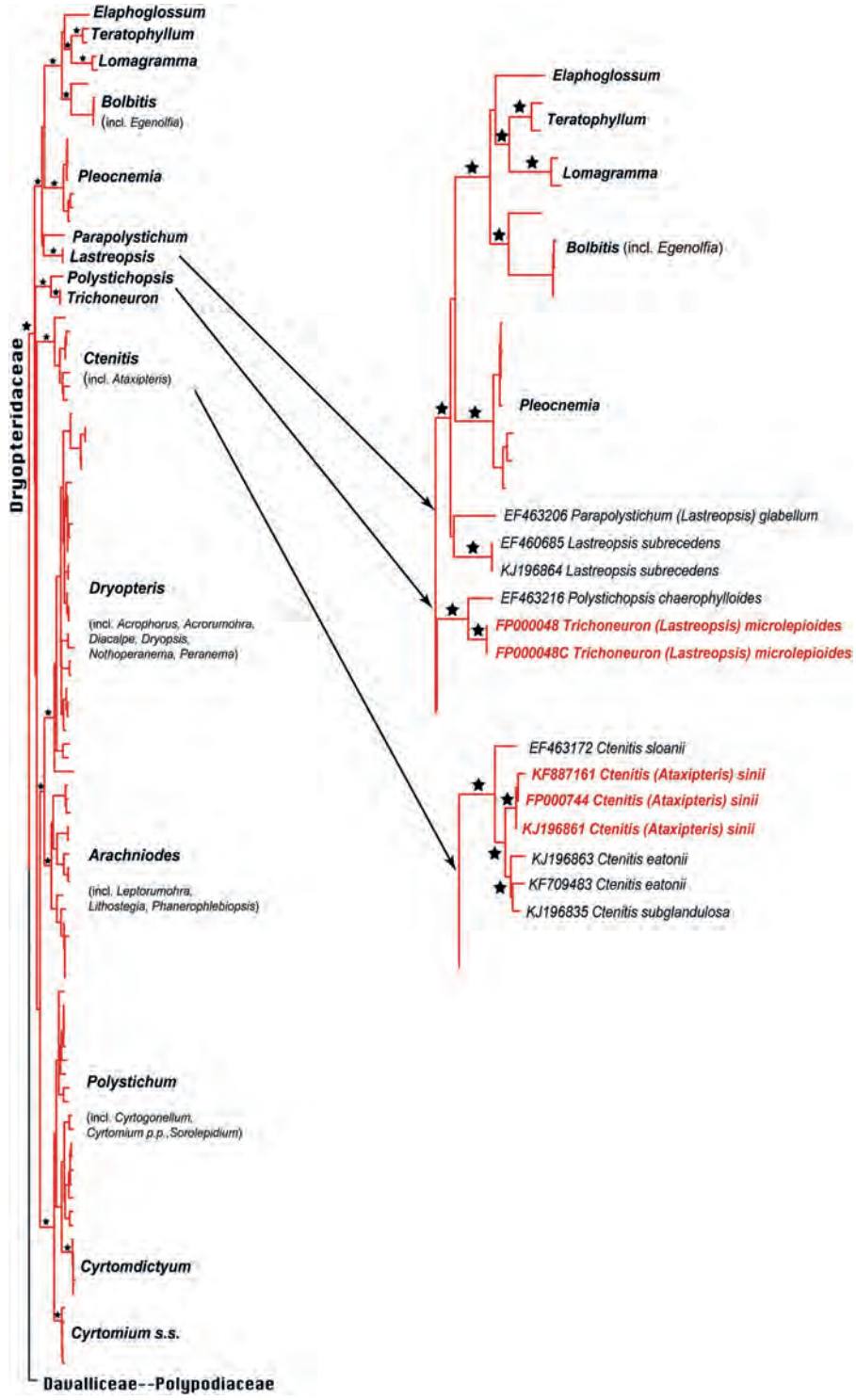


Fig. 6. Phylogram of Dryopteridaceae. Species names are not shown with the exception of lastreopioid ferns. Generic concept and species treatment follows Zhang et al. (2013c) with modifications according to Labiak et al. (2014), with synonyms in brackets.

the following orders are accepted now as defined by Ching: Equisetales, Isoëtales, Lycopodiales, Marattiales, Ophioglossales, Osmundales, Psilotales, and Selaginellales (Ching, 1978; Smith et al., 2006; Liu et al., 2008; Christenhusz et al., 2011; Zhang et al., 2013c). In contrast, the definition of the orders within the leptosporangiate ferns was substantially changed as the results of the replacement of a broadly defined order Polypodiales by six orders namely Cyatheales, Gleicheniales, Hymenophyllales, Salviniales, and Schizaeales besides the now narrowly defined Polypodiales (Smith et al., 2006;

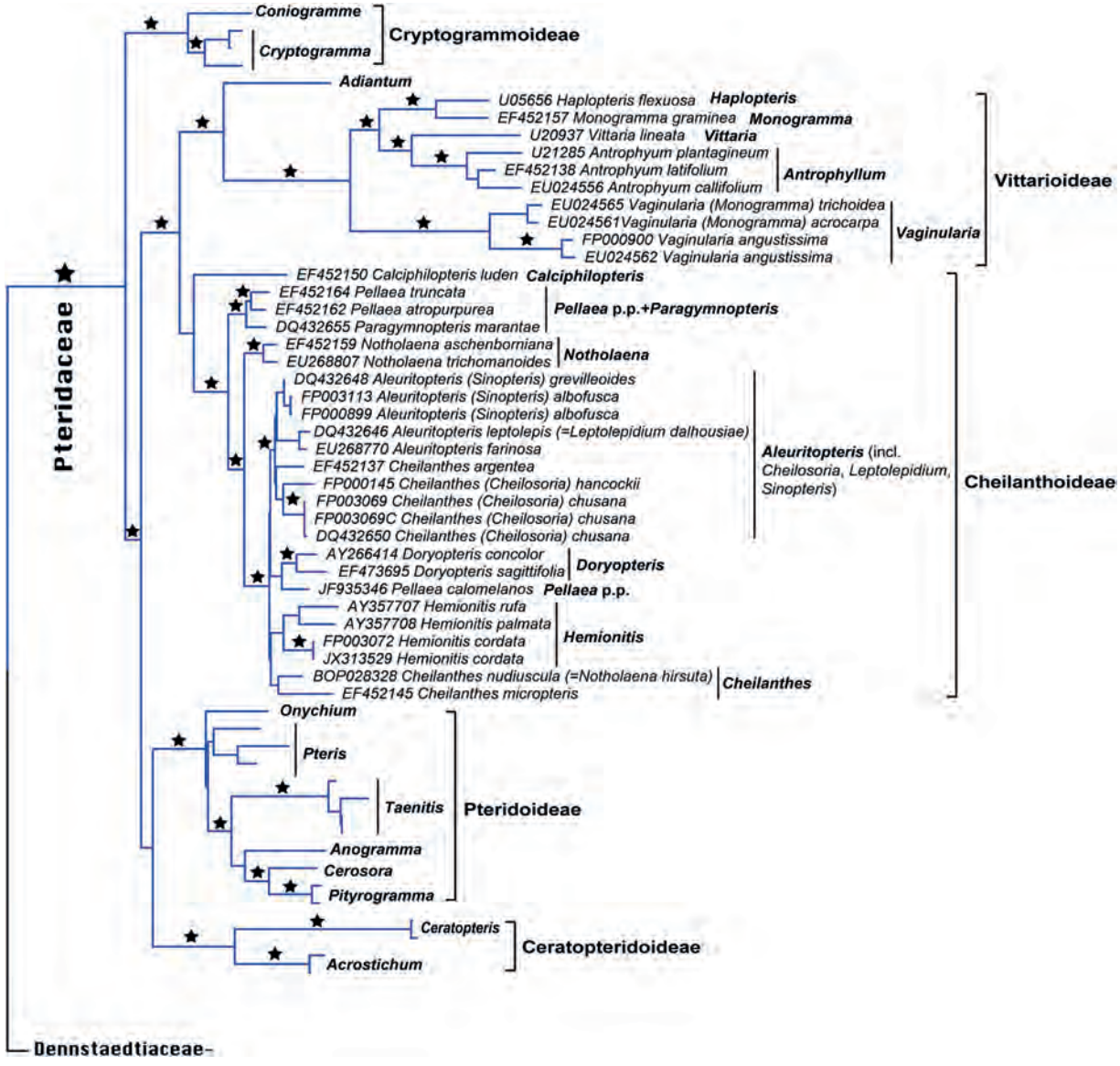


Fig. 7. Phylogenetic relationships within the family Pteridaceae. Detail of three subfamilies Ceratopteridoideae, Cryptogrammoideae, and Pteridoideae are collapsed, whereas the detailed relationships of the species are shown within the subfamilies Cheilanthoideae and Vittarioideae. Generic concept and species treatment follows Zhang et al. (2013b), with Ching's concept (1978) is shown in parentheses.

Liu et al., 2008; Christenhusz et al., 2011; Zhang et al., 2013c; Zhang & Gilbert, 2015).

Approximately half of the families (36 out of 64, Table 3) of ferns and lycophytes recognized by pre-cladistic classifications (e.g., Ching, 1978) are supported by molecular evidence (Fig. 2). Similar observations were reported by Zhang & Gilbert (2015). Among them, 20 families (including Dicksoniaceae, substituted with Cibotiaceae, and Didymochlaenaceae, newly recorded in China) have identical delimitations. These families have been widely accepted in the past based on morphological homogeneity (see Table 2, and table S1 of Zhang & Gilbert, 2015). Another 16 families are also supported by molecular data; however, their circumscription has changed, by either more inclusive definitions, such as Polypodiaceae comprising Drynariaceae, Grammitidaceae, Gymnogrammitidaceae, Loxogrammaceae, Pleurosoriopsidaceae, and Platyceriaceae, and Pteridaceae comprising Acrostichaceae, Adiantaceae, Antrophyaceae, Hemionitidaceae, Parkeriaceae, Sinopteridaceae, and Vittariaceae, or replacement by narrower definitions as result of recognition of additional families or exclusion of members, such as Athyriaceae and Tectariaceae. Phylogenetic results support Ching's segregation of Athyriaceae, Onocleaceae, and Tectariaceae rather than the widely accepted concept merging these ferns into a broadly defined Dryopteridaceae (e.g., Kramer, 1990a). For example, Tectariaceae as defined in Ching's classification (Ching, 1978) are a natural group comprising *Pteridrys* and several segregates from *Tectaria*, whereas some genera, like *Ctenitis*, *Lastreopsis*,

Table 3 Familial concepts of Chinese ferns and lycophytes in different classifications

Ching (1978)	Smith et al. (2006)	Christenhusz et al. (2011)	Zhang et al. (2013c)/This study
Acrostichaceae	Pteridaceae	Pteridaceae	Pteridaceae
Adiantaceae	Pteridaceae	Pteridaceae	Pteridaceae
Angiopteridaceae	Marattiaceae	Marattiaceae	Marattiaceae
Antrophyaceae	Pteridaceae	Pteridaceae	Pteridaceae
Aspidiaceae	**Tectariaceae**	**Tectariaceae**	**Tectariaceae**
Aspleniaceae	**Aspleniaceae**	**Aspleniaceae**	**Aspleniaceae**
Athyriaceae	**Woodsiaceae**	**Athyriaceae**	**Athyriaceae**
Athyriaceae	Woodsiaceae	Cystopteridaceae	Cystopteridaceae
Athyriaceae	Woodsiaceae	Diplaziopsidaceae	Diplaziopsidaceae
Athyriaceae	Woodsiaceae	Rhachidosoraceae	Rhachidosoraceae
Azollaceae	Salviniaceae	Salviniaceae	Salviniaceae
Blechnaceae	**Blechnaceae**	**Blechnaceae**	**Blechnaceae**
Bolbitidaceae	Dryopteridaceae	Dryopteridaceae	Dryopteridaceae
Botrychiaceae	Ophioglossaceae	Ophioglossaceae	Ophioglossaceae
Cheiropleuriaceae	Dipteridaceae	Dipteridaceae	Dipteridaceae
Christenseniaceae	Marattiaceae	Marattiaceae	Marattiaceae
Cyatheaceae	**Cyatheaceae**	**Cyatheaceae**	**Cyatheaceae**
Davalliaceae	**Davalliaceae**	**Davalliaceae**	**Davalliaceae**
Dennstaedtiaceae	**Dennstaedtiaceae**	**Dennstaedtiaceae**	**Dennstaedtiaceae**
Dicksoniaceae	**Cibotiaceae**	**Cibotiaceae**	**Cibotiaceae**
Didymochlaenaceae[†]	Dryopteridaceae	Hypodematiaceae	**Didymochlaenaceae**
Dipteridaceae	**Dipteridaceae**	**Dipteridaceae**	**Dipteridaceae**
Drynariaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Dryopteridaceae	**Dryopteridaceae**	**Dryopteridaceae**	**Dryopteridaceae**
Elaphoglossaceae	Dryopteridaceae	Dryopteridaceae	Dryopteridaceae
Equisetaceae	**Equisetaceae**	**Equisetaceae**	**Equisetaceae**
Gleicheniaceae	**Gleicheniaceae**	**Gleicheniaceae**	**Gleicheniaceae**
Grammitaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Gymnogrammitidaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Helminthostachyaceae	Ophioglossaceae	Ophioglossaceae	Ophioglossaceae
Hemionitidaceae	Pteridaceae	Pteridaceae	Pteridaceae
Huperziaceae	Lycopodiaceae	Lycopodiaceae	Lycopodiaceae
Hymenophyllaceae	**Hymenophyllaceae**	**Hymenophyllaceae**	**Hymenophyllaceae**
Hypodematiaceae	**Dryopteridaceae**	**Hypodematiaceae**	**Hypodematiaceae**
Hypolepidaceae	Dennstaedtiaceae	Dennstaedtiaceae	Dennstaedtiaceae
Isoëtaceae	**Isoëtaceae**	**Isoëtaceae**	**Isoëtaceae**
Lindsaeaceae	**Lindsaeaceae**	**Lindsaeaceae**	**Lindsaeaceae**
Lomariopsidaceae	**Lomariopsidaceae**	**Lomariopsidaceae**	**Lomariopsidaceae**
Loxogrammaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Lycopodiaceae	**Lycopodiaceae**	**Lycopodiaceae**	**Lycopodiaceae**
Lygodiaceae	**Lygodiaceae**	**Lygodiaceae**	**Lygodiaceae**
Marattiaceae	**Marattiaceae**	**Marattiaceae**	**Marattiaceae**
Marsileaceae	**Marsileaceae**	**Marsileaceae**	**Marsileaceae**
Monachosoraceae	Dennstaedtiaceae	Dennstaedtiaceae	Dennstaedtiaceae
Nephrolepidaceae	Lomariopsidaceae	Nephrolepidaceae	**Nephrolepidaceae**
Nephrolepidaceae	Lomariopsidaceae	Nephrolepidaceae	**Arthropteridaceae**
Oleandraceae	**Oleandraceae**	**Oleandraceae**	**Oleandraceae**
Onocleaceae	**Onocleaceae**	**Onocleaceae**	**Onocleaceae**
Ophioglossaceae	**Ophioglossaceae**	**Ophioglossaceae**	**Ophioglossaceae**
Osmundaceae	**Osmundaceae**	**Osmundaceae**	**Osmundaceae**
Parkeriaceae	Pteridaceae	Pteridaceae	Pteridaceae
Peranemaceae	Dryopteridaceae	Dryopteridaceae	Dryopteridaceae
Plagiogyriaceae	**Plagiogyriaceae**	**Plagiogyriaceae**	**Plagiogyriaceae**
Platyceriaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Pleurosoriopsidaceae	Polypodiaceae	Polypodiaceae	Polypodiaceae
Polypodiaceae	**Polypodiaceae**	**Polypodiaceae**	**Polypodiaceae**

Continued

Table 3 Continued

Ching (1978)	Smith et al. (2006)	Christenhusz et al. (2011)	Zhang et al. (2013c)/This study
Psilotaceae	**Psilotaceae**	**Psilotaceae**	**Psilotaceae**
Pteridaceae	**Pteridaceae**	**Pteridaceae**	**Pteridaceae**
Pteridiaceae	Dennstaedtiaceae	Dennstaedtiaceae	Dennstaedtiaceae
Salviniaceae	**Salviniaceae**	**Salviniaceae**	**Salviniaceae**
Schizaeaceae	**Schizaeaceae**	**Schizaeaceae**	**Schizaeaceae**
Selaginellaceae	**Selaginellaceae**	**Selaginellaceae**	**Selaginellaceae**
Sinopteridaceae	Pteridaceae	Pteridaceae	Pteridaceae
Stenochlaenaceae	Blechnaceae	Blechnaceae	Blechnaceae
Taenitidaceae	Pteridaceae	Pteridaceae	Pteridaceae
Thelypteridaceae	**Thelypteridaceae**	**Thelypteridaceae**	**Thelypteridaceae**
Vittariaceae	Pteridaceae	Pteridaceae	Pteridaceae
Woodsiaceae	**Woodsiaceae**	**Woodsiaceae**	**Woodsiaceae**

Families accepted in phylogeny-based classifications are shown in bold. †Didymochlaenaceae was recognized in Ching (1940).

and *Pleocnemia*, were transferred to Dryopteridaceae (Liu et al., 2007a, 2007c, 2014).

The Dryopteridaceae constitute a further example of a taxon recognized in pre-cladistic classifications that was re-defined to incorporate the results of phylogenetic studies. It is especially important to note that Ching's definition of Dryopteridaceae looks much more similar to the current concept (Smith et al., 2006; Liu et al., 2007c, 2008, 2016; Zhang et al., 2013c) than other broadly defined concepts, such as Copeland's Aspidiaceae (Copeland, 1947) or Dryopteridaceae according to Tryon & Tryon (1982) and Kramer (1990a). The family has been expanded by the incorporation of three families (Bolbitidaceae, Elaphoglossaceae, and Peranemaceae) and several genera considered by Ching (1978) as members of his Aspidiaceae (= Tectariaceae) such as *Ctenitis*, *Dryopsis*, *Lastreopsis*, and *Pleocnemia* (Fig. 6; see also Li & Lu, 2006; Liu et al., 2007c, 2008, 2014, 2016). It is worth mentioning that the family includes the enigmatic genus *Trichoneuron* (Fig. 6, see detail in Liu et al., 2016), which was tentatively placed in Thelypteridaceae by Ching (1965, 1978) or treated as synonym of *Lastreopsis* (Shing et al., 1999; Chu & He, 2000; Dong & Christenhusz, 2013). This genus was only recently recorded in Vietnam and reinstated as an independent genus (Liu et al., 2016). Only one exception recognized by Ching, the genus *Cyclopeltis*, is moved to another family (Smith et al., 2006; Liu et al., 2007c, 2008; Schuettpelz & Pryer, 2007; Christenhusz et al., 2011). In this context, it is also interesting to note that Ching (1978) recognized the small family Hypodematiaceae including only one genus whereas current classifications confirmed this family with the expansion from one to two genera as a result of the transfer of *Leucostegia* from Davalliaceae to Hypodematiaceae (Tsutsumi & Kato, 2006; Liu et al., 2007c, 2008; Christenhusz et al., 2011; Zhang et al., 2013c).

The monotypic Didymochlaenaceae are an example of a family new to China as a result of newly recorded fern species (Tan et al., 2015). It includes the single pantropical species *Didymochlaena truncatula*, which is distinguished from any other extant fern by the combination of elliptic-oblong sori and dimidiate pinnules. The genus was seen as a member of Dryopteridaceae until recently (see detail in Zhang & Zhang, 2015), although Ching (1940) considered the segregation of this family based on his integrative approach to fern systematics.

Another example of the consistency of Ching's classification and the phylogenetic hypothesis is provided by homosporous lycophytes (Fig. 8). Ching (1978) proposed two families instead of the widely accepted one family concept (Øllgaard,

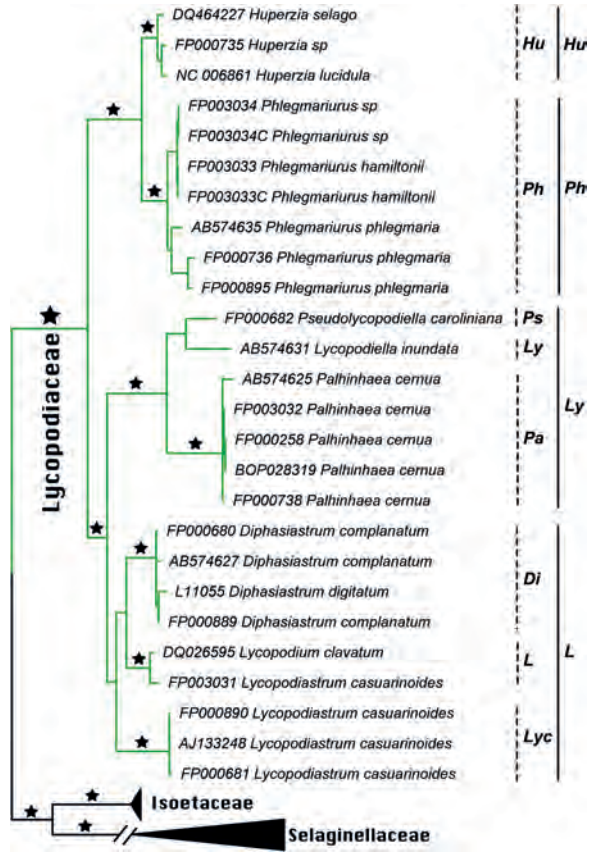


Fig. 8. Phylogeny of Lycopodiaceae. The classification follows Ching (1978); alternative genera concepts are shown at right. The branch length of sister group Selaginellaceae is shortened as indicated by "//". *Di*, *Diphasiastrum*; *Hu*, *Huperzia*; *L*, *Lycopodium*; *Ly*, *Lycopodiella*; *Lyc*, *Lycopodiastrum*; *Pa*, *Palhinhaea*; *Ph*, *Phlegmariurus*; *Ps*, *Pseudolycopodiella*.

1990; Christenhusz et al., 2011). The segregation of Huperziaceae from Lycopodiaceae is consistent with the phylogeny (Fig. 8) but the separation of the two clades may be arguably better represented by the recognition of subfamilies. The separation of two lineages of homosporous lycophytes is supported by differences in the branching pattern and arrangement of fertile leaves. More importantly, the generic classification of the group is still controversial. The *Huperzia* clade (Fig. 8) comprises two lineages that correspond to Ching's concept of two genera in his Huperziaceae, *Huperzia* and *Phlegmariurus* (Ching, 1978; Øllgaard, 2012; Zhang & Iwatsuki, 2013). The latter is widely treated as a member of *Huperzia* (Christenhusz et al., 2011; Zhang et al., 2013c) but some arguments—especially the separation of the enigmatic *Phylloglossum* (Øllgaard, 1990)—support the recognition of the two clades as separate genera. A similar pattern was found in the *Lycopodium* clade, namely Lycopodiaceae by Ching (1978). The inclusion of *Palhinhaea* in *Lycopodium* as treated in FOC (Zhang & Iwatsuki, 2013) was not supported (Fig. 8), and instead *Palhinhaea* and *Pseudolycopodiella* were part of the *Lycopodiella* lineage (Fig. 8, see also Wikström & Kenrick, 2001). This lineage may be either treated as a single genus *Lycopodiella* (Øllgaard, 1990) or divided into three genera, namely *Palhinhaea, Lycopodiella*, and *Pseudolycopodiella* (Ching, 1978; Zhang & Kung, 2000; Øllgaard, 2012). Alternative solutions also exist in the group consisting of *Diphasiastrum, Lycopodiastrum*, and *Lycopodium* (Fig. 8), by either accepting a broadly defined *Lycopodium* (Øllgaard, 1990) or recognizing two to three genera. However, a denser sampling is required to resolve these issues in homosporous lycophytes.

The filmy ferns (Hymenophyllales) are another group of interest. The obtained phylogeny (results not shown) is highly consistent with the phylogenetic hypotheses obtained in previous studies (e.g., Ebihara et al., 2006, 2007; Hennequin et al., 2010). At the generic level, the post-cladistic generic classification of filmy ferns (Ebihara et al., 2006) differs substantially from Ching's arrangement (Ching, 1978). However, a closer look reveals an astonishing consistency between Ching's classification and the post-cladistic classification. Most of his genera are now recognized either as genera, subgenera, or sections. One exception is the genus *Meringium*, which is now recognized as a synonym of *Hymenophyllum* subgenus *Hymenophyllum* (Ebihara et al., 2006; Hennequin et al., 2010).

A further example is the recognition of the vittarioid genus *Vaginularia*. This genus comprises several species of epiphytes found in tropical climate and has frequently been treated as a synonym of *Monogramma* based on the shared reduction of the lamina complexity (Kramer, 1990b; Smith et al., 2006; Zhang & Gilbert, 2013). Consistent with previously published evidence (Lindsay, 2003; Ruhfel et al., 2008), *Vaginularia* was supported as a distinct clade separated from *Monogramma*, and the latter was shown to be embedded in *Haplopteris* (Fig. 7). Thus, the reported evidence supports the recognition of *Vaginularia* to be separated from *Monogramma* (Ching, 1978; Lindsay, 2003), with the true *Monogramma* not occurring in China.

Classifications inconsistent with modern phylogeny

As mentioned above, despite a general agreement between Ching's classification(s) and the obtained phylogenetic hypotheses, several differences are found between his treatment and the current phylogeny-based classification. At the family level, 28 Chinese pteridophyte families recognized by Ching's classification (Ching, 1978) are not recognized and/or supported by molecular evidence. These families are now considered to be part of broadly defined families (Table 3; see also table S1 of Zhang & Gilbert, 2015), for example, Adiantaceae in Pteridaceae (Schuettpelz et al., 2007), Bolbitidaceae in Dryopteridaceae (Liu et al., 2007c; Zhang et al., 2013c), Botrychiaceae in Ophioglossaceae (Hauk et al., 2003), and Loxogrammaceae in Polypodiaceae (Schneider et al., 2004c); some of them correspond to well-supported phylogenetic lineages, such as Botrychiaceae to the *Botrychium* clade of the Ophioglossaceae (Hauk et al., 2003; Williams & Waller, 2012; Shinohara et al., 2013). Dryopteridaceae, Polypodiaceae, and Pteridaceae are three examples where circumscription has been expanded in both species number and morphological disparity compared to Ching's treatments. They now comprise genera previously recognized either as segregate families or as members of other families (Table 2, see also table S1 of Zhang & Gilbert, 2015). Athyriaceae are arguably one of the opposite examples. The family as treated by Ching (1978) now comprises four independent families, namely Athyriaceae s.s., Cystopteridaceae, Diplaziopsidaceae, and Rhachidosoraceae (Christenhusz et al., 2011; Rothfels et al., 2012; Zhang, 2012; Zhang et al., 2013c).

Asplenium and *Dryopteris* are the two most species-rich fern genera occurring in China. It is therefore not unexpected that some disagreements exist about the definition of these two genera. Current classifications recognize the family Aspleniaceae in the same circumscription as Ching (1978) but accept only two separate genera, *Asplenium* and *Hymenasplenium* (Schneider et al., 2004a; Christenhusz et al., 2011; Lin & Viane, 2013; Zhang et al., 2013c). The segregates *Camptosorus*, *Ceterach*, *Ceterachopsis*, and *Neottopteris* were previously shown to be nested in *Asplenium* (e.g., Murakami, 1995; Schneider et al., 2004a), whereas the segregate *Sinephropteris* was confirmed in this study to belong to *Asplenium* (Fig. 5). The only segregate accepted by Ching (1978) not belonging to *Asplenium*, the genus *Boniniella* is now accepted as a synonym of *Hymenasplenium* (Murakami, 1995; Schneider et al., 2004a). Similar to the expansion of *Asplenium*, the genus *Dryopteris* as accepted in current treatment (Zhang & Zhang, 2012) includes six segregates that were recognized not only as separate genera but as members of different families by Ching (1978), namely *Acrorumohra* and *Nothoperanema* as Dryopteridaceae, *Acrophorus*, *Diacalpe*, and *Peranema* as Peranemaceae, and *Dryopsis* as Tectariaceae. The inclusion of these formerly separated genera resulted in the expansion of the morphological disparity of both the family Dryopteridaceae and the genus *Dryopteris* by incorporating species with unique morphological characters, such as the sorus of the former segregate *Peranema*. The genus *Polystichum* is another example of genera with a circumscription that was expanded by including the segregates of *Cyrtogonellum, Cyrtomidictyum*, part *Cyrtomium* members (*Cyrtomium* subseri. *Balansana*), and *Sorolepidium* (Li & Lu, 2006; Liu et al., 2007b; Lu et al., 2007; Zhang & Barrington, 2013) when compared with Ching's concept (Ching, 1978). The monophyly of *Polystichum* s.l. (as by Zhang & Barrington, 2013) is not

supported in the current phylogeny (Fig. 6) and alternative treatment may need to be considered to achieve a natural classification of these ferns (Liu et al., 2010).

Cyclosorus, *Davallia*, *Diplazium*, and *Lepisorus* are further examples of genera that underwent substantial changes in their circumscription as the result of the improved understanding of the phylogenetic relationships of these ferns (Sano et al., 2000; Smith et al., 2006; Tsutsumi et al., 2008; Wang et al., 2010; He & Zhang, 2012; Liu & Schneider, 2013; Wei et al., 2013). However, the newly obtained insights into the phylogeny of ferns may not always correspond to the establishment of broadly defined taxa but may also result in the acceptance of smaller, biologically meaningful taxa as illustrated in the studies on the polygrammoid genus *Grammitis* (detailed discussion below).

New names, name changes, and/or placement of taxa since Ching

Several taxa discussed here were not known to occur in China during Ching's lifetime. These taxa can be classified into four categories. The first group comprises genera only recognized in recent years, such as *Calciphilopteris* (Yesilyurt & Schneider, 2010), *Caobangia* (Smith & Zhang, 2002), and *Osmolindsaea* (Lehtonen et al., 2010). The second group comprises genera previously unknown to occur in China, such as *Argyrochosma* (Wang et al., 2015), *Didymochlaena* (Tan et al., 2015), *Didymoglossum* (Chu & Zhou, 1994), and *Teratophyllum* (Dong & Zhang, 2005). The third group of genera is formed of those only now recognized to form rather distinct lineages, such as the genus *Himalayopteris* (Shao & Lu, 2011) and *Kaulinia* (Zhang et al., 2013c). The fourth and arguably the most complicated group comprises genera that were redefined based on phylogenetic or taxonomic studies, such as *Haplopteris* instead of *Vittaria*, *Odontosoria* rather than *Stenoloma*, *Ptisana* rather than *Marattia*, *Oreopteris* rather than *Lastrea*, and *Vandenboschia* rather than *Trichomanes* (e.g., Crane, 1997; Ebihara et al., 2006; Murdock, 2008; Lehtonen et al., 2010; see summary in Table 2). In comparison with Ching's treatment (1978), the names of several genera were changed based on the improved understanding of the typification of genera, such as *Paragymnopteris* rather than *Gymnopteris*, and *Parahemionitis* rather than *Hemionitis* (Zhang, 2003a; Zhang et al., 2013c). The grammitid ferns are a special case that needs to be discussed. These ferns were recognized as part of the Polypodiaceae (Schneider et al., 2004c; Smith et al., 2006) but their reclassification continues at the generic level (see Sundue et al., 2014). Six genera were recognized by Ching (1978) as occurring in China and later *Micropolypodium* was added to the list (Zhang, 2000); in contrast, the current classification recognizes up to 12 genera for China, namely *Calymmodon*, *Chrysogrammitis*, *Ctenopterella*, *Dasygrammitis*, *Micropolypodium*, *Oreogrammitis*, *Prosaptia*, *Radiogrammitis*, *Scleroglossum*, *Themelium*, *Tomophyllum*, and *Xiphopterella* (Moore & Parris, 2013). Only three of these genera recognized by Ching are still recorded in China (Moore & Parris, 2013).

A similar rearrangement of genera is expected in the cheilanthoid group given the already known fact that the broadly defined genera *Cheilanthes* and *Pellaea* (as in Tryon & Tryon, 1982; Tryon et al., 1990) are polyphyletic (e.g., Zhang et al., 2007; Eiserhardt et al., 2011). Some of the differences in the generic definition are caused by the interpretation of the type species. For example, the genus *Notholaena* is restricted to the New World (Rothfels et al., 2008), whereas Ching (1978) considered this genus to occur in China because of the incorrect acceptance of the Australian species *Notholaena distans* as the type of *Notholaena* (Yatskievych & Smith, 2003). Current classifications instead recognize the Chinese species previously assigned to *Notholaena* as part of the polyphyletic genus *Cheilanthes* and thus the species name *Notholaena chinensis* and *N. hirsuta* are now replaced with *Cheilanthes chinensis* and *C. nudiuscula*, respectively (Zhang & Yatskievych, 2013). Similarly, none of the species occurring in China treated as *Pellaea* belongs to the genus *Pellaea* (e.g., Kirkpartick et al., 2007; Eiserhardt et al., 2011; Wang et al., 2015) but the segregate *Paragymnopteris* is nested within the *Pellaea* clade (Zhang et al., 2007; Kirkpatrick et al., 2007; see also Fig. 7). Most Chinese cheilanthoid species belong to a single clade that includes all species sampled so far from the genera *Aleuritopteris* (including *Leptolepidium*), *Cheilosoria* (respectively *Cheilanthes*), *Sinopteris*, and some species treated as *Pellaea*, which was formerly recognized by Zhang et al. (2007). The clade includes also a small group of species previously treated as *Cheilanthes* occurring from Pakistan throughout Mediterranean Europe towards the Macaronesian islands. This group includes the type of the genus *Allosorus* (Eiserhardt et al., 2011), which is the oldest generic name available for the clade (Christenhusz, 2012). The present study recovered evidence for some phylogenetic structure within this clade that may lead to the recognition of several small genera comparable to the recent studies on Neotropical cheilanthoids such as *Gaga* (Li et al., 2012) and *Myriopteris* (Grusz & Windham, 2013).

New knowledge benefits from current large-scale sampling

Blechnaceae and Tectariaceae are two groups within the eupolypods that still require comprehensive study, although some recent progress has been made through phylogenetic investigations (e.g., Liu et al., 2007a; Ding et al., 2014; Perrie et al., 2014; Wang et al., 2014). Besides some clades of the pteridoid ferns, these two families are arguably the lineages with the most urgent need to be addressed in the context of generic classification (Smith et al., 2006; Liu et al., 2008; Christenhusz et al., 2011; Wang et al., 2013b).

With a critical sampling to include all putative segregates as well as *Stenochlaena*, Blechnaceae as defined by Ching (1978) were strongly supported as a monophylum including several well-supported clades (Fig. 3; Cranfill, 2001; Shepherd et al., 2007; Perrie et al., 2014). One clade comprised *Woodwardia* and the putative segregate *Chieniopteris*, whereas the other clade comprised several genera widely recognized, namely *Blechnum*, *Brainea*, and *Stenochlaena*, and three segregates recognized by Ching (1978), namely *Blechnidium*, *Diploblechnum*, and *Struthiopteris*. The segregation of *Chieniopteris* from *Woodwardia* as well as *Blechnidium*, *Diploblechnum*, and *Struthiopteris* from *Blechnum* as treated in FOC (Wang et al., 2013a) will generate a paraphyletic *Woodwardia* and *Blechnum*; however, the monophyly of the *Blechnum* s.l. (as in Fig. 3) was recognized yet not well supported in the current phylogeny. To uncover the natural generic classification of the family requires a dense taxon sampling with focus on the genus *Blechnum* (Perrie et al., 2014).

Tectariaceae are a pantropical family within eupolypods I. The delimitation of the family, especially the distinction from Dryopteridaceae, has long been controversial until phylogenetic studies provided convincing evidence for its segregation (Li & Lu, 2006; Liu et al., 2007a, 2007c, 2013, 2014; Schuettpelz & Pryer, 2007). However, the definition of the Tectariaceae is still not fully resolved, as illustrated by the exclusion of several genera including *Ctenitis, Dryopsis, Lastreopsis*, and *Pleocnemia* (Ching, 1978; see also tectarioid group in Kramer, 1990a). These genera are now recognized as part of the Dryopteridaceae (Li & Lu, 2006; Liu et al., 2007a, 2007c, 2013, 2014; Ding et al., 2014). As shown in Fig. 9 and earlier studies (Liu et al., 2007a; Ding et al., 2014; Wang et al., 2014), two clades were resolved in the family, *Pteridrys* and *Tectaria* s.l., the latter including segregates *Ctenitopsis* (=*Heterogonium* which is an older name according to Holttum, 1983), *Hemigramma*, and *Quercifilix*. Acceptance of these segregates will result in a paraphyletic genus *Tectaria*.

Among the three segregates that were included in the current large-scale phylogenetic sampling for the first time, *Saxiglossum* was shown to be nested in *Pyrrosia* (Fig. 4), confirming the conclusion achieved by comparative morphological analysis (see detail in Ravensberg & Hennipman, 1986). Similar to the former segregate *Drymoglossum*, *Saxiglossum* was recognized by overvaluing the reduction of the density of the stellate hairs on the lamina. The placement of two Asian endemic genera, *Gymnogrammitis* and *Pleurosoriopsis*, in the family Polypodiaceae is confirmed by the current independent and multiple accessions of molecular data (Fig. 4; Schneider et al., 2002, 2004c); both were formerly recognized as independent families by Ching (1978).

Some notes on the importance of Ching's contribution as reflected in the current phylogenetic hypothesis

As discussed in the introduction, Ching (1940) spearheaded the replacement of a broadly defined Polypodiaceae with

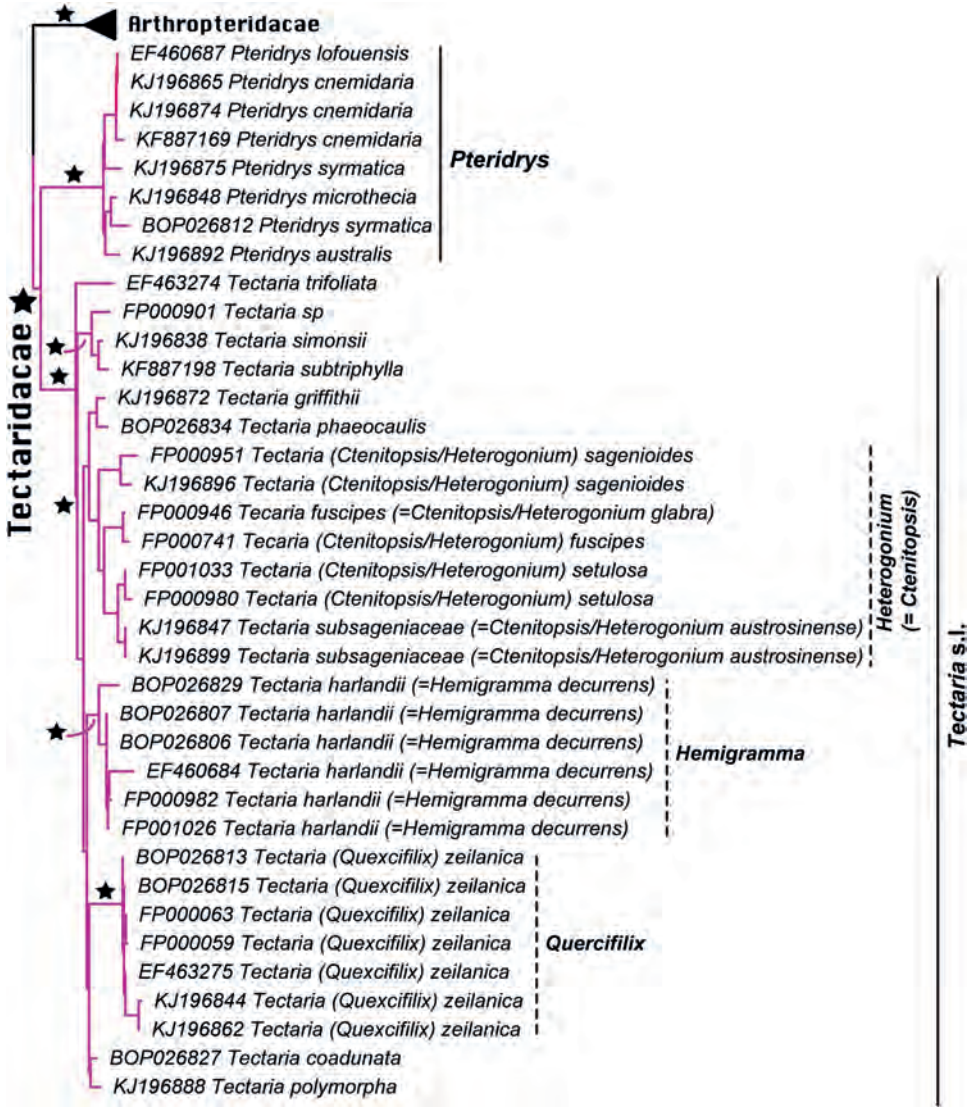


Fig. 9. Phylogeny of Tectariaceae. Generic concept and species treatment follows Xing et al. (2013), with synonyms in brackets.

narrower but more informative defined families. Phylogenetic studies confirmed the monophyly of most of the orders, half of the families, and a number of genera introduced by Ching (Tables 2, 3). Several families and genera first recognized by Ching have long been ignored; they have only recently been reconfirmed by molecular phylogenetic studies and are now accepted in recent classifications, e.g., Athyriaceae, Didymochlaenaceae, Dryopteridaceae, Hypodematiaceae, Tectariaceae, and Woodsiaceae. In turn, it is not surprising to see that several families defined by Ching (1978) were found to be paraphyletic and/or polyphyletic. Given the fact that paraphyly was necessarily seen as a major issue in the past, this is expected (Hörandl & Stuessy, 2010, 2014; Schmidt-Lebuhn, 2012). Despite these issues, Ching continues to have a strong influence on the classification and taxonomy of pteridophytes. The majority of current studies on ferns and lycophytes will continue to follow Ching by aiming to elucidate taxonomic units that are biologically meaningful as shown here in the context of phylogenetic evidence, despite some authors who argue to re-erect broadly defined taxa in the tradition of Hooker (Christenhusz & Chase, 2014).

Phylodiversity of Chinese ferns and lycophytes

With more than 2300 species, the pteridophyte flora of China is one of the most species-rich pteridophyte floras in the world and comprises a considerable proportion of the Asian diversity of ferns and lycophytes (Zhang, 2003b, 2012). This species richness is reflected by the occurrence of more than 151 genera in 40 families (Table S2; Zhang, 2012; Wu et al., 2013; Zhang et al., 2013c; Zhang & Zhang, 2015) out of 276 genera in 50 families occurring globally (Smith et al., 2006; Christenhusz et al., 2011; Liu et al., 2013; Zhang & Zhang, 2015). Thus, 55% of pteridophyte genera and 80% of pteridophyte families occur in China. Fern families absent from China are either monotypic or consist of only a few species, such as Culcitaceae (two species), Cystodiaceae (one species), Lonchitidaceae (two species), Loxomataceae (two species), Matoniaceae (four species), Metaxyaceae (two species), and Thyrsopteridaceae (one species), with the notable exceptions of the families Anemiaceae (ca. 100 species) and Dicksoniaceae (ca. 30 species). The latter occurs predominantly in the Southern Hemisphere with the exception of Hawaii and some occurrences in western parts of Malesia. The other exception, the family Anemiacae, has its diversity center in the Neotropics as well as some occurrences in Africa, Madagascar, and some islands in the Indian Ocean.

Similar trends are visible at the generic level. Approximately 55% of the currently accepted fern genera occurs in China. Genera absent from China show three biogeographic patterns. First, genera having a species diversity center in the Neotropics are either absent—such as *Anemia*, *Campyloneurum*, *Cyathea*, *Megalastrum*, *Lellingeria*, *Pecluma*, *Pleopeltis*, *Serpocaulon*, and *Terpsichore*—or species poor, such as *Elaphoglossum*. Second, genera with a species center in the Southern Hemisphere, especially New Zealand and Australia, are absent, such as *Gleichenia*. Finally, genera with a species center in the tropical climates of Malesia are absent, such as *Matonia*, *Phanerosorus*, and *Thylacopteris*. Absent genera are contributed by different families and orders, suggesting that the absence is not the result of a single evolutionary event. In this context, it is worth noting that the highest proportion of missing genera comprises those occurring exclusively or preferably in the Neotropics. Thus, the generic pattern reflects the widely recognized division between Paleotropical and Neotropical fern diversity. The main lineages of Polypodiaceae diversified in the Paleotropics, for example, loxogrammoids, drynarioids, selligueoids, and microsoroids, with the exception of the clade comprising *Polypodium* and relatives as well as grammitids (see Schneider et al., 2010). The latter group shows an initial diversification in the Neotropics followed by several colonization events leading to radiations in the Old World tropics (Schneider et al., 2010; Sundue et al., 2014). Similar patterns have now been documented for many lineages of ferns including Aspleniaceae (Schneider et al., 2004a), Athyriaceae (Wei et al., 2013, 2015), Dryopteridaceae (e.g., Sessa et al., 2012; Zhang et al., 2012; Labiak et al., 2014), and Pteridaceae (e.g., Lu et al., 2011; Schneider et al., 2013; Chao et al., 2014; Zhang et al., 2015).

Conclusion

As a result of the effort to sample all recorded families and genera of Chinese pteridophytes in a phylogenetic framework, this study allowed the re-evaluation of Ching's contributions to the classification of pteridophytes, and correspondingly resolved some uncertainties in the phylogeny of Chinese ferns and lycophytes. Among the 223 genera considered by Ching, 122 are recognized (three genera are now confirmed as not occurring in China) by the phylogeny. Some are substantially changed in the context of circumscription, nine are supported but replaced with new names, two (*Craspedosorus* and *Emodiopteris*) are still recognized as ambiguous and requiring further study, and 90 are now accepted as synonyms. Of the 45 new generic names introduced after Ching as either new taxa, new records or name replacements, 32 are recognized by the phylogeny. Among them, 10 are new name replacements, one (*Himalayopteris*) is ambiguous, and 12 are reduced to synonyms. Considering the current understanding and progress of phylogenetic relationships at familial and generic levels, this paper tentatively accepts 40 families and 151 genera of ferns and lycophytes occurring in China (Tables 2, 3, S2). This list comprises two newly recognized/reconfirmed families Arthropteridaceae and Didymochlaenaceae. Several monotypic genera, like *Craspedosorus* and *Emodiopteris*, are still treated as ambiguous in the current checklist, although they may be reduced to synonyms in studies exploring their phylogenetic relationships using DNA sequence data. Further changes are expected to be introduced in future studies on families that still require more exhaustive analyses, such as Blechnaceae, Polypodiaceae, Tectariaceae, and Thelypteridaceae.

Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (Grant No. 31300183), the Scientific Research Project of Shenzhen Urban Management Bureau (Grant No. 201312), and the State-Sponsored Overseas Academic Visiting Scholarship Program (Grant No.

201308440224). The author is grateful to An-Ming Lu, Zhi-Duan Chen, and Shou-Zhou Zhang for encouragement to prepare this paper, Harald Schneider (Natural History Museum, UK) for inspirational discussions, valuable suggestions, and comments to the manuscript. The author thanks the following colleagues for sharing materials (Xian-Chun Zhang, Bao-Dong Liu, Yue-Hong Yan, Yi-Han Zhang, Fa-Guo Wang, Ralf Knapp, Hui Shang, Ri-Hong Jiang, Wen-Bin Xu, and Ce-Hong Li), assistance with the analyses (Liang-Sheng Zhang, Yang Liu, Miao Sun, and Hong-Lei Li), and support in the laboratory (Ping Yang, Mei-Ping Wang, and En-Xu Liu).

References

Bauret L, Grall A, Senterre B, Rouhan G, Hennequin S, Ebihara A, Dubuisson J-Y. 2015. New circumscription of *Trichomanes cupressoides* Desvaux (Hymenophyllaceae), an endemic filmy fern from the Seychelles (Indian Ocean), and new insights into the genus *Abrodictyum* C. Presl in the western Indian Ocean. *Phytotaxa* 202: 1–14.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2012. GenBank. *Nucleic Acids Research* 41: D36–D42.

Chang Y-F, Li J, Lu S-G, Schneider H. 2013. Species diversity and reticulate evolution in the *Asplenium normale* complex (Aspleniaceae) in China and adjacent areas. *Taxon* 62: 673–687.

Chao Y-S, Rouhan G, Amoroso VB, Chiou W-L. 2014. Molecular phylogeny and biogeography of the fern genus *Pteris* (Pteridaceae). *Annals of Botany* 114: 109–124.

Chen ZD, Yang T, Lu LM, Lin L, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, Cao ZY, Liu HM, Wang XM, Wang W, Zhang JB, Meng Z, Cao W, Li JH, Wu SD, Zhao HL, Liu ZJ, Du ZY, Wang QF, Guo J, Tan XX, Su JX, Zhang LJ, Yang LL, Liao YY, Li MH, Zhang GQ, Chung SW, Zhang J, Xiang KL, Li RQ, Soltis DE, Soltis PS, Zhou SL, Ran JH, Wang XQ, Jin XH, Chen YS, Gao TG, Li JH, Zhang SZ, Lu AM, China Phylogeny Consortium. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54: 277–306.

Ching R-C. 1940. On natural classification of the "Polypodiaceae". *Sunyatsenia* 5: 201–268.

Ching R-C. 1954. Systematic arrangements of families and genera of Chinese pteridophytes with corresponding names in Chinese. *Acta Phytotaxonomica Sinica* 3: 93–99.

Ching R-C. 1965. Two new fern genera from China. *Acta Phytotaxonomica Sinica* 10: 115–122.

Ching R-C. 1978. The Chinese fern families and genera: Systematic arrangement and historical origin. *Acta Phytotaxonomica Sinica* 16: 1–19, 16–37.

Ching R-C, Fu S-H, Wang C-H, Shing G-H. 1959. Ophioglossaceae through Oleandraceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Ching R-C, Shing K-H. 1983. *Lepidomicrosorum* Ching et Shing, a new fern genus of Polypodiaceae from China. *Botanical Research: Contributions from the Institute of Botany, Academica Sinica* 1: 1–14.

Ching R-C, Shing K-H, Lin Y-X, WU S-H, Wu S-K. 1990. Pteridaceae through Parkeriaceae (Ceratopteridaceae). In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Ching R-C, Wang Z-R. 1982. On the systematic position of *Athyrium crenulato-serrulatum* Makino from north-eastern Asia. *Acta Phytotaxonomica Sinica* 20: 73–77.

Ching R-C, Wu S-K. 1980. *Platygyria* Ching et S.K. Wu, an unique new genus of the Polypodiaceae from China. *Acta Botanica Yunnanica* 2: 67–74.

Christ KH. 1897. *Die Farnkräuter der Erde*. Jena: Gustav Fischer Verlag.

Christenhusz MJM. 2012. Pteridaceae. In: Greuter W, von Raab-Straube E eds. *Euro+Med Nutulae*, 6. *Willdenowia* 42: 283–285.

Christenhusz MJM, Chase MW. 2014. Trends and concepts in fern classification. *Annals of Botany* 113: 571–594.

Christenhusz MJM, Zhang X-C, Schneider H. 2011. A linear sequence of extant families and genera of lycophytes and ferns. *Phytotaxa* 19: 7–54.

Christensen C. 1938. Filicinae. In: Verdoorn F ed. *Manual of Pteridology*. Hague: Nijhoff. 522–550.

Chu W-M, He Z-R. 2000. Taxonomic notes on some pteridophytes from Yunnan (II). *Acta Botanica Yunnanica* 22: 255–262.

Chu W-M, Wang Z-R, Zhang X-C, He Z-R, Hsieh Y-T. 1999. Antrophyaceae, Vittariaceae & Athyriaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Chu W-M, Zhou H-G. 1994. Some pteridophytes new to Hainan and China. *Acta Botanica Yunnanica* 16: 123–130.

Copeland EB. 1947. *Genera Filicum—the genera of ferns*. Massachusetts: The Chronica Botanica Company.

Crane EH. 1997. A revised circumscription of the genera of the fern family Vittariaceae. *Systematic Botany* 22: 509–517.

Cranfill RB. 2001. *Phylogenetic studies in the Polypodiales (Pteridophyta) with an emphasis on the family Blechnaceae*. PhD Dissertation. Berkeley: University of California.

Diels L. 1899/1900. Filicales. In: Engler A, Prantl K eds. *Die natürlichen pflanzenfamilien*. I4. Leipzig: Engelmann.

Ding H-H, Chao Y-S, Callado JR, Dong S-Y. 2014. Phylogeny and character evolution of the fern genus *Tectaria* (Tectariaceae) in the Old World inferred from chloroplast DNA sequences. *Molecular Phylogenetics and Evolution* 80: 66–78.

Dong S-Y, Christenhusz MJM. 2013. *Lastreopsis*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 628–629.

Dong S-Y, Zhang X-C. 2005. *Teratophyllum hainanense* (Lomariopsidaceae), a new species from Hainan Island, China. *Novon* 15: 104–108.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 1–15.

Ebihara A, Dubuisson J-Y, Iwatsuki K, Hennequin S, Ito M. 2006. A taxonomic revision of Hymenophyllaceae. *Blumea* 51: 221–280.

Ebihara A, Iwatsuki K, Hennequin S, Ito M, Dubuisson J-Y. 2007. A global molecular phylogeny of the fern genus *Trichomanes* (Hymenophyllaceae) with special reference to stem anatomy. *Botanical Journal of the Linnean Society* 155: 1–27.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Eiserhardt WL, Rohwer JG, Russell SJ, Yesilyurt JC, Schneider H. 2011. Evidence for radiations of cheilanthoid ferns in the Greater Cape Floristic region. *Taxon* 60: 1269–1283.

Feldberg K, Schneider H, Stadler T, Schäfer-Verwimp A, Schmidt AR, Heinrichs J. 2014. Epiphytic leafy liverworts diversified in angiosperm-dominated forests. *Scientific Reports* 4: 5974.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.

Fiz-Palacios O, Schneider H, Heinrichs J, Savolainen V. 2011. Diversification of land plants: Insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology* 11: 341.

Grusz AM, Windham MD. 2013. Toward a monophyletic *Cheilanthes*: The resurrection and recircumscription of *Myriopteris* (Pteridaceae). *PhytoKeys* 32: 49–64.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Hauk WD, Parks CR, Chase MW. 2003. Phylogenetic studies of Ophioglossaceae: Evidence from *rbcL* and *trnL-F* plastid DNA sequences and morphology. *Molecular Phylogenetics and Evolution* 28: 131–151.

He L-J, Zhang X-C. 2012. Exploring generic delimitation within the fern family Thelypteriaceae. *Molecular Phylogenetics and Evolution* 65: 757–764.

Hennequin S, Ebihara A, Dubuisson J-Y, Schneider H. 2010. Chromosome number evolution in *Hymenophyllum* (Hymenophyllaceae), with special reference to the subgenus *Hymenophyllum*. *Molecular Phylogenetics and Evolution* 55: 47–59.

Hennequin S, Kessler M, Lindsay S, Schneider H. 2014. Evolutionary patterns in the assembly of fern diversity on the oceanic Mascarene Islands. *Journal of Biogeography* 41: 1651–1663.

Holttum RE. 1947. A revised classification of leptosporangiate ferns. *Botanical Journal of the Linnean Society* 53: 123–258.

Holttum RE. 1983. The fern-genera *Tectaria, Heterogonium* and *Ctenitis* in the Mascarene Islands. *Kew Bulletin* 38: 107–130.

Hörandl E, Stuessy TF. 2010. Paraphyletic groups as natural units of biological classification. *Taxon* 59: 1641–1653.

Hörandl E, Stuessy TF. 2014. The importance of comprehensive phylogenetic (evolutionary) classification—a response to Schmidt-Lebuhn's commentary on paraphyletic taxa. *Cladistics* 30: 291–293.

Jiang R-H, Zhang X-C, Liu Y. 2011. *Asplenium cornutissimum* (Aspleniaceae), a new species from karst caves in Guangxi, China. *Brittonia* 63: 83–86.

Kirkpatrick REB. 2007. Investigating the monophyly of *Pellaea* (Pteridaceae) in the context of a phylogenetic analysis of cheilanthoid ferns. *Systematic Botany* 32: 504–518.

Kramer KU. 1990a. Dryopteridaceae. In: Kubitzki K ed. *The families and genera of vascular plants—Pteridophytes and gymnosperms*. Berlin: Springer. 1: 101–144.

Kramer KU. 1990b. *Monogramma*. In: Kubitzki K ed. *The families and genera of vascular plants—Pteridophytes and gymnosperms*. Berlin: Springer. 1: 276.

Kramer KU. 1990c. Notes on the higher level classification of the recent ferns. In: Kubitzki K ed. *The families and genera of vascular plants—Pteridophytes and gymnosperms*. Berlin: Springer. 1: 49–52.

Kramer KU. 1995. *Farne und Farnverwandte*. New York: Georg Thieme Verlag.

Kreier HP, Zhang X-C, Muth H, Schneider H. 2008. The microsoroid ferns: Inferring the relationships of a highly diverse lineage of paleotropical epiphytic ferns (Polypodiaceae, Polypodiopsida). *Molecular Phylogenetics and Evolution* 48: 1155–1167.

Kung H-S, Zhang L-B, Chu W-M, He Z-R, Hsieh Y-T. 2001. Dryopteridaceae (2). In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Labiak PH, Sundue M, Rouhan G, Hanks JG, Mickel JT, Moran RC. 2014. Phylogeny and historical biogeography of the lastreopsid ferns (Dryopteridaceae). *American Journal of Botany* 101: 1207–1228.

Laenen B, Shaw B, Schneider H, Goffinet B, Paradis E, Désamoré A, Heinrichs J, Villarreal JC, Gradstein SR, McDaniel SF, Long DG, Forrest LL, Hollingsworth ML, Crandall-Stotler B, Davis EC, Engel J, Konrat MV, Cooper ED, Patiño J, Cox CJ, Vanderpoorten A, Shaw AJ. 2014. Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts. *Nature Communications* 5: 6134.

Lehtonen S, Tuomisto H, Rouhan G, Christenhusz MJM. 2010. Phylogenetics and classification of the pantropical fern family Lindsaeaceae. *Botanical Journal of the Linnean Society* 163: 305–359.

Lehtonen S. 2011. Towards resolving the complete fern tree of life. *PLoS ONE* 6: e24851.

Li C-X, Lu S-G. 2006. Phylogenetic analysis of Dryopteridaceae based on chloroplast *rbcL* sequences. *Acta Phytotaxonomica Sinica* 44: 503–515.

Li F-W, Pryer KM, Windham MD. 2012. Gaga, a new fern genus segregated from *Cheilanthes* (Pteridaceae). *Systematic Botany* 37: 845–860.

Li Q-J. 1990. *Photinopteris* J. SM., a new recorded genus of Drynariaceae in China. *Acta Phytotaxonomica Sinica* 28: 412–413.

Lin Y-X, Viane R. 2013. Aspleniaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 267–316.

Lin Y-X, Zhang X-C, Shi L, Lu S-G. 2000. Dipteridaceae through Azollaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Lindsay S. 2003. Considerations for a revision of the fern family Vittariaceae for Flora Malesiana. *Telopea* 10: 99–112.

Liu H-M, He L-J, Schneider H. 2014. Towards the natural classification of tectarioid ferns: Confirming the phylogenetic relationships of *Pleocnemia* and *Pteridrys* (eupolypods I). *Journal of Systematics and Evolution* 52: 161–174.

Liu H-M, Jiang R-H, Guo J, Hovenkamp P, Perrie LR, Shepherd L, Hennequin S, Schneider H. 2013. Towards a phylogenetic classification of the climbing fern genus *Arthropteris*. *Taxon* 62: 688–700.

Liu H-M, Schneider H. 2013. Evidence supporting *Davallia canariensis* as a Late Miocene relict endemic to Macaronesia and Atlantic Europe. *Australian Systematic Botany* 26: 378–385.

Liu H-M, Wang L, Zhang X-C, Zeng H. 2008. Advances in the studies of lycophytes and monilophytes with reference to systematic arrangement of families distributed in China. *Acta Phytotaxonomica Sinica* 46: 808–829.

Liu H-M, Zhang X-C, Chen Z-D, Dong S-Y, Qiu Y-L. 2007a. Polyphyly of the fern family Tectariaceae *sensu* Ching: Insights from cpDNA sequence data. *Science in China Series C: Life Sciences* 50: 789–798.

Liu H-M, Zhang X-C, Chen Z-D, Qiu Y-L. 2007b. Inclusion of the eastern Asia endemic genus *Sorolepidium* in *Polystichum* (Dryopteridaceae): Evidence from the chloroplast *rbcL* gene and morphological characteristics. *Chinese Science Bulletin* 52: 631–638.

Liu H-M, Zhang X-C, Wang M-P, Shang H, Zhou S-L, Yan Y-H, Wei X-P, Xu W-B, Schneider H. 2016. Phylogenetic placement of the enigmatic fern genus *Trichoneuron* informs on the infra-familial relationship of Dryopteridaceae. *Plant Systematics and Evolution* 302: 319–332.

Liu H-M, Zhang X-C, Wang W, Qiu Y-L, Chen Z-D. 2007c. Molecular phylogeny of the fern family Dryopteridaceae inferred from chloroplast *rbcL* and *atpB* genes. *International Journal of Plant Sciences* 689: 1311–1323.

Liu H-M, Zhang X-C, Wang W, Zeng H. 2010. Molecular phylogeny of the endemic fern genera *Cyrtomidictyum* and *Cyrtogonellum* (Dryopteridaceae) from East Asia. *Organisms, Diversity and Evolution* 10: 57–86.

Liu H-M, Zhang X-C. 2004. *Selaginella decipiens* Warb. (Selaginellaceae), a new record from China. *Acta Phytotaxonomica Sinica* 42: 89–92.

Lu J-M, Barrington DS, Li D-Z. 2007. Molecular phylogeny of the polystichoid ferns in Asia based on *rbcL* sequences. *Systematic Botany* 32: 26–33.

Lu J-M, Li D-Z, Lutx S, Soejima A, Yi T-S, Wen J. 2011. Biogeography disjunction between eastern Asia and North America in the *Adiantum pedatum* complex (Pteridaceae). *American Journal of Botany* 98: 1680–1693.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the 2010 Gateway Computing Environments Workshop* (GCE), New Orleans, 14 Nov. 2010. LA. 1–8.

Moore S-J, Parris BS. 2013. *Scleroglossum* to *Tomophyllum*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 840–850.

Moran RC. 2008. Diversity, biogeography, and floristics. In: Ranker TA, Haufler CH eds. *Biology and evolution of ferns and lycophytes*. New York: Cambridge University Press. 367–383.

Murakami N. 1995. Systematics and evolutionary biology of the fern genus *Hymenasplenium* (Aspleniaceae). *Journal of Plant Research* 108: 257–268.

Murdock AG. 2008. A taxonomic revision of the ensporangiate fern family Marattiaceae, with description of a new genus *Ptisana*. *Taxon* 57: 737–755.

Nagalingum NS, Marshall CR, Quental TB, Rai HS, Little DP, Mathews S. 2011. Recent synchronous radiation of a living fossil. *Science* 334: 796–799.

Øllgaard B. 1990. Lycopodiaceae. In: Kubitzki J ed. *The families and genera of vascular plants—Pteridophytes and gymnosperms*. Berlin: Springer. 1: 31–39.

Øllgaard B. 2012. New combinations in Neotropical Lycopodiaceae. *Phytotaxa* 57: 10–22.

Perrie LR, Wilson RK, Shepherd LD, Ohlsen DJ, Batty EL, Brownsey PJ, Bayly MJ. 2014. Molecular phylogenetics and generic taxonomy of Blechnaceae ferns. *Taxon* 63: 745–758.

Pichi-Sermolli REG. 1977. Tentamen pteridophytorum genera in taxonomicum ordinem redigendi. *Webbia* 31: 313–512.

Posada D, Crandall KA. 1998. ModelTest: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409: 618–622.

Rai HS, Graham SW. 2010. Utility of a large, multigene plastid data set in inferring higher-order relationships in ferns and relatives (monilophytes). *American Journal of Botany* 97: 1444–1456.

Ravensberg WJ, Hennipman E. 1986. The *Pyrrosia* species formerly referred to as *Drymoglossum* and *Saxiglossum* (Filicales, Polypodiaceae). *Leiden Botanical Series* 9: 281–310.

Rothfels CJ, Larsson A, Kuo L-Y, Korall P, Chiou W-L, Pryer KM. 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic Biology* 61: 490–509.

Rothfels CJ, Li F-W, Sigel EM, Huiet L, Larsson A, Burge DO, Ruhsam M, Deyholos M, Soltis DE, Stewart CN, Shaw SW, Pokorny L, Chen T, dePamphilis C, DeGironimo L, Chen L, Wei X-F, Sun X, Korall P, Stevenson DW, Graham SW, Wong K-S, Pryer KM. 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *American Journal of Botany* 102: 1–19.

Rothfels CJ, Windham MD, Grusz AL, Gastony GJ, Pryer KM. 2008. Toward a monophyletic *Notholaena* (Pteridaceae): Resolving patterns of evolutionary convergence in xeric-adapted ferns. *Taxon* 57: 712–724.

Ruhfel B, Lindsay S, Davis CC. 2008. Phylogenetic placement of *Rheopteris* and the polyphyly of *Monogramma* (Pteridaceae s.l.): Evidence from *rbcL* sequence data. *Systematic Botany* 33: 37–43.

Sano R, Takamiya M, Ito M, Kurita S, Hasebe M. 2000. Phylogeny of the lady fern group, tribe Physematieae (Dryopteridaceae), based on chloroplast *rbcL* gene sequences. *Molecular Phylogenetics and Evolution* 15: 403–413.

Schmidt-Lebuhn AN. 2012. Fallacies and false premises—a critical assessment of the arguments for the recognition of paraphyletic taxa in botany. *Cladistics* 28: 174–187.

Schneider H, He L-J, Hennequin S, Zhang X-C. 2013. Towards a natural classification of Pteridaceae: Inferring the relationships of enigmatic pteridoid fern species occurring in the Sino-Himalaya and Afro-Madagascar. *Phytotaxa* 77: 49–60.

Schneider H. 2007. Plant morphology as the cornerstone to the integration of fossils and extant taxa in phylogenetic analyses. *Species, Phylogeny and Evolution* 1: 65–71.

Schneider H, Kreier HP, Janssen T, Otto E, Muth H, Heinrichs J. 2010. Key innovations versus key opportunities: Identifying causes of rapid radiations in derived ferns. In: Glaubrecht M ed. *Evolution in action*. Berlin: Springer. 61–75.

Schneider H, Russell SJ, Cox CJ, Bakker F, Henderson S, Rumsey F, Barrett J, Gibby M, Vogel JC. 2004a. Chloroplast phylogeny of asplenioid ferns based on *rbcL* and *trnL-F* spacer sequences (Polypodiidae, Aspleniaceae) and its implications for biogeography. *Systematic Botany* 29: 260–274.

Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallón S, Lupia R. 2004b. Ferns diversified in the shadow of angiosperms. *Nature* 428: 553–557.

Schneider H, Smith AR, Cranfill R, Haufler CH, Ranker TA, Hildebrand TE. 2002. *Gymnogrammitis dareiformis* is a polygrammoid fern (Polypodiaceae)—resolving an apparent conflict between morphological and molecular data. *Plant Systematics and Evolution* 234: 121–136.

Schneider H, Smith AR, Cranfill R, Hildebrand TE, Haufler C, Ranker TA. 2004c. Unraveling the phylogeny of polygrammoid ferns (Polypodiaceae and Grammitidaceae): Exploring aspects of the diversification of epiphytic plants. *Molecular Phylogenetics and Evolution* 31: 1041–1063.

Schneider H, Smith AR, Pryer KM. 2009. Is morphology really at odds with molecules in estimating fern phylogeny? *Systematic Botany* 34: 455–475.

Schuettpelz E, Pryer KM. 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* 56: 1037–1050.

Schuettpelz E, Pryer KM. 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences USA* 106: 11200–11205.

Schuettpelz E, Schneider H, Huiet L, Windham MD, Pryer KM. 2007. A molecular phylogeny of the fern family Pteridaceae: Assessing overall relationships and the affinities of previously unsampled genera. *Molecular Phylogenetics and Evolution* 44: 1172–1185.

Sessa EB, Zimmer EA, Givnish TJ. 2012. Reticulate evolution on a global scale: A nuclear phylogeny for New World *Dryopteris* (Dryopteridaceae). *Molecular Phylogenetics and Evolution* 64: 563–581.

Shao W, Lu S-G. 2009. First recognition of the genus *Crypsinus* (Polypodiaceae) in China. *Acta Botanica Yunnanica* 31: 29–31.

Shao W, Lu S-G. 2011. *Himalayopteris*, a new fern genus from India and the adjacent Himalayas (Polypodiaceae, Polypodioideae). *Novon* 21: 90–93.

Shen Y-Y, Chen X, Murphy RW. 2013. Assessing DNA barcoding as a tool for species identification and data quality control. *PLoS ONE* 8: e57125.

Shepherd LD, Perrie LR, Parris BS, Brownsey PJ. 2007. A molecular phylogeny for the New Zealand Blechnaceae ferns from analyses of chloroplast *trnL-trnF* DNA sequences. *New Zealand Journal of Botany* 45: 67–80.

Shing K-H, Lin Y-X, Chiu P-S, Yao G-H. 1999. Hypodematiaceae and Thelypteridaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Shinohara W, Nakato N, Yatabe-Kakugawa Y, Oka T, Kim JK, Murakami N, Noda H, Sahashi N. 2013. The use of *matK* in Ophioglossaceae phylogeny and the determination of *Mankyua* chromosome number shed light on chromosome number evolution in Ophioglossaceae. *Systematic Botany* 38: 564–570.

Smith AR, Pryer KM, Schuettpelz E, Korall P, Schneider H, Wolf PG. 2006. A classification for extant ferns. *Taxon* 55: 705–731.

Smith AR, Zhang X-C. 2002. *Caobangia*, a new genus and species of Polypodiaceae from Vietnam. *Novon* 12: 546–550.

Smith SA, Beaulieu JM, Donoghue MJ. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences USA* 107: 5897–5902.

Smith SA, Beaulieu J, Stamatakis A, Donoghue MJ. 2011. Understanding angiosperm diversification using large and small phylogenies. *American Journal of Botany* 98: 404–414.

Sundue M, Parris BS, Ranker TA, Smith AR, Fujimoto EL, Zamora-Crosby D, Morden CW, Chiou W-L, Chen C-W, Rouhan G, Hirai RY, Prado J. 2014. Global phylogeny and biogeography of grammitid ferns (Polypodiaceae). *Molecular Phylogenetics and Evolution* 81: 195–206.

Tan Y-H, Wei R, Li J-W, Zhang X-C. 2015. *Didymochlaena* Desv. (Hypodematiaceae): A newly recorded fern genus from China. *Plant Diversity and Resources* 37: 135–138.

Tryon RM, Tryon AF. 1982. *Ferns and allied plants, with special reference to tropical America*. New York: Springer.

Tryon RM, Tryon AF, Kramer KU. 1990. Pteridaceae. In: Kubitzki J ed. *The families and genera of vascular plants–Pteridophytes and gymnosperms*. Berlin: Springer. 1: 230–256.

Tsutsumi C, Kato M. 2006. Evolution of epiphytes in Davalliaceae and related ferns. *Botanical Journal of the Linnean Society* 151: 495–510.

Tsutsumi C, Zhang X-C, Kato M. 2008. Molecular phylogeny of Davalliaceae and implications for generic classification. *Systematic Botany* 33: 44–48.

Wang F-G, Barratt S, Falcon W, Fay MF, Lehtonen S, Tuomisto H, Xing F-W, Christenhusz MJM. 2014. On the monophyly of subfamily Tectarioideae (Polypodiaceae) and the phylogenetic placement of some associated fern genera. *Phytotaxa* 164: 1–16.

Wang F-G, Xing F-W, Kato M. 2013a. *Blechnum* through *Woodwarida*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 411–417.

Wang F-H, Li D-Z, Xie C-Y, Lu J-M. 2013b. A comparison of major classifications of Chinese lycophytes and ferns. *Plant Diversity and Resources* 35: 791–809.

Wang L, Qi X-P, Xiang Q-P, Heinrichs J, Schneider H, Zhang X-C. 2010. Phylogeny of the paleotropical fern genus *Lepisorus* (Polypodiaceae, Polypodiopsida) inferred from four chloroplast DNA regions. *Molecular Phylogenetics and Evolution* 54: 211–225.

Wang L, Schneider H, Zhang X-C, Xiang Q-P. 2012. The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *BMC Plant Biology* 12: 210.

Wang P-S. 1992. *Pteridophytes from Fanjing Mountain*. Guiyang: Guizhou Science and Technology Press.

Wang W, Yang W-L, Mao X-X, Zhao R-R, Dou P, Zhang G-M. 2015. The phylogenetic affinities of *Pellaea connectens*, a rare endemic Chinese fern. *Phytotaxa* 220: 30–42.

Wei R, Schneider H, Zhang X-C. 2013. Toward a new circumscription of the twinsorus-fern genus *Diplazium* (Athyriaceae): A molecular phylogeny with morphological implications and infrageneric taxonomy. *Taxon* 62: 441–457.

Wei R, Xiang Q-P, Schneider H, Sundue MA, Kessler M, Kamau PW, Hidayat A, Zhang X-C. 2015. Eurasian origin, boreotropical migration and transoceanic dispersal in the pantropical fern genus *Diplazium* (Athyriaceae). *Journal of Biogeography* 42: 1809–1819.

Wei R, Zhang X-C. 2014. Rediscovery of *Cystoathyrium chinense* Ching (Cystopteridaceae): Phylogenetic placement of the critically endangered fern species endemic to China. *Journal of Systematics and Evolution* 52: 450–457.

Wesche P, Gaffney D, Keightley P. 2004. DNA sequence error rates in GenBank records estimated using the mouse genome as a reference. *DNA Sequence* 15: 362–364.

Wikström N, Kenrick P. 2001. Relationships of *Lycopodium* and *Lycopodiella* based on combined plastid *rbcL* gene and *trnL* intron sequence data. *Systematic Botany* 25: 495–510.

Williams EW, Waller DM. 2012. Phylogenetic placement of species within the genus *Botrychium* s.s. (Ophioglossaceae) on the basis of plastid sequences, amplified fragment length polymorphisms, and flow cytometry. *International Journal of Plant Sciences* 173: 516–531.

Wu C-H. 1999. Aspleniaceae throught Peranemaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Wu C-H, Wang C-H. 1999. Aspidiaceae through Gymnogrammtidaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Wu S-K. 1979. Leptolepidium, a new genus of Sinopteridaceae. *Acta Botanica Yunnania* 1: 115–118.

Wu S-K, Hsieh Y-T, Lu S-G. 2000. Dryopteridaceae (1). In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Wu Z-H, Ching R-C. 1991. *Fern Families and Genera of China*. Beijing: Science Press.

Wu Z-Y, Raven PH, Hong D-Y. 2013. Lycopodiaceae through Polypodiaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3.

Xi Z-X, Ruhfel BR, Schaefer H, Amorim AM, Sugumaran M, Wurdack KJ, Endress PK, Matthews ML, Stevens PF, Mathews S, Davis CC. 2012. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proceedings of the National Academy of Sciences USA* 109: 17519–17524.

Xing F-W, Yan Y-H, Dong S-Y, Wang F-G, Christenhusz MJM, Hovenkamp PH. 2013. Tectariaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 730–746.

Xu W-B, Liang Y-Y, Zhang X-C, Liu Y. 2008. *Caobangia* A. R. Smith & X. C. Zhang (Polypodiaceae), a newly recorded fern genus from China. *Journal of Systematics and Evolution* 46: 916–918.

Yan Y-H, Qi X-P, Serizawa S. 2013. *Dennstaedtia*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 155.

Yatskievych G, Smith AR. 2003. Typification of *Notholaena* R. Br. (Pteridaceae). *Taxon* 52: 331–336.

Yesilyurt JC, Schneider H. 2010. The new fern genus *Calciphilopteris* (Pteridaceae). *Phytotaxa* 7: 52–59.

Zhang C-F. 1993. *Ataxipteris*. In: *Fern Flora of Zhejiang*. Hangzhou: Zhejiang Science and Technology Press. 1: 288–289.

Zhang G-M, Kato M, Yatskievych G. 2013a. Woodsiaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 404.

Zhang G-M, Liao W-B, Ding M-Y, Lin Y-X, Wu Z-H, Zhang X-C, Dong S-Y, Prado J, Gilbert MG, Yatskievych G, Ranker TA, Hooper EA, Alverson ER, Metzgar JS, Funston AM, Masuyama S, Kato M. 2013b. Pteridaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 169–256.

Zhang G-M, Yatskievych G. 2013. *Cheilanthes*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 218–222.

Zhang G-M, Zhang X-C, Chen Z-D, Liu H-M, Yang W-L. 2007. First insights in the phylogeny of Asian cheilanthoid ferns based on sequences of two chloroplast markers. *Taxon* 56: 369–378.

Zhang L, Rothfels CJ, Ebihara A, Schuettpelz E, Péchon TL, Kamau P, He H, Zhou X-M, Prado J, Field A, Yatskievych G, Gao X-F, Zhang L-B. 2015. A global plastid phylogeny of the brake fern genus *Pteris* (Pteridaceae) and related genera in the Pteridoideae. *Cladistics* 31: 406–423.

Zhang L-B, Barrington D. 2013. *Polystichum*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 629–713.

Zhang L-B, Gilbert MG. 2015. Comparison of classifications of vascular plants of China. *Taxon* 64: 17–26.

Zhang L-B, Iwatsuki K. 2013. Lycopodiaceae. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 13–34.

Zhang L-B, Kung H-S. 2000. The reclassification of Lycopodiaceae (s. str.) in China. *Acta Phytotaxonomica Sinica* 38: 266–275.

Zhang L-B, Zhang L. 2012. The inclusion of *Acrophorus, Diacalpe, Nothoperanema*, and *Peranema* in *Dryopteris*: The molecular phylogeny, systematics, and nomenclature of *Dryopteris* subg. *Nothoperanema* (Dryopteridaceae). *Taxon* 61: 1199–1216.

Zhang L-B, Zhang L. 2015. Didymochlaenaceae: A new fern family of eupolypods I (Polypodiales). *Taxon* 64: 27–38.

Zhang L-B, Zhang L, Dong S-Y, Sessa EB, Gao X-F, Ebihara A. 2012. Molecular circumscription and major evolutionary lineages of the fern genus *Dryopteris* (Dryopteridaceae). *BMC Evolutionary Biology* 12: 180.

Zhang X-C. 2000. Grammitidaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press. 6: 298–305.

Zhang X-C. 2003a. New combinations in *Haplopteris* (Pteridophyta: Vittariaceae). *Annals Botanici Fennici* 40: 459–461.

Zhang X-C. 2003b. The east Asiatic genera and endemic genera of the pteridophytes in China. *Korean Journal of Plant Taxonomy* 33: 309–338.

Zhang X-C. 2012. *Lycophytes and ferns of China*. Beijing: Peking University Press.

Zhang X-C, Gilbert MG. 2013. *Monogramma*. In: Wu Z-Y, Raven PH, Hong D-Y eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 2–3: 256.

Zhang X-C, Liu H-M. 2004. Miscellaneous notes on pteridophytes from China and neighbouring regions (III)—the onocleoid ferns, *Pentarhizidium* Hayata and *Matteuccia* (L.) Todaro. In: Srivastava PC ed. *Vistas in paleobotany and plant morphology: Evolutionary and environmental perspectives: Professor D.D. Pant Memorial Volume*. Lucknow: Offset. 361–366.

Zhang X-C, Wei R, Liu H-M, He L-J, Wang L, Zhang G-M. 2013c. Phylogeny and classification of the extant lycophytes and ferns from China. *Chinese Bulletin of Botany* 48: 119–137.

Zhang X-C, Zhang L-B. 2004. Huperziaceae through Cyatheaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12184/suppinfo:

Table S1. For each accession the following information is included: species name, voucher information including collectors and locality, and GenBank accession numbers for each regions gained. In addition, lab DNA numbers are provided for accessions with newly generated DNA sequences (BOPXXXXXX/FPXXXXXX). Concepts of family, genera, and type species follow Ching (1978) with the exception of those having conflict with current views. These are given in brackets and the information was obtained from Tropicos (http://www.tropicos.org). In addition, each taxon is also given with the corresponding taxonomic treatments as in FOC if generic and/or species concepts were changed (Wu et al., 2013). The following symbols correspond to: – = sequences were not available; // = no treatments in FOC; (C.) = cultivated materials.

Table S2. Families and genera checklist of ferns and lycophytes. Those occurring in China are marked as 1, otherwise marked as 0.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12090

Research Article

Accelerated evolution of early angiosperms: Evidence from ranunculalean phylogeny by integrating living and fossil data

Wei Wang[1], David L. Dilcher[2,3,4]*, Ge Sun[2,3], Hong-Shan Wang[2,5], and Zhi-Duan Chen[1]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[2]Paleontological Institute of Shenyang Normal University, Shenyang 110034, China
[3]Research Center of Paleontology, Jilin University, Changchun 130026, China
[4]Department of Geology/Biology, Indiana University, Bloomington, IN 47405, USA
[5]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA
*Author for correspondence. E-mail: dilcher@indiana.edu. Tel.: 1-812-856-0618. Fax: 1-812-855-7899.
Received 6 December 2013; Accepted 10 March 2014; Article first published online 15 April 2014

Abstract The new discovery of angiosperm remains in the Jehol Biota of northeastern China contributes to our understanding of the origin and early evolution of flowering plants. The earliest eudicot genus with reproductive organs, *Leefructus*, was recently documented from the Lower Cretaceous Yixian Formation at 125.8–123.0 Ma, and was reconsidered to be close to the extant family Ranunculaceae based on gross morphology. However, this hypothesis has not been tested using a cladistic approach. To determine the possible allies of *Leefructus* within extant eudicots, we constructed a 66 morphological data matrix. Molecular and morphological analyses of extant Ranunculales combined with the fossil suggest that it has an affinity with the Ranunculaceae. The earliest fossil record of the eudicots is 127–125 Ma based on tricolpate pollen grains. Thus, we suggest a hypothesis that the basal eudicots might have experienced an accelerated evolution and diversification during the latest Barremian and earliest Aptian, leading to the stem groups of at least six extant families or lineages, 10–15 Myr earlier than currently documented. Angiosperms have undergone multiple uneven pulses of radiation since their origin. Many key character innovations occurred in different stages that could have triggered those radiations in concert with various biotic and abiotic factors.

Key words: angiosperms, Cretaceous, diversification, paleobotany, phylogeny, Ranunculales.

Angiosperms are of exceptional evolutionary interest because of their diversity of over 250 000 species (Palmer et al., 2004) and their abundance as the dominant vegetation in most terrestrial ecosystems today. Their evolutionary history has been filled with uneven pulses of radiation since their Early Cretaceous origin (140–135 Ma) and their rapid radiations during the mid-Cretaceous (107–93 Ma) that has been documented in many research papers (e.g., Crepet et al., 2004; Anderson et al., 2005; Moore et al., 2007) and textbooks (e.g., Stewart & Rothwell, 1993; Willis & McElwain, 2002; Taylor et al., 2009). This was followed by further radiations as the result of fruit evolution in concert with radiations of birds and mammals (Dilcher, 2010).

The current molecular systematics of angiosperms (Palmer et al., 2004; Soltis et al., 2009, 2011; Qiu et al., 2010) recognizes the basalmost angiosperms and five major angiosperm lineages. The five lineages include the sister lineages Chloranthaceae and magnoliids, the monocots, *Ceratophyllum*, and the eudicots consisting of the basal eudicots and the core eudicots (Moore et al., 2007; Soltis et al., 2011). The eudicots now constitute more than 70% of all extant angiosperm species.

Our current understanding of the natural relationships of angiosperms is based mainly on molecular phylogenetic data (e.g., Qiu et al., 1999, 2006, 2010; Soltis & Soltis, 2004; Doyle et al., 2008; Qiu & Estabrook, 2008; APG III, 2009; Wang et al., 2009b; Soltis et al., 2011). These data are most often presented in the form of cladograms showing stem and branch lineages that are derived from the analysis of large datasets including DNA sequences from chloroplast, mitochondrial, and nuclear genomes (Qiu et al., 1999; Palmer et al., 2004; APG III, 2009). Such cladograms have come to represent the evolutionary history of angiosperms while the branching points between the clades are considered to be constrained in time (Crepet et al., 2004; Anderson et al., 2005). The origin of flowering plants and the origins of major clades as constrained by time (Bell et al., 2005, 2010) are best revealed by the fossil record when available. It is the fossil record that holds the key to understanding the sequences and timing of the multiple radiations involved in flowering plant evolution (Crepet et al., 2004; Gandolfo et al., 2008). Whenever possible the angiosperm fossil record should be used to ground truth the systematic relationships and the time of the divergence and radiation of major clades of extant angiosperms as understood from molecular and morphological data (Crepet et al., 2004; Anderson et al., 2005; Gandolfo et al., 2008). However, the earliest known fossil records of angiosperms are rare, often incomplete, and difficult to interpret.

Recently, Sun et al. (2011) reported a fossil eudicot genus with reproductive organs, *Leefructus*, from the Lower Cretaceous Yixian Formation at 125.8–123.0 Ma (Meng et al., 2008), and placed it as an extinct eudicot on the stem lineage of the extant family Ranunculaceae based on gross morphology (Fig. 1). However, *Leefructus* has not been included in a cladistic analysis or been shown to possess any previously defined synapomorphies for the eudicot total group, crown group, or any clade within the crown group, which prevents its use as a calibration constraint for a molecular clock (Clarke et al., 2011). In this study, 14 characters of *Leefructus* were coded based on the description of the fossil genus (Sun et al., 2011), and added into our previous established 65 morphological data matrix (Wang et al., 2009b) to determine the positions of the fossil *Leefructus* within eudicots.

Material and Methods

We added one beak character (present vs. absence) into our previous 65 morphological data matrix (Wang et al., 2009b). Fourteen characters of *Leefructus mirus* Sun et al. (2011) were coded based on the description of the fossil genus: beak (present), growth form (herbaceous), stipules (absent), leaf arrangement (spiral), major venation (palmate), blade shape (ovate), inflorescence (solitary), stamen arrangement (irregular), stamen fusion (free), carpel number (more than 3), carpel form (ascidiate up to stigma), carpel fusion (pseudo-syncarpous), ovule number (more than two), and fruit wall (dry). When a character was poorly known or unavailable for the species, it was coded as missing or inapplicable.

Phylogenetic analyses were carried out using maximum parsimony (MP) and Bayesian inference (BI) methods were used in PAUP* version 4.0b10 (Swofford, 2003) and MrBayes version 3.0b4 (Ronquist & Huelsenbeck, 2003), respectively. For MP analyses, we used the backbone constraint tree approach as in Endress & Doyle (2009). The tree is based primarily on the combined analysis of morphology, *rbcL*, *matK*, *trnL-F*, and 26S rDNA by Wang et al. (2009b), but with change in the position of Ceratophyllaceae based on more recent analyses. The plastid phylogenomic analyses of the whole angiosperm (Moore et al., 2007, 2010) and 17 genes from three genomes (Soltis et al., 2011) found *Ceratophyllum* sister to eudicots. The MP heuristic searches were carried out with 1000 random sequence addition replicates, tree bisection–reconnection branch swapping, MulTrees in effect, and steepest descent off. Bootstrapping was carried out with 1000 replicates, using a heuristic search strategy (five random addition replicates, saving five trees per replicate).

For BI analyses, the combined morphological and four DNA matrix by Wang et al. (2009b) was reconstructed, where the molecular data of *L. mirus* were coded as missing. The detailed analysis approach was described in Wang et al. (2009b).

Results and Discussion

The MP and BI analyses resulted in identical topology at the familial level (Fig. 2). Within Ranunculales, Eupteleaceae is the earliest-diverging family, followed by Papaveraceae; the other five families form a clade. These results are congruent with our previous study (Wang et al., 2009b). Significantly, all our

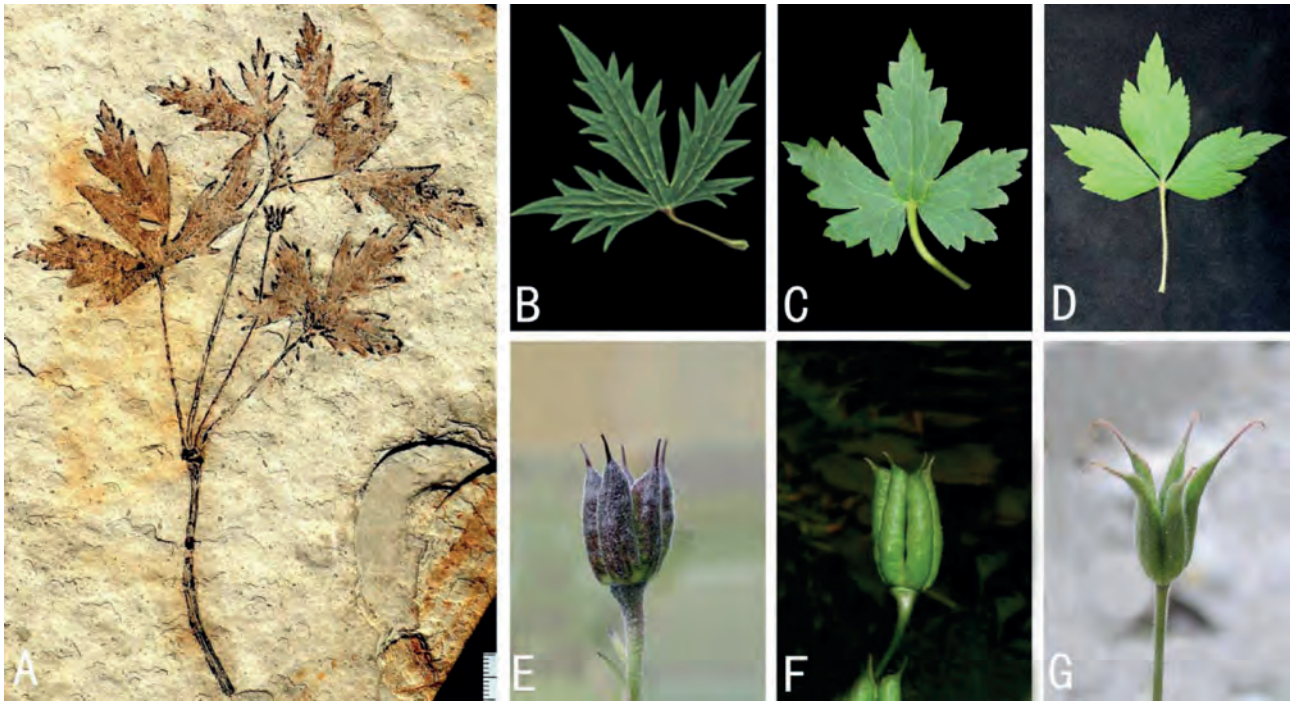


Fig. 1. Leaves and fruits of fossil (**A**) and extant Ranunculaceae species (**B–G**). **A,** *Leefructus mirus* Sun, Dilcher, Wang et Chen. **B,** *Delphinium* sp. **C,** *Aconitum hemsleyanum* Pritz. **D,** *Anemone virginiana* L. **E,** *Delphinium glaucum* S. Watson. **F,** *Aconitum kusnezoffii* Reichb. **G,** *Aquilegia einseleana* F. W. Schultz. Photograph A is from Sun et al. (2011). Photographs B–G were taken by S.-X. YU and used with permission.

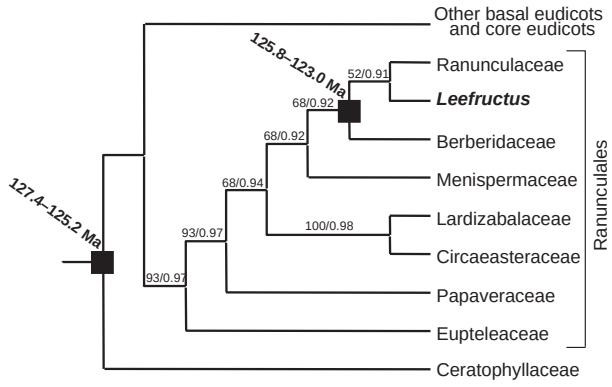


Fig. 2. Bootstrap tree obtained after addition of *Leefructus* to the tree of Wang et al. (2009b). Numbers at the nodes are bootstrap percentages and Bayesian posterior probabilities, respectively.

analyses indicate that *Leefructus* is sister to the Ranunculaceae (bootstrap support = 52%; posterior probability = 0.91). *Leefructus* shares many characters with the extant Ranunculaceae, including presence of herbaceous habit, palmate venation, free stamens, more than two ovules, stamens of irregular arrangement, and fruits with beaks. The leaves of *Leefructus* resemble the venation patterns and forms of *Delphinium* leaves (Fig. 1: B), typical of the Ranunculaceae, and the fruits of *Leefructus* resemble grossly those of *Aconitum* and *Delphinium* (Fig. 1: E, F). But *Leefructus* has some unique characters, such as leaves in spiral arrangement and syncarpous carpels (also in *Nigella* and some *Helleborus*). Thus, combining phylogenetic analyses, we are in agreement with Sun et al. (2011) that *Leefructus* belongs to the stem group of Ranunculaceae.

The antiquity of *Leefructus* requires a revision of some ages suggested for the age of the basal eudicots. An estimate for the diversification of the eudicots, a major clade of flowering plants, is 124.8 (± 3.2) Ma based on the molecular analysis of 61 plastid genomes (Moore et al., 2007). Based on 36 minimum age constraints treated as exponential distributions, Bell et al. (2010) suggested that the crown age of the eudicots is 129 Ma (123–134 Ma). Based on the plastid *rbcL* data, Anderson et al. (2005) used multiple fossil calibration points and placed the diversification of Ranunculales at an age of 114 Ma and the Ranunculaceae/Berberidaceae divergence from the Menispermaceae at 105 Ma. In this study, we suggest that the Ranunculaceae/Berberidaceae divergence from the Menispermaceae took place nearly 20 Myr earlier.

The term eudicot was proposed in 1991 (Doyle & Hotton, 1991) as a "putatively monophyletic group" using tricolpate pollen to define the clade. This clade was recognized at the Barremian–Hauterivian boundary from sediments in Gabon (Doyle et al., 1977) and from the early Albian of the Potomac Group in North America (Endress & Doyle, 2009). Hughes & McDougall (1990) and Hughes (1994) recognized tricolpate pollen from Bed 35 at the base of the Vectis Formation that Hughes considered "Phase 4," which was at the Barremian–Aptian boundary. This current understanding of the fossil record placed the earliest fossils of the eudicots at this Barremian–Aptian boundary age of ca. 125 Ma (http://www.stratigraphy.org/ICSchart/ChronostratChart2013-01.pdf). The distribution of this pollen type from cores taken off the coast of Gabon and from England suggests that plants belonging to this eudicot pollen type were already distributed widely by ca. 125 Ma and might have had an earlier origin. *Prototricolpites* pollen was described from the Jianshangou Bed (127.4–125.2 Ma) of the Yixian Formation, which extended both the age and distribution of the basal eudicots (Wang et al., 2000).

If the first appearance of basal eudicots based on the occurrence of tricolpate pollen is at ca. 127–125 Ma or slightly earlier and the morphological characters for *Leefructus* have placed it firmly as an extinct form on the stem lineage of the Ranunculaceae at 125.8–123.0 Ma, then there might have been a time of accelerated evolution of the basal eudicots during the latest Barremian and early Aptian. This evolution needed to require an accelerated rate of evolution within the basal eudicot clade during the Early Cretaceous. The early fossil record of basal eudicot evolution is incomplete so each fossil that can be placed in this clade provides new and important information about early angiosperms. Previous fossil angiosperms reported from the Yixian Formation included *Archaefructus* and *Hyrcantha decussata* (Sun et al., 1998, 2002; Dilcher et al., 2007). It is possible that *Hyrcantha* could also be considered as a stem lineage of the Ranunculaceae (Leng & Friis, 2003, 2006; Dilcher et al., 2007). Additionally, Jud & Hickey (2013) recently reported a fossil eudicot genus with leaf organs, *Potomacapnos*, from Aptian sediments of the Potomac Group exposed at the Dutch Gap locality in Virginia, USA, and tentatively placed it in Papaveraceae of Ranunculales based on leaf architecture.

The cladogram of Wang et al. (2009b) is presented in Fig. 2 with the addition of the fossil data. We accept the presence of the earliest tricolpate pollen in the fossil record as the time when eudicots can be firmly documented (Doyle et al., 1977, 2008; Hughes & McDougall, 1990; Doyle & Hotton, 1991; Doyle, 1992; Hughes, 1994) (mid-Barremian, Isle of Wright) and is placed at ca. 127–125 Ma. *Leefructus* is recognized at 125.8–123.0 Ma. This means less than 5 Myr from the initial recognition of the basal eudicots to the recognition of the stem lineage of the Ranunculaceae. The minimum age node mapping method (Crepet et al., 2004; Hermsen & Hendricks, 2008), as applied to a cladogram containing fossils, proposes that each node of the cladogram is at least as old as the oldest descendant. This leads to us concluding that the extant families of the basal eudicots came into existence between 127–125 and 125–123 Ma. In order to emphasize this rapid evolution, we suggest the hypothesis of an "accelerated angiosperm evolution" for the basal eudicots at a time much earlier than previously recognized by Anderson et al. (2005).

This "accelerated angiosperm evolution" between 127–125 and 125–123 Ma is presented as a hypothesis here because there are limited data at present to support it. The presence of *Leefructus*, *Hyrcantha*, and *Potomacapnos*, which have been suggested as having possible Ranunculales/Ranunculaceae affinities, documents basal eudicot evolution early in angiosperm history. To date, more than 12 putative early eudicot megafossils reported from the Aptian to mid-Albian have been compared with or assigned to eudicots or Ranunculales (Jud & Hickey, 2013, and references therein). Why did such rapid radiations take place so quickly?

The "accelerated angiosperm evolution" hypothesis occurred at a post-Jurassic and Early Cretaceous time when there

was rapid evolution of a variety of insect pollinators active with gymnospermous seed plants (Ren et al., 2009). These pollinators must have transferred to the new angiospermous plant sources of pollen and nectar easily and rapidly. They were also joined by diverse lineages of insect pollinators that could play key roles in flowering plant evolution during the Barremian, Aptian, and Albian. The coevolutionary nature of the early angiosperms was well documented in the Albian when pollen clumps occurred (Hu et al., 2007), indicating continuation and further accommodation of flowering plants to insect pollination.

Accelerated angiosperm evolution has also been observed for particular clades such as the rosids, which might have diversified rapidly in perhaps as little as in 4–5 Myr (Wang et al., 2009a), and the Saxifragales, which diversified rapidly in as little as 6 Myr (Jian et al., 2008). Chaloner (1970) showed that the occurrence of spores preceded the finding of abundant megafossils of early land plants by at least 10–15 Myr. This suggests that spores and pollens might be found long before abundant angiosperm megafossils or mesofossils, and were common in sediments such as the Barremian, Aptian, and Albian angiosperm explosive diversity. Martínez-Millán et al. (2009) noted that the presence of the Ericales in the Turonian in the Late Cretaceous followed their earlier diversification during the Early Cretaceous by a few million years. The presence of tricolpate pollen followed by the burst of evolution of basal eudicots during the earliest Barremian preceded the conspicuous radiation of angiosperm diversity by approximately 20 Myr, during the late Albian (Dilcher & Eriksen, 1983; Pedersen et al., 1994; Magallón-Puebla et al., 1997; Mohr & Friis, 2000; Crepet et al., 2004; Mohr & Bernardes-de-Oliveira, 2004; Friis et al., 2006a, 2006b, 2009; Mohr et al., 2008; Dilcher & Wang, 2009; Taylor et al., 2009). Some of the basal eudicot fossil taxa known from the Albian have very modern features and are identified with extant taxa such as Platanaceae (Dilcher & Eriksen, 1983; Pedersen et al., 1994; Magallón-Puebla et al., 1997), Cabombaceae (Wang & Dilcher, 2006), Buxaceae (Pedersen et al., 2007), Priscaceae (Retallack & Dilcher, 1981), and Lauraceae (Drinnan et al., 1990), as well as sister taxa to the basal eudicots such as Ceratophyllaceae (Dilcher & Wang, 2009). The initial major lineages of basal eudicot evolution took place rapidly beginning during the Barremian and continued during the Aptian and early Albian. This is a time of insect radiation (Labandeira & Sepkoski, 1993; Grimaldi & Engel, 2005; Hu et al., 2007; Labandeira & Conrad, 2013) and early modifications for wind pollination. Based on the fossil record presented by Sun et al. (2011), it is evident that the Barremian and the early Aptian were important times of accelerated angiosperm evolution.

During the evolution of angiosperms, many important character changes occurred over many millions of years. The closed carpel has been proposed as the defining feature of angiosperms (Sun et al., 1998), which allows for biochemical incompatibility of pollen and ovule. Bisexual axis (fertile shoot) with ovules and pollens (male = pollen, ovule = female), occurred at ca. 125 Ma (Sun et al., 1998, 2002, 2011). The flowers of angiosperms with four whorl organs and fragrances and/or nectar first occur later, and the shift from radial to bilateral flowers happened at 70 Ma, which functioned for attracting potential pollinators (Dilcher, 2010). Fruits of angiosperms that are attractive to birds and mammals evolved in the Late Cretaceous and early Cenozoic (Dilcher, 2010), which aided dispersal. Flowering plants have experienced multiple radiation pulses since their origin. The appearance of many important character innovations of the flowering plants, associated with coevolution in a biotic environment, is responsible for their radiations during different epochs. Thus, the accelerated evolution of the early eudicots, documented in this study, is just the initial example of the various radiation bursts that have continued throughout their long history.

Acknowledgements

We sincerely thank Dr. Sheng-Xiang YU for permission to use the photographs of extant plants. This research was partially funded by the National Basic Research Program of China (Grant No. 2014CB954101), the National Natural Science Foundation of China (Grant Nos. 31270269 and 31270268), and the Youth Innovation Promotion Association Foundation of CAS.

References

Anderson CL, Bremer K, Friis EM. 2005. Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. *American Journal of Botany* 92: 1737–1748.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161: 105–121.

Bell CD, Soltis DE, Soltis PS. 2005. The age of the angiosperms: A molecular timescale without a clock. *Evolution* 59: 1245–1258.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* 97: 1296–1303.

Chaloner WG. 1970. The rise of the first land plants. *Biological Reviews* 45: 353–377.

Clarke JT, Warnock RCM, Donoghue PCJ. 2011. Establishing a time-scale for plant evolution. *New Phytologist* 192: 266–301.

Crepet WL, Nixon KC, Gandolfo MA. 2004. Fossil evidence and phylogeny: The age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *American Journal of Botany* 91: 1666–1682.

Dilcher DL. 2010. Major innovations in angiosperm evolution. In: Gee CT ed. *Plants in Mesozoic time: Morphological innovations, phylogeny, ecosystems*. Bloomington: Indiana University Press. 96–116.

Dilcher DL, Eriksen L. 1983. Sycamores are ancient trees. *The Museum of Western Colorado Quarterly* 8–11.

Dilcher DL, Sun G, Ji Q, Li H. 2007. An early infructescence *Hyrcantha decussata* (comb. nov.) from the Yixian Formation in northeastern China. *Proceedings of the National Academy of Sciences USA* 104: 9370–9374.

Dilcher DL, Wang H. 2009. An Early Cretaceous fruit with affinities to Ceratophyllaceae. *American Journal of Botany* 96: 2256–2269.

Doyle JA. 1992. Revised palynological correlations of the lower Potomac Group (USA) and the Cocobeach sequence of Gabon (Barremian–Aptian). *Cretaceous Research* 13: 337–349.

Doyle JA, Biens P, Doerenkamp A, Jardiné S. 1977. Angiosperm pollen from the pre-Albian Lower Cretaceous of equatorial Africa. *Bulletin des Centres de Recherches Exploration–Production Elf Aquitaine* 1: 451–473.

Doyle JA, Endress PK, Upchurch GR. 2008. Early Cretaceous monocots: A phylogenetic evaluation. *Acta Musei Nationalis Pragae, Series B: Historia Naturalis* 64: 59–87.

Doyle JA, Hotton CL. 1991. Diversification of early angiosperm pollen in a cladistic context. In: Blackmore S, Barnes SH eds. *Pollen and spores: Patterns of diversification*. Oxford: Clarendon Press. 169–195.

Drinnan AN, Crane PR, Friis EM, Pedersen KR. 1990. Lauraceous flowers from the Potomac Group (Mid-Cretaceous) of eastern North America. *Botanical Gazette* 151: 370–384.

Endress ME, Doyle JA. 2009. Reconstructing the ancestral angiosperm flower and its initial specializations. *American Journal of Botany* 96: 22–66.

Friis EM, Pedersen KR, Crane PR. 2006a. Cretaceous angiosperm flowers: Innovation and evolution in plant reproduction. *Palaeogeography, Palaeoclimatology, Palaeoecology* 232: 251–293.

Friis EM, Pedersen KR, Schönenberger J. 2006b. Normapolles plants: A prominent component of the Cretaceous rosid diversification. *Plant Systematics and Evolution* 206: 107–140.

Friis EM, Pedersen KR, von Balthazar M, Grimm GW, Crane PR. 2009. *Monetianthus mirus* gen. et sp. nov., a Nymphaealean flower from the early Cretaceous of Portugal. *International Journal of Plant Sciences* 170: 1086–1101.

Gandolfo MA, Nixon KC, Crepet WL. 2008. Selection of fossils for calibration of molecular dating models. *Annals of the Missouri Botanical Garden* 95: 34–42.

Grimaldi D, Engel MS. 2005. *Evolution of the insects*. Cambridge: Cambridge University Press. 772.

Hermsen EJ, Hendricks JR. 2008. W(h)ither fossils? Studying morphological character evolution in the age of molecular sequences. *Annals of the Missouri Botanical Garden* 95: 72–100.

Hu S, Dilcher DL, Jarzen DM, Taylor WD. 2007. Early steps of angiosperm–pollinator coevolution. *Proceedings of the National Academy of Sciences USA* 105: 240–245.

Hughes NF. 1994. *The enigma of angiosperm origins*. Cambridge: Cambridge University Press.

Hughes NF, McDougall AB. 1990. Barremian–Aptian angiospermid pollen records from southern England. *Review of Palaeobotany and Palynology* 65: 145–151.

Jian S, Soltis PS, Gitzendanner MA, Moore MJ, Li R, Hendry TA, Qiu YL, Dhingra A, Bell CD, Soltis DE. 2008. Resolving an ancient, rapid radiation in Saxifragales. *Systematic Biology* 57: 38–57.

Jud NA, Hickey LJ. 2013. *Potomacapnos apeleutheron* gen. et sp. nov., a new Early Cretaceous angiosperm from the Potomac Group and its implications for the evolution of eudicot leaf architecture. *American Journal of Botany* 100: 2437–2449.

Labandeira CC, Conrad C. 2013. A paleobiologic perspective on plant–insect interactions. *Current Opinion in Plant Biology* 16: 414–421.

Labandeira CC, Sepkoski JJ. 1993. Insect diversity in the fossil record. *Science* 261: 310–315.

Leng Q, Friis EM. 2003. *Sinocarpus decussates* gen. et sp. nov., a new angiosperm with basally syncarpous fruits the Yixian Formation of Northeast China. *Plant Systematics and Evolution* 241: 77–88.

Leng Q, Friis EM. 2006. Angiosperm leaves associated with *Sinocarpus* infructescences from the Yixian Formation (mid-Early Cretaceous) of NE China. *Plant Systematics and Evolution* 262: 173–187.

Magallón-Puebla S, Herendeen PS, Crane PR. 1997. *Quadriplatanus georgianus* gen. et sp. nov.: Staminate and pistillate platanaceous flowers from the late Cretaceous (Coniacian–Santonian) of Georgia, USA. *International Journal of Plant Sciences* 162: 963–983.

Martínez-Millán M, Crepet WL, Nixon KC. 2009. *Pentapetalum trifasciculandricus* gen. et sp. nov., a thealean fossil flower from the Raritan Formation, New Jersey, USA (Turonian, Late Cretaceous). *American Journal of Botany* 96: 933–949.

Meng FX, Gao S, Liu XM. 2008. U-Pb zircon geochronology and geochemistry of volcanic rocks of the Yixian Formation in the Lingyuan area, western Liaoning, China. *Geological Bulletin of China* 27: 364–373.

Mohr BAR, Bernardes-de-Oliveira MEC. 2004. *Endressinia brasiliana*, a magnolialean angiosperm from the Lower Cretaceous Crato Formation (Brazil). *International Journal of Plant Sciences* 165: 1121–1133.

Mohr BAR, Bernardes-de-Oliveira MEC, Taylor WD. 2008. *Pluricarpellatia*, a nymphaealean angiosperm from the lower Cretaceous of northern Gondwana (Crato Formation, Brazil). *Taxon* 57: 1147–1158.

Mohr BAR, Friis EM. 2000. Early angiosperms from the Lower Cretaceous Crato Formation (Brazil), a preliminary report. *International Journal of Plant Sciences* 161: S155–S167.

Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences USA* 104: 19363–19368.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences USA* 107: 4623–4628.

Palmer JD, Soltis DE, Chase MW. 2004. The plant tree of life: An overview and some points of view. *American Journal of Botany* 91: 1437–1445.

Pedersen KR, Friis EM, Crane PR, Drinnan AN. 1994. Reproductive structures of an extinct platanoid from the early Cretaceous (latest Albian) of eastern North America. *Review of Palaeobotany and Palynology* 80: 291–303.

Pedersen KR, von Balthazar M, Crane PR, Friis EM. 2007. Early cretaceous floral structures and in situ tricolpate-striate pollen: New early eudicots from Portugal. *Grana* 46: 176–196.

Qiu YL, Estabrook GF. 2008. Inference of phylogenetic relationships among key angiosperm lineages using a compatibility method on a molecular data set. *Journal of Systematics and Evolution* 46: 130–141.

Qiu YL, Lee JH, Bernasconi-Quadroni F, Soltis DE, Soltis PS, Zanis M, Zimmer EA, Chen ZD, Savolainen V, Chase MW. 1999. The earliest angiosperms: Evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402: 404–407.

Qiu YL, Li L, Hendry TA, Li Q, Taylor DW, Issa MJ, Ronen AJ, Vekaria ML, White AM. 2006. Reconstructing the basal angiosperm phylogeny: Evaluating information content of mitochondrial genes. *Taxon* 55: 837–856.

Qiu YL, Li L, Wang B, Xue JY, Hendry TA, Li RQ, Brown JW, Liu Y, Hudson GT, Chen ZD. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* 48: 391–425.

Ren D, Labandeira CC, Santiago-Blay JA, Rasnitsyn A, Shih C, Bashkuev A, Logan MA, Hotton CL, Dilcher D. 2009. A probable pollination mode before angiosperms: Eurasian, long-proboscid scorpionflies. *Science* 326: 840–847.

Retallack G, Dilcher DL. 1981. Early angiosperm reproduction: *Prisca reynoldsii*, gen. et sp. nov. from mid-Cretaceous coastal deposits in Kansas, USA. *Palaeontographica Abteilung B* 179: 103–137.

Ronquist F, Huelsenbeck J. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez N, Walker JB, Moore MJ, Carlsward BS. 2011.

Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis PS, Brockington SF, Yoo MJ, Piedrahita A, Latvis M, Moore MJ, Chanderbali AS, Soltis DE. 2009. Floral variation and floral genetics in basal angiosperms. *American Journal of Botany* 96: 110–128.

Soltis PS, Soltis DE. 2004. The origin and diversification of angiosperms. *American Journal of Botany* 91: 1614–1626.

Stewart WN, Rothwell GW. 1993. *Paleobotany and the evolution of plants*. 2nd ed. Cambridge: Cambridge University Press.

Sun G, Dilcher DL, Wang HS, Chen ZD. 2011. A eudicot from the Early Cretaceous of China. *Nature* 471: 625–628.

Sun G, Dilcher DL, Zheng S, Zhou ZK. 1998. In search of the first flower: A Jurassic angiosperm, *Archaefructus*, from Northeast China. *Science* 282: 1601–1772.

Sun G, Ji Q, Dilcher DL, Zheng S, Nixon KC, Wang X. 2002. Archaefructaceae, a new basal angiosperm family. *Science* 296: 899–904.

Swofford DL. 2003. PAUP*: Phylogenetic analysis using parsimony (*and other methods), v4.0b10. Sunderland: Sinauer Associates.

Taylor TN, Taylor EL, Krings M. 2009. *Paleobotany: The biology and evolution of fossil plants*. 2nd ed. Burlington: Elsevier/Academic Press.

Wang H, Dilcher DL. 2006. Aquatic angiosperms from the Dakota Formation (Albian, Lower Cretaceous), Hoisington III Locality, Kansas, USA. *International Journal of Plant Sciences* 167: 385–401.

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington S, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE. 2009a. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences USA* 106: 3853–3858.

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009b. Phylogeny and classification of Ranunculales: Evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology, Evolution and Systematics* 11: 81–90.

Wang XZ, Ren D, Wang YF. 2000. First discovery of angiospermous pollen from Yixian Formation in Western Liaoning. *Acta Geologica Sinica* 74: 265–272. (in Chinese)

Willis KJ, McElwain J. 2002. *The evolution of plants*. Oxford: Oxford University Press.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12182

Research Article

Phylogenetic tree of vascular plants reveals the origins of aquatic angiosperms

Zhi-Yuan Du, Qing-Feng Wang*, and China Phylogeny Consortium†

Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China
†Members of the China Phylogeny Consortium are listed in the Appendix.
*Author for correspondence. E-mail: qfwang@wbgcas.cn. Tel./Fax: 86-27-87510526.
Received 5 July 2015; Accepted 5 September 2015; Article first published online 26 October 2015

Abstract Although aquatic plants are discussed as a unified biological group, they are phylogenetically well dispersed across the angiosperms. In this study, we annotated the aquatic taxa on the tree of vascular plants, and extracted the topology of these aquatic lineages to construct the tree of aquatic angiosperms. We also reconstructed the ancestral areas of aquatic families. We found that aquatic angiosperms could be divided into two different categories: the four aquatic orders and the aquatic taxa in terrestrial orders. Aquatic lineages evolved early in the radiation of angiosperms, both in the orders Nymphaeales and Ceratophyllales and among basal monocots (Acorales and Alismatales). These aquatic orders do not have any extant terrestrial relatives. They originated from aquatic habitats during the Early Cretaceous. Asia would have been one of the centers for early diversification of aquatic angiosperms. The aquatic families within terrestrial orders may originate from other areas besides Asia, such as America or Australia. The lineages leading to extant angiosperms diversified early in underexploited freshwater habitats. The four extant aquatic orders were relicts of an early radiation of angiosperm in aquatic environments. Their extinct ancestors might be aquatic early angiosperms.

Key words: ancestral area, aquatic plant, early angiosperm, fossil age, origin.

Amborella, Nymphaeales, and Austrobaileyales represent the three earliest splits in angiosperm phylogeny (ANA grade angiosperms), followed by the five lineages of mesangiosperms, Magnoliids, Chloranthales, monocots, Ceratophyllales, and eudicots (Cantino et al., 2007; Soltis et al., 2008). The rapid diversification of angiosperms in the Early Cretaceous is well documented in the fossil record (Feild & Arens, 2005). Due to the diversity of fossils, it is impossible to draw unequivocal conclusions on the life form of early angiosperms (Löhne, 2006). There are two divergent views on the general habit of the earliest angiosperms: woody and terrestrial or herbaceous and aquatic (Soltis et al., 2005). The hypothesis that the earliest angiosperms were woody is supported by the evidence that most basal angiosperms are woody and all gymnosperms are woody (Soltis et al., 2008).

An aquatic origin of angiosperms is supported by the evidence that several of the earliest known fossil angiosperms were aquatic. *Archaefructus* represents one of the oldest, most complete angiosperm fossils (Sun et al., 2002). It is estimated to be approximately 125 million years old, and on the basis of morphology, it clearly was aquatic. The phylogenetic placement of *Archaefructus* as sister to all extant angiosperms (Sun et al., 2002), and the near-basal phylogenetic position of Nymphaeales, support the hypothesis that the aquatic habit arose early in angiosperms and the earliest angiosperms might be aquatic (Coiffard et al., 2007; Soltis et al., 2008).

Aquatic plants are plants that have adapted to live in aquatic environments (freshwater or saltwater). These plants require special adaptations for living submerged in water, or at the water's surface (Sculthorpe, 1967; Cook, 1990). Although aquatic plants are discussed as a unified biological group, the ways that species have evolved to live in the aquatic environment are diverse (Sculthorpe, 1967; Philbrick & Les, 1996). Aquatic plants are phylogenetically well dispersed across the angiosperms, with at least 50 independent origins, although they comprise less than 2% of the angiosperm species (Cook, 1990; Les et al., 1997).

Traditional systematic studies proposed that all aquatic plants evolved from terrestrial relatives (Sculthorpe, 1967; Cook, 1990). However, some recent phylogenetic and paleobotanical studies suggested the aquatic origin of angiosperms. Therefore, it is necessary for researchers of aquatic plants to examine the phylogenetic tree of vascular plants to find whether all the aquatic angiosperms were evolved from terrestrial relatives, and to explore the possibility of an aquatic origin of angiosperms. In this study, we annotated the aquatic taxa on the tree of vascular plants, and extracted the topology of these aquatic lineages to construct the tree of aquatic angiosperms. We also reconstructed the ancestral areas of aquatic families. Our aim is to study the origins of aquatic angiosperms and find whether there is any possibility that some aquatic lineages might originate from aquatic ancestors.

Material and Methods

We annotated the aquatic taxa on the tree of vascular plants (Fig. S1), and extracted the topology of these aquatic lineages to construct the tree of aquatic angiosperms. These 40 families used in this study represent the majority of aquatic angiosperms that are obligately living in water (Cook, 1990). Amphibious plants are distinct from aquatic species that live constantly in water, and most amphibious plants were not included in this study. The known fossil ages of aquatic families and orders were annotated on the tree (Lumbert et al., 1984; Crabtree, 1987; Friis et al., 2001, 2004; Gandolfo et al., 2004; Riley & Stockey, 2004; Sille et al., 2006; Taylor et al., 2008). A fossil species *Sinocarpus decussatus* Leng & Friis of the Late Barremian–Early Aptian age (125 Ma) was chosen as the oldest eudicot fossil (Leng & Friis, 2003). We used the statistical dispersal–vicariance analysis (S-DIVA) option in RASP (Yu et al., 2011) to construct the ancestral areas of aquatic families and orders. The distributions were categorized into the following areas: Asia (A), Africa (B), North America (C), South America (D), Australia (E), and Europe (F).

Results

Annotation of aquatic taxa on the tree of vascular plants

The overall phylogeny of angiosperms (see Fig. S1) was congruent with other angiosperm phylogenetic studies (Jansen et al., 2007; Moore et al., 2007; Graham & Iles, 2009). Nymphaeales diverged from the near basal node of the angiosperm tree. Monocots and eudicots were well supported as monophyletic. Acorales and Alismatales were successive sister groups of the remaining monocots. *Ceratophyllum* was placed sister to eudicots. From the annotated tree (see Fig. S1), we could find that aquatic plants evolved from different ecological backgrounds at different times. Some old ones are aquatic at the level of order or family, while other recent ones are isolated genera or species in the terrestrial family. When studying the origins of aquatic plants, the aquatic orders and the aquatic taxa in terrestrial orders should be treated differently. The fossil ages suggested that the aquatic orders (Nymphaeales, Alismatales, and Ceratophyllales) dated from the Early Cretaceous (Fig. 1).

Fossil ages and ancestral areas of Nymphaeales

The order Nymphaeales was considered to be among the oldest independent lineages of angiosperms. Recent molecular study placed the family Hydatellaceae (*Trithuria*) in Nymphaeales, as sister to traditional Nymphaeales (Saarela et al., 2007). The stem lineage to Nymphaeales is old, based on a fossil attributed to Nymphaeales from the Early Cretaceous (Friis et al., 2001). *Pluricarpellatia peltata* B. Mohr, Bernardes-de-Oliveira & David W. Taylor is another nymphaealean fossil from the Early Cretaceous (late Aptian to earliest Albian). It was an extinct member of Nymphaeales and branched off early in this lineage (Mohr et al., 2008). The Cretaceous age of the Cabombaceous fossils was supported by *Scutifolium jordanicum* Taylor, Brenner & Basha, a fossil species of Cabombaceae from the Albian (100 Ma; Taylor et al., 2008). The oldest fossil in Nymphaeaceae was some fossil flowers from the Turonian (90 Ma; Gandolfo et al., 2004).

The diversification of the Nymphaeales crown group started in Asia and Australia (Fig. 1). The ancestors of core Nymphaeales distributed in Asia and North America. There were two distinct radiation events in core Nymphaeales, a rapid first differentiation into two major lineages (Cabombaceae and Nymphaeaceae) and the radiation of Nymphaeaceae (*Nuphar*, *Barclaya*, *Nymphaea*, *Ondined*, *Victoria*, and *Euryale*). Cabombaceae were ancestrally distributed in the American continents (Fig. 1). After the two genera diverged, *Cabomba* diversified on the American continents. *Brasenia* occurred on all continents except Europe and Antarctica. The fossil record supported that *Brasenia* was distributed in Europe and extinct during glaciations (Yoo et al., 2005). The second radiation probably started in the Northern Hemisphere. *Nuphar* retained its ancestral distribution in the Northern Hemisphere. Four of the remaining five genera in Nymphaeaceae were distributed in small geographic areas. *Barclaya* and *Euryale* were found only in Southeast Asia. *Ondinea* occurred only in Western Australia, and *Victoria* was native to South America. *Nymphaea* had a cosmopolitan distribution, due to its ability to adapt to a wider range of temperatures than other genera of Nymphaeaceae (Yoo et al., 2005).

Fossil ages and ancestral areas of Alismatales

The order Alismatales includes core Alismatales and two other families, Araceae and Tofieldiaceae. Core Alismatales consists of 12 families and 56 genera. Species of Alismatales are wetland or aquatic herbs, most of which have a completely submerged seedling phase. All marine angiosperms and most water-pollinated angiosperms are confined to this order. Previous studies proposed that Alismatales originated in the Early Cretaceous (Janssen & Bremer, 2004; Magallon & Castillo, 2008), with all the families presented in the Late Cretaceous and Early Tertiary periods (Les et al., 2003; Janssen & Bremer, 2004; Chen et al., 2012a, b). A fossil species *Mayoa portugallica* Friis, Pedersen & Crane of Late Barremian–Early Aptian age (125 Ma) was the oldest fossil found in Alismatales (Friis et al., 2004).

The RASP results suggested that the ancestor of core Alismatales probably occurred in Asia in the Early Cretaceous (Fig. 1). The core Alismatales split into two lineages. The lineage comprising Butomaceae, Alismataceae, and Hydrocharitaceae originated in Asia. Its first clade (Butomaceae and Alismataceae) dispersed from Asia to South America. Alismataceae originated in Asia and South America, and dispersed to Europe and Africa. The oldest fossil in Alismataceae was a fossil similar to *Limnocharis* from the Late Cretaceous (Riley & Stockey, 2004). The second clade (Hydrocharitaceae) dispersed from Asia to Africa. In Hydrocharitaceae, the seagrasses diversified in Asia and dispersed to other regions by ocean currents. The oldest fossil in Hydrocharitaceae was a fossil of *Stratiotes*, which was 0.1 Ma younger than the Paleocene–Eocene boundary (55.9 Ma; Sille et al., 2006).

Another lineage comprising the remaining families dispersed from Asia to Australia (Fig. 1). Australia played an important role in the diversification of this lineage. Several seagrass families (Posidoniaceae, Cymodoceaceae, and Zosteraceae) originated from Australia. Previous studies proposed that seagrasses had independently arisen from their freshwater relatives in the course of habitat alteration from

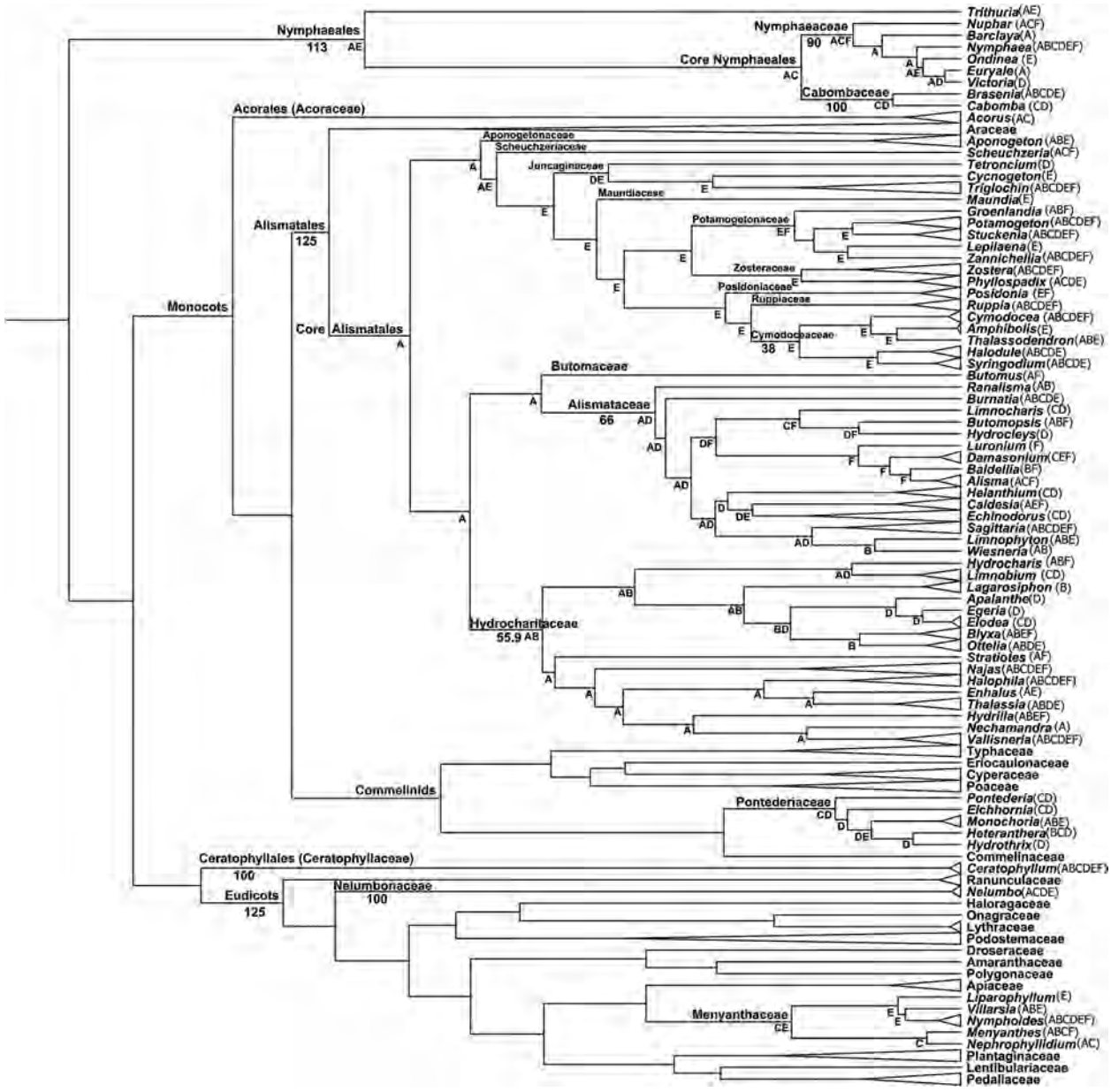


Fig. 1. Fossil ages and ancestral areas of aquatic families and orders. The known fossil ages of aquatic families and orders are below their names. Because the fossil records of some aquatic families are problematic, they are not used in this study. The timescale is in million years ago. Ancestral areas of aquatic families are below the nodes. We categorized the distributions into the following areas: Asia (A), Africa (B), North America (C), South America (D), Australia (E), and Europe (F). The ancestral areas of terrestrial families that have aquatic taxa were not reconstructed.

fresh water to salty water (Les et al., 1997; Kato et al., 2003). The fossil records of these families were problematic, except a *Thalassodendron* fossil of Cymodoceaceae from the Middle Eocene (Lumbert et al., 1984).

Ancestral areas of aquatic families in terrestrial orders

In monocots and eudicots, there are some families in terrestrial orders composed entirely of aquatic plants, e.g., Pontederiaceae (Commelinales), Nelumbonaceae (Proteales), and Menyanthaceae (Asterales) (see Fig. S1). Their geographical distributions are annotated in Fig. 1. Nelumbonaceae had a disjunct distribution in Asia and North America. The fossil record of Nelumbonaceae can date back to the Albain (100 Ma; Crabtree, 1987). The ancestral areas of two aquatic families Pontederiaceae and Menyanthaceae were reconstructed. Pontederiaceae originated and diversified in American continents (Fig. 1). The five genera of Menyanthaceae split into two clades. One clade originated in North America and another clade in Australia (Fig. 1). These aquatic families within terrestrial orders might originate from other areas besides Asia. Some terrestrial families also contained aquatic genera, e.g., *Hygroryza* (Poaceae), *Batrachium* (Ranunculaceae),

Myriophyllum (Haloragidaceae), *Utricularia* (Lentibulariaceae), and *Trapella* (Pedaliaceae). These aquatic genera evolved from their terrestrial ancestors that adapted to the aquatic environments.

Discussion

Origins of extant aquatic orders

The aquatic angiosperms split into three lineages, and the aquatic orders are at the basal node of each lineage with Early Cretaceous origins. Nymphaeales diverged from the near-basal node of the extant angiosperm phylogenetic tree. Considering the diversity and the nearly global distribution of its members, Nymphaeales stands out as the first globally diverse clade in the tree of extant angiosperms (Löhne et al., 2008). Other basal angiosperm lineages, such as Amborellales or Austrobaileyales, are restricted to much narrower relict ranges (Löhne et al., 2008). The placement of *Trithuria* (Hydatellaceae) in the Nymphaeales indicates that water lilies are part of a larger lineage than previously recognized (Saarela et al., 2007). It greatly expands the morphological diversity encompassed by the Nymphaeales. The adaptation of Nymphaeales to seasonal drying pools triggered the evolution of rapid growth via herbaceousness and condensation of the reproductive axes, which underlay the ecological success of angiosperms (Feild & Arens, 2007).

Beyond the all-aquatic Nymphaeales, basal monocots (*Acorus* and Alismatales), *Ceratophyllum* and basal eudicot lineages (*Nelumbo* and some *Ranunculus* species) are other examples of near-basal aquatic angiosperms (Les & Schneider, 1995; Soltis et al., 2005). Monocots may be relicts of an early radiation of herbaceous angiosperms (Doyle, 1973). Acorales and Alismatales are successive sister groups of the remaining monocots. Acorales distributed widely in north temperate wetlands. Alismatales globally distributed in the aquatic environments. These two lineages differentiated early in the monocots. Therefore, monocots are likely to be of aquatic origin, followed by gradual evolution to adapt to the terrestrial environment (Les & Schneider, 1995). This hypothesis is consistent with the fossil records of monocots (Doyle, 1973).

In another study, we used Mesquite to reconstruct the ancestral character state of life form in aquatic plants (Du & Wang, 2014). The submerged life form was suggested as the progenitorial state of Nymphaeales and Alismatales, which gave rise to floating-leaved, free-floating, and emergent life forms. This is in accordance with their aquatic origins. Most of the aquatic families in terrestrial orders evolved from emergent life forms, which is consistent with their terrestrial origins. For example, the ancestral state of Menyanthaceae was emergent, which gave rise to the floating-leaved life form.

In the past, the genera *Ceratophyllum* and *Nelumbo* were assumed to be closely related to water lilies and were included in the Nymphaeales. However, similarities among these taxa were due to parallel adaptations to aquatic environments. *Ceratophyllum*, the only extant genus in the Ceratophyllaceae, possesses unique morphological characters, which suggests that it has an isolated position and an early divergence from other angiosperms (Les, 1988; Dilcher & Wang, 2009). *Ceratophyllum* species are submerged aquatic plants widely distributed in freshwater habitats around the world. A fossil species of Ceratophyllaceae, *Donlesia dakotensis* Dilcher & Wang was described from the Albian (100 Ma; Dilcher & Wang, 2009). Paleobotanical evidence supports the hypothesis that Ceratophyllaceae are relicts of early angiosperms (Les, 1988; Dilcher & Wang, 2009). Phylogenetic studies indicate that Ceratophyllales are one of the five lineages of mesangiosperms (Soltis et al., 2008).

Ancestral areas and dispersal of aquatic angiosperms

The plants in aquatic orders do not have any extant terrestrial relatives. They originated from aquatic habitats during the Early Cretaceous. Asia would have been one of the centers for early diversification of aquatic angiosperms, which is in accordance with the geographic origin of angiosperms. In terrestrial orders, aquatic families and genera evolved from terrestrial ancestors that adapted to aquatic habitats. They may originate from other areas besides Asia, such as America or Australia.

Although aquatic plants comprise less than 2% of the angiosperm species, they represent a disproportionately large number of taxa with global distributions including all sorts of intercontinental disjunctions (Les et al., 2003). The origin times of most aquatic species are far too recent to implicate continental drift as the major determinant of their discontinuous distributions. The modern continents (Africa, Eurasia, Australia, South America, and North America) have been separated by oceans since at least 105 Ma (Davis et al., 2002). Therefore, the transoceanic distribution of most aquatic species might have resulted from dispersals through land bridges, island chains, or long-distance dispersal (Les et al., 2003).

Early radiation of aquatic angiosperms

Molecular phylogeny indicates that angiosperms are not closely related to any other extant seed plant group. Therefore, the information from fossils might provide the basis for reconstructing the origins of angiosperms (Friis et al., 2003). Prior to the angiosperms, the ferns and gymnosperms had been widely distributed on land. The fossil records indicate that stoneworts and green algae were already present in the Early Palaeozoic, but plant evolution in water bypassed two of the main steps in terrestrial plant evolution: the dominance of the ferns during the Palaeozoic, and the dominance of the gymnosperms during the Mesozoic (Martín-Closas, 2003). The late arrival of aquatic ferns, and complete failure to develop aquatic species in gymnosperms, differentiated the evolution of aquatic plants from terrestrial plants (Martín-Closas, 2003).

Angiosperms colonized freshwater until the Late Mesozoic. Angiosperm ancestors may have adopted the aquatic lifestyle to escape competition on land (Aquatic hypothesis; Feild & Arens, 2005), or early angiosperms may have escaped crowded habitats on land by adaptation to water (Terrestrial hypothesis; Feild & Arens, 2007) (Fig. 2). The early appearance of aquatic angiosperms supports the view that early angiosperms had a growth habit of rhizomatous herbs, which could be easily transformed into aquatic habit (Doyle, 2012). Taylor & Hickey (1992) also suggested that early angiosperms were perennial rhizomatous herbs that competed with ferns in stream margins. Because of their rhizomatous growth habit and efficient seedlings, these herbaceous species invaded the aquatic habitats in the Early Cretaceous.

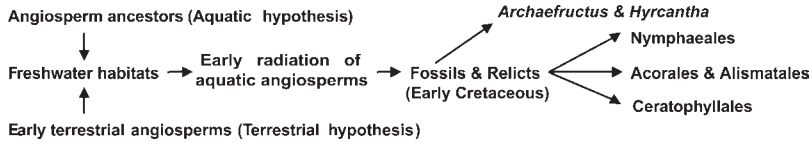


Fig. 2. Early radiation of aquatic angiosperms.

Fossil records from Asia, Europe, and America all show that angiosperms occurred in aquatic environments (freshwater lakes, swamps, and floodplains) during the Early Cretaceous (Coiffard et al., 2012). Recently, an early angiosperm fossil genus *Hyrcantha* that lived in Asia 125–122 Ma was discovered (Dilcher et al., 2007). The genus appear to be aquatic, living in shallow water (20–40 cm deep), with the terminal fruiting axes exposed above the water. The earliest well-documented fossil records of angiosperms from Asia, such as *Archaefructus* and *Hyrcantha*, are significant in our understanding of the primitive characters (aquatic and herbaceous) of early angiosperms (Sun et al., 2008), and the early ecological radiation of angiosperms, especially the existence of multiple early trends toward an aquatic habit (Friis et al., 2003).

These early angiosperm fossils and the near-basal phylogenetic position of aquatic orders suggest that the aquatic habit arose early (Fig. 2). Whether the extinct ancestors of these early aquatic lineages were aquatic or terrestrial primitive angiosperms remains unclear. But the lineages leading to extant angiosperms diversified early in underexploited freshwater habitats (Fig. 2). The four extant aquatic orders are relicts of an early radiation of angiosperms in aquatic environments.

Conclusions

When we annotated the aquatic taxa on the tree of vascular plants, we found that aquatic angiosperms could be divided into two different categories: the four aquatic orders and the aquatic taxa in terrestrial orders. Traditional systematic studies proposed that all aquatic plants evolved from terrestrial relatives. However, molecular phylogeny has suggested that the phylogenetic positions of some aquatic lineages have changed. In the past, *Ceratophyllum* and *Nelumbo* were included in the Nymphaeales, which was considered a eudicot order. More recently, Ceratophyllales are considered one of the five lineages of mesangiosperms, and Nymphaeales are one of the three clades of basal angiosperms. Acorales and Alismatales are the basal clades of monocots. Because of new findings in phylogeny and paleobotany, the origin times of aquatic orders are thought to be earlier than most terrestrial angiosperms. The plants in aquatic orders do not have any extant terrestrial relatives. They were relicts of an early radiation of angiosperms in aquatic habitats, and might originate from aquatic ancestors. There is some possibility that angiosperms had an aquatic origin.

Acknowledgements

The authors thank Prof. Zhi-Duan Chen and Jin-Ming Chen for their helpful discussion on the manuscript. We are grateful to the members in the Laboratory of Prof. Zhi-Duan Chen for their assistance in the laboratory and data analysis. The National Natural Science Foundation of China (Grant No. 31100175) and the strategic pilot science and technology projects of the Chinese Academy of Sciences (Grant No. XDA05090305) supported this research.

References

Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis DE, Soltis PS, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56: 822–846.

Chen LY, Chen JM, Gituru RW, Temam TD, Wang QF. 2012a. Generic phylogeny and historical biogeography of Alismataceae, inferred from multiple DNA sequences. *Molecular Phylogenetics and Evolution* 63: 407–416.

Chen LY, Chen JM, Gituru RW, Wang QF. 2012b. Generic phylogeny, historical biogeography and character evolution of the cosmopolitan aquatic plant family Hydrocharitaceae. *BMC Evolutionary Biology* 12: 30. Doi: 10.1186/1471-2148-12-30.

Coiffard C, Gomez B, Daviero-Gomez V, Dilcher DL. 2012. Rise to dominance of angiosperm pioneers in European Cretaceous environments. *Proceedings of the National Academy of Sciences USA* 109: 20955–20959.

Coiffard C, Gomez B, Thevenard F. 2007. Early Cretaceous angiosperm invasion of Western Europe and major environmental changes. *Annals of Botany* 100: 545–553.

Cook CDK. 1990. *Aquatic plant book*. Amsterdam: SPB Academic Publishing.

Crabtree DR. 1987. Angiosperms of the Northern Rocky Mountains: Albian to Campanian (Cretaceous) megafossil floras. *Annals of the Missouri Botanical Garden* 74: 707–747.

Davis CC, Bell CD, Mathews S, Donoghue MJ. 2002. Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae. *Proceedings of the National Academy of Sciences USA* 99: 6833–6837.

Dilcher DL, Sun G, Ji Q, Li H. 2007. An early infructescence *Hyrcantha decussata* (comb. nov.) from the Yixian Formation in northeastern China. *Proceedings of the National Academy of Sciences USA* 104: 9370–9374.

Dilcher DL, Wang H. 2009. An Early Cretaceous fruit with affinities to Ceratophyllaceae. *American Journal of Botany* 96: 2256–2269.

Doyle JA. 1973. Fossil evidence on early evolution of the Monocotyledons. *Quarterly Review of Biology* 48: 399–413.

Doyle JA. 2012. Molecular and fossil evidence on the origin of angiosperms. *Annual Review of Earth and Planetary Sciences* 40: 301–326.

Du ZY, Wang QF. 2014. Correlations of life form, pollination mode and sexual system in aquatic angiosperms. *PLoS ONE* 9: e115653. Doi: 10.1371/journal.pone.0115653.

Feild TS, Arens NC. 2005. Form, function and environments of the early angiosperms: Merging extant phylogeny and ecophysiology with fossils. *New Phytologist* 166: 383–408.

Feild TS, Arens NC. 2007. The ecophysiology of early angiosperms. *Plant, Cell and Environment* 30: 291–309.

Friis EM, Doyle JA, Endress PK, Leng Q. 2003. *Archaefructus* – angiosperm precursor or specialized early angiosperm? *Trends in Plant Science* 8: 369–373.

Friis EM, Pedersen KR, Crane PR. 2001. Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. *Nature* 410: 357–360.

Friis EM, Pedersen KR, Crane PR. 2004. Araceae from the Early Cretaceous of Portugal: Evidence on the emergence of monocotyledons. *Proceedings of the National Academy of Sciences USA* 101: 16565–16570.

Gandolfo MA, Nixon KC, Crepet WL. 2004. Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. *Proceedings of the National Academy of Sciences USA* 101: 8056–8061.

Graham SW, Iles WJ. 2009. Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny. *American Journal of Botany* 96: 216–227.

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee SB, Peery R, McNeal JR, Kuehl JV, Boore JL. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences USA* 104: 19369–19374.

Janssen T, Bremer K. 2004. The age of major monocot groups inferred from 800+ *rbc*L sequences. *Botanical Journal of the Linnean Society* 146: 385–398.

Kato Y, Aioi K, Omori Y, Takahata N, Satta Y. 2003. Phylogenetic analyses of *Zostera* species based on *rbc*L and *mat*K nucleotide sequences: Implications for the origin and diversification of seagrasses in Japanese waters. *Genes and Genetic Systems* 78: 329–342.

Leng Q, Friis EM. 2003. *Sinocarpus decussatus* gen. et sp. nov., a new angiosperm with basally syncarpous fruits from the Yixian Formation of Northeast China. *Plant Systematics and Evolution* 241: 77–88.

Les DH. 1988. The origin and affinities of the Ceratophyllaceae. *Taxon* 37: 326–345.

Les DH, Cleland MA, Waycott M. 1997. Phylogenetic studies in Alismatidae, II: Evolution of marine angiosperms (seagrasses) and hydrophily. *Systematic Botany* 22: 443–463.

Les DH, Crawford DJ, Kimball RT, Moody ML, Landolt E. 2003. Biogeography of discontinuously distributed hydrophytes: A molecular appraisal of intercontinental disjunctions. *International Journal of Plant Sciences* 164: 917–932.

Les DH, Schneider EL. 1995. The Nymphaeales, Alismatidae, and the theory of an aquatic monocotyledon origin. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ eds. *Monocotyledons: Systematics and evolution*. Richmond: Royal Botanic Gardens, Kew. 23–42.

Löhne C. 2006. *Molecular phylogenetics and historical biogeography of basal angiosperms: A case study in Nymphaeales*. Ph.D. Dissertation. Bonn: Universität Bonn.

Löhne C, Yoo MJ, Borsch T, Wiersema J, Wilde V, Bell CD, Barthlott W, Soltis DE, Soltis PS. 2008. Biogeography of Nymphaeales: Extant patterns and historical events. *Taxon* 57: 1123–1146.

Lumbert SH, Hartog CD, Phillips RC, Olsen FS. 1984. The occurrence of fossil seagrasses in the Avon Park Formation (late Middle Eocene), Levy County, Florida (USA). *Aquatic Botany* 20: 121–129.

Magallon S, Castillo A. 2008. Angiosperm diversification through time. *American Journal of Botany* 96: 349–365.

Martín-Closas C. 2003. The fossil record and evolution of freshwater plants: A review. *Geologica Acta* 1: 315–338.

Mohr BAR, Bernardes-de-Oliveira MEC, Taylor DW. 2008. *Pluricarpellatia*, a nymphaealean angiosperm from the Lower Cretaceous of northern Gondwana (Crato Formation, Brazil). *Taxon* 57: 1147–1158.

Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences USA* 104: 19363–19368.

Philbrick CT, Les DH. 1996. Evolution of aquatic angiosperm reproductive systems. *Bioscience* 46: 813–826.

Riley M, Stockey R. 2004. *Cardstonia tolmanii* gen. et sp. nov. (Limnocharitaceae) from the Upper Cretaceous of Alberta, Canada. *International Journal of Plant Sciences* 165: 897–916.

Saarela JM, Rai HS, Doyle JA, Endress PK, Mathews S, Marchant AD, Briggs BG, Graham SW. 2007. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature* 446: 312–315.

Sculthorpe CD. 1967. *The biology of aquatic vascular plants*. New York: St. Martin's Press.

Sille NP, Collinson ME, Kucera M, Hooker JJ. 2006. Morphological evolution of *Stratiotes* through the paleogene in England: An example of microevolution in flowering plants. *Palaios* 21: 272–288.

Soltis DE, Bell CD, Kim S, Soltis PS. 2008. Origin and early evolution of angiosperms. *Annals of the New York Academy of Sciences* 1133: 3–25.

Soltis DE, Soltis PS, Endress PK, Chase MW. 2005. *Phylogeny and evolution of the angiosperms*. Sunderland: Sinauer Associates.

Sun G, Dilcher DL, Zheng S-L. 2008. A review of recent advances in the study of early angiosperms from northeastern China. *Palaeoworld* 17: 166–171.

Sun G, Ji Q, Dilcher DL, Zheng SL, Nixon KC, Wang XF. 2002. Archaefructaceae, a new basal angiosperm family. *Science* 296: 899–904.

Taylor DW, Brenner GJ, Basha SH. 2008. *Scutifolium jordanicum* gen. et sp. nov. (Cabombaceae), an aquatic fossil plant from the Lower Cretaceous of Jordan, and the relationships of related leaf fossils to living genera. *American Journal of Botany* 95: 340–352.

Taylor DW, Hickey LJ. 1992. Phylogenetic evidence for the herbaceous origin of angiosperms. *Plant Systematics and Evolution* 180: 137–156.

Yoo MJ, Bell CD, Soltis PS, Soltis DE. 2005. Divergence times and historical biogeography of Nymphaeales. *Systematic Botany* 30: 693–704.

Yu Y, Harris AJ, He XJ. 2011. RASP (Reconstruct Ancestral State in Phylogenies) 1.1 [online]. Available from http://mnh.scu.edu.cn/soft/blog/RASP [accessed 12 September 2013].

Appendix

Members of the China Phylogeny Consortium

Zhi-Duan CHEN, An-Ming LU, Hong-Zhi KONG, Xiao-Quan WANG, Yin-Zheng WANG, Shi-Liang ZHOU
Institute of Botany, Chinese Academy of Sciences, Beijing, China
Shou-Zhou ZHANG, Xiao-Ming WANG
Fairylake Botanical Garden, Shenzhen, China
Zhong-Jian LIU
The Orchid Conservation and Research Center of Shenzhen, Shenzhen, China

Qing-Feng WANG
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
Jian-Hui LI
Computer Network Information Center, Chinese Academy of Sciences, Beijing, China
De-Zhu LI, Ting-Shuang YI
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Hong MA
Fudan University, Shanghai, China
Douglas E. SOLTIS, Pamela S. SOLTIS
University of Florida, Gainesville, USA
Jian-Hua LI
Hope College, Holland, USA
Cheng-Xin FU
Zhejiang University, Hangzhou, China
Qi-Xin LIU
Nanjing Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12182/suppinfo:

Fig. S1. Annotation of aquatic taxa on the tree of vascular plants. This tree is edited from Zhi-Duan Chen et al. (pers. comm.). Aquatic families are in red. Terrestrial families that contain aquatic taxa are in blue.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12187

Research Article

A molecular phylogeny of Chinese orchids

Ming-He Li[1,2,3†], Guo-Qiang Zhang[1,2,3†], Si-Ren Lan[1,3*], Zhong-Jian Liu[1,2*], and China Phylogeny Consortium[§]

[1]College of Landscape, Fujian Agriculture and Forestry University, Fuzhou 350002, China

[2]Shenzhen Key Laboratory for Orchid Conservation and Utilization, The National Orchid Conservation Center of China and The Orchid Conservation and Research Center of Shenzhen, Shenzhen 518114, China

[3]Orchid Conservation and Research Center of Fujian Agriculture and Forestry University, Fuzhou 350002, China

[†]These authors contributed equally to this work.

[§]Members of the China Phylogeny Consortium are listed in the Appendix.

*Authors for correspondence. S.-R. Lan. E-mail: lkzx@fafu.edu.cn. Tel.: 86-591-83789201. Fax: 86-591-83768188. Z.-J. Liu. E-mail: liuzj@sinicaorchid.org. Tel.: 86-755-25712359. Fax: 86-755-25711928.

Received 28 July 2015; Accepted 6 November 2015; Article first published online 24 January 2016

Abstract We estimated the molecular phylogenetic relationships of the Chinese members of the family Orchidaceae. Within the Tree of Life for the Genera of Chinese Vascular Plants using *atpB, rbcL, matK, ndhF*, and *matR*, the currently delimited subfamilies, tribes, and subtribes are highly supported as monophyletic except for the perplexing Epidendroideae. Five genes (*rbcL*, *matK*, *psaB*, *ycf1*, and *Xdh*), which are more universally used in Orchidaceae, were further analyzed to reconstruct the phylogeny of Epidendroideae. The reconstructed trees were in strong agreement and showed significant support for the tribal and subtribal clades. Based on the highly supported circumscription and arrangement of the suprageneric levels in the Tree of Life and reconstructed trees, we have proposed a new phylogenetic classification of Chinese Orchidaceae that includes five subfamilies, 17 tribes, and 21 subtribes.

Key words: Anthogoniinae, *Arundina*, China, Epidendroideae, orchid phylogenetics.

Orchidaceae represent a large and diverse taxon of flowering plants and include over 800 genera and 26 000 species (Govaerts et al., 2014). They are particularly diverse in wet tropical regions but absent in polar regions and the driest deserts (Chase et al., 2006), accounting for approximately 10% of seed plants (Roberts & Dixon, 2008). This complex family presents a considerable challenge to systematists interested in phylogenetic reconstruction and classification. All of the pre-DNA era classifications of Orchidaceae were based on a relatively small set of morphological aspects and features, particularly on the column and pollinarium. The morphological analyses of Freudenstein & Rasmussen (1999) indicated that the high degree of hierarchical structure indicated in previous classifications of Orchidaceae was not warranted because the cladistic analyses of the morphologic data showed limited resolution at lower taxonomic levels.

Orchids have recently been the focus of a greater number of published DNA phylogenetic studies than any other family of angiosperms. In terms of the general patterns of orchid relationships, published DNA phylogenetic studies have shown remarkably similar results. Recent molecular phylogenetic analyses of the representative groups of monocotyledons have revealed that Orchidaceae are related to the order Asparagales (APG III, 2009). DNA analyses of orchids (Chase et al., 2003; Cameron, 2004; Górniak et al., 2010; Chase et al., 2015) have provided surprising findings for orchid taxonomists and supported the monophyly of the orchid family, including the apostasioids and cypripedioids. The analyses collectively support the following relationships: (Apostasioideae (Vanilloideae (Cypripedioideae (Orchidoideae, Epidendroideae)))), with Spiranthoideae nested within a more broadly defined Orchidoideae and Tropidieae transferred to Epidendroideae from Orchidoideae.

Chinese Orchidaceae contain approximately 200 genera with over 1450 species (Zhang et al., 2015b), accounting for approximately 25% and 5% of the genera and species worldwide, respectively; they are one of the largest families of Chinese seed plants. In the last century, the frame arrangement of pre-DNA era Chinese orchid classifications, *Florae Reipublicae Popularis Sinicae* (FRPS) (Chen et al., 1999; Lang et al., 1999; Tsi et al., 1999), were primarily constructed based on the system of Seidenfaden (1992) and Seidenfaden et al. (1992), although the components of the subtribes were primarily determined using the Dressler system (1993). The Dressler system has periodically been modified to include increasing data on the morphology and anatomy of orchids, which has been arranged into five subfamilies (Apostasioideae, Cypripedioideae, Orchidoideae, Spiranthoideae, and Epidendroideae), 22 tribes, and 70 subtribes. Although Chinese orchids were divided into three subfamilies (Apostasioideae, Cypripedioideae, and Orchidoideae) in the pre-DNA era, these plants covered all of the five subfamilies according to Dressler (1993). Spiranthoideae and Epidendroideae were classified as the subtribe Spiranthinae and the tribe Epidendreae, respectively, in Orchidoideae. Both Apostasioideae and Cypripedioideae contain two

genera, and Orchidoideae, which include the bulk of the taxa, were divided into four tribes: Neottieae, containing eight subtribes and 28 genera; Orchideae, containing three subtribes and 21 genera; Epidendreae, containing 31 subtribes and 73 genera; and Vandeae, containing one subtribe and 46 genera.

A tree of life for Chinese orchids is a pivotal platform for studying the phylogenetic diversity, the species relationships of community, and the distribution pattern of biodiversity of orchids in the Chinese area (Lu et al., 2014; Chen et al., 2016). In this century, new information from molecular studies has been added. The *Flora of China* (Chen et al., 2009) divided Chinese orchids into five subfamilies and redefined 194 genera and 1388 species based on molecular data. However, the tribes and subtribes remained undefined. The newly described orchids included 23 more genera (29 new genera and six combined genera) and 141 more species than FRPS. Subsequently, more than 10 combined genera and 20 new or newly recorded genera (e.g., Liu & Chen, 2009; Chen & Liu, 2010; Li et al., 2011, 2014a, 2015; Liu et al., 2011; Huang et al., 2012; Xiang et al., 2012; Yang et al., 2013; Zhai et al., 2013, 2014; Zhang et al., 2013, 2015a; Jin et al., 2014, 2015; Tang et al., 2015a), and 60 species (e.g., Hu et al., 2013, 2014; Jin et al., 2013; Li & Yan, 2013; Peng et al., 2013; Guan et al., 2014; Li et al., 2014b; Meng et al., 2014; Tian et al., 2014; Xu et al., 2014; Zou et al., 2014; Su et al., 2015) have been reported in recent years. The completion of the FRPS, *Flora of China*, and the updated taxa of orchids have contributed greatly to the understanding of Chinese orchid classification and diversification. Although molecular phylogenetic studies in the past 30 years have greatly improved our understanding of the classification systems (e.g., Chase et al., 2003, 2015) and evolution (e.g., Freudenstein & Chase, 2015; Givnish et al., 2015) of the whole orchid family, the Chinese orchids have never been rigorously tested using molecular data.

Here, we present a large phylogenetic analysis that represents nearly all of the genera of the Orchidaceae clade *sensu* Chen et al. (2009) in the Tree of Life for the Genera of Chinese Vascular Plants (hereafter referred to as "TL"), in which an analysis of the chloroplast genes *atpB*, *rbcL*, *matK*, and *ndhF*, and mitochondrial gene *matR* is carried out (see Chen et al., 2016). Due to the unclear relationships for the Epidendroideae clade in TL, a reconstructed phylogenetic tree of Epidendroideae was presented and used to infer more robust relationships based on the chloroplast genes *rbcL*, *matK*, *psaB*, and *ycf1*, and the nuclear gene *Xdh*. The aim of the present study was to elucidate the circumscription and arrangement of the suprageneric levels of Chinese orchids.

Material and Methods

Tree of Life for the Genera of Chinese Vascular Plants

In TL, a total of 6106 species representing 3118 genera (including 169 genera and 264 species of orchids) were sampled to reconstruct the tree of life of Chinese vascular plants at the generic level, to develop it as a useful platform for ecological and biological studies of plants in China. The details of the Material and Methods have been described in Chen et al. (2016).

Epidendroideae reconstruction

Current studies based on DNA sequence analyses in Orchidaceae have been carried out using two approaches. At the family level, studies have primarily focused on the coding plastid genes *rbcL* (Cameron et al., 1999), *matK* (Freudenstein et al., 2004), *psaB* (Cameron, 2004), and *ycf1* (Neubig et al., 2009), and the low-copy nuclear coding gene *Xdh* (Górniak et al., 2010). Conversely, at a lower taxonomic level (tribes and below), non-coding plastid markers have been used, including the *trnL* intron, the *trnL-F* intergenic spacer, and the nuclear ribosomal DNA internal transcribed spacer (ITS). Although there are various non-coding markers, they are also difficult or impossible to align with confidence across the family. Because the selected markers (*atpB*, *rbcL*, *matK*, *ndhF*, and *matR*) in TL are not all universally used in Orchidaceae, tribal and subtribal delimitation and relationships, especially in Epidendroideae, remain unresolved.

To reconstruct a robust phylogeny for high-level Chinese epidendroids, we sampled as many taxa as possible from all of the subtribes recognized within Epidendroideae (except for Epipogiinae), with a total of 73 genera and 91 species sampled. The outgroup included two genera from the subfamily Cypripedioideae and seven genera from the subfamily Orchidoideae. Five markers (*rbcL*, *matK*, *psaB*, *ycf1*, and *Xdh*) were used to assemble the data matrix. The sequence information and GenBank accession numbers are listed in Table S1. DNA extraction, polymerase chain reaction amplification, sequencing, sequence editing, and assembly were carried out according to Zhang et al. (2013). The primers used for the polymerase chain reaction analysis are listed in Table S2. The phylogenetic analyses were carried out using the maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI) methods. The MP analysis was carried out in paup version 4.0b10 (Swofford, 2003), and the ML analysis was undertaken using the CIPRES Science Gateway web server (RAxML-HPC2 on XSEDE 8.1.11) with 100 bootstrap replicates and settings that are described in Stamatakis et al. (2008) (Miller et al., 2010). The BI analysis was carried out using MrBayes version 3.2.6 (Ronquist & Huelsenbeck, 2003) and the CIPRES Science Gateway web server (MrBayes 3.2.6 on XSEDE) (Miller et al., 2010). The following settings were used: sampling frequency = 1000; tem = 0.1; burn-in = 2000; and number of Markov chain Monte Carlo generations = 10 000 000.

Results

Tree of Life for the Genera of Chinese Vascular Plants Analysis

The TL shows strong support for the monophyly of Chinese orchids and that this clade is sister to the six remaining Asparagales (Amaryllidaceae, Asparagaceae, Hypoxidaceae, Iridaceae, Ixioliriaceae, and Xanthorrhoeaceae), as shown in Chen et al. (2016). Within Orchidaceae (Fig. 1), the analysis highlights a clear differentiation between the five monophyletic sister subfamilies of Apostasioideae, Vanilloideae, Cypripedioideae, Orchidoideae, and Epidendroideae. The circumscriptions of the currently recognized tribes and subtribes were all supported except for that of the paraphyletic Calypsoinae (two supported groups) and

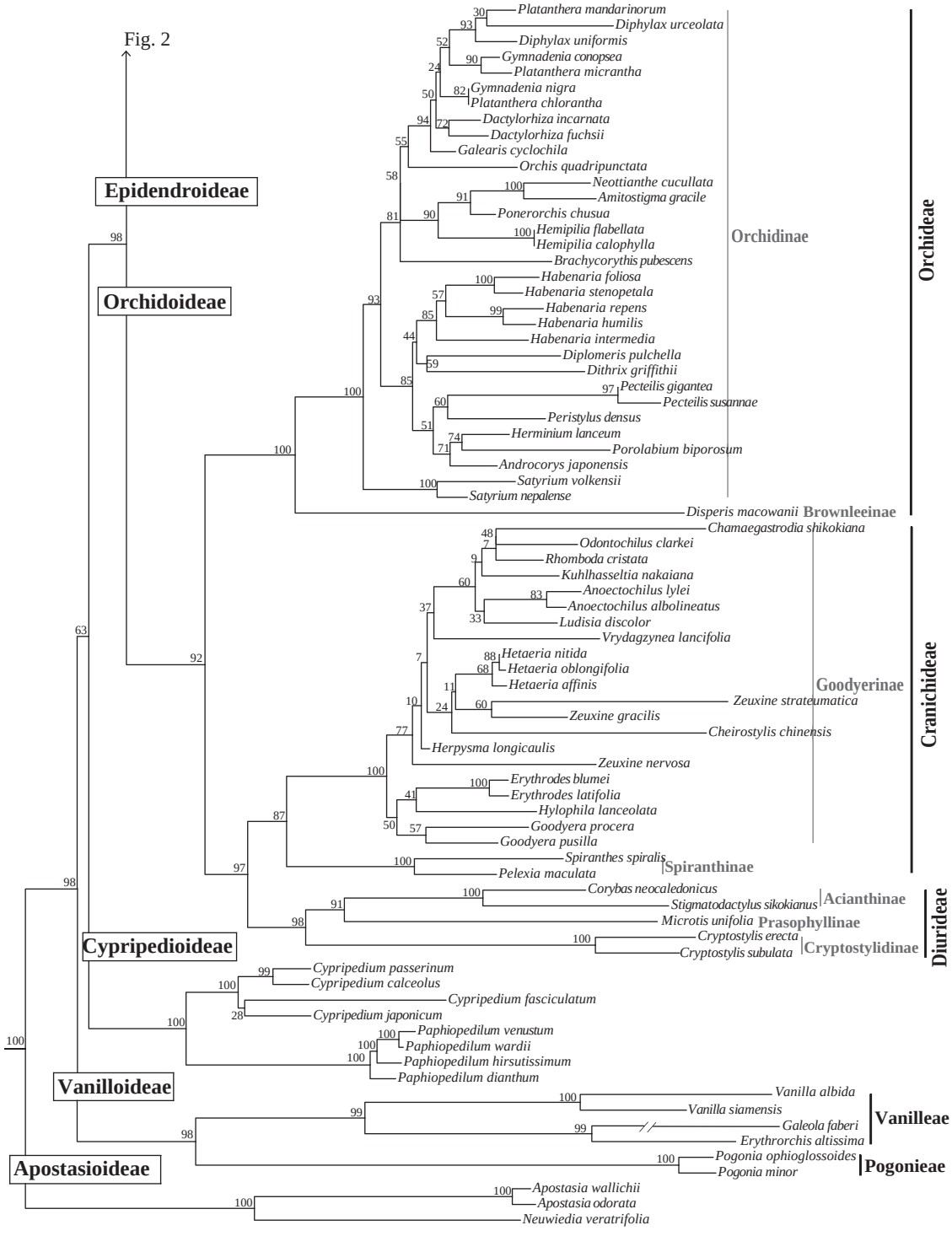


Fig. 1. Bootstrap consensus tree for the Chinese orchids (Apostasioideae, Vanilloideae, Cypripedioideae, and Orchidoideae) in the Tree of Life for the Genera of Chinese Vascular Plants. The subfamilies, tribes, and subtribes of the classification of Chase et al. (2015) are indicated. Bootstrap proportions are indicated near the nodes.

Epidendreae (including Agrostophyllinae) (Fig. 2). The relationships within the tribes and subtribes of Epidendroideae were difficult to infer, reflecting weak support (Fig. 2).

The first clade Apostasioideae, including *Neuwiedia* and *Apostasia* (bootstrap proportion (BS) = 100), is sister to all other Orchidaceae (100). The second clade Vanilloideae, including Vanilleae and Pogonieae (98), is sister to the remaining three subfamilies (98). The third clade Cypripedioideae, including two genera (100), shows moderate support (63). The fourth clade contains the majority of terrestrial monandrous orchids that are morphologically assigned to Orchidoideae and Spiranthoideae. The spiranthoid orchids

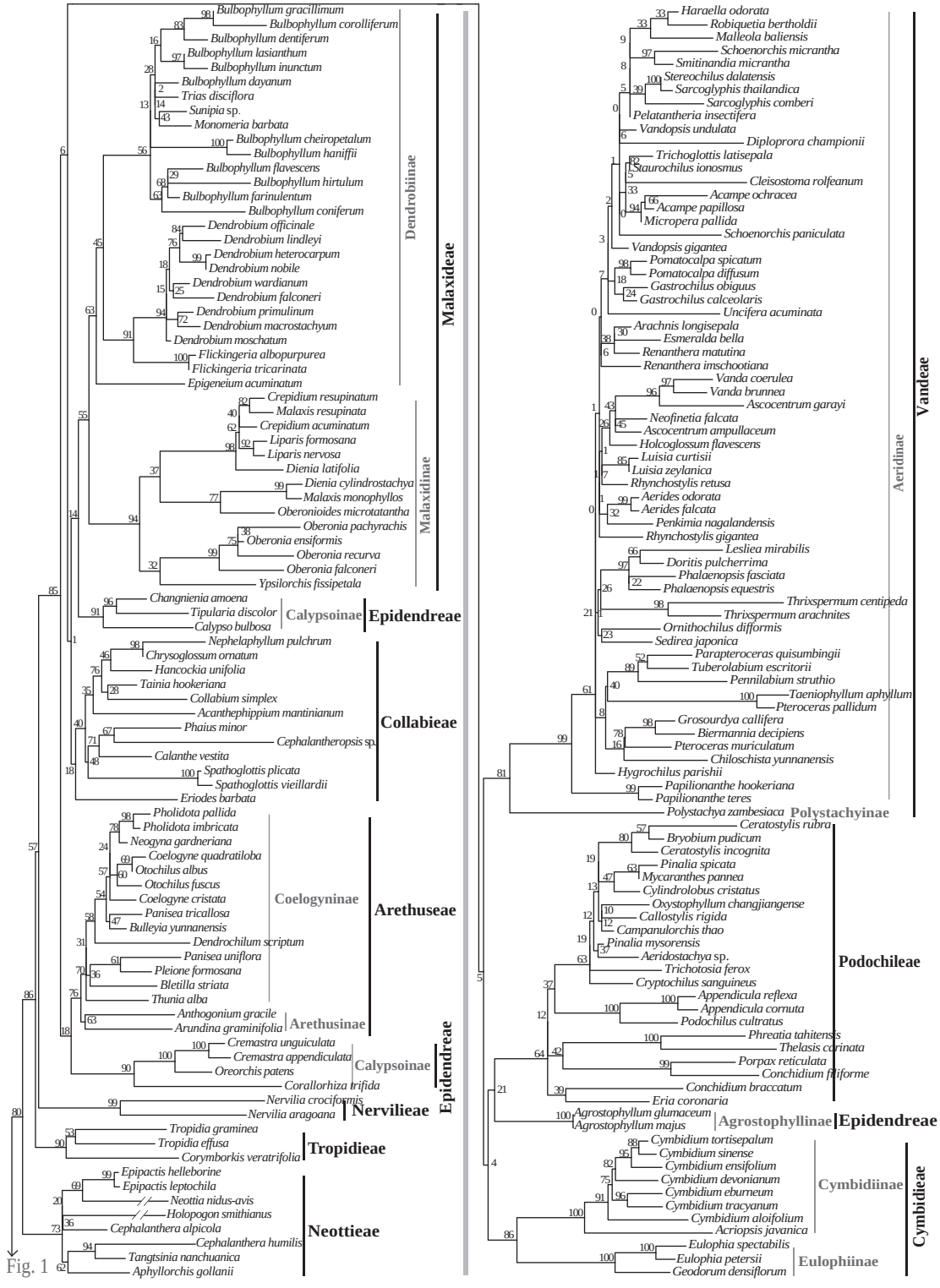


Fig. 1

Fig. 2. Bootstrap consensus tree for the Chinese orchids (Epidendroideae) in the Tree of Life for the Genera of Chinese Vascular Plants. The tribes and subtribes of the classification of Chase et al. (2015) are indicated. Bootstrap proportions are indicated near the nodes.

and Goodyerinae (87) are monophyletic and sisters (97) to Diurideae, and these orchids are sisters (92) to the remaining Orchidoideae (Orchideae). The last clade, Epidendroideae, contains taxa that are typically classified within this subfamily, and Neottieae and Tropidieae are also included, which had been morphologically classified as Neottioideae and Spiranthoideae, respectively. In addition, Vandeae are also embedded in this clade, and they have been morphologically

classified into the subfamily Vandoideae. Markedly low support has been observed; thus, the relationships of the tribes/subtribes are difficult to infer.

Epidendroideae reconstruction

Here, we reconstructed the relationships at the subtribe level of the Chinese epidendroids taxa to infer a more robust phylogenetic tree. The matrix included five relatively low substitution rate genes and consisted of 7580 aligned nucleotides (*rbcL* 1349 bp, *matK* 1674 bp, *psaB* 1673 bp, *ycf1* 1924 bp, and *Xdh* 960 bp), with more than 26% (1994 bp) of the parsimony-informative characters (*rbcL* 193 bp, *matK* 565 bp, *psaB* 212 bp, *ycf1* 599 bp, and *Xdh* 425 bp). A total of 381 DNA sequences (97 *rbcL*, 100 *matK*, 79 *psaB*, 50 *ycf1*, and 55 *Xdh*) were obtained.

The reconstructed trees (RT) are shown in Fig. 3 and presented as ML trees. The bootstrap proportions (BS_{ML}/BS_{MP}) and posterior probabilities (PP) are indicated near the nodes. The BS and PP analyses recovered similar patterns at the tribe/subtribe level except for two incongruent tribes (Nervilieae and Tropidieae) of lower epidendroids in the BI tree and three incongruent tribes (Collabieae, Podochileae, and Vandeae) of higher epidendroids in the MP tree. Most of the relationships among the subtribes and genera were strongly supported except for that of the genera of Aeridinae.

In the lower epidendroids, highly supported clades included a monophyletic clade ($BS_{ML} = 91/BS_{MP} = 56/PP = 1.00$) consisting of Neottieae sister to the remaining Epidendroideae, followed by Gastrodieae (60/33/1.00) and two other tribes (Nervilieae and Tropidieae) with relatively weak BS and PP. The last clade was Thaieae, which is a sister group to the higher epidendroids with high support (93/93/1.00).

In the higher epidendroids, the first clade, Arethuseae, was monophyletic as sister to all of the other higher epidendroids with strong support (87/58/1.00), followed by Malaxideae with strong support (87/56/1.00), and then Cymbidieae with moderate support (61/32/1.00). The next supported clade included Epidendreae (including Agrostophyllinae and Calypsoinae) with moderate support (58/62/0.94). The "last" clade included Collabieae, Podochileae, and Vandeae, which have relatively strong support and unstable topologies.

Discussion

The present study provides a comprehensive representation of the genetic relationships between nearly all of the genera of Chinese orchids (Figs. 1–3). The overall results of the phylogenetic analysis of Chinese orchids are consistent with previously published results and support the relationships among the five subfamilies Apostasioideae, Vanilloideae, Cypripedioideae, Orchidoideae, and Epidendroideae. The classification of Chase et al. (2015) supports the division of the latter three subfamilies into 20 tribes. In China, there are 15 tribes, and their monophyletic classifications are supported in the present study. The phylogenetic analyses included in the present study are based on the coding regions of the genes, and the production of a highly detailed and hierarchical classification within subtribes would be premature; therefore, the following discussion is focused on the suprageneric levels.

Apostasioideae

Based on the results of the present study, Apostasioideae, including *Neuwiedia* and *Apostasia*, are sister to all other Orchidaceae with high support. The basal position of this subfamily is consistent with its morphological characteristics. Namely, the taxa of Apostasioideae lack a developed endosperm, possess a protocorm, and have a vascularized funiculus; these traits suggest a loss of function because the development of a protocorm and non-vascularization of the funiculus may reflect a lack of endosperm development (Pridgeon et al., 1999).

Vanilloideae

The results indicated that Vanilloideae, containing Vanilleae and Pogonieae, are sister to three subfamilies with high support, which is consistent with the results of previous nuclear and mitochondrial studies (Cameron & Chase, 2000; Freudenstein & Chase, 2001; Górniak et al., 2010). The definition of the vanilloid subfamily is relatively unknown to most botanists because the incumbent anther and poorly organized pollinia of the subfamily support a relationship within Epidendroideae (Dressler, 1993; Freudenstein & Rasmussen, 1999). The near-basal position of Vanilloideae within this family is supported by similarities in the crustose seed among Vanilleae, *Apostasia*, and *Curculigo* (Hypoxidaceae) (Sokoloff et al., 2008). Other unusual vanilloid features, such as a reticulate leaf venation, are highly derived and atypical in the context of Asparagales and orchids.

Cypripedioideae

Scientific consensus has been reached on Cypripedioideae, with the TL indicating that this subfamily is a third independent line with moderate support, which is consistent with the positions disputed between Cypripedioideae and Vanilloideae in previous studies. The *rbcL* analysis (Cameron et al., 1999) showed that Vanilloideae are sister to Orchidoideae and Epidendroideae, which is consistent with the distribution of a single anther. The 18S rDNA (Cameron & Chase, 2000) analysis showed that Vanilloideae are sister to Cypripedioideae and all of the other monandrous orchids, and the results of the mitochondrial *nad1*b-c intron (Freudenstein & Chase, 2001) and *psaB* (Cameron, 2004) analyses showed that Cypripedioideae and Vanilloideae formed a single clade as sister to Orchidoideae and Epidendroideae. However, relatively new evidence (e.g., Chomicki et al., 2014; Chase et al., 2015) has shown the same subfamily arrangement as described in the present study.

Orchidoideae

Orchidoideae consist of approximately 190 genera and 3600 species that are well represented in the northern temperate and tropical zones of both the Old and New Worlds. There are 47 genera and over 340 species in China, with four endemic genera and more than 150 endemic species (Chen et al., 2009). The topological patterns of the suprageneric levels resolved in the present study (Fig. 1) are largely consistent with those previously reported (e.g., Kores et al., 2001; Cameron, 2004; Górniak et al., 2010). Three major clades are evident, and the relationships are supported, with Diurideae sister to the Cranichideae clade (BS = 97), and this combined clade sister to Orchideae (92). Orchideae include Brownleeinae

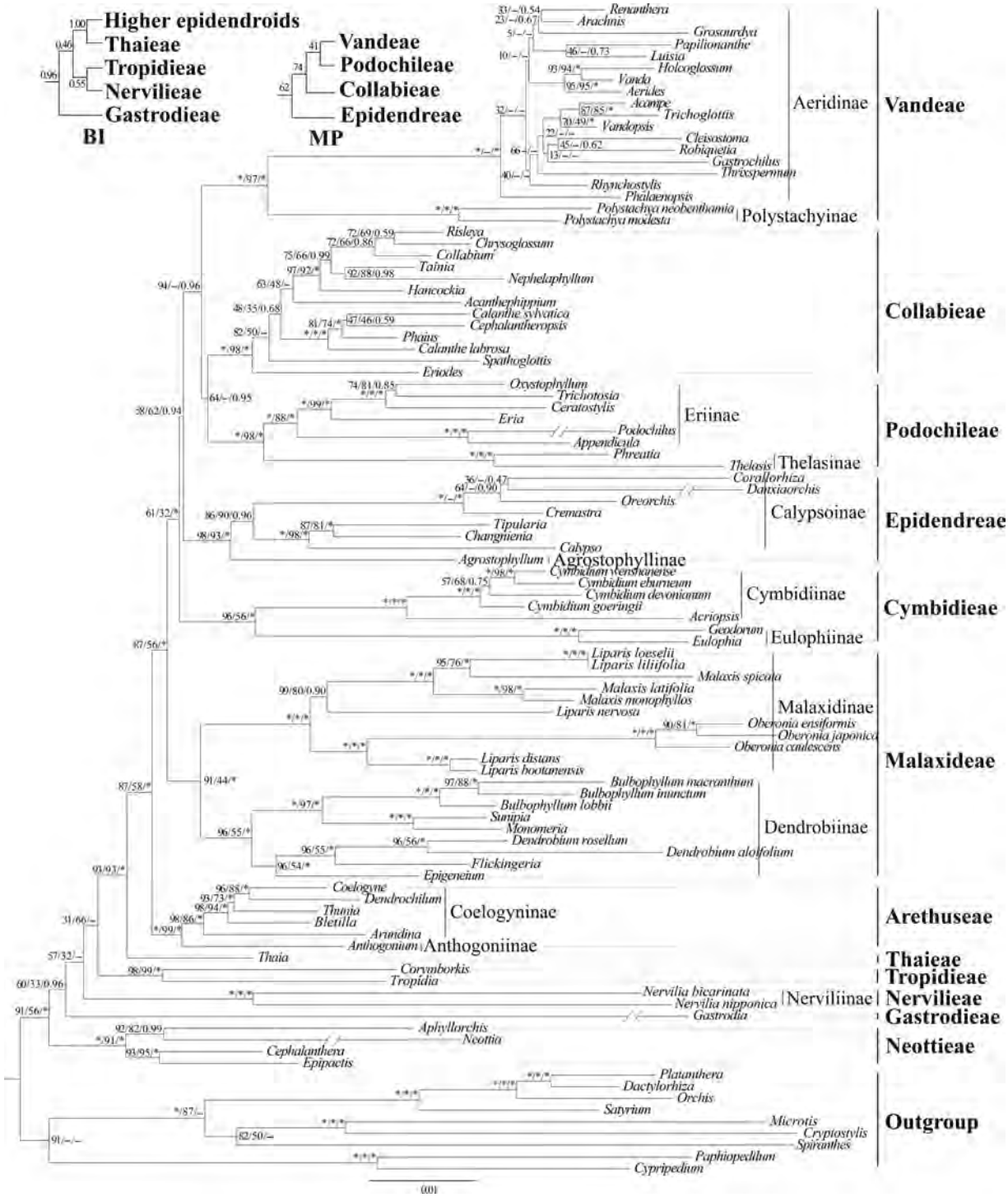


Fig. 3. Phylogenetic tree of the Chinese Epidendroideae reconstructed based on a combined analysis of the genes *rbcL*, *matK*, *psaB*, *ycf1*, and *Xdh*. Tribes and subtribes used here are indicated on the right. The three numbers near the nodes are bootstrap percentages according to maximum likelihood (BS_{ML}) and maximum parsimony (BS_{MP}) analyses, and Bayesian posterior probabilities (PP), respectively. – Node is inconsistent between the topology of the MP/ML trees and the Bayesian tree. *Node is 100% supported.

(only *Disperis*) and Orchidinae (100). Cranichideae include Spiranthinae and Goodyerinae (87). The third major clade consists of Diurideae, which include four genera in China with strong support (BS > 90).

Focusing on Orchideae, four strongly supported clades (*Disperis*, *Satyrium*, *Habenaria* clade, and *Orchis* clade) are apparent (Fig. 1). Pridgeon et al. (2001) recognized Orchideae as two tribes, Diseae and Orchideae. Although only two

genera (*Disperis* and *Satyrium*) of Diseae have been identified in China, these genera form two single basal clades with high support (100), which is consistent with previous molecular studies (Kores et al., 2001; Cameron, 2004; Górniak et al., 2010; Waterman et al., 2009). *Habenaria* and *Orchis* clades are two major groups of Chinese Orchidoideae. The genera and relationships within these two clades (93) are supported, although the relationships within these two clades do not have high PP, and the topologies are not entirely consistent with the results obtained in recent phylogenetic studies (Jin et al., 2014; Raskoti et al., 2015; Tang et al., 2015b), in which the phylogenetic relationships of these clades, including additional Asian groups, were constructed based on additional DNA markers. A number of changes in the Chinese generic circumscription have been recognized, with *Galearis* expanded to include *Aceratorchis* and *Chondradenia*, and *Platanthera* expanded to include *Diphylax* and *Smithorchis*. In addition, two new genera (*Hsenhsua* and *Tsaiorchis*) have been added (Jin et al., 2014). *Herminium* has been expanded to include *Androcorys*, *Bhutanthera*, *Frigidorchis*, and *Porolabium* (Raskoti et al., 2015); *Hemipilia* has been expanded to include *Amitostigma*, *Hemipiliopsis*, *Neottianthe*, *Ponerorchis*, and *Tsaiorchis* (Tang et al., 2015b). Moreover, one new monotypic genus, *Shizhenia* (*Amitostigma pinguicula*), has been added (Jin et al., 2015). In addition, all of the studies carried out thus far have indicated that the *Habenaria* group is paraphyletic to several genera, including *Pecteilis* and parts of *Platanthera* (Bateman et al., 2009; Inda et al., 2012; Jin et al., 2014). Thus, additional studies are required before the generic limits are adjusted in *Habenaria* and *Orchis* clades. Based on the reviews of Chase et al. (2003, 2015), we recognize *Disperis* as a member of Brownleeinae and the other three clades as members of Orchidinae.

Diurideae consist of approximately 40 genera and 1000 species that are well represented in Australia (Pridgeon et al., 2001) and four genera and nine species represented in China (Chen et al., 2009). This tribe was recognized as nine subtribes (Pridgeon et al., 2001; Chase et al., 2015); however, the relationships of the subtribes have not been completely resolved or supported. The relationships among the Chinese genera were supported as the following set: (*Cryptostylis* (*Microtis* (*Corybas*, *Stigmatodactylus*))), which is consistent with the results of Kores et al. (2001). According to Pridgeon et al. (2001) and the reviews of Chase et al. (2015), we recognize the Chinese taxa as three subtribes, with *Cryptostylis* a member of Cryptostylidinae, *Microtis* a member of Prasophyllinae, and the remaining two genera members of Acianthinae.

In the recent molecular phylogenetic study of Cranichideae (Salazar et al., 2014), the subtribal relationships were supported as the following set: (Chloraeinae (Achlydosinae, Pterostylidinae) (Goodyerinae (Galeottiellinae (Manniellinae (Spiranthinae (Discyphinae, Cranichidinae)))))). In China, Cranichideae only include two subtribes, Goodyerinae and Spiranthinae. Goodyerinae consist of 34 genera and approximately 750 species that are widely distributed in both the Old and New Worlds and are particularly diverse in tropical and subtropical regions (Pridgeon et al., 2003). Spiranthinae consist of approximately 34 genera and 640 species that are well represented in America, and only *Spiranthes* extends to Asia, Africa, Australasia, and Europe (the Asian *Pelexia* is introduced). Goodyerinae consist of approximately 17 genera and more than 100 species, and represent one of the largest subtribes of Chinese Orchidoideae (Chen et al., 2009). The results support a monophyletic origin for the members of Goodyerinae, which is consistent with Chase et al. (2015). However, for other subtribes, only two genera of Spiranthinae were included because these analyses were not specifically designed as rigorous tests for the monophyly of the subtribe. Previous molecular analyses of Goodyerinae have been largely focused on defining the circumscription of other subtribes (e.g., Salazar et al., 2003, Spiranthinae; Álvarez-Molina & Cameron, 2009, Prescottiinae; Salazar et al., 2009, Cranichidinae and Prescottiinae). These studies only included eight genera within Goodyerinae. The taxonomic and molecular phylogenetic relationships of the subtribes are poorly understood. The results of the present study propose that the 12 genera of Goodyerinae in China should be subdivided into two strongly supported clades (100), although the support is not high within the clades (Fig. 1). Therefore, additional taxonomic and phylogenetic studies on Goodyerinae are required.

Epidendroideae

The subfamily Epidendroideae include approximately 650 genera and 18 000 species and represent a greater number of genera and species than the other four subfamilies combined (Pridgeon et al., 2005; Freudenstein & Chase, 2015), and they are classified into 16 tribes and 28 subtribes (Chase et al., 2015). Numerous DNA phylogenetic studies have now been published for this subfamily, including studies covering the subfamilies, tribes, subtribes, and genera. The most widely used Epidendroideae classification is based on the DNA classification of Chase et al. (2015) and the suprageneric phylogenetic analysis of Freudenstein & Chase (2015). In the subfamily Epidendroideae (Cameron, 2004; van den Berg et al., 2005; Górniak et al., 2010; Xiang et al., 2012), there are nine tribes (Gastrodieae, Neottieae, Nervilieae, Sobralieae, Thaieae, Triphoreae, Tropidieae, Wullschlaegelieae, and Xerorchideae) that fall at the basal nodes defined as "Lower Epidendroideae" (Cameron et al., 1999). It is surprising that no fully resolved hypothesis of historical relationships has been presented for these orchids in nearly all of the molecular phylogenetic studies of the Orchidaceae, namely, most of the tribes are paraphyletic assemblages or have interrelationships that are weakly supported through bootstrapping (Chase et al., 2015; Freudenstein & Chase, 2015). Except for the lower Epidendroideae, seven tribes (Arethuseae, Malaxideae, Cymbidieae, Epidendreae, Collabieae, Podochileae, and Vandeae) are defined as "Higher Epidendroideae" and the topology of these tribes is better than that of the lower Epidendroideae, although certain relationships among these tribes remain unresolved.

In China, over 138 genera and 950 species are defined into 12 tribes *sensu* Chase et al. (2015), except for the four tribes of lower epidendroids that contain one or several genera (Sobralieae, Tropidieae, Wullschlaegelieae and Xerorchideae). In the present study, a total of 119 genera and 185 species of Epidendroideae representing 10 tribes (except for Gastrodieae and Thaieae) were sampled in the TL. There are three tribes within the lower epidendroids, and seven tribes within the higher epidendroids (Figs. 1, 2). The basal node (80)

consists of neottioid orchids, including *Aphyllorchis*, *Cephalanthera*, *Epipactis*, *Holopogon*, *Neottia*, and *Tangtsinia*, as well as two groups (Tropidieae and Nervilieae) considered primitive by most orchid phylogeneticists. In the RT (Fig. 3), all five Chinese tribes of the lower epidendroids are included, and the same topologies are presented in the ML, MP, and BI analyses except for that of the poorly supported nodes (Nervilieae and Tropidieae). The basal position of Neottieae in the TL is consistent with the RT and the report of Freudenstein & Chase (2015) based on eight DNA sequence datasets. Thaieae are sister to the higher epidendroids with high support, which is consistent with the report of Xiang et al. (2012) based upon the combined analyses of *rbcL*, *matK*, and *psaB*.

The phylogenetic relationships of the holomycotrophic *Epipogium* and *Stereosandra* pose one of most controversial questions in orchid systematics. Unfortunately, the molecular markers used in this study for these plants could not be obtained. Based on the nuclear ITS, the study of Molvray et al. (2000) showed that *Epipogium* is allied to Nervilieae, but the analyses of Rothacker (2007) showed it is sister to Triphoreae with high jackknife support (90). However, based on the mitochondrial *nad1*, the analyses of Rothacker (2007) showed that *Epipogium aphyllum* is sister to *Nervilia shirensis* with high jackknife support (99). These suggest *Epipogium* may be close to Nervilieae. So far, there has been no relevant molecular evidence for the phylogenetic analysis for *Stereosandra*. Additional taxonomic and phylogenetic studies are required to examine the relationships of them at the tribe level within lower epidendroids and/or the subtribe level among Nervilieae, Triphoreae, and Tropidieae. Here, *Epipogium* and *Stereosandra* are tentatively defined as the subtribe Epipogiinae and classified into the tribe Nervilieae (including the other subtribe, Nerviliinae) according to Chase et al. (2015).

Within the higher epidendroids, support for the relationships of these clades is not high, and the sequences of certain genera are lacking; thus, the relationships among the tribes and monophylies of Epidendreae are not clearly refuted on the basis of the TL results. The topology of the RT is nearly consistent with previous molecular analyses (Górniak et al., 2010; Luo et al., 2014; Freudenstein & Chase, 2015; Givnish et al., 2015; Zhang et al., 2015a) and the classification of Chase et al. (2015) and the tribe and subtribe levels are successively sister to the next; thus, we chose the results of the RT to elucidate the arrangement of the suprageneric levels.

Seven major clades are evident based on the ML (BS $>$ 55) and BI (PP $>$ 0.90) analyses, and the relationships are supported as the following set: (Arethuseae (Malaxideae (Cymbidieae (Epidendreae (Vandeae (Podochileae, Collabieae)))))). Arethuseae include the subtribes Coelogyninae and Anthogoniinae (including *Anthogonium*). *Anthogonium* and *Arundina* were formerly considered as members of the subtribe Arethusinae based on plastid and/or mitochondrial evidence (van den Berg et al., 2005) and the classifications of Chase et al. (2003, 2015). Although only *Anthogonium* and *Arundina* of Arethusinae were included among the genera in China, these genera formed two single subclades with high support (100/99/1.00 and 98/86/1.00). Studies of *Xdh* (Górniak et al., 2010; Zhang et al., 2015a) have indicated that *Anthogonium* is sister to the remaining higher epidendroids. Because of the observed discrepancy between the plastid and nuclear *Xdh* results, hybridization and/or organism-level processes might have been involved in the evolution of this set of species (Wendel & Doyle, 1998). The position of *Arundina* in this tribe appears clear, although in which subtribe this genus should be included remains unresolved. The best-sampled analysis in terms of data was carried out by Freudenstein & Chase (2015), who provided a mixed result depending on the type of analysis (ML vs. MP). Based on the combined chloroplasts of *rbcL*, *matK*, and *psaB*, and the nuclear *Xdh*, the results of Zhang et al. (2015a) showed that *Arundina* is sister to Coelogyninae.

Anthogonium presents different characteristics from that of the other taxa in Arethusinae, and the morphological characteristics are also clearly distinctive among genera in the tribe Arethuseae as *Anthogonium* has lateral inflorescences and a distinct flower bend, whereas the other taxa have terminal inflorescences. Holttum (1964) stated that *Arundina* is closely related to *Dilochia* and *Thunia* because these plants have elongated stems with numerous linear to lanceolate leaves and inflorescences with large floral bracts. Dressler (1993) placed *Arundina* and *Dilochia* in a distinct subtribe, Arundinae. The RT and analyses by Zhang et al. (2015a) showed that *Arundina* is sister to Coelogyninae (including *Bletilla*, *Coelogyne*, *Dilochia*, *Glomera*, *Pleione*, and *Thunia*). Further generic studies or nomenclatural changes have not been undertaken on Coelogyninae since the studies of Gravendeel et al. (2001, 2004). However, the results of this study showed that substantial changes are needed to the circumscription of *Coelogyne* and related genera (e.g., *Dendrobium* and *Pholidota*). In the present study, the Chinese subtribal circumscription was consistent with that delimited in Chase et al. (2015), except that *Arundina* was placed in Coelogyninae and Anthogoniinae S. C. Chen, Z. H. Tsi & G. Zhu were recognized (Chen et al., 1999). Anthogoniinae (only *Anthogonium*) are tentatively classified into the tribe Arethuseae, although additional taxonomic and phylogenetic studies are required to examine the relationships of Anthogoniinae at the tribe level within Epidendroideae and/or the subtribe level within Arethuseae.

The second clade is represented as two major groups of orchids, Dendrobiinae and Malaxidinae (Malaxideae), which have both held debated positions in Epidendroideae (Chase et al., 2015). According to the reed stem, upper lateral inflorescences, and spherical silica bodies, Dressler (1990) placed the Dendrobieae and Malaxideae taxa into two separate groups, cymbidioid phylad and epidendroid phylad, respectively. The traditional classification system of the tribes that uses characters primarily related to floral morphology does not reflect the evolutionary history of these taxa. Molecular analyses have not indicated consistent positions within this subfamily for these two groups. Based on the classification of Chase et al. (2003), and analyses of *rbcL* and *matK* combined (Freudenstein et al., 2004) and *rbcL* and *psaB* combined (Cameron, 2004), these two groups were classified as *incertae sedis* within Epidendroideae. Based on the ITS and four chloroplast sequences, van den Berg et al. (2005) held that Dendrobiinae and Malaxideae were sisters to vandoid orchids and Collabieae, respectively. Dendrobiinae species are similar to Malaxideae in the

synapomorphic state of the "naked" pollinium (but see Li & Yan, 2013). Similar to other analyses based on the *Xdh* gene (Górniak et al., 2010), the four chloroplast genes *accD*, *ccsA*, *matK*, and *ycf1* (Luo et al., 2014) and seven loci (Freudenstein & Chase, 2015), the analyses of the present study confirm that Dendrobiinae and Malaxideae are sister relatives. Here, we recognize that Chinese Malaxideae include two subtribes, Dendrobiinae and Malaxidinae.

The third highly supported clade consists of Cymbidieae, which are divided into 10 subtribes in Chase et al. (2015), and two subtribes Cymbidiinae (only including *Acriopsis* and *Cymbidium*) and Eulophiinae (only *Eulophia* and *Geodorum*), which are represented in China with strong support (96/58/1.00). Here, we recognize the subtribal circumscriptions of Chinese Cymbidieae as delimited in Chase et al. (2015).

The fourth supported clade consists of Epidendreae. The classification of Chase et al. (2015) supports the division of Epidendreae into six subtribes, Agrostophyllinae (only including *Agrostophyllum* and *Earina*), Calypsoinae, Bletiinae, Ponerinae, Pleurothallidinae, and Laeliinae, with the former two subtribes represented in China and the latter four subtribes (Epidendreae s.s.) distributed in America. The TL indicated that the Chinese groups were divided into three supported groups with weak support (BS $<$ 0.50). The results of this paraphyletic analysis are consistent with those obtained in previous studies based on plastid analyses (Cameron, 2004; Freudenstein et al., 2004; Zhang et al., 2015a). The RT showed that the Calypsoinae taxa are monophyletic (86/90/0.96), and *Agrostophyllum* (Agrostophyllinae) is a sister to these taxa with high support (98/93/1.00), which is consistent with previous studies (Freudenstein & Chase, 2015; Zhang et al., 2015a).

In Chase et al. (2003), Agrostophyllinae are listed as unplaced, and Calypsoinae are defined as the tribe Calypsoeae. The sister relationship of Agrostophyllinae and Calypsoeae has been demonstrated in Freudenstein et al. (2004) and Zhang et al. (2015a). Using plastid analyses, these authors reported that the Calypsoeae are divided into two supported groups, and *Agrostophyllum* and/or *Earina* are sisters with weak support to one group of Calypsoeae. Additionally, Agrostophyllinae and Calypsoeae were shown to be related to Epidendreae s.s. in the results of Neubig et al. (2009), which was based on the plastid gene *ycf1*, Górniak et al. (2010), which was based on the nuclear gene *Xdh*, and Zhang et al. (2015a), which was based on the combined three plastid genes (*rbcL*, *matK*, and *psaB*) and *Xdh*. Additionally, van den Berg et al. (2005) indicated the Calypsoeae were sister to Epidendreae s.s. (PP 0.70), whereas Agrostophyllinae were sister to Vandeae. There are no morphological data confirming the sister relationship between Agrostophyllinae, Calypsoeae, and Epidendreae s.s. (Górniak et al., 2010). In addition, these studies only included the *Xdh* gene of *Earina*. Because the *Xdh* gene could be a paralog, Pridgeon et al. (2009) suggested that the peculiar position of *Earina* may be spurious, and they recognized Agrostophyllinae as a subtribe of Vandeae. The present study added the *Xdh* gene of *Agrostophyllum*, and the results (Fig. 3) show a sister relationship between Agrostophyllinae and Calypsoeae. Here, we recognize Chinese *Agrostophyllum* and Calypsoeae as the two subtribes Agrostophyllinae and Calypsoinae of Epidendreae.

The "last" clade of the tree consists of Collabieae, Podochileae, and Vandeae with moderate support. Dressler (1993) used the common characteristics of spherical silica bodies (Møller & Rasmussen, 1984) to place Dendrobieae, Podochileae, and Vandeae in one clade. The molecular analyses of Cameron (2004), Freudenstein et al. (2004), van den Berg et al. (2005), Luo et al. (2014), and Xiang et al. (2014) did not show close relationships among Collabieae, Podochileae, and Vandeae. The combined analyses of the three plastid genes *rbcL*, *psaB*, and *matK* (Xiang et al., 2014) indicated that Collabieae are nested within a superclade that includes Epidendreae, Podochileae, Cymbidieae, and Vandeae. The combined analyses of four plastid genes (Luo et al., 2014) indicated that Collabieae are related to Podochileae, whereas Vandeae are closely related to Epidendreae and Cymbidieae. In the ML and BI analyses carried out in this study, Vandeae form a sister clade to Podochileae and Collabieae. These results suggested that the spherical silica bodies of Podochileae and Vandeae could have been inherited from a common ancestor. Additionally, except for certain species of *Calanthe* and *Polystachya*, all of the Collabieae, Podochileae, and Vandeae taxa are distributed in the Old World. These common geographical distributions also support the close relationships among these three tribes.

Vandeae have been found to contain four subtribes (Polystachyinae, Adrorhizinae, Angraecinae, and Aeridinae), 139 genera and approximately 2600 species. In China, Vandeae include two subtribes, Polystachyinae and Aeridinae. Chinese Aeridinae consist of approximately 50 genera and are one of the largest subtribes in Chinese Orchidaceae (Chen et al., 2009). The results presented here support a monophyletic origin for the members of Aeridinae, which is consistent with Pridgeon et al. (2014) and Chase et al. (2015). Within this subtribe, low support is observed, and the relationships are difficult to infer. Recent molecular phylogenetic analyses of Aeridinae (Zou et al., 2015) are based on five DNA sequences (ITS, *atpI-H*, *matK*, *psbA-trnH*, and *trnL-F*), and most Chinese taxa indicate that the subtribe is primarily grouped into 10 clades. However, the relationships between most of these clades remain unresolved. Moreover, several genera are polyphyletic as currently circumscribed. Here, we recognize Chinese Vandeae as two subtribes, Polystachyinae and Aeridinae.

The tribe, Collabieae, is a medium-sized group with 20 genera and approximately 450 species that are primarily distributed in the Old World tropics, with several species extending into northern temperate Asia and Mesoamerica (Pridgeon et al., 2005). There are approximately 14 genera and 100 species in China (Chen et al., 2009; Xiang et al., 2014). The results of the present study show that the relationships within this tribe are well supported, and the paraphyletic relationship of *Calanthe* with *Cephalantheropsis* and *Phaius* was observed, which is consistent with the reports of Xiang et al. (2014) and Zhai et al. (2014). In Xiang et al. (2014), the authors circumscribed *Calanthe* in a broad sense to include *Calanthe* s.s., *Cephalantheropsis*, *Gastrorchis*, and *Phaius*. However, in Zhai et al. (2014), the authors circumscribed *Calanthe* to include six genera: *Calanthe*, *Cephalantheropsis*, *Paraphaius*, *Phaius*, *Preptanthe*, and *Styloglossum*. In addition, *Ania* is a distinct genus supported

Table 1 Phylogenetic classification of Chinese Orchidaceae

Subfamily	Tribe	Subtribe	Genera
Apostasioideae			*Apostasia*, *Neuwiedia*
Vanilloideae	Pogonieae		*Pogonia*
	Vanilleae		*Cyrtosia*, *Erythrorchis*, *Galeola*, *Lecanorchis*, *Vanilla*
Cypripedioideae			*Cypripedium*, *Paphiopedilum*
Orchidoideae	Orchideae	Brownleeinae	*Disperis*
		Orchidinae	*Amitostigma*, *Androcorys*, *Bhutanthera*, *Brachycorythis*, *Dactylorhiza*, *Diphylax*, *Diplomeris*, *Frigidorchis*, *Galearis*, *Gennaria**, *Gymnadenia*, *Habenaria*, *Hemipilia*, *Hemipiliopsis*, *Herminium*, *Hsenhsua**, *Neottianthe*, *Nujiangia**, *Orchis*, *Pecteilis*, *Peristylus*, *Platanthera*, *Ponerorchis*, *Porolabium*, *Satyrium*, *Sirindhornia**, *Shizhenia**, *Smithorchis*, *Tsaiorchis*
	Diurideae	Cryptostylidinae	*Cryptostylis*
		Prasophyllinae	*Microtis*
		Acianthinae	*Corybas*, *Stigmatodactylus*
	Cranichideae	Goodyerinae	*Anoectochilus*, *Chamaegastrodia*, *Cheirostylis*, *Cystorchis**, *Erythrodes*, *Goodyera*, *Herpysma*, *Hetaeria*, *Hylophila*, *Kuhlhasseltia*, *Ludisia*, *Myrmechis*, *Odontochilus*, *Rhomboda*, *Vrydagzynea*, *Zeuxine*, *Zeuxinella**
		Spiranthinae	*Pelexia*, *Spiranthes*
Epidendroideae	Neottieae		*Aphyllorchis*, *Cephalanthera*, *Diplandrorchis*, *Epipactis*, *Holopogon*, *Neottia*, *Tangtsinia*
	Gastrodieae		*Didymoplexiella*, *Didymoplexis*, *Didymoplexiopsis*, *Gastrodia*
	Nervilieae	Epipogiinae	*Epipogium*, *Stereosandra*
		Nerviliinae	*Nervilia*
	Tropidieae		*Corymborkis*, *Tropidia*
	Thaieae		*Thaia**
	Arethuseae	Anthogoniinae	*Anthogonium*
		Coelogyninae	*Arundina*, *Bletilla*, *Bulleyia*, *Coelogyne*, *Dendrochilum*, *Ischnogyne*, *Neogyna*, *Otochilus*, *Panisea*, *Pholidota*, *Pleione*, *Thunia*, *Thuniopsis**
	Malaxideae	Dendrobiinae	*Bulbophyllum*, *Dendrobium*, *Epigeneium*, *Flickingeria*, *Monomeria*, *Sunipia*, *Trias**
		Malaxidinae	*Cestichis**, *Crepidium*, *Dienia*, *Diteilis**, *Empusa**, *Liparis*, *Malaxis*, *Oberonia*, *Oberonioides*, *Platystyliparis**, *Ypsilorchis*
	Cymbidieae	Cymbidiinae	*Acriopsis*, *Cymbidium*
		Eulophiinae	*Eulophia*, *Geodorum*
	Epidendreae	Agrostophyllinae	*Agrostophyllum*
		Calypsoinae	*Calypso*, *Changnienia*, *Corallorhiza*, *Cremastra*, *Danxiaorchis**, *Oreorchis*, *Tipularia*, *Yoania*, *Yunorchis**
	Vandeae	Polystachyinae	*Polystachya*
		Aeridinae	*Acampe*, *Aerides*, *Arachnis*, *Ascocentrum*, *Biermannia*, *Chamaeanthus*, *Chiloschista*, *Chroniochilus**, *Cleisostoma*, *Cleisostomopsis*, *Diploprora*, *Doritis*, *Esmeralda*, *Gastrochilus*, *Grosourdya*, *Gunnaria**, *Haraella*, *Holcoglossum*, *Hygrochilus*, *Lesliea**, *Luisia*, *Malleola*, *Micropera*, *Microtatorchis*, *Neofinetia*, *Nothodoritis*, *Ornithochilus*, *Papilionanthe*, *Paraholcoglossum**, *Parapteroceras*, *Pelatantheria*, *Pendulorchis**, *Penkimia*, *Pennilabium*, *Phalaenopsis*, *Pomatocalpa*, *Pteroceras*, *Renanthera*, *Rhynchostylis*, *Robiquetia*, *Saccolabiopsis*, *Sarcoglyphis*, *Sarcophyton*, *Schoenorchis*, *Sedirea*, *Singchia**, *Smitinandia*, *Staurochilus*, *Stereochilus*, *Taeniophyllum*, *Thrixspermum*, *Trichoglottis*, *Tsiorchis**, *Tuberolabium*, *Uncifera*, *Vanda*, *Vandopsis*
	Collabieae		*Acanthephippium*, *Ania**, *Calanthe*, *Cephalantheropsis*, *Chrysoglossum*, *Collabium*, *Diglyphosa*, *Eriodes*, *Hancockia*, *Nephelaphyllum*, *Pachystoma*, *Paraphaius**, *Phaius*, *Preptanthe**, *Risleya*, *Spathoglottis*, *Styloglossum**, *Tainia*
	Podochileae	Thelasinae	*Phreatia*, *Thelasis*
		Eriinae	*Aeridostachya*, *Appendicula*, *Bryobium*, *Callostylis*, *Campanulorchis*, *Ceratostylis*, *Conchidium*, *Cryptochilus*, *Cylindrolobus*, *Dendrolirium*, *Eria*, *Mycaranthes*, *Oxystophyllum*, *Pinalia*, *Podochilus*, *Porpax*, *Trichotosia*

*Newly described or recognized genera after Chen et al. (2009).

by both morphological and molecular evidence (Li et al., 2014a).

Podochileae consist of approximately 28 genera and approximately 1280 species that are well represented in Asia and Australia. The single genus *Stolzia* has been identified in tropical Africa. There are 19 genera and approximately 70 species in China. The RT showed that Podochileae consist of two monophyletic subtribes, Eriinae and Thelasinae, with high support (100/98/1.00). This relationship is nearly consistent with the review of Pridgeon et al. (2005) and the results of van den Berg et al. (2005). Since Chase et al. (2003) was published, substantial changes have been made to the generic circumscription of *Eria* s.l., which includes approximately 370 species and 44 species recorded in China, and the genus has been classified as polyphyletic. Many of the genera recognized in Pridgeon et al. (2005) did not have published combinations for the putative species, although the classifications described herein have subsequently been published by several authors (e.g., Chen et al., 2009; Chase et al., 2015). Here, we recognize Chinese Podochileae as two subtribes, Eriinae and Thelasinae.

Conclusions and Perspectives

In the classification of Chase et al. (2015), Orchidaceae were defined into five subfamilies, 22 tribes, and 49 subtribes. In China, there are approximately 200 genera with over 1450 species representing all the subfamilies, 17 tribes (apart from five tribes that contain one or few genera, namely Codonorchideae, Sobralieae, Tropidieae, Wullschlaegelieae, and Xerorchideae), and 19 subtribes. The present study provides the first phylogenetic analyses of Chinese orchids that represent nearly all genera based on chloroplast genes *atpB*, *rbcL*, *matK*, and *ndhF*, and mitochondrial gene *matR*, and we have recognized Chinese Epidendroideae, which represent all subtribes (except for Epipogiinae) based on the chloroplast genes *rbcL*, *matK*, *psaB*, and *ycf1*, and nuclear gene *Xdh*.

The phylogenetic trees provide support for previous hypotheses of subfamilial relationships within Orchidaceae and also indicate several new patterns that include the rearrangement of particular tribes and minor changes to the placement of several genera. Chinese orchids should be defined as five monophyletic subfamilies, 17 tribes, and 21 subtribes. Based on the highly supported circumscription and arrangement of suprageneric levels, a new phylogenetic classification of Chinese Orchidaceae has been proposed in conjunction with the *Flora of China* (Chen et al., 2009) and the later newly described and recorded genera (Table 1).

Acknowledgements

The authors would like to thank Yi-Bo Luo for providing helpful comments and thank Qiu-Xiang Guan and Xin-Yi Wu for performing an analysis of the molecular sequence data. This work was financially supported through grants from the Scientific Research Foundation of Graduate School of Fujian Agriculture and Forestry University (Grant No. 1122yb006) and the Innovation Fund for Forestry Technology of Guangdong (Grant No. 2011KJCX009).

References

Álvarez-Molina A, Cameron KM. 2009. Molecular phylogenetics of Prescottiinae s.l. and their close allies (Orchidaceae, Cranichideae) inferred from plastid and nuclear ribosomal DNA sequences. *American Journal of Botany* 82: 247–277.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 85: 531–553.

Bateman RM, James KE, Luo YB, Lauri RK, Fulcher T, Cribb PJ, Chase MW. 2009. Molecular phylogenetics and morphological reappraisal of the *Platanthera* clade (Orchidaceae: Orchidinae) prompts expansion of the generic limits of *Galearis* and *Platanthera*. *Annals of Botany* 104: 431–445.

Cameron KM. 2004. Utility of plastid *psaB* gene sequences for investigating intrafamilial relationships within Orchidaceae. *Molecular Phylogenetics and Evolution* 31: 1157–1180.

Cameron KM, Chase MW. 2000. RDNA sequences of Orchidaceae confirm the subfamilial status and circumscription of Vanilloideae. In: Wilson KL, Morrison DA eds. *Monocots: Systematics and evolution 18*. Melbourne: CSIRO. 457–464.

Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Yukawa T, Hills HG, Goldman DH. 1999. A phylogenetic analysis of the Orchidaceae: Evidence from *rbcL* nucleotide sequences. *American Journal of Botany* 86: 208–224.

Chase MW, Cameron KM, Barrett RL, Freudenstein JV. 2003. DNA data and Orchidaceae systematics: A new phylogenetic classification. In: Dixon KW, Kell SP, Barrett RL, Cribb PJ eds. *Orchid conservation*. Kota Kinabalu: Natural History Publications (Borneo). 69–89.

Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, van den Berg C, Schuiteman A. 2015. An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society* 177: 151–174.

Chase MW, Fay MF, Devey DS, Maurin O, Rønsted N, Davies TJ, Pillon Y, Petersen G, Seberg O, Tamura MN. 2006. Multigene analyses of monocot relationships: A summary. *Aliso* 22: 63–75.

Chen LJ, Liu ZJ. 2010. *Chroniochilus sinicus*, a new species of Orchidaceae from Yunnan, China. *Novon* 20: 252–255.

Chen SC, Tsi ZH, Lang KY, Zhu GH. 1999. Orchidaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press. 18: 1–412.

Chen XQ, Liu ZJ, Zhu GH, Lang KY, Ji ZH, Luo YB, Jin XH, Cribb PJ, Wood JJ, Gale SW. 2009. Orchidaceae. In: Wu ZG, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 25: 1–506.

Chen ZD, Yang T, Lu LM, Lin L, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, Cao ZY, Liu HM, Wang XM, Wang W, Zhang JB, Meng Z, Cao W, Li JH, Wu SD, Zhao HL, Liu ZJ, Du ZY, Wang QF, Guo J, Tan XX, Su JX, Zhang LJ, Yang LL, Liao YY, Li MH, Zhang GQ, Chung SW, Zhang J, Xiang KL, Li RQ, Soltis DE, Soltis PS, Zhou SL, Ran JH, Wang XQ, Jin XH, Chen YS, Gao TG, Li JH, Zhang SZ, Lu AM, China Phylogeny Consortium. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54: 277–306.

Chomicki G, Bidel LPR, Ming F, Coiro M, Zhang X, Wang YF, Baissac Y, Jay-Allemand C, Renner SS. 2014. The velamen protects photosynthetic orchid roots against UV-B damage, and a large dated phylogeny implies multiple gains and losses of this function during the Cenozoic. *New Phytologist* 205: 1330–1341.

Dressler RL. 1990. The major clades of the Orchidaceae-Epidendroideae. *Lindleyana* 5: 117–125.

Dressler RL. 1993. *Phylogeny and classification of the orchid family*. Cambridge: Harvard University Press.

Freudenstein JV, Chase MW. 2001. Analysis of mitochondrial *nad1*b-c intron sequences in Orchidaceae: Utility and coding of length-change characters. *Systematic Botany* 26: 643–657.

Freudenstein JV, Chase MW. 2015. Phylogenetic relationships in Epidendroideae (Orchidaceae), one of the great flowering plant radiations: Progressive specialization and diversification. *Annals of Botany* 115: 665–681.

Freudenstein JV, Rasmussen FN. 1999. What does morphology tell us about orchid relationships?—A cladistic analysis. *American Journal of Botany* 86: 225–248.

Freudenstein JV, van den Berg C, Goldman DH, Kores PJ, Molvray M, Chase MW. 2004. An expanded plastid DNA phylogenetic analysis of Orchidaceae and analysis of jackknife clade support strategy. *American Journal of Botany* 91: 149–157.

Givnish TJ, Spalink D, Ames M, Lyon SP, Hunter SJ, Zuluaga A, Iles WJD, Clements MA, Arroyo MTK, Leebens-Mack J, Endara L, Kriebe R, Neubig KM, Whitten WM, Williams NH, Cameron KM. 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceeding of the Royal Society B: Biological Sciences* 282: 20151553.

Górniak M, Paun O, Chase MW. 2010. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, *Xdh*: Congruence with organellar and nuclear ribosomal DNA results. *Molecular Phylogenetics and Evolution* 56: 784–795.

Govaerts R, Dransfield J, Zona SF, Hodel DR, Henderson A. 2014. *World Checklist of Orchidaceae* [online]. Kew: The Board of Trustees of the Royal Botanic Gardens. Available from http://apps.kew.org/wcsp [accessed 15 December 2014].

Gravendeel B, Chase MW, de Vogel E, Roos MC, Mes THM, Bachmann K. 2001. Molecular phylogeny of *Coelogyne* (Epidendroideae: Orchidaceae) based on plastid RFLPs, *matK*, and nuclear ribosomal ITS sequences: Evidence for polyphyly. *American Journal of Botany* 88: 1915–1927.

Gravendeel B, Smithson A, Slik FJW, Schuiteman A. 2004. Epiphytism and pollinator specialization: Drivers for orchid diversity? *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 1523–1535.

Guan QX, Chen GZ, Li MH, Chen SP. 2014. *Goodyera malipoensis* (Cranichideae, Orchidaceae), a new species from China: Evidence from morphological and molecular analyses. *Phytotaxa* 186: 51–60.

Holttum RE. 1964. *A revised flora of Malaya. Vol. 1, Orchids of Malaya*. 3rd ed. Singapore: Government Printing Office.

Hu AQ, Hsu TC, Liu Y. 2014. *Gastrodia damingshanensis* (Orchidaceae: Epidendroideae): A new myco-heterotrophic orchid from China. *Phytotaxa* 175: 256–262.

Hu GH, Long CL, Motley TJ. 2013. *Cremastra malipoensis* (Orchidaceae), a new species from Yunnan, China. *Systematic Botanay* 38: 64–68.

Huang YF, Wu L, Liu Y. 2012. *Zeuxinella*, a newly recorded genus of Orchidaceae from China. *Journal of Tropical and Subtropical Botany* 20: 132–134.

Inda LA, Pimentel M, Chase MW. 2012. Phylogenetics of tribe Orchideae (Orchidaceae: Orchidoideae) based on combined DNA matrices: Inferences regarding timing of diversification and evolution of pollination syndromes. *Annals of Botany* 110: 71–90.

Jin WT, Jin XH, Schuiteman A, Li DZ, Xiang XG, Huang WC, Li JW, Huang LQ. 2014. Molecular systematics of subtribe Orchidinae and Asian taxa of Habenariinae (Orchideae, Orchidaceae) based on plastid *matK*, *rbcL* and nuclear ITS. *Molecular Phylogenetics and Evolution* 77: 41–53.

Jin WT, Xiang XG, Jin XH. 2015. Generic delimitation of Orchidaceae from China: Current situation and perspective. *Biodiversity Science* 23: 237–242.

Jin WT, Zhou HL, Jin XH. 2013. *Platanthera yadongensis* (Orchidaceae, Orchideae), a new species from Tibet, China. *Systematic Botany* 38: 982–986.

Kores PJ, Molvray M, Weston PH, Hopper SD, Brown AP, Cameron KM, Chase MW. 2001. A phylogenetic analysis of Diurideae (Orchidaceae) based on plastid DNA sequence data. *American Journal of Botany* 88: 1903–1914.

Lang KY, Chen SC, Luo YB, Zhu GH. 1999. Orchidaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press. 17: 1–499.

Li JW, Tao GD, Liu Q. 2011. *Lesliea*, a newly recorded genus of Orchidaceae from China. *Plant Diversity and Resources* 33: 643–644.

Li L, Yan HF. 2013. A remarkable new species of *Liparis* (Orchidaceae) from China and its phylogenetic implications. *PLoS ONE* 8: e78112.

Li L, Yan HF, Niu M, Tu TY, Li SJ, Xing FW. 2014a. Re-establishment of the genus *Ania* Lindl. (Orchidaceae). *PLoS ONE* 9: e103129.

Li L, Ye DP, Niu M, Yan HF, Wen TL, Li SJ. 2015. *Thuniopsis*: A new orchid genus and phylogeny of the tribe Arethuseae (Orchidaceae). *PLoS ONE* 10: e0132777.

Li MH, Zhang GQ, Liu ZJ, Lan SR. 2014b. Revision of *Hygrochilus* (Orchidaceae: Epidendroideae: Aeridinae) and a molecular phylogenetic analysis. *Phytotaxa* 159: 256–268.

Liu ZJ, Chen LJ. 2009. *Singchia* and *Gunnaria*, two new genera of Orchidaceae. *Journal of Systematics and Evolution* 47: 599–604.

Liu ZJ, Chen LJ, Chen SC, Cai J, Tsai WC, Hsiao YY, Rao WH, Ma XY, Zhang GQ. 2011. *Paraholcoglossum* and *Tsiorchis*, two new orchid genera established by molecular and morphological analyses of the *Holcoglossum* alliance. *PLoS ONE* 6: e24864.

Lu LM, Sun M, Zhang JB, Li HL, Lin L, Yang T, Chen M, Chen ZD. 2014. Tree of life and its applications. *Biodiversity Science* 22: 3–20.

Luo J, Hou BW, Niu ZT, Liu W, Xue QY, Ding XY. 2014. Comparative chloroplast genomes of photosynthetic orchids: Insights into evolution of the Orchidaceae and development of molecular markers for phylogenetic applications. *PLoS ONE* 9: e99016.

Meng YF, Hu AQ, Yi QF, Xing FW. 2014. *Malaxis malipoensis* (Malaxideae: Orchidaceae), a new species from Yunnan, China. *Phytotaxa* 167: 195–200.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA, 14 November, 2010. doi: 10.1109/GCE2010.5676129.

Møller JD, Rasmussen H. 1984. Stegmata in Orchidales: Character state distribution and polarity. *Botanical Journal of the Linnean Society* 89: 53–76.

Molvray M, Kores PJ, Chase MW. 2000. Polyphyly of mycoheterotrophic orchids and functional influences on floral and molecular characters. In: Wilson KL, Morrison DA eds. *Monocots: Systematics and evolution*. Collingwood: CSIRO. 441–448.

Neubig KM, Whitten WM, Carlsward BS, Blanco MA, Endara L, Williams NH, Moore M. 2009. Phylogenetic utility of *ycf1* in orchids: A plastid gene more variable than *matK*. *Plant Systematics and Evolution* 277: 75–84.

Peng DH, Liu ZJ, Zhai JW. 2013. *Changnienia malipoensis*, a new species from China (Orchidaceae: Epidendroideae: Calypsoeae). *Phytotaxa* 115: 55–58.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 1999. *Genera Orchidacearum Vol. 1. General introduction, Apostasioideae, Cypipedioideae*. Oxford: Oxford University Press.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2001. *Genera Orchidacearum Vol. 2. Orchidoideae (Part 1)*. Oxford: Oxford University Press.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2003. *Genera Orchidacearum Vol. 3. Orchidoideae (Part 2), Vanilloideae*. Oxford: Oxford University Press.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2005. *Genera Orchidacearum Vol. 4. Epidendroideae (Part 1)*. Oxford: Oxford University Press.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2009. *Genera Orchidacearum Vol. 5. Epidendroideae (Part 2)*. Oxford: Oxford University Press.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. 2014. *Genera orchidacearum Vol. 6. Epidendroideae (Part 3)*. Oxford: Oxford University Press.

Raskoti BB, Jin WT, Xiang XG, Schuiteman A, Li DZ, Li JW, Huang WC, Jin XH, Huang LQ. 2015. A phylogenetic analysis of molecular and morphological characters of *Herminium* (Orchidaceae, Orchideae): Evolutionary relationships, taxonomy, and patterns of character evolution. *Cladistics*. doi: 10.1111/cla.12125.

Roberts DL, Dixon KW. 2008. Orchids. *Current Biology* 18: R325–R329.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Rothacker EP. 2007. *The primitive Epidendroideae (Orchidaceae): Phylogeny, character evolution and the systematics of Psilochilus (Triphoreae)*. Ph.D. Dissertation. Columbus: Graduate School of the Ohio State University.

Salazar GA, Cabrera LI, Madriñán S, Chase MW. 2009. Phylogenetic relationships of Cranichidinae and Prescottiinae (Orchidaceae, Cranichideae) inferred from plastid and nuclear DNA sequences. *Annals of Botany* 104: 403−416.

Salazar GA, Chase MW, Arenas MAS, Ingrouille M. 2003. Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): Evidence from plastid and nuclear DNA sequences. *American Journal of Botany* 90: 777–795.

Salazar GA, van den Berg C, Popovkin A. 2014. Phylogenetic relationships of *Discyphus scopulariae* (Orchidaceae, Cranichideae) inferred from plastid and nuclear DNA sequences: Evidence supporting recognition of a new subtribe, Discyphinae. *Phytotaxa* 173: 127–139.

Seidenfaden G. 1992. The orchids of Indochina. *Opera Botanica* 114: 1–502.

Seidenfaden G, Wood JJ, Holttum RE. 1992. *The orchids of peninsular Malaysia and Singapore*. Fredensborg: Olsen & Olsen.

Sokoloff DD, Remizowa MV, Macfarlane TD, Rudall PJ. 2008. Classification of the early-divergent angiosperm family Hydatellaceae: One genus instead of two, four new species and sexual dimorphism in dioecious taxa. *Taxon* 57: 179−200.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57: 758–771.

Su YY, Huang YL, Chen LJ, Zhang PW, Liu ZJ, Zhang GQ. 2015. *Liparis wenshanensis*, a new orchid species from China: Evidence from morphological and molecular analyses. *Phytotaxa* 204: 253–264.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4. Sunderland: Sinauer Associates.

Tang GD, Zhang GQ, Hong WJ, Liu ZJ, Zhuang XY. 2015a. Phylogenetic analysis of Malaxideae (Orchidaceae: Epidendroideae): Two new species based on the combined nrDNA ITS and chloroplast *matK* sequences. *Guihaia* 35: 447–463.

Tang Y, Yukawa T, Bateman RM, Jiang H, Peng H. 2015b. Phylogeny and classification of the East Asian *Amitostigma* alliance (Orchidaceae: Orchideae) based on six DNA markers. *BMC Evolutionary Biology* 15: 96.

Tian HZ, Liu QX, Chen ZQ, Hu AQ, Jiang H. 2014. *Anoectochilus longilobus* (Orchidoideae: Orchidaceae), a new species from Yunnan, China. *Phytotaxa* 164: 276–280.

Tsi ZH, Chen SC, Luo YB, Zhu GH. 1999. Orchidaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press. 19: 1–505.

van den Berg C, Goldman DH, Freudenstein JV, Pridgeon AM, Cameron KM, Chase MW. 2005. An overview of the phylogenetic relationships within Epidendroideae inferred from multiple DNA regions and recircumscription of Epidendreae and Arethuseae (Orchidaceae). *American Journal of Botany* 92: 613–624.

Waterman RJ, Pauw A, Barraclough TG, Savolainen V. 2009. Pollinators underestimated: A molecular phylogeny reveals widespread floral convergence in oil-secreting orchids (subtribe Coryciinae) of the Cape of South Africa. *Molecular Phylogenetics and Evolution* 51: 100–110.

Wendel JF, Doyle JJ. 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. In: Soltis DE, Soltis PS, Doyle J eds. *Molecular systematics of plants II*. Boston: Kluwer Academic Publishers. 265–296.

Xiang XG, Jin WT, Li DZ, Schuiteman A, Huang WC, Li JW, Jin XH, Li ZY. 2014. Phylogenetics of tribe Collabieae (Orchidaceae, Epidendroideae) based on four chloroplast genes with morphological appraisal. *PLoS ONE* 9: e87625.

Xiang XG, Li DZ, Jin WT, Zhou HL, Li JW, Jin XH. 2012. Phylogenetic placement of the enigmatic orchid genera *Thaia* and *Tangtsinia*: Evidence from molecular and morphological characters. *Taxon* 61: 45–54.

Xu Q, Zhang GQ, Liu ZJ, Luo YB. 2014. Two new species of *Dendrobium* (Orchidaceae: Epidendroideae) from China: Evidence from morphology and DNA. *Phytotaxa* 174: 129–143.

Yang HJ, Xu H, Cui DF, Li YD. 2013. *Cystorchis*, a newly recorded genus of Orchidaceae from Hainan Island in China. *Subtropical Plant Science* 42: 177–180.

Zhai JW, Zhang GQ, Chen LJ, Xiao XJ, Liu KW, Tsai WC, Hsiao YY, Tian HZ, Zhu JQ, Wang MN, Wang FG, Xing FW, Liu ZJ. 2013. A new orchid genus, *Danxiaorchis*, and phylogenetic analysis of the tribe Calypsoeae. *PLoS ONE* 8: e60371.

Zhai JW, Zhang GQ, Li L, Wang MN, Chen LJ, Chung SW, Rodríguez FJ, Francisco-Ortegaf J, Lan SR, Xing FW, Liu ZJ. 2014. A new phylogenetic analysis sheds new light on the relationships in the *Calanthe* alliance (Orchidaceae) in China. *Molecular Phylogenetics and Evolution* 77: 216–222.

Zhang GQ, Li MH, Su YY, Chen LJ, Lan SR, Liu ZJ. 2015a. A new mycoheterotrophic genus, *Yunorchis*, and molecular phylogenetic relationships of the tribe Calypsoeae (Epidendroideae, Orchidaceae) inferred from plastid and nuclear DNA sequences. *PLoS ONE* 10: e0123382.

Zhang GQ, Liu KW, Chen LJ, Xiao XJ, Zhai JW, Li LQ, Cai J, Hsiao YY, Rao WH, Huang J, Ma XY, Chung SW, Huang LQ, Tsai WC, Liu ZJ. 2013. A new molecular phylogeny and a new genus, *Pendulorchis*, of the *Aerides–Vanda* alliance (Orchidaceae: Epidendroideae). *PLoS ONE* 8: e60097.

Zhang ZJ, Yan YJ, Tian Y, Li JS, Hea JS, Tang ZY. 2015b. Distribution and conservation of orchid species richness in China. *Biological Conservation* 181: 64–72.

Zou LH, Huang JX, Zhang GQ, Liu ZJ, Zhuang XY. 2015. A molecular phylogeny of Aeridinae (Orchidaceae: Epidendroideae) inferred from multiple nuclear and chloroplast regions. *Molecular Phylogenetics and Evolution* 85: 247–254.

Zou LH, Liu ZJ, Huang JX. 2014. *Vanda malipoensis*, a new species of *Vanda* (Orchidaceae: Epidendroideae; Vandeae) from China: Evidence from morphological and molecular phylogenetic analyses. *Phytotaxa* 186: 87–96.

Appendix

Members of the China Phylogeny Consortium

Zhi-Duan CHEN, An-Ming LU, Hong-Zhi KONG, Xiao-Quan WANG, Yin-Zheng WANG, Shi-Liang ZHOU
Institute of Botany, Chinese Academy of Sciences, Beijing, China
Shou-Zhou ZHANG, Xiao-Ming WANG
Fairylake Botanical Garden, Shenzhen, China
Zhong-Jian LIU
The Orchid Conservation and Research Center of Shenzhen, Shenzhen, China
Qing-Feng WANG
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
Jian-Hui LI
Computer Network Information Center, Chinese Academy of Sciences, Beijing, China
De-Zhu LI, Ting-Shuang YI
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Hong MA
Fudan University, Shanghai, China
Douglas E. SOLTIS, Pamela S. SOLTIS
University of Florida, Gainesville, USA
Jian-Hua LI
Hope College, Holland, USA
Cheng-Xin FU
Zhejiang University, Hangzhou, China
Qi-Xin LIU
Nanjing Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12187/suppinfo:

Table S1. Taxa and GenBank accession numbers for the Epidendroideae.

Table S2. Primers used for amplification and sequencing.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12211

Research Article

Phylogeny of the *Rosidae*: A dense taxon sampling analysis

Miao Sun[1,2,3,4], Rehan Naeem[5], Jun-Xia Su[6], Zhi-Yong Cao[1], J. Gordon Burleigh[3,7], Pamela S. Soltis[4,7], Douglas E. Soltis[3,4,7]*, and Zhi-Duan Chen[1]*

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Department of Biology, University of Florida, Gainesville, FL 32611, USA
[4]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
[5]Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology, Kohat 26000, Pakistan
[6]College of Life Science, Shanxi Normal University, Linfen 041004, Shanxi, China
[7]Genetics Institute, University of Florida, Gainesville, FL 32608, USA
*Authors for correspondence. Douglas E. Soltis. E-mail: dsoltis@ufl.edu. Tel.: 1-352-273-1963. Fax: 1-352-846-2154. Zhi-Duan Chen. E-mail: zhiduan@ibcas.ac.cn. Tel.: 86-10-62836434. Fax: 86-10-62590843.
Received 18 January 2016; Accepted 18 May 2016; Article first published online 22 July 2016

Abstract *Rosidae*, a clade of approximately 90 000 species of angiosperms, exhibits remarkable morphological diversity and extraordinary heterogeneity in habitats and life forms. Resolving phylogenetic relationships within *Rosidae* has been difficult, in large part due to nested radiations and the enormous size of the clade. Current estimates of phylogeny contain areas of poor resolution and/or support, and there have been few attempts to synthesize the available data into a comprehensive view of *Rosidae* phylogeny. We aim to improve understanding of the phylogeny of *Rosidae* with a dense sampling scheme using both newly generated sequences and data from GenBank of the chloroplast *rbcL*, *atpB*, and *matK* genes and the mitochondrial *matR* gene. We combined sequences from 9300 species, representing 2775 genera, 138 families, and 17 orders into a supermatrix. Although 59.26% of the cells in the supermatrix have no data, our results generally agree with previous estimates of *Rosidae* phylogeny and provide greater resolution and support in several areas of the topology. Several noteworthy phylogenetic relationships are recovered, including some novel relationships. Two families (Euphorbiaceae and Salvadoraceae) and 467 genera are recovered as non-monophyletic in our sampling, suggesting the need for future systematic studies of these groups. Our study shows the value of a botanically informed bioinformatics approach and dense taxonomic sampling for resolving rosid relationships. The resulting tree provides a starting point for large-scale analyses of the evolutionary patterns within *Rosidae*.

Key words: phylogeny, rapid radiation, *Rosidae*, supermatrix.

With approximately 90 000 species (estimated from Hinchliff et al., 2015), 135–140 families, and 17 orders (Soltis et al., 2005; APG III, 2009; APG IV, 2016), *Rosidae* contains at least one quarter of all angiosperm species and approximately 39% of eudicot species diversity (Magallón et al., 1999; Wang et al., 2009). Molecular dating indicates that *Rosidae* originated in the Early to Late Cretaceous, between 115 and 93 million years ago (Mya), followed by rapid diversification resulting in the *Fabidae* and *Malvidae* crown groups approximately 112 to 91 Mya (Albian to Coniacian) and 109 to 83 Mya (Cenomanian to Santonian), respectively (Wang et al., 2009; Bell et al., 2010), with major lineages diversifying in perhaps as little as 4 to 5 million years (Wang et al., 2009; Soltis et al., 2010). The radiations in *Rosidae* also represent the rapid rise of angiosperm-dominated forests and associated co-diversification events that have profoundly shaped much of the current terrestrial biodiversity (Wang et al., 2009).

This extraordinarily diverse clade exhibits enormous heterogeneity in habitats and life forms, including herbs, shrubs, trees, vines, aquatics, succulents, and parasites. Species of *Rosidae* generally have bitegmic, crassinucellate ovules, distinguishing them from *Asteridae*, which are generally characterized by unitegmic, tenuinucellate ovules. Moreover, some members possess novel biochemical pathways, such as glucosinolate production and cyanogenic glycosides in Brassicales (Rodman et al., 1998; Soltis et al., 2005; Edger et al., 2015). Symbioses with nitrogen-fixing bacteria also characterize this clade (Soltis et al., 1995; Li et al., 2015). Many important crops and economic plants, including legumes (Fabaceae) and fruit crops (Rosaceae), are also members of *Rosidae*.

Recent developments in software and strategies for phylogeny reconstruction make it feasible to assemble and analyze far larger numbers of terminals than ever before (e.g., de Queiroz & Gatesy, 2007; Goloboff et al., 2009;

Smith et al., 2009, 2011; Bazinet et al., 2014; Stamatakis, 2014; Hinchliff et al., 2015; Xu et al., 2015). As a result, much progress has been made in establishing a phylogenetic framework of angiosperms using plastid, mitochondrial, and nuclear gene sequence data (e.g., Qiu et al., 2010; Soltis et al., 2011; Zhang et al., 2012; Ruhfel et al., 2014; Xi et al., 2014; Zeng et al., 2014). These recent studies reveal a well-supported *Rosidae*, in which Vitaceae are generally placed as sister to eurosids, which in turn comprise two major clades, *Fabidae* (i.e., eurosids I, fabids) and *Malvidae* (i.e., eurosids II, malvids). Within *Fabidae*, Zygophyllales are sister to two major subclades, the nitrogen-fixing clade (Cucurbitales, Fagales, Fabales, Rosales; Soltis et al., 1995) and Celastrales–Oxalidales–Malpighiales (the COM clade; Matthews & Endress, 2006; Zhu et al., 2007). Within *Malvidae*, Crossosomatales are sister to ((((Malvales + Brassicales) + Huerteales) + Sapindales) + Picramniales), and this whole clade above is sister to (Geraniales + Myrtales) (Soltis et al., 1999, 2000, 2005, 2007, 2011; Hilu et al., 2003; Judd & Olmstead, 2004; Jansen et al., 2007; Wang et al., 2009; Moore et al., 2010, 2011; Ruhfel et al., 2014).

However, nested radiations and complex evolutionary histories within *Rosidae* (Soltis et al., 2004) make phylogenetic inferences difficult, and the placements of certain subclades of *Rosidae* (e.g., the COM clade, Zygophyllales, relationship of Geraniales and Myrtales) have varied (Hilu et al., 2003; Judd & Olmstead, 2004; Jansen et al., 2007; Zhu et al., 2007; Burleigh et al., 2009; Qiu et al., 2010; Zhang et al., 2012; Ruhfel et al., 2014; Xi et al., 2014; Zeng et al., 2014; Yang et al., 2015). An intriguing example of a problematic group is the COM clade, which is strongly placed in *Fabidae* based on datasets dominated by plastid genes (e.g., Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014), but it is placed in *Malvidae* based on studies using nuclear and mitochondrial gene sequence data (Zhu et al., 2007; Qiu et al., 2010; Zhang et al., 2012; Wang et al., 2014; Xi et al., 2014). Certain morphological features (e.g., the inner integument of the ovule, contorted petals) also support the association of the COM clade and *Malvidae* (Matthews & Endress, 2006). Our recent study has detected both phylogenetic signals of COM clade + *Fabidae* and COM clade + *Malvidae* from nuclear genome data, suggesting such phylogenetic discord may be due to an ancient reticulation or lineage sorting event (Sun et al., 2015).

Hence, despite rapid progress in elucidating the major branches of rosid phylogeny, relationships among many orders and families (sensu APG III, 2009) and patterns of diversification of *Rosidae* are still unknown. In the present study, we expanded on the sampling in the phylogeny of *Rosidae* to evaluate relationships among orders and the monophyly of families and many of the large genera. We used a botanically informed bioinformatics approach to synthesize available and new sequence data, carefully evaluating and curating the data in light of our understanding of phylogeny in *Rosidae*, and we explored in detail the resulting synthesis and how it related to past phylogenetic studies.

Material and Methods

Data assembly and alignment

A sampling strategy with an aim of maximizing taxon sampling while minimizing missing data in the supermatrix was designed. Three chloroplast genes (*rbcL*, *atpB*, and *matK*) and one mitochondrial gene (*matR*) were selected for this study, because they are single copy with no paralogy issues, easy to align across *Rosidae*, and have extensive sampling in GenBank. We combined 328 newly generated sequences (Table S1) with sequences retrieved from GenBank (released December, 2012). The 328 newly generated sequences will be submitted to GenBank together with sequences newly generated in Chen et al. (2016) as an integrated contribution of the China Phylogeny Consortium (Chen et al., 2016). Voucher information for these new sequences and the methods of DNA extraction, gene amplification, and sequencing, as well as primer sequences, are reported in Chen et al. (2016).

To assemble the data from GenBank, we first downloaded the DNA sequences from the nucleotide database for *Rosidae* as well as Saxifragales, Proteales, and Trochodendrales, which were used as outgroups. We used BLAST (Altshul et al., 1990) for nucleotides to compare existing alignments of our genes of interest (i.e., *rbcL*, *atpB*, *matK*, and *matR*) with the nucleotide sequences in GenBank, using an e-value cut-off of 1.0×10^{-5}. For each gene, we then used a custom Perl script to process the BLAST output, putting the sequences in the correct orientation and trimming off parts of the sequences that extended beyond the original alignments. We removed any sequences that were less than 300 bp in length, and any sequences that were not associated with a formal species name. We removed all subspecies designations (e.g., "subsp.", "var.", "f.", "cf.", and "aff."), and then pruned the data so that each species had at most one sequence per gene. When there were multiple sequences per gene from a species, we kept the longest sequence.

For each gene, we combined the new and GenBank sequences for alignment using MUSCLE (Edgar, 2004). For *matK* and *rbcL*, we first broke the sequences into four and five clusters, respectively, of roughly equal size, aligned each cluster with MUSCLE, and then made a profile alignment of the cluster alignments, also using MUSCLE. We visually inspected each alignment and edited them with Geneious version R6 (http://www.geneious.com, Kearse et al., 2012).

In order to identify erroneous sequences and minimize possible effects of alignment error or missing data from partial sequences, we checked the topologies of gene trees to identify and remove problematic sequences. We built maximum likelihood trees from each gene alignment with RAxML version 8.1.12 (Stamatakis, 2014) using the GTRCAT model. We also used RAxML to identify rogue taxa with the "dropsets" algorithm (Pattengale et al., 2010), and these rogue taxa were removed (Table S2). Furthermore, the gene tree topologies were visually examined, and sequences from taxa in obviously spurious locations in the tree were removed from the alignment.

The edited and pruned alignments of each gene were concatenated into a single supermatrix using FASconCAT version 1.0 (Kück & Meusemann, 2010). In the concatenated matrix, at least one gene was sampled per species; if a gene sequence was not available, the FASconCAT software treated it as missing data in the supermatrix.

Phylogenetic analyses

Maximum likelihood analyses for each individual gene alignment and the combined supermatrix were carried out using RAxML version 8.1.12 (Stamatakis, 2014) with 20, 100, 500, and to 1000 replicates, respectively, under the unpartitioned GTRCAT model as implemented on HiPerGator 1.0 Research Computing at the University of Florida (Gainesville, FL, USA), a Linux platform from National Science and Technology Infrastructure — National Specimen Information Infrastructure (Institute of Botany, Chinese Academy of Sciences, Beijing, China), and the Cyberinfrastructure for Phylogenetic Research (CIPRES) cluster (Miller et al., 2010). When the supermatrix was optimized and stable, the best RAxML trees and support values from both analyses of 100, 500, and 1000 bootstrap replicates varied little. Thus, to reduce computing time, we only ran 100 bootstrap replicates in all subsequent analyses.

Mitochondrial genes (e.g., *matR*) and plastid data occasionally have different phylogenetic signals (Zhu et al., 2007; Qiu et al., 2010; Sun et al., 2015). Therefore, although we emphasize the phylogenetic results from the concatenated four-gene tree, we also carefully evaluated and discussed strongly supported differences between trees inferred from *matR* and the plastid genes.

All resultant trees were manipulated for display and visual examination by Newick utilities (Junier & Zdobnov, 2010), Dendroscope (Huson & Scornavacca, 2012), MEGA 6.0 (Tamura et al., 2013), and FigTree version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). In our discussion of phylogenetic relationships, we mainly follow Angiosperm Phylogeny Group (APG) III (2009) for names of orders and families, unless other studies with significant changes after APG III (2009) are available. For major supraordinal clades, we follow Soltis et al. (2011) and Cantino et al. (2007).

Results and Discussion

The final supermatrix is composed of 9300 species-level operational taxonomic units, including 445 taxa from Saxifragales, Proteales, and Trochodendrales that were used as outgroups (Table S1). In our final matrix, the *atpB* alignment was 1491 bp long with 1293 sequences, *rbcL* was 1466 bp long with 6131 sequences, *matK* was 2104 bp long with 6978 sequences, and *matR* was 2424 bp long with 717 sequences. The total length of the supermatrix was 6846 bp, and 59.26% of the cells contained no data.

The topologies of *Rosidae* inferred using single-gene or combined-gene datasets are generally consistent, but with a few differences, particularly when comparing trees made from chloroplast loci with the *matR* tree (Figs. S1–S3). The three-chloroplast-gene tree largely agrees with the overall four-gene tree (Fig. S2). Finally, a well-supported phylogeny of *Rosidae* is obtained when all four genes are combined (Fig. 1 for overview; Fig. S3 for the complete tree). A summary tree at the family level is provided with bootstrap support (BS) values $\geq$50% (Fig. 2).

Among 9297 overall nodes throughout the 4-gene tree, 6071 (65.3%) nodes have $\geq$50% BS support, 4007 (43.1%) nodes have $\geq$75% BS support, 3245 (34.9%) nodes $\geq$85% BS support, and 1001 (10.8%) nodes have 100% BS support. The monophyly of most orders and families is well supported (Fig. 2; Table 1). Approximately 88.4% (122/138) of all of the sampled families are monophyletic with strong support (BS $\geq$ 85%); 10.1% (14/138) of them are monophyletic with low to moderate support (57% $\leq$ BS $<$ 85%). Two families (Euphorbiaceae and Salvadoraceae) are non-monophyletic in the four-gene tree (Figs. 2, S3; Table 1). Moreover, within our four-gene-combined analysis as well as the *matR* analysis (Figs. 2A, S1), we support the recognition of Peraceae family as suggested in APG III (2009) and proposed in APG IV (2016). Cynomoriaceae is recognized in Saxifragales (outgroup) in the *matR* tree (Fig. S1), and Betulaceae, Euphorbiaceae, Simaroubaceae, and Hypericaceae are non-monophyletic in the *matR* tree (Fig. S1). Salvadoraceae are also non-monophyletic when the three chloroplast genes are combined (Fig. S2), but Euphorbiaceae are monophyletic in this tree (Fig. S2). There are other placements that differ between *matR* (Fig. S1) and the three-chloroplast-gene tree (Fig. S2). The placements of the COM clade (Zhu et al., 2007; Qiu et al., 2010; Zhang et al., 2012; Wang et al., 2014; Sun et al., 2015), Vitales (Qiu et al., 2010; Moore et al., 2011; Zhang et al., 2012), Crossosomatales (Zhu et al., 2007; Qiu et al. 2010; Sun et al., 2015), and Zygophyllales (Qiu et al. 2010) also vary when comparing the *matR* tree (Fig. S1) with the trees from the three chloroplast genes and four combined genes (Figs. S2, S3), but the BS supports for these differences on the *matR* tree are $<$50%.

The *Superrosidae* (Moore et al. 2010; Soltis et al., 2011) is strongly supported as a clade in the four-gene tree (BS = 94%; Fig. 2B) with Vitaceae sister to eurosids. In addition, within eurosids, the two major sister clades *Malvidae* and *Fabidae*, as recognized in previous studies (Hilu et al., 2003; Judd & Olmstead, 2004; Soltis et al., 2005, 2007, 2011; Jansen et al., 2007; APG III, 2009; Burleigh et al., 2009; Moore et al., 2010, 2011; Wang et al., 2009; Soltis & Soltis, 2013; Xi et al., 2014), also receive strong support (BS = 86%), except a weakly supported Geraniales + Myrtales clade (BS = 58%; see discussion below) is isolated from *Malvidae*, as sister to *Fabidae* and the remaining *Mavidae*. However, the sister relationship between these two orders and *Fabidae* + the remaining *Mavidae* was also observed in previous studies (Zhu et al., 2007; Qiu et al., 2010; Zhang et al., 2012; Wang et al., 2014; Xi et al., 2014; Zeng et al., 2014; Sun et al., 2015; Yang et al., 2015).

Below we provide an overall clade-by-clade summary of *Rosidae* relationships recovered in our four-gene analysis.

Phylogenetic analysis

Relationship between Vitales and the rest of *Rosidae*

Vitales

With Saxifragales, Trochodendrales, and Proteales as outgroups, Vitales (Vitaceae) are monophyletic with high support (BS = 97%; Fig. 2B), as sister to the remaining *Rosidae* (eurosids), which have been recognized in most studies (APG III, 2009; Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014). However, various different topologies have been reported in other studies (Burleigh et al., 2011; Moore et al., 2011; Morton, 2011; Zhang et al., 2012; Zeng et al., 2014).

Vitaceae. Within Vitaceae, *Leea* (BS = 100%) is sister to all remaining genera in this family (BS = 97%; Fig. S3). In the second clade (BS = 96%), *Ampelocissus* is sister to the rest of the clade. Only four genera are monophyletic (*Ampelocissus*, *Tetrastigma*,

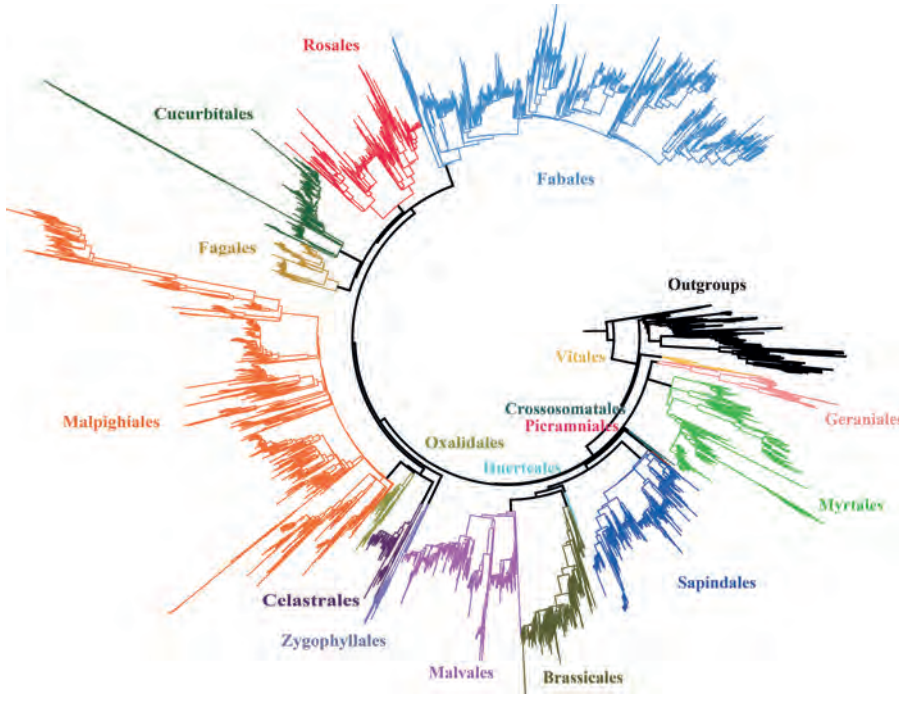


Fig. 1. Overview of the maximum likelihood phylogenetic tree of 9300 taxa of *Rosidae*. Each clade is color-coded, representing each order according to APG III (2009).

Cayratia, and *Cissus*), whereas *Parthenocissus* and *Vitis* are paraphyletic (Fig. S3). Previous studies have reported that other genera from this family are not monophyletic (Rossetto et al., 2007; Trias-Blasi et al., 2012; Lu et al., 2013; Zhang et al., 2015).

Relationships within *Fabidae*

The capability of symbiotic nitrogen fixation via root nodules is present in a single lineage of angiosperms known as the "nitrogen-fixing clade" (Soltis et al., 1995; Li et al., 2015), which includes Fabales (BS = 92%), Cucurbitales (BS = 96%), Fagales (BS = 99%), and Rosales (BS = 99%). *Fabidae* (eurosids I) is resolved as comprising the nitrogen-fixing clade (BS = 83%), the COM clade (BS = 52%), and Zygophyllales (BS = 98%). Zygophyllales are sister to the nitrogen-fixing clade plus the COM clade (BS = 79%), which agrees with previous findings using plastid sequence data (Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014). However, recent studies using nuclear genes have suggested that the COM clade is more closely related to *Malvidae* than *Fabidae* (Wang et al., 2014; Wickett et al., 2014; Sun et al., 2015; also see discussion below). The different signals from plastid and nuclear + mitochondrial genes suggest a complex history for the COM clade, perhaps involving ancient hybridizations (Sun et al., 2015).

Nitrogen-fixing clade

Within the nitrogen-fixing clade, Cucurbitales and Fagales are sisters with 79% BS support. This clade is in turn sister to Rosales (BS = 76%), and then all three are sister to Fabales (BS = 91%) (Fig. 2A). The topology of the nitrogen-fixing clade agrees with most previous studies (Zhu et al., 2007; APG III, 2009; Burleigh et al., 2009; Wang et al., 2009; Qiu et al., 2010; Soltis et al., 2011; Ruhfel et al., 2014 (for the all nucleotide positions analysis); Li et al., 2015, 2016), but different topologies were also proposed in a few other studies (Lee et al., 2011; Zhang et al., 2012; Ruhfel et al., 2014 (for the amino acid analysis); Xi et al., 2014).

Fagales

Fagales are strongly supported (BS = 99%), and include Nothofagaceae (BS = 99%), Fagaceae (BS = 100%), Juglandaceae (BS = 100%), Myricaceae (BS = 99%), Casuarinaceae (BS = 100%), Ticodendraceae (BS = 100%), and Betulaceae (BS = 100%). Nothofagaceae are sister to the rest of Fagales with high support (BS = 99%). Fagaceae and then Juglandaceae are subsequent sisters to a clade of Myricaceae + (Casuarinaceae + (Betulaceae + Ticodendraceae)) with strong support (Fig. 2A). These relationships differ from published topologies concerning the position of Myricaceae (Li et al., 2004; Herbert et al., 2006; Zhu et al., 2007; Soltis et al., 2011). For example, in Soltis et al. (2011), Myricaceae + Juglandaceae are sister to Betulaceae + Casuarinaceae. In our *matR* gene analysis, Myricaceae are sister to Casuarinaceae (BS = 87%), and Betulaceae are not monophyletic (Fig. S1). Overall, our four-gene-combined results agree with Sauquet et al. (2012) and Xiang et al. (2014).

Nothofagaceae. Nothofagaceae are a monogeneric family. Relationships recovered within the family concur with the recognized taxonomic sections (see *Nothofagus* website: nothofagus.free.fr).

Fagaceae. Internal relationships within Fagaceae are not fully resolved. *Fagus* (BS = 100%) is sister to other Fagaceae (BS = 100%; Sauquet et al., 2012), and then *Trigonobalanus* is sister to the rest of Fagaceae (BS = 100%; also see Manos et al., 2001; Xiang et al., 2014). Additionally, *Quercus* remains paraphyletic in Fagaceae (Fig. S3; also see Manos et al., 2001; Xiang et al., 2014).

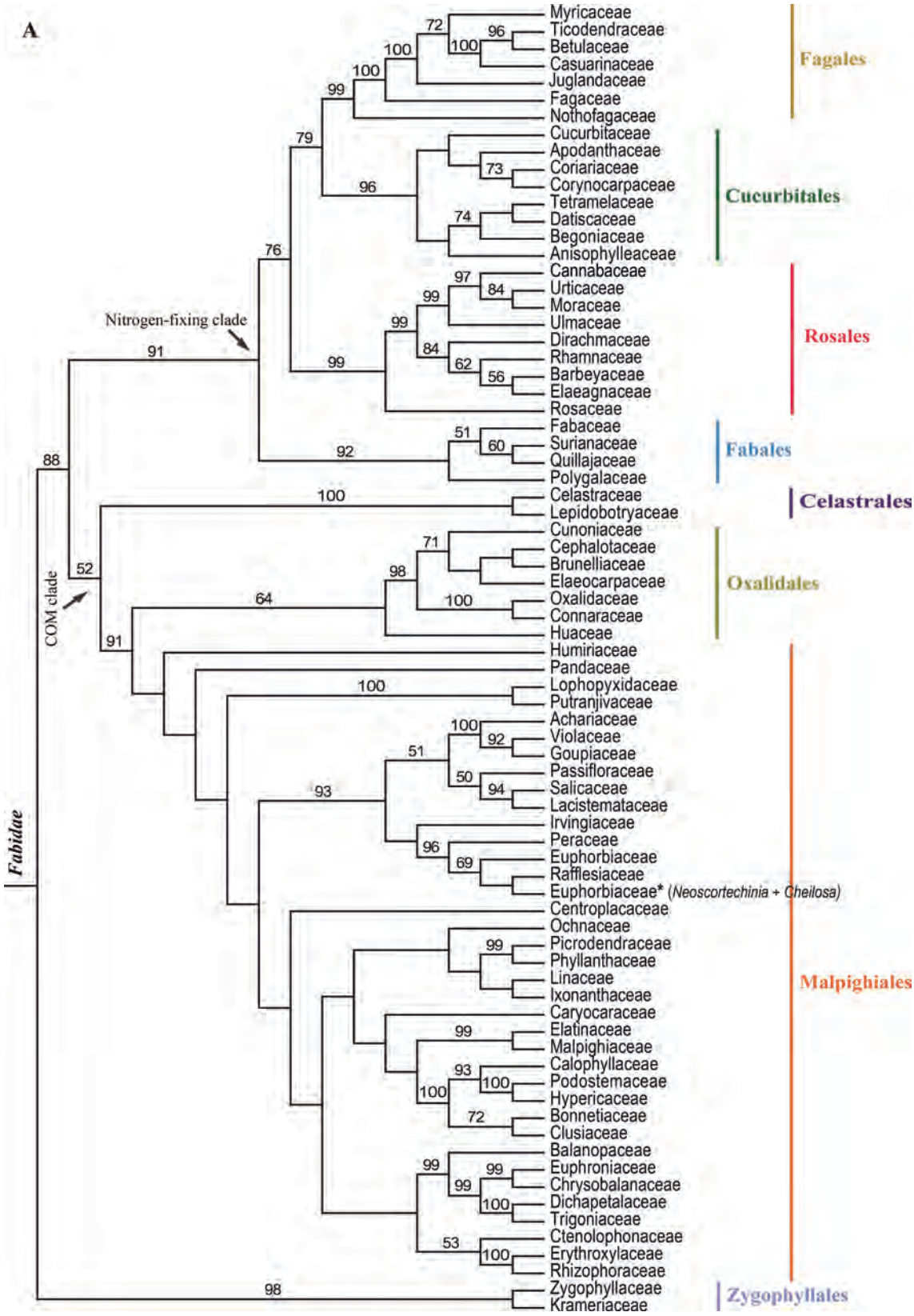


Fig. 2. Summary family tree of *Rosidae* with bootstrap (BS) values on most nodes from the maximum likelihood analysis. **A**, *Fabidae*. **B**, *Malvidae* and outgroups. All familial and ordinal names follow APG III (2009), Cantino et al. (2007), and Soltis et al. (2011). Numbers above branches are BS values; BS < 50% are not shown. *The family is resolved as non-monophyletic; those family names in brackets indicate new changes occurred according to APG IV (2016). All rosid orders and most families are well supported.

Continued

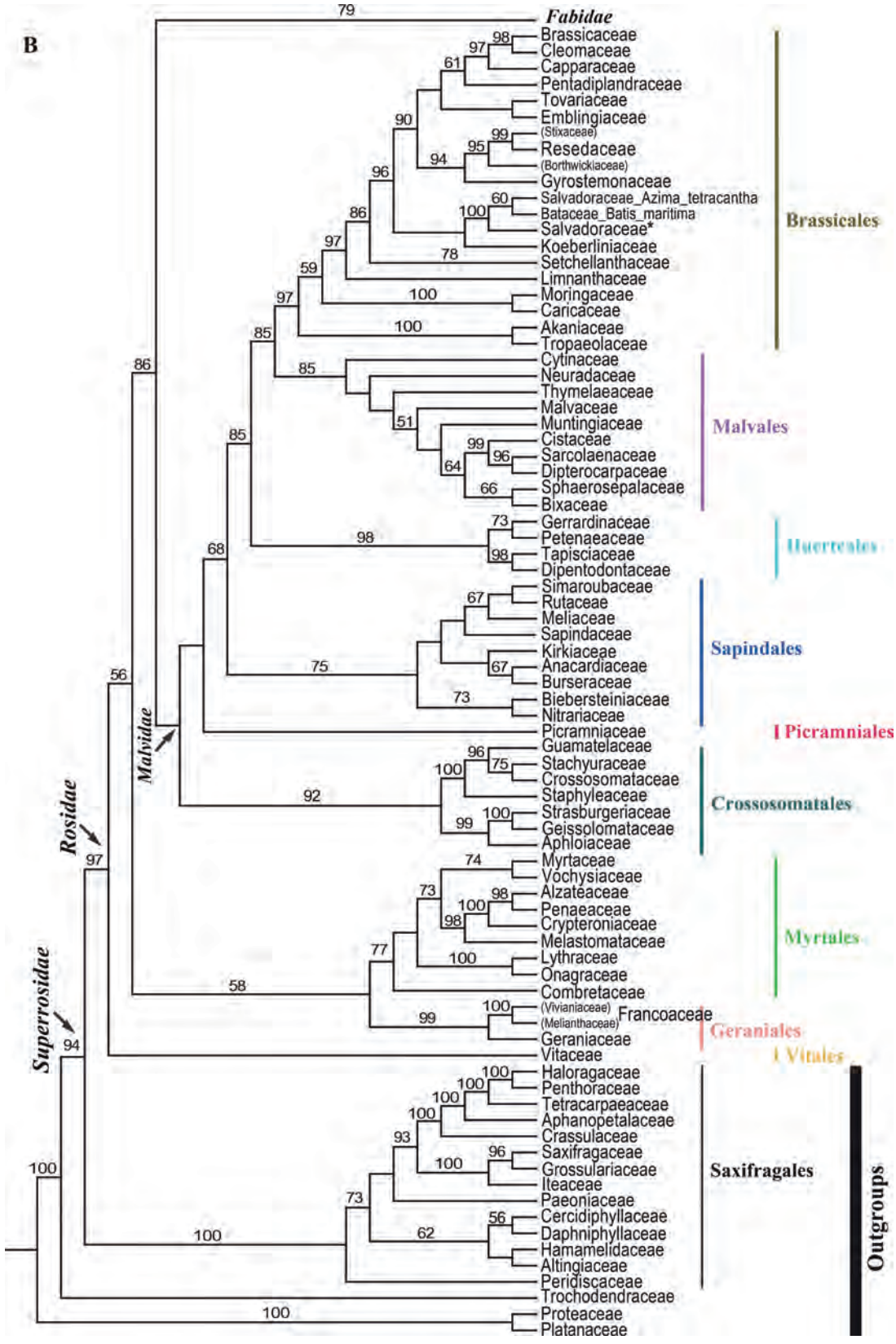


Fig. 2. Continued

Table 1 Support for monophyly of rosid families

Order	Family	BS%	Order	Family	BS%
Fagales	Nothofagaceae	99	Zygophyllales	Krameriaceae	100
	Juglandaceae	100		Zygophyllaceae	89
	Ticodendraceae	100	Brassicales	Limnanthaceae	100
	Betulaceae	100		Bataceae	100
	Casuarinaceae	100		Salvadoraceae*	–
	Myricaceae	99		Koeberliniaceae	100
	Fagaceae	100		Gyrostemonaceae	100
Cucurbitales	Corynocarpaceae	100		(Borthwickiaceae)	100
	Coriariaceae	100		(Stixaceae)	73
	Cucurbitaceae	99		Resedaceae	99
	Datiscaceae	97		Emblingiaceae	100
	Begoniaceae	99		Tovariaceae	100
	Tetramelaceae	100		Pentadiplandraceae	100
	Anisophylleaceae	100		Cleomaceae	97
	Apodanthaceae	100		Brassicaceae	98
Rosales	Ulmaceae	100		Capparaceae	95
	Cannabaceae	85		Setchellanthaceae	100
	Moraceae	100		Tropaeolaceae	100
	Urticaceae	99		Akaniaceae	94
	Rhamnaceae	99		Moringaceae	100
	Barbeyaceae	100		Caricaceae	100
	Elaeagnaceae	100	Malvales	Cytinaceae	100
	Dirachmaceae	100		Neuradaceae	99
	Rosaceae	100		Malvaceae	94
Fabales	Polygalaceae	100		Muntingiaceae	100
	Fabaceae	92		Cistaceae	100
	Quillajaceae	100		Sarcolaenaceae	97
	Surianaceae	100		Dipterocarpaceae	64
Malpighiales	Irvingiaceae	100		Sphaerosepalaceae	99
	Calophyllaceae	90		Bixaceae	89
	Hypericaceae	100		Thymelaeaceae	94
	Podostemaceae	100	Huerteales	Dipentodontaceae	83
	Bonnetiaceae	100		Tapisciaceae	74
	Clusiaceae	100		Petenaeaceae	100
	Elatinaceae	100		Gerrardinaceae	100
	Malpighiaceae	99	Sapindales	Biebersteiniaceae	100
	Caryocaraceae	100		Nitrariaceae	99
	Centroplacaceae	99		Anacardiaceae	57
	Linaceae	100		Burseraceae	84
	Ixonanthaceae	100		Kirkiaceae	99
	Picrodendraceae	100		Sapindaceae	78
	Phyllanthaceae	99		Meliaceae	98
	Ochnaceae	98		Rutaceae	99
	Goupiaceae	100		Simaroubaceae	69
	Violaceae	100	Picramniales	Picramniaceae	100
	Achariaceae	81	Crossosomatales	Aphloiaceae	100
	Lacistemataceae	100		Geissolomataceae	100
	Salicaceae	94		Strasburgeriaceae	100
	Passifloraceae	99		Guamatelaceae	100
	Peraceae	78		Crossosomataceae	100
	Euphorbiaceae*	–		Stachyuraceae	100
	Rafflesiaceae	100		Staphyleaceae	100
	Pandaceae	100	Myrtales	Combretaceae	98
	Balanopaceae	100		Onagraceae	99
	Trigoniaceae	100		Lythraceae	97
	Dichapetalaceae	100		Penaeaceae	62

Continued

Table 1 Continued

Order	Family	BS%	Order	Family	BS%
	Euphroniaceae	100		Alzateaceae	100
	Chrysobalanaceae	99		Crypteroniaceae	100
	Ctenolophonaceae	100		Melastomataceae	98
	Rhizophoraceae	100		Vochysiaceae	96
	Erythroxylaceae	100		Myrtaceae	79
	Humiriaceae	100	Geraniales	Geraniaceae	100
	Putranjivaceae	100		(Vivianiaceae)	100
	Lophopyxidaceae	100		(Melianthaceae)	100
Celastrales	Lepidobotryaceae	100	Vitales	Vitaceae	97
	Celastraceae	100		Total	
Oxalidales	Huaceae	100			
	Connaraceae	100		Support (BS%)	No.
	Oxalidaceae	95		99–100	99
	Cephalotaceae	100		85–100	122
	Brunelliaceae	100		57–84	14
	Cunoniaceae	80		<57	0
	Elaeocarpaceae	69		Non-monophyletic	2

*The family is resolved as non-monophyletic; family names in brackets indicate the family is not recognized in APG IV (2016). −, Not applicable. BS% = bootstrap support percentage.

Juglandaceae. For Juglandaceae, we follow the treatment of APG III (2009), which has an expanded Juglandaceae with inclusion of the monogeneric *Rhoiptelea*. In our analyses, *Rhoiptelea* is sister to the remainder of the family. Within core Juglandaceae, two clades are resolved, corresponding to subfamilies Engelhardioideae and Juglandoideae. We corroborate the previous treatment of retaining *Alfaropsis* in *Engelhardia* (Fig. S3; Manos & Stone, 2001). *Juglans* is non-monophyletic.

Myricaceae. Within Myricaceae, *Canacomyrica* and *Comptona* are successive sisters to *Myrica* + *Morella* with strong support, in agreement with previous studies (Herbert et al., 2006; Xiang et al., 2014; Fig. S3).

Casuarinaceae. Four genera of Casuarinaceae were sampled in our study, and internal relationships within the family are strongly supported (BS = 100%); *Gymnostoma* and then *Ceuthostoma* are the subsequent sisters to *Casuarina* + *Allocasuarina*, congruent with Xiang et al. (2014).

Betulaceae. In Betulaceae, *Betula* is sister to the other genera, and then *Alnus*, *Corylus*, and *Ostryopsis* are successive sisters to *Ostrya* and *Carpinus* with 54%, 95%, and 63% BS support, respectively. Additionally, the monophyly of each of these genera is well supported, with the exception of *Ostrya*, which is paraphyletic (Fig. S3; also see Xiang et al., 2014).

Cucurbitales

Cucurbitales are monophyletic with strong support and include Apodanthaceae (BS = 100%), Tetramelaceae (BS = 100%), Datiscaceae (BS = 97%), Begoniaceae (BS = 99%), Anisophylleaceae (BS = 100%), Coriariaceae (BS = 100%), Corynocarpaceae (BS = 100%), and Cucurbitaceae (BS = 99%). The relationships among families in this order remain to be resolved.

Anisophylleaceae. The relationships among genera within Anisophylleaceae are resolved with strong support. *Combretocarpus* and *Polygonanthus* are collectively sister to *Poga* + *Anisophyllea* (BS ≥ 94%).

Begoniaceae. Within Begoniaceae, *Hillebrandia* is sister to *Begonia* and *Symbegonia* (BS = 99%); however, *Symbegonia* is nested within *Begonia*. Hence, we prefer to reduce *Symbegonia* as a section of *Begonia* (see Forrest & Hollingsworth, 2003).

Cucurbitaceae. The relationships within Cucurbitaceae are generally not well supported. *Neoalsomitra* and *Gynostemma* are closely related (BS = 74%), and *Gomphogyne* is nested in *Hemsleya* (BS = 100%), causing *Hemsleya* to be paraphyletic. These two groups are then allied with each other (BS = 63%), followed by *Bayabusua* (BS = 85%). *Pteropepon* is sister to *Sicydium* (BS = 87%), followed by *Fevillea* (BS = 87%). Likewise, *Zanonia* is sister to *Siolmatra* (BS = 100%) and then in turn sister to *Xerosicyos* (BS = 68%). The rest of Cucurbitaceae forms a clade with 50% BS support. Within this clade, *Actinostemma* + *Bolbostemma* (BS = 100%) and *Siraitia* are successive sisters to the remaining taxa of this group. *Cogniauxia* is sister to *Ampelosicyos–Odosicyos* with strong support. *Austrobryonia* is sister (BS = 90%) to *Ecballium* + *Bryonia* (BS = 100%). Similarly, *Ctenolepis* is sister (BS = 97%) to *Trochomeria* + *Dactyliandra* (BS = 88%). *Tecunumania*, *Schizocarpum*, and *Cionosicyos* are subsequent sisters to *Abobra* + *Cayaponia* (BS = 68%). *Trichosanthes* (in part), *Echinocystis*, and *Marah* are successive sisters to an unresolved group. Within this group, *Parasicyos* and *Sechium* + *Sicyosperma* are successive sisters to *Sechiopsis* + *Microsechium–Sicyos* complex. *Echinopepon* is not monophyletic with *Frantzia* nested within it. *Hanburia* + *Cyclanthera* (BS = 56%) and *Rytidostylis* + *Pseudocyclanthera* (BS = 98%) are sisters with 72% BS support. Additionally, *Baijiania* and *Thladiantha* are sisters (BS = 59%), as well as *Bambekea* and *Eureiandra* (BS = 83%). *Psiguria*, *Gurania*, and *Helmontia* are closely related (BS = 94%), and *Schizopepon* and *Herpetospermum* are sisters (BS = 89%).

Indomelothria is sister to *Melothria–Cucumeropsis* (BS = 80%). The relationships among the remaining genera are unresolved. *Microsechium* is nested in *Sicyos*. Similarly, both *Hodgsonia* and *Gymnopetalum* are nested in *Trichosanthes*, and *Neoachmandra* and *Zehneria* are interdigitated. Likewise, *Ampelosicyos*, *Melothria*, *Psiguria*, and *Trichosanthes* are not monophyletic.

Rosales

Rosales are monophyletic with strong support, and relationships among the families are well supported. The families within the order are also monophyletic (Fig. 2A; Table 1). Rosaceae are monophyletic (BS = 100%) and sister to the rest of Rosales (BS = 99%). The remaining Rosales form two clades, with 99% and 84% BS support, respectively. One is (Rhamnaceae + (Elaeagnaceae + Barbeyaceae)) + Dirachmaceae, with 56%, 62%, and 84% BS support, respectively. The other is ((Urticaceae + Moraceae) + Cannabaceae) + Ulmaceae, with 84%, 97%, and 99% BS support, respectively. Relationships among families in Rosales recovered here agree with previous studies (e.g., Wang et al., 2009; Soltis et al., 2011).

Rosaceae. Rosaceae comprise three subclades, corresponding to Rosoideae (BS = 100%), Dryadoideae (BS = 100%), and Amygdaloideae (BS = 100%), following Stevens (2001 onwards). Rosoideae plus Dryadoideae (BS = 68%) are sister to Amygdaloideae (BS = 100%), in agreement with Potter (2003) and Chin et al. (2014).

Within Rosoideae, *Filipendula* is sister to the rest of Rosoideae (Rosodae; Potter et al., 2007a) with 100% BS support. Within Rosodae, Sanguisorbeae are sister to Potentileae + *Rosa*, with strong support. Within Colurieae, *Taihangia* is nested within *Geum*, causing it to be non-monophyletic, and the phylogenetic placement of *Rubus* is not resolved. It may be allied with Colurieae (also see Eriksson et al., 2003; Potter et al., 2007a). Within Sanguisorbeae, *Leucosidea* + *Agrimonia* and *Hagenia* are successive sisters to a clade of *Sanguisorba*–(*Polylepis* + (*Acaena* + *Cliffortia*)) with ≥67% BS support; *Sanguisorba* is non-monophyletic. For relationships within Potentileae (Potter et al., 2007a), *Potentilla* (including *Ivesia* and *Duchesnea*) is sister to subtribe Fragariinae with 100% BS support (also see Eriksson et al., 2003; Potter et al., 2007a; Dobeš & Paule, 2010). Additionally, this supports the expansion of *Potentilla* to include *Ivesia* and *Duchesnea* (Potter et al., 2007a). In subtribe Fragariinae, *Sibbaldia* is well supported as sister to two sister groups, *Fragaria* + (*Drymocallis* + *Dasiphora*) and *Comarum* + *Alchemilla–Aphanes* (BS ≥ 65%). However, Eriksson et al. (2015), using internal transcribed spacer (ITS), *trnL-F* spacer, and *trnL* intron DNA data, showed that *Sibbaldia* is polyphyletic, falling into five separate clades scattered within Fragariinae and *Potentilla*. Moreover, *Alchemilla* is not monophyletic unless *Aphanes* is included (also see Gehrke et al., 2008).

Within Dryadoideae, *Dryas* and then *Purshia* are successive sisters to *Chamaebatia* + *Cercocarpus* with 100% and 72% BS support, respectively.

Within Amygdaloideae (Spiraeaoideae; see Potter et al., 2007b), *Lyonothamnus* is resolved as sister to the rest of this subfamily (BS = 100%). The remaining genera form six clades (corresponding to tribes and supertribes) whose circumscriptions largely agree with the taxonomic treatment of Potter et al. (2007b): Neillieae (BS = 100%), Kerriodae (BS = 93%), Spiraeeae (BS = 100%), Pyrodae (BS = 100%), Amygdaleae (BS = 100%), and Sorbarieae (BS = 100%). Among these six clades, Amygdaleae and Sorbarieae are sisters (BS = 81%), and this clade is then sister to supertribe Pyrodae (BS = 55%), followed successively by Spiraeeae (BS = 54%), and Kerriodae + Neillieae (BS = 94%). Within Neillieae, *Physocarpus* and *Neillia* are sisters (BS = 100%). Supertribe Kerriodae are composed of two sister groups, tribe Osmaronieae (BS = 100%) and tribe Kerrieae (BS = 100%). The topology of tribe Osmaronieae is (*Exochorda* + *Prinsepia*) + *Oemleria* (BS ≥ 60%; also see Lee & Wen, 2001; Potter et al., 2002). The topology of tribe Kerrieae is ((*Neviusia* + *Kerria*) + *Coleogyne*) + *Rhodotypos* (BS ≥ 99%). Within Spiraeeae, (*Luetkea* + *Aruncus*) + *Holodiscus* (BS = 100%) is sister to ((*Petrophytum* + *Kelseya*) + *Spiraea–Padus*) + *Pentactina* (BS ≥ 71%) with high support. The relationship between *Spiraea* and *Padus* is not resolved here. Within supertribe Pyrodae, *Gillenia* is sister to tribe Pyreae (BS = 100%), and within tribe Pyreae, *Lindleya*, *Kageneckia*, and then *Vauquelinia* are successive sisters to subtribe Pyrinae (BS ≥ 75%, but with BS < 50% for the relationship of *Kageneckia* to the others). The relationships within subtribe Pyrinae are poorly resolved, except *Chaenomeles* and *Pseudocydonia* are sisters (BS = 87%), and *Sorbus* appears to be polyphyletic (also see Lo & Donoghue, 2012), with species closely associated with *Aria* and *Torminalis*. Within tribe Amygdaleae, *Pygeum*, *Laurocerasus*, *Cerasus*, and *Armeniaca* are scattered in *Prunus*, causing the latter to be non-monophyletic. This suggests that these genera might be included in *Prunus* to maintain its monophyly (also see Potter et al., 2007b). Within tribe Sorbarieae, *Sorbaria* is sister to *Chamaebatiaria* (BS = 91%), and these are then collectively sister to *Adenostoma* (BS = 100%).

Elaeagnaceae. Within Elaeagnaceae, *Shepherdia* and *Hippophae* are sisters (BS = 100%), and then collectively sister to *Elaeagnus* (BS = 100%).

Rhamnaceae. The relationships recognized in Rhamnaceae roughly agree with those of Richardson et al. (2000), with the "Rhamnoid" and "Ziziphoid" groups recovered, but with BS < 50%. In the "Rhamnoid" group, tribe Maesopsideae, represented by *Maesopsis*, are sister to tribe Rhamneae (BS = 81%). Tribe Rhamneae form a clade (BS = 70%) with the topology ((((*Rhamnella* + *Rhamnidium–Karwinskia*) + (*Krugiodendron* + *Condalia*)–*Reynosia*) + *Berchemia*) + *Sageretia*) + (*Scutia* + (*Frangula* + *Rhamnus*)), and *Berchemia* is non-monophyletic. Within the "Ziziphoid" group, tribe Paliureae is strongly supported (BS = 96%), with the topology of *Hovenia* + (*Paliurus* + *Ziziphus*). Relationships in tribe Colletieae are recovered as *Discaria* + (*Colletia* + *Retanilla*); *Alphitonia* is sister to *Granitites* (BS = 64%). Within tribe Phyliceae, *Noltea* is sister to *Trichocephalus*–(*Phylica* + *Nesiota*) with 61% BS support, but the BS support for *Trichocephalus* to *Phylica* + *Nesiota* < 50%. Within tribe Gouanieae, *Pleuranthodes*–(*Crumenaria* + *Reissekia*) is sister to *Gouania* + *Helinus* with 61% BS support.

Ulmaceae. Ulmaceae are strongly supported (BS = 100%), and the relationships among genera within this family is (*Holoptelea* + *Ampelocera*) + (*Hemiptelea* + (*Zelkova* + *Ulmus*)), with 100% BS support at each node (also see Sytsma et al., 2002).

Moraceae. The phylogeny of Moraceae is not well resolved. *Artocarpus* is sister to an unresolved group (BS = 71%) including *Trophis*, *Taxotrophis*, *Bagassa*, *Sorocea*, *Malaisia*, *Morus*, *Streblus*, and *Prainea*, with 98% BS support. The rest of the genera form a clade with 73% BS support. Within this clade, *Dorstenia* is sister to two groups (BS = 98%), *Trymatococcus* + *Brosimum–Clarisia* (BS = 100%) and *Malaisia–Broussonetia* (BS = 76%). Likewise, *Ficus* is sister to another group (BS = 96%) including *Naucleopsis*, *Perebea*, *Maquira*, *Poulsenia*, *Pseudolmedia*, *Trophis*, *Streblus*, *Castilla*, *Antiaris*, and *Helicostylis* (BS = 95%). Additionally, *Brosimum*, *Broussonetia*, *Maquira*, *Morus*, *Streblus*, *Trophis*, and *Perebea* are non-monophyletic.

Urticaceae. Four clades are recovered within Urticaceae, congruent with Wu et al. (2013), which correspond to Urticeae, Elatostemateae, the Boehmerieae–Forsskaoleeae–Parietarieae clade, and the "Cecropieae" clade (see Hadiah et al., 2008; Kim et al., 2015). Cecropieae are embedded in Urticaceae, and they are polyphyletic, with *Poikilospermum* nested in Urticeae (BS = 70%; also see Hadiah et al., 2008; Wu et al. 2013; Kim et al., 2015). Within Urticeae, *Didymodoxa* + (*Obetia* + *Urera–Poikilospermum*) are sister to an unresolved group including *Laportea*, *Dendrocnide*, *Boehmeria*, *Urtica*, *Hesperocnide*, and *Girardinia* (BS = 97%). Within Elatostemateae, *Myriocarpa*–(*Lecanthus* + *Pilea*) is sister to *Pellionia–Elatostema* + *Procris–Pellionia* (BS = 72%), which indicates that *Pellionia* is paraphyletic (see Hadiah et al., 2003, 2008; Wu et al., 2013). Within the Boehmerieae–Forsskaoleeae–Parietarieae clade, Boehmerieae are sister to Forsskaoleeae and Parietarieae (BS = 97%). Within the "Cecropieae" clade, *Leucosyke* is sister to Cecropieae (excluding *Poikilospermum*; BS = 83%) with 81% BS support, and the topology of Cecropieae is *Cecropia* + (*Coussapoa* + *Pourouma*) with 83% and 61% BS support, respectively (also see Wu et al., 2013). *Urera*, *Urtica*, *Boehmeria*, and *Pouzolzia* are not monophyletic.

Cannabaceae. The monophyly of Cannabaceae is strongly supported. Within this family, *Aphananthe*, *Gironniera*, and *Lozanella* are successive sisters to the rest of the family, with BS support for these nodes of 85%, 90%, and 52%, respectively (Fig. S3; see van Velzen et al., 2006; Yang et al., 2013). The remaining genera form two groups: *Pteroceltis* + *Chaetachme* (BS = 99%) are sister to *Humulus* + *Cannabis*. This clade is then sister to a clade of *Celtis*, *Trema*, and *Parasponia*, with *Parasponia* nested within *Trema*, causing *Trema* to be non-monophyletic, in agreement with previous studies (Sytsma et al., 2002; van Velzen et al., 2006; Yang et al., 2013).

Fabales

In our analyses, Fabales are monophyletic (BS = 92%) and comprise Fabaceae (BS = 92), Polygalaceae (BS = 100%), Quillajaceae (BS = 100%), and Surianaceae (BS = 100%). Polygalaceae are sister to the rest of Fabales (BS = 92%), which form a clade of Fabaceae + (Surianaceae + Quillajaceae) with weak support (Fig. 2A). This topology is generally consistent with previous studies (Wojciechowski et al., 2004; Bruneau et al., 2008; Bello et al., 2009; Cardoso et al., 2013; Koenen et al., 2013).

Polygalaceae. Within Polygalaceae, *Xanthophyllum* is sister to the rest (BS = 100%; see Forest et al., 2007; Bello et al., 2012). Among the remaining genera, (*Atroxima* + *Carpolobia*) + *Eriandra*, *Bredemeyera*, and *Securidaca* are successive sisters to an unresolved group of (*Comesperma* + *Monnina*)–*Polygala*–*Muraltia*, with BS support of 98%, 85%, and 96%, respectively. The relationship between *Comesperma* + *Monnina* (BS = 64%) and *Polygala* has <50% BS support, and *Polygala* is not monophyletic, with *Muraltia* nested within it.

Surianaceae. The phylogeny of Surianaceae is resolved as (*Suriana* + (*Guilfoylia* + *Stylobasium*)) + (*Recchia* + *Cadellia*) with ≥95% BS support at each node (Fig. S3), congruent with Forest et al. (2007) and Bello et al. (2009).

Fabaceae. The phylogenetic relationships within Fabaceae generally agree with the schematic phylogeny of the Legume Phylogeny Working Group (LPWG, 2013). The traditional subfamily Caesalpinioideae are not monophyletic; we recovered 11 separate clades, (i.e., Cercideae, Deterieae s.l., *Duparquetia*, Dialiinae s.l., *Umtiza*, *Cassia*, *Caesalpinia*, *Tachigali*, *Peltophorum*, and *Dimorphandra* groups A and B), with subfamilies Mimosoideae and Papilionoideae embedded among them (Fig. S4; also see Bruneau et al., 2008; LPWG, 2013; Li et al., 2016). The relationships among these 11 clades are not all strongly supported (Fig. S4).

Within Cercideae (exclude *Gigasiphon*), *Cercis* and *Adenolobus* are strongly placed as sister to the rest of the tribe, which are composed of two major clades: one is *Griffonia* + (*Bauhinia* + (*Brenierea* + *Piliostigma*)) (BS = 80%), and the other includes *Tylosema*, *Barklya*, *Lysiphyllum*, and *Phanera*, with the remaining species of *Bauhinia* scattered among them (also see Sinou et al., 2009).

Gigasiphon is sister to tribe Deterieae s.l. (Bruneau et al., 2001; MacKinder, 2005), but with <50% BS support. Within tribe Deterieae s.l., *Macrolobium* is sister to the rest of the tribe (BS = 73%), which is composed of the Amherstieae clade (BS = 59%), the resin-producing Detarieae, and an unresolved group of *Schotia* and *Barnebydendron* + *Goniorrhachis*. The resin-producing Detarieae (Fougère-Danezan et al., 2007) are composed of the *Prioria* clade (BS = 72%), Detarieae s.s. clade (BS = 57%), and the *Daniellia–Brandzeia* group, and the generic relationships within this tribe are largely consistent with Bruneau et al. (2008) and Li et al. (2016), except *Guibourtia* and *Sindora* are non-monophyletic. Within the Amherstieae clade, several well-resolved and strongly supported subclades and sister groups are identified (also see Bruneau et al., 2008; LPWG, 2013), namely, the *Saraca* clade (BS = 96%), the *Afzelia* clade (BS = 92%), the *Berlinia* clade (BS = 67%), the *Brownea* clade (BS = 75%), *Dicymbe* + *Polystemonanthus* (BS = 56%), *Paramacrolobium* + *Cryptosepalum* (BS = 83%), and *Loesenera* + (*Talbotiella* + *Leonardoxa* + *Hymenostegia*) (BS ≥ 66%). Additionally, *Hymenostegia*, *Cynometra*, *Brownea*, *Paloue*, *Macrolobium*, *Gilbertiodendron*, *Bikinia*, *Julbernardia*, *Berlinia*, and *Isoberlinia* are non-monophyletic in our current sampling.

Duparquetia is sister to the rest of Fabaceae, but BS support is <50%.

Within Dialiinae s.l., *Poeppigia* is resolved as sister to the remaining tribe (BS = 100%), but the generic relationships are not fully resolved, in agreement with Li et al. (2016).

Papilionoideae. The topology of subfamily Papilionoideae is similar to that found in recent analyses (Cardoso et al., 2013; Li et al., 2016), although the monophyly of this subfamily

received BS support <50% (Fig. S4). Most major clades are recovered here, in agreement with recent legume studies (Wojciechowski et al., 2004; Cardoso et al., 2012, 2013; LPWG, 2013), but the relationships among these clades are not fully resolved here, except for the Angylocalyceae, Dipterygeae, and Amburanae (ADA) clade, non-protein–amino-acid-accumulating (NPAAA) clade (=Old World clade, see LPWG, 2013), and Baphieae. We corroborate that the ADA clade is potentially the first-branching clade in Papilionoideae (Cardoso et al., 2012, 2013), and among the rest, Swartzieae (BS = 97%) and then the *Cladrastis* clade (BS = 91%) appear as successive sisters to the large 50-kb inversion clade, which comprises Exostyleae (BS = 96%), the Vataireoid clade (BS = 100%), Dalbergioids s.l., Genistoids s.l. (BS = 71%), the *Andira* clade (BS = 98%), *Aldina* (BS = 93%), and a clade (BS = 76%) of Baphieae (BS = 100%) + the NPAAA clade (BS = 52%).

Within the ADA clade, Dipterygeae (BS = 65%) and Amburanae are weakly supported as sisters (BS = 53%; also see Cardoso et al., 2012, 2013) and then allied with Angylocalyceae (BS = 65%). Within Amburanae, *Amburana* and *Cordyla* are forming a sister-group (BS = 90%), and *Myrospermum*, *Myroxylon*, and *Myrocarpus* are interdigitated. Neither *Myrospermum* nor *Myrocarpus* is monophyletic. Dipterygeae are resolved as (*Taralea* + (*Pterodon* + *Dipteryx*)) + *Monopteryx*. Similarly, the topology of Angylocalyceae is (*Castanospermum* + *Alexa*) + (*Xanthocercis* + *Angylocalyx*).

Swartzieae are resolved as two sister clades (Cardoso et al., 2013), the *Ateleia* clade (BS = 100%) and the *Swartzia* clade (BS = 99%). Within the *Ateleia* clade, *Bocoa* + *Trischidium* is sister to *Cyathostegia* + *Ateleia*. However, *Bocoa* is paraphyletic, with some members also allied with *Candolleodendron* in the *Swartzia* clade, in which *Bobgunnia* is resolved as sister to the remainder (BS = 99%).

Within the *Cladrastis* clade, *Pickeringia* is sister to the remainder (BS = 91%), and *Cladrastis* is paraphyletic.

Exostyleae (corresponds to the Lecointeoid clade; Cardoso et al., 2013) is sister to the remainder of the large 50-kb inversion clade, but the intergeneric relationships within this clade are unclear, although *Harleyodendron* and *Uribea* are weakly supported as sister-group (BS = 56%), and *Lecointea* is non-monophyletic.

The Vataireoid clade comprises *Vataireopsis* + (*Sweetia* + (*Luetzelburgia* + *Vatairea*)) (see Cardoso et al., 2013), except that *Vataireopsis surinamensis* H. C. Lima is nested in *Hymenolobium* within the *Andira* clade (but see Cardoso et al., 2012, 2013). Additionally, *Calia arizonica* (S. Watson) Yakovlev + *Sophora secundiflora* (Ortega) Lag. ex DC is sister to the Vataireoid clade, but with BS < 50%.

The Dalbergioid s.l. clade consists of two tribes, Amorpheae (BS = 99%) and Dalbergieae (Cardoso et al., 2013). Relationships in Dalbergieae are poorly supported. We recovered three clades, the *Adesmia* clade (BS = 98%), the *Pterocarpus* clade (BS = 83%), and the *Dalbergia* clade. Within the *Adesmia* clade, *Adesmia* is sister to two strongly supported subclades: one is (*Zornia* + *Amicia*) + *Poiretia*, and the other is a non-monophyletic *Chaetocalyx* with *Nissolia* nested within. Additionally, *Zornia* and *Chaetocalyx* are also non-monophyletic. Within the *Dalbergia* clade, *Soemmeringia* and *Cyclocarpa* are sisters. *Steinbachiella* is nested in *Machaerium*, forming a group with 90% BS support, and is sister to *Dalbergia* (BS = 98%). *Diphysa* and *Pictetia* are successive sisters to an unresolved clade of *Ormocarpum*, *Ormocarpopsis*, and *Peltiera*. Within the *Pterocarpus* clade, *Riedeliella* + *Discolobium* and *Platymiscium* are successive sisters to the rest of this clade, which in turn consists of three subclades: (i) *Geoffroea* + *Cascaronia* is strongly sister to (*Fissicalyx* + *Fiebrigiella*) + ((*Stylosanthes* + *Arachis*) + *Chapmannia*); (ii) *Brya* + *Cranocarpus* is weakly supported as sister to the third clade; and (iii) a poorly resolved clade in which a few relationships are recovered with weak support, namely (*Tipuana* + *Inocarpus*) + (*Ramorinoa* + *Centrolobium*). The topology of tribe Amorpheae is largely congruent with Cardoso et al. (2013). *Amorpha* and *Errazurizia* are not monophyletic here, and members of *Psorothamnus*, *Dalea*, and *Marina* are interdigitated.

The Genistoids s.l. clade forms a large and well-resolved clade including Ormosieae (BS = 92%), Brongniartieae (BS = 63%), Leptolobieae (BS = 100%), Camoensieae (*Camoensia*), *Bolusanthus*, and the Core Genistoids (BS = 66%). Within the Genistoids s.l. clade, Ormosieae, Brongniartieae, Leptolobieae, Camoensieae, and *Bolusanthus* are successive sisters to the Core Genistoids (BS ≥ 59%). Here, *Bolusanthus* is sister to the Core Genistoids with 96% BS support; however, only using *matK* sequence, Cardoso et al. (2013) identified *Bolusanthus* as a member of Sophoreae, sister to *Dicraeopetalum*. The specific reason for this difference in placement is unclear. Within Ormosieae, *Panurea* and *Ormosia* are successive sisters to *Clathrotropis* + *Spirotropis* with strong support. In our study, *Pericopsis* is resolved as sister to the rest of Brongniartieae (BS = 63%), in agreement with Doyle et al. (1997). However, *Pericopsis* was suggested to have an affinity to Leptolobieae (Cardoso et al., 2012). The remainder of Brongniartieae is further split into two well-supported clades (BS = 85%): one is *Poecilanthe* + (*Brongniartia* + (*Hovea* + *Callistachys*)), and the other is (*Poecilanthe* + *Harpalyce*) + ((*Hovea* + *Brongniartia*) + *Cyclolobium*). Both *Brongniartia* and *Hovea* are paraphyletic. Relationships within Leptolobieae are consistent with Cardoso et al. (2013), but *Leptolobium* is non-monophyletic. The Core Genistoids are composed of Podalyrieae (BS = 95%), Crotalarieae (BS = 99%), Genisteae (BS = 97%), and Sophoreae (excluding *Bolusanthus*; BS = 88%). Within the Core Genistoids, Crotalarieae and Genisteae are sisters (BS = 95%) and then sister to Podalyrieae (BS = 93%). This latter clade is then collectively sister to Sophoreae (BS = 66%). Podalyrieae are poorly resolved, and both *Cadia* and *Amphithalea* are non-monophyletic. Likewise, within Crotalarieae, a well-supported clade of *Euchlora* + (*Crotalaria* + *Bolusia*) is sister to the rest (BS = 99%), which forms a polytomy. Within this poorly resolved clade, *Lebeckia* and *Pearsonia* are non-monophyletic. Within Genisteae, *Melolobium*, *Dichilus*, *Anarthrophyllum*, *Polhillia* + *Argyrolobium*, and *Lupinus* are successive sisters to the rest (BS ≥ 55%), which forms an unresolved clade with 74% BS support. *Cytisus* is paraphyletic. Sophoreae comprise two clades, Thermopsideae (BS = 83%) and Sophoreae s.s. (BS = 73%). Relationships within Thermopsideae are resolved as ((*Piptanthus* + *Anagyris*) + *Thermopsis*) + *Baptisia* with moderate BS support. However, based on analysis of ITS sequences, Wang et al. (2006) suggested Thermopsideae is not a monophyletic group, with some species of *Sophora* s.s. nested within it. Sophoreae s.s. are not fully resolved; *Salweenia* + *Maackia* are sister to *Sophora*, with *Echinosophora*, *Euchresta*, and *Ammodendron* embedded in *Sophora*.

Baphieae are resolved as *Airyantha* + (*Baphia* + (*Baphiopsis* + *Leucomphalos*)). *Aldina* has traditionally been placed within Baphieae, but here is resolved as sister to Baphieae + the NPAAA clade, with BS support <50%.

The NPAAA clade includes the monotypic Hypocalypteae (*Hypocalyptus*; Schutte & van Wyk, 1998), the Mirbelioid clade (BS = 79%), the Hologalegina clade (BS = 97%), and the Indigoferoid clade (also Millettioids, BS = 91%). The monotypic *Hypocalyptus* is sister to the rest of the NPAAA clade, in agreement with Wojciechowski et al. (2004). Among the rest, the Hologalegina clade and Indigoferoid clade are sisters (BS = 86%), and then together are sister to the Mirbelioid clade (BS = 61%).

Within the Indigoferoid clade, Indigofereae (BS = 99%) are sister to the rest of the Indigoferoid clade with strong support. Within Indigofereae, *Phylloxylon* is not monophyletic, but forms two groups successively sister to an unresolved clade that includes *Indigofera*, *Microcharis*, and *Cyamopsis*. In the rest of the Indigoferoid clade, *Xeroderris* + *Craibia*, *Dalbergiella* + (*Centrosema* + *Clitoria*), and *Austrosteenisia* are resolved as successive sisters to the remainder, which includes the core Millettieae clade (Hu et al., 2002) and the Desmodieae, Phaseoleae, and Psoraleeae (DPP) complex (see also LPWG, 2013). *Clitoria* is not monophyletic, with some of *Clitoria* nested within *Chamaecrista* in the Cassia clade. *Dewevrea* and *Platycyamus* are successive sisters to the DPP complex (BS = 100%). The monophyly of the DPP complex is well supported, but generic relationships are often unclear. Within the DPP complex, *Apois* is sister to the rest of the clade, among which Desmodieae (BS = 100%) and Psoraleeae (BS = 92%) are interdigitated by members of a polyphyletic Phaseoleae. In our study, a monospecific *Disynstemon*, Abreae (*Abrus*), and Diocleinae (BS = 100%) are recovered as successive sisters to the core Millettieae clade (BS = 76%) with strong BS support, which is different from the placement suggested by Schrire et al. (2009) that *Disynstemon* is sister to tribe Indigofereae. Relationships within Diocleinae are resolved as (*Dioclea* + (*Rhodopis* + *Galactia*)) + *Canavalia*. *Dioclea megacarpa* Rolfe is nested in *Mucuna* (Phaseoleae). Within the core Millettieae clade, the *Philenoptera* clade (BS = 74%) is sister to the rest of the core Millettieae clade (BS = 76%), within which relationships are not fully resolved. Both *Millettia* and *Apurimacia* are non-monophyletic.

The Mirbelioid clade is composed of two well-resolved subclades: *Bossiaea* + (*Gompholobium* + (*Isotropis* + (*Gastrolobium* + *Aotus*))) and *Goodia* + (*Daviesia* + (*Chorizema* + *Isotropis*)). *Isotropis* is paraphyletic with species falling in both subclades.

The Hologalegina clade contains two major clades: the Robinioid clade (BS = 99%) and the IRLC clade (refering to the loss of one copy of the inverted repeat in the plastid genome; Wojciechowski et al., 2000; also see LPWG, 2013). The Robinioid clade is further split into three subclades, corresponding largely to tribes Sesbanieae (*Sesbania*), Loteae (BS = 100%), and Robinieae (BS = 100%). Generic relationships within all of these subclades are well supported, but relationships among the subclades need to be further clarified. *Hippocrepis*, *Lotus*, *Gliricidia*, and *Coursetia* are not monophyletic. The IRLC clade includes the traditional tribes Galegeae s.l. (BS = 75%, except *Galega*), Cicereae (*Cicer*), Fabeae (BS = 100%), and Trifolieae (BS = 95%, except *Trifolium*), as well as the *Callerya* clade, including *Afgekia*, *Endosamara*, *Wisteria*, and *Millettia*. However, neither Galegeae nor Trifolieae are monophyletic in our analysis. *Vicia*, *Caragana*, and *Callerya* are also non-monophyletic here (Fig. S4; also see LPWG, 2013).

The topology of the *Umtiza* clade is the same as in Li et al. (2016), but BS support for this clade is <50%. *Gymnocladus* is not monophyletic.

The *Caesalpinia* clade consists of three subclades, but relationships among them are unclear. *Cordeauxia* + *Stuhlmannia* (BS = 87%) is the first subclade. The second is composed of *Caesalpinia*, *Coulteria*, *Haematoxylum*, *Pterolobium*, *Moullava*, and *Mezoneuron*, but internal relationships are not fully resolved, with *Caesalpinia* scattered throughout the clade. In the third clade, *Pomaria* + *Erythrostemon*–*Caesalpinia* is sister to an unresolved group (BS = 74%) that includes *Zuccagnia*, *Stahlia*, *Balsamocarpon*, and *Hoffmannseggia*.

The *Cassia* clade consists of two poorly supported subclades. The first subclade includes *Vouacapoua* + *Andira* (BS = 61%), *Melanoxylon*, and *Senna* + *Cassia*–*Bauhinia*, but relationships among these groups are unclear, and *Cassia* is not monophyletic. Likewise, relationships are uncertain in the second subclade; *Chamaecrista* is not monophyletic, with *Aeschynomene*, *Clitoria*, and *Bauhinia* embedded within it, and the placement of *Batesia* is uncertain. *Pterogyne* is outside of the *Cassia* clade, but without BS > 50%.

Mimosoideae. Mimosoideae are not monophyletic, and several clades within it are also non-monophyletic (Figs. S3, S4; also see LPWG, 2013). Additionally, certain circumscriptions from traditional classifications (e.g., Parkieae, Mimoseae, Acacieae, and Ingeae) are not monophyletic (Figs. S3, S4; also see LPWG, 2013). The monophyly of the *Adenanthera* group is not supported here (see Luckow et al., 2005), but has been split into two clades, *Pseudoprosopis* + (*Xylia* + *Calpocalyx*) (BS = 99%) and *Adenanthera*, *Tetrapleura*, and *Amblygonocarpus* (BS = 98%). *Adenanthera* is not monophyletic, with some species nested within Parkieae. The *Dichrostachys* group is recovered with 96% BS support; however, *Dichrostachys* is non-monophyletic, with *Acacia*, *Neptunia*, and *Alantsilodendron* nested within it. The *Leucaena* group is resolved as (*Desmanthus* + *Kanaloa*) + *Schleinitzia*, but *Leucaena* itself is not actually placed in this clade. *Parkia* is allied with *Pseudopiptadenia* (BS = 59%), but is not monophyletic. *Prosopis*, *Piptadeniopsis*, and *Mimozyganthus* form a clade with weak support (BS = 57%), but *Prosopis* is not monophyletic, with some species of *Prosopis* scattered in different clades. *Adenopodia* and *Piptadenia* are sisters (BS = 54%). *Vachellia*, *Acacia*, *Albizia*, *Senegalia*, *Acaciella*, *Calliandra*, *Zapoteca*, *Mariosousa*, and *Mimosa* are all non-monophyletic, with species frequently nested within other genera.

The *Tachigali* clade is generally resolved as ((*Tachigali* + *Jacqueshuberia*) + *Arapatiella*) + *Campsiandra* (also see Li et al., 2016); however, some species of *Tachigali*, *Jacqueshuberia*, and *Arapatiella* are also interdigitated with *Bussa* and *Peltophorum* in the *Peltophorum* clade. *Schizolobium* is sister to the rest of the *Peltophorum* clade with 70% BS support. *Lemuropisum* and *Colvillea* are nested in *Delonix*, which are collectively sister to *Conzattia* + *Heteroflorum* (BS = 78%) with

moderate support. This clade is then sister to *Parkinsonia–Cercidium* (BS = 86%), in which *Cercidium* is embedded in *Parkinsonia*. The *Dimorphandra* group is not monophyletic (see also Bruneau et al., 2008; and LPWG, 2013) but is divided into *Dimorphandra* group A (BS = 88%; excluding), with a topology of (*Burkea* + (*Stachyothyrsus* + *Mora*)) + *Dimorphandra*, and *Dimorphandra* group B, including *Recordoxylon*, *Moldenhawera*, and *Diptychandra*, but without BS > 50%.

Celastrales–Oxalidales–Malpighiales clade

As in other studies based completely or in large part on plastid genes, the COM clade is resolved here as part of *Fabidae*, sister to the nitrogen-fixing clade (e.g., Wang et al., 2009; Moore et al., 2010, 2011; Soltis et al. 2011; Ruhfel et al., 2014; Fig. 2A). However, the topology of COM + *Malvidae* is recovered in the *mat*R gene tree (Fig. S1), in agreement with the analyses of Sun et al. (2015). Within the COM clade, Oxalidales (BS = 64%) are sister to Malpighiales (BS < 50%) with strong support (BS = 91%), and this clade is in turn sister to Celastrales (BS = 100%) with weak support (BS = 52%). The topology (Fig. 2A) obtained here is consistent with most previous studies (Zhu et al., 2007; APG III, 2009; Wang et al., 2009; Wurdack & Davis, 2009; Qiu et al., 2010; Soltis et al., 2011; Ruhfel et al., 2014 from the first and second codon positions analysis; APG IV, 2016), but alternative topologies also have been inferred (Hilu et al., 2003; Zhang & Simmons, 2006; Soltis et al., 2007; Ruhfel et al., 2014 from all nucleotide positions, amino acid and RY-coded analyses).

Celastrales

Celastrales are monophyletic with strong support (BS = 100%), containing two sister families Lepidobotryaceae (BS = 100%) and Celastraceae (BS = 100%), in agreement with previous studies (Fig. 2A; Zhang & Simmons, 2006; Wang et al., 2009; Soltis et al., 2011).

Celastraceae. Within Celastraceae, *Mortonia* is sister to the rest of Celastraceae with strong support (BS = 100%), and the remaining families are clustered into six major clades (Fig. S3; also see Coughenour et al., 2011; McKenna et al., 2011; Simmons et al., 2012; Simmons & Cappa, 2013).

In Clade I, *Parnassia–Lepuropetalon* is sister to (*Microtropis* + *Quetzalia*) + *Zinowiewia* with 64% BS support. *Parnassia* and *Lepuropetalon* form a strongly supported clade (BS = 100%).

In Clade II, *Brassiantha* and *Hedraianthera* are sisters (BS = 100%), as are *Cassine* and *Denhamia* (BS = 96%), and then these two sister groups together are sister to (*Tripterococcus* + *Stackhousia*) + *Psammomoya* with high support (BS = 89%). This entire clade is then sister to *Crossopetalum* and then to *Siphonondon* with 66% and 98% BS support, respectively.

Within Clade III, only a few clades are recovered. *Mystroxylon*, *Maytenus*, *Robsonodendron*, and *Pseudosalacia* form a strongly supported group (BS = 100%), as do *Celastrus* and *Tripterygium* (BS = 81%), but *Maytenus* and *Tripterygium* are not monophyletic. *Rzedowskia* is sister to *Orthospenia* (BS = 68%). This clade is in turn sister to *Gyminda* (BS = 86%), followed by *Schaefferia* (BS = 88%). *Glyptopetalum* is nested in *Euonymus*, forming a clade with 87% BS support. This clade is sister to *Wimmeria* + (*Acanthothamnus* + *Canotia*-part) with 62% BS support, and this larger clade is collectively sister to *Paxistima* (BS = 52%). *Canotia* and *Euonymus* is non-monophyletic in our analysis.

In Clade IV, *Fraunhofera* and *Plenckia* are sisters (BS = 100%) nested within the *Maytenus* clade (BS = 62%).

In Clade V, *Salaciopsis* is sister to an unresolved clade (BS = 71%) that includes *Gloveria*, *Putterlickia*, *Gymnosporia*, *Lydenburgia*, *Lauridia*, *Robsonodendron*, *Catha*, *Cassine*, *Maurocenia*, and *Allocassine*.

Clade VI is composed of two subclades, with BS support of 73% and 84%, respectively. In the first group (BS = 84%), *Brexia* is sister to *Elaeodendron* + (*Pleurostylia* + *Hartogiopsis*) with 77%, 93%, and 84% BS support, respectively. Within the second group, *Salacia* and *Plagiopteron* are successive sisters to the rest of the clade with 73% and 91% BS support, respectively.

The traditional subfamilial circumscriptions of Celastraceae are not supported. Clades III, IV, V, and VI form a clade, but relationships among them remain to be resolved. The successive sister groups to this large clade are *Canotia* (part; BS = 90%), Clade II (BS = 93%), *Monimopetalum* (BS = 96%), and then Clade I (BS = 57%). *Cassine*, *Canotia*, *Euonymus*, *Maytenus*, *Tripterygium*, *Pristimera*, *Gymnosporia*, and *Elaeodendron* are not monophyletic.

Oxalidales

Oxalidales are weakly supported (BS = 64%) and include Huaceae (BS = 100%), Cunoniaceae (BS = 80%), Cephalotaceae (BS = 100%), Brunelliaceae (BS = 100%), Elaeocarpaceae (BS = 69%), Connaraceae (BS = 100%), and Oxalidaceae (BS = 95%). Huaceae are sister to the rest of Oxalidales with low support (BS = 64%). The remaining families in Oxalidales are further split into two subclades. In the first, Cunoniaceae are sister to an unresolved group of Elaeocarpaceae, Cephalotaceae, and Brunelliaceae, with moderate support (BS = 71%), and the second comprises Connaraceae + Oxalidaceae with 100% BS support (also see Wurdack & Davis, 2009; Soltis et al., 2011).

Huaceae. Within Huaceae, *Afrostyrax* and *Hua* are sisters with 100% BS support.

Cunoniaceae. The relationships among genera in Cunoniaceae are not well resolved.

Elaeocarpaceae. Within Elaeocarpaceae, two clades are recognized. Within the first, *Aceratium* and *Elaeocarpus* are closely related (BS = 63%), but *Elaeocarpus* is not monophyletic in our analyses. This clade is subsequently sister to *Platytheca* (BS = 93%) and then *Crinodendron* (BS = 94%). The second clade is (*Aristotelia* + *Vallea*) + *Sloanea*.

Oxalidaceae. Within Oxalidaceae, *Oxalis* and *Sarcotheca* are successive sisters to *Averrhoa* + *Dapania* (BS = 98%), with 95% and 83% BS support, respectively.

Connaraceae. Four genera in Connaraceae, *Agelaea*, *Connarus*, *Rourea*, and *Byrsocarpus*, are interdigitated, forming a non-monophyletic complex.

Malpighiales

Malphigiales are a large and heterogeneous clade and remain as one of the most poorly resolved orders in *Rosidae* (Wurdack & Davis, 2009). In our study, Malphigiales are well represented, and the monophyly of most families is recovered, including Achariaceae (BS = 81%), Balanopaceae (BS = 100%),

Bonnetiaceae (BS = 100%), Calophyllaceae (BS = 90%), Caryocaraceae (BS = 100%), Centroplacaceae (BS = 99%), Chrysobalanaceae (BS = 99%), Clusiaceae (BS = 100%), Ctenolophonaceae (BS = 100%), Dichapetalaceae (BS = 100%), Elatinaceae (BS = 100%), Erythroxylaceae (BS = 100%), Euphroniaceae (BS = 100%), Goupiaceae (BS = 100%), Hypericaceae (BS = 100%), Humiriaceae (BS = 100%), Irvingiaceae (BS = 100%), Ixonanthaceae (BS = 100%), Lacistemataceae (BS = 100%), Linaceae (BS = 100%), Lophopyxidaceae (BS = 100%), Malpighiaceae (BS = 99%), Ochnaceae (BS = 98%), Pandaceae (BS = 100%), Passifloraceae (BS = 99%), Peraceae (BS = 78%), Phyllanthaceae (BS = 99%), Picrodendraceae (BS = 100%), Podostemaceae (BS = 100%), Putranjivaceae (BS = 100%), Rafflesiaceae (BS = 100%), Rhizophoraceae (BS = 100%), Salicaceae (BS = 94%), Trigoniaceae (BS = 100%), and Violaceae (BS = 100%). However, internal relationships are poorly resolved. Putranjivaceae and Lophopyxidaceae are resolved as sister groups (BS = 100%), as are Phyllanthaceae and Picrodendraceae (BS = 99%). Similarly, Malpighiaceae and Elatinaceae are also sister groups (BS = 98%; also see Wurdack & Davis, 2009; Xi et al., 2012). Several small clades are also resolved. Lacistemataceae are sister to Salicaceae (BS = 94%), with this pair then sister to Passifloraceae (BS = 50%); Goupiaceae are sister to Violaceae (BS = 92%), and this clade is sister to Achariaceae (BS = 100%). These two small clades are sisters with 51% BS support (Fig. 2A; also see Soltis et al., 2011). Bonnetiaceae and Clusiaceae are sisters (BS = 72%) and together are sister to a small clade of Calophyllaceae + (Hypericaceae + Podostemaceae) with >93% BS support at each node. Chrysobalanaceae + Euphroniaceae and Dichapetalaceae + Trigoniaceae form a clade, which is sister to Balanopaceae with ≥99% BS support at each node. Rhizophoraceae and Erythroxylaceae are sisters (BS = 100%) with Centroplacaceae as their sister (BS = 53%). Rafflesiaceae are nested in Euphorbiaceae (BS = 69%), which are sister to Peraceae (sensu APG IV, 2016; BS = 96%); this topology was obtained in our four-gene-combined analysis as well as the *matR* analysis (Fig. S1), but not in our chloroplast analysis (Fig. S2) or in previous studies (e.g, Wurdack & Davis, 2009; Xi et al., 2012), which find Rafflesiaceae as sister to Malpighiaceae, with this pair sister to Peraceae (see Stevens, 2001 onwards). The relationships among the remaining families of Malpighiales are not resolved here, and previous studies of the relationships within Malpighiales largely agree with our current topology.

Irvingiaceae. Within Irvingiaceae, *Klainedoxa* is sister to *Irvingia* and *Desbordesia* with strong support at each node.

Calophyllaceae. Within Calophyllaceae, *Mesua* and *Endodesmia* are sister to the rest of the family (also see Ruhfel et al., 2011), but *Mesua* is non-monophyletic. The remaining genera form two groups (BS = 98%). The first group strictly corresponds to the New World clade (BS = 58%; see Ruhfel et al., 2011), but only the clade of *Kielmeyera* + (*Haploclathra* + *Caraipa*) is recovered with BS support ≥87% at each node. Support for the second group is <50%, including *Calophyllum*, *Mesua*, *Mammea*, *Poeciloneuron*, and *Kayea*.

Hypericaceae. Within Hypericaceae, *Eliea* + *Cratoxylum* (corresponding to Cratoxyleae; BS = 100%) are sister to *Hypericum* + (*Vismia*–*Psorospermum* + *Harungana*–*Vismia*) with BS support ≥95%. *Vismia* is not monophyletic (see Ruhfel et al., 2011): *Harungana* is nested in *Vismia*, and *Vismia guineensis* (L.) Choisy is embedded in *Psorospermum*.

Podostemaceae. Podostemaceae are resolved as Tristichoideae + (Weddellinoideae + Podostemoideae) with strong support (see Koi et al., 2012). There are two groups within Tristichoideae (BS = 100%). Most species of *Dalzellia* are sister to *Indotristicha* (BS = 54%), subsequently followed by *Dalzellia gracilis* C. J. Mathew, Jäger-Zürn & Nileena (BS = 100%) and *Tristicha* (BS = 100%); these results agree with Koi et al. (2009), who treated *Dalzellia gracilis* as representive of a new genus, *Indodalzellia*. The second group is composed of *Malaccotristicha australis* (C. Cusset & G. Cusset) M. Kato, Y. Kita & Koi and *Terniopsis* (BS = 100%); *Malaccotristicha australis* is nested in *Terniopsis* causing *Terniopsis* to be non-monophyletic, likely making the taxonomic change from *T. australis* to *M. australis* (Kato et al., 2003) premature. *Weddellina squamulosa* Tul. is the only member in Weddellinoideae, sister to Podostemoideae with 100% BS support. Within Podostemoideae, *Diamantina* is sister to all other Podostemoideae (BS = 100%; also see Ruhfel et al., 2011; Koi et al., 2012), and the relationships among the remaining genera largely agree with Koi et al. (2012) with a few exceptions: (i) *Rhyncholacis* and *Marathrum* are not monophyletic, nor is *Apinagia*, which is closely related to *Jenmaniella* (BS = 75%); (ii) *Oserya* (*Oserya perpusilla* (Went) P. Royen) is sister to *Castelnavia* (BS = 55%), whereas *Noveloa* (syn. *Oserya* Tul. and Wedd. *pro parte*; Tippery et al., 2011) is close to *Marathrum*–*Vanroyenella* (see Ruhfel et al., 2011); (iii) *Ledermanniella* is polyphyletic, scattered throughout a clade composed of *Inversodicraea*, *Monandriella*, *Saxicolella*, *Dicraeanthus*, *Djinga*, *Stonesia*, *Letestuella*, *Macropodiella*, *Winklerella*, and *Leiothylax*, corroborating the results of Thiv et al. (2009) and Schenk et al. (2015) regarding the treatment of *Ledermanniella*; (iv) the placement of *Podostemum* is not resolved; (v) *Torrenticola queenslandica* (Domin) (Domin) is nested in *Cladopus*, in agreement with Cook & Rutishauser (2001) regarding the treatment of *Torrenticola queenslandica* by transferring it to *Cladopus* as *C. queenslandicus*; (vi) Koi & Kato (2010) transferred *Diplobryum koyamae* M. Kato & Fukuoka to *Hydrobryum* based on molecular evidence; our results suggest that *Diplobryum vientianense* M. Kato & Fukuoka and *Diplobryum ramosum* C. Cusset also need to be transferred to *Hydrobryum* to maintain the monophyly of *Hydrobryum* and that *Diplobryum* may need further study; and (vii) our results also support the treatment of Cook & Rutishauser (2001), who transferred *Hydrobryopsis sessilis* (Willis) Engl. to *Zeylanidium* (as *Z. sessile* (Willis) C. D. K. Cook & Rutish), but *Zeylanidium* is still not monophyletic here (see Ruhfel et al., 2011; Koi et al., 2012). *Zeylanidium subulatum* (Gardner) C. Cusset is sister to (*Farmeria* + *Griffithella*) + *Polypleurum* (BS = 97%), whereas the rest of the genus is sister to *Willisia* (BS = 76%).

Bonnetiaceae. Within Bonnetiaceae, *Bonnetia* is sister to *Ploiarium* + *Archytaea*, with 100% BS support at all nodes (see Ruhfel et al., 2011).

Clusiaceae. Within Clusiaceae, Clusieae are a well-supported clade sister to the rest of the family (BS = 100%), including *Clusia*, *Chrysochlamys*, *Tovomita*, *Tovomitopsis*, and then *Dystovomita*, as subsequent sister to the rest of Clusieae

(BS = 99%). The remaining Clusiaceae form a clade with 100% BS support, but do not correspond to Symphonieae and Garcinieae (see Ruhfel et al., 2011). *Symphonia* and *Garcinia* are both non-monophyletic, and they may be closely related to *Montrouziera* (BS = 100%) and *Tripetalum* (BS = 94%), respectively.

Malpighiaceae. The phylogeny of Malpighiaceae obtained here roughly agrees with Davis & Anderson (2010), and most clades recovered here are the same as in Anderson et al. (2006). The *Galpimia* clade (BS = 100%), *Acmanthera* clade (BS = 100%), and *Byrsonima* clade (BS = 95%) together form a large clade (BS = 96%) that is sister to the rest of Malpighiaceae with 99% BS support. The *Galpimia* clade (BS = 100%) is sister to the *Acmanthera* clade (BS = 100%; *Coleostachys* excluded) plus the *Byrsonima* clade, with strong support. However, here *Coleostachys* is not in the *Acmanthera* clade, but sister to the *Byrsonima* clade, also with strong support. The *Acridocarpus* clade (BS = 95%), *Mcvaughia* clade (BS = 100%), *Barnebya* clade (including only *Barnebya*), *Ptilochaeta* clade (BS = 100%), *Bunchosia* clade (BS = 94%), and the *Ectopopterys* clade (only including *Ectopopterys*) are all recovered, with strong suppor. The *Acridocarpus* clade + *Mcvaughia* clade (BS = 83%), *Barnebya* clade, *Ptilochaeta* clade, *Bunchosia* clade, and *Ectopopterys* clade are successive sisters to the remaining members of the family with 99%, 68%, 98%, 97%, and 94% BS support, respectively. Within the *Acridocarpus* clade, *Acridocarpus* and *Brachylophon* are sisters, and the topologies resolved within the *Mcvaughia* clade and the *Ptilochaeta* clade are identical to those of Davis & Anderson (2010). Within the *Bunchosia* clade, relationships also match Davis & Anderson (2010) except that *Thryallis* is non-monophyletic here. Finally, the generic relationships among the remaining Malpighiaceae are not well resolved except for the presence of the *Stigmaphyllon* clade (BS = 96%), *Malpighia* clade (BS = 85%), and a sister grouping of *Tetrapterys* + (*Flabellariopsis* + *Hiptage*). Within the *Stigmaphyllon* clade, *Diplopterys* and then *Bronwenia* are successive sisters to the rest of the *Stigmaphyllon* clade, whose internal relationships are not well resolved. *Mionandra* and *Gallardoa* are sisters (BS = 62%), followed by *Cordobia* (BS = 100%). *Aspicarpa*, *Gaudichaudia*, *Camarea*, and *Janusia* form a clade with 100% BS support, but neither *Aspicarpa*, *Gaudichaudia*, nor *Janusia* is monophyletic (also see Davis & Anderson, 2010). *Ryssopterys* is nested within *Stigmaphyllon*, causing *Stigmaphyllon* to be non-monophyletic (see Davis & Anderson, 2010). To maintain the monophyly of *Stigmaphyllon*, Anderson (2011) transferred *Ryssopterys* as a subgenus under *Stigmaphyllon*, which is also supported by our phylogeny. *Sphedamnocarpus* is not monophyletic, with some species placed as sister to *Philgamia* with strong support. Within the *Malpighia* clade, *Calcicola* is sister to the rest of the *Malpighia* clade (BS = 85%). *Madagasikaria* + *Microsteira* form a clade (BS = 64%). *Mascagnia* is not monophyletic, with species found scattered among the *Hiraea* clade, the *Malpighia* clade, and *Carolus*.

Linaceae. Linaceae comprise two major clades, corresponding to subfamilies Linoideae (BS = 100%) and Hugonioideae (BS = 92%; Dressler et al., 2014), except *Hebepetalum* and *Roucheria* are resolved as sisters to Linoideae, but BS support is <50%. The rest of the Hugonioideae form a clade with 92% BS support. *Indorouchera* and *Philbornea* are sisters (BS = 98%), and *Hugonia* is non-monophyletic. Within Linoideae, *Anisadenia*, *Reinwardtia*, and *Tirpitzia* form a clade (BS = 67%), sister to the remaining genera, including *Linum*, *Cliococca*, *Sclerolinon*, *Hesperolinon*, *Millegrana*, *Radiola*, and *Adenolinum*. *Millegrana* and *Radiola* are sisters (BS = 100%), and *Linum* is paraphyletic (see McDill & Simpson, 2011) relative to *Adenolinum* and *Cliococca* (Fig. S3).

Ixonanthaceae. Within Ixonanthaceae, *Ixonanthes* is sister to *Ochthocosmus* + *Cyrillopsis*, all with 100% BS support, concordant with Xi et al. (2012).

Picrodendraceae. Within Picrodendraceae, *Podocalyx* is sister to the rest of the family (BS = 100%), which is composed of two well-resolved sister clades. The first is *Petalostigma* + (*Scagea* + (*Stachystemon* + *Micrantheum*) + (*Dissiliaria* + *Austrobuxus*)), and the other is *Tetracoccus* + (*Picrodendron* + (*Hyaenanche* + (*Aristogeitonia* + (*Oldfieldia* + *Androstachys*)))). This topology generally agrees with Wurdack & Davis (2009).

Phyllanthaceae. Two well-supported sister clades are identified in Phyllanthaceae, corresponding to subfamilies Phyllanthoideae (BS = 100%) and Antidesmatoideae (BS = 100%), in general agreement with previous studies (Wurdack et al., 2004; Kathriarachchi et al., 2005; Samuel et al., 2005; Hoffmann et al., 2006).

Within Antidesmatoideae, *Bischofia* is sister to the rest (BS = 100%), which is composed of two major clades (Hoffmann et al., 2006). Within the first clade, *Celianella* + *Jablonskia* (representing Jablonskieae, BS = 100%) are sister to Antidesmateae (BS = 90%), but with only 50% BS support; *Uapaca* is sister to Jablonskieae (BS < 50%). Within Antidesmateae, only three sister-groups are recovered, *Antidesma* + *Thecacoris* (BS = 85%), *Apodiscus* + *Martretia* (BS = 100%), and *Didymocistus* + *Hymenocardia* (BS = 100%). Within the second clade, *Spondianthus* is sister to Scepeae (BS = 100%) with 76% BS support. Within Scepeae, *Protomegabaria* is strongly supported as sister to an unresolved clade of *Aporosa*, *Richeria*, *Maesobotrya*, and *Baccaurea*.

Within subfamily Phyllanthoideae, four major clades are identified, representing tribes Phyllantheae (BS = 99%), Bridelieae (BS = 100%), Wielandieae (BS = 73%), and Poranthereae (BS = 100%) (See Hoffmann et al., 2006). Phyllantheae and Bridelieae are successively sisters to Wielandieae and Poranthereae, with 100% and 59% BS support, respectively. Within Phyllantheae, (*Margaritaria* + *Phyllanthus diandrus* Pax)–*Lingelsheimia* and *Savia* (*S. bahamensis* Britton) + (*Flueggea* + *Richeriella*) are sequential sisters to an unresolved clade (BS = 94%) composed of *Reverchonia*, *Glochidion*, *Sauropus*, *Breynia*, and *Phyllanthus* (see Samuel et al., 2005; Kathriarachchi et al, 2006). *Phyllanthus* is not monophyletic, with *Reverchonia*, *Glochidion*, *Sauropus*, and *Breynia* nested within it; hence, Hoffmann et al. (2006) enlarged *Phyllanthus* by including all of these genera. Within Bridelieae, subtribe Keayodendrinae (*Keayodendron*) and subtribe Amanoinae (*Amanoa*) are sisters (BS = 85%) that are followed as sequential sisters by subtribe Pseudolachnostylidinae (BS = 100%), Saviinae (BS = 100%), and Securineginae (BS = 98%), with 52%, 100%, and 100% BS support, respectively. Within Pseudolachnostylidinae, *Cleistanthus*, *Pseudolachnostylis*, and *Pentabrachion* are successive sisters to *Bridelia* + *Cleistanthus* with strong support, but *Cleistanthus* appears

non-monophyletic (Li et al., 2009). Within Saviinae, *Gonatogyne* + *Savia* are strongly supported as sister to *Discocarpus* + *Tacarcuna–Croizatia*, with *Tacarcuna* nested in *Croizatia*. In our study, both *Savia* and *Croizatia* are non-monophyletic. Within Securineginae, *Lachnostylis* and *Securinega* are sisters (BS = 98%). Wielandieae are composed of two sister subtribes, Astrocasiinae (BS = 98%) and Wielandiinae (BS = 99%), with 73% BS support. Within Astrocasiinae, *Heywoodia* is the strongly supported sister to *Chascotheca* + *Astrocasia* (BS = 100%). Within Wielandiinae, *Dicoelia* + *Chorisandrachne* are strongly supported as sister to a well-supported but internally unresolved clade of *Petalodiscus*, *Blotia*, and *Wielandia*. Hoffmann et al. (2006) transferred *Petalodiscus* and *Blotia* to *Wielandia* for nomenclatural reasons. Within tribe Poranthereae, internal relationships remain to be addressed further (see Vorontsova & Hoffmann, 2008). ((*Zimmermannia* + *Zimmermanniopsis*) + *Meineckia*) + *Andrachne* is sister to the rest; *Poranthera* and *Actephila* are sisters (BS = 100%). *Flueggea suffruticosa* (Pall.) Baill. is embedded in *Leptopus* (BS = 97%); *Andrachne*, *Poranthera*, and *Zimmermannia* are non-monophyletic in our study.

Ochnaceae. Three main clades are recognized in Ochnaceae, corresponding to Ochnoideae (BS = 95%), Medusagynoideae, and Quiinoideae (BS = 98%), based on Stevens (2001 onwards) and Schneider et al. (2014). Ochnoideae are strongly supported as sister to Medusagynoideae + Quiinoideae (BS = 66%; see Xi et al., 2012; Schneider et al., 2014). Within Quiinoideae, *Froesia* and *Quiina* are successively sisters to *Touroulia* + *Lacunaria*, with ≥94% BS support at each node (see Schneider et al., 2014). Within Ochnoideae, *Luxemburgia* (*Testulea* not sampled) is strongly supported as sister to tribe Sauvagesieae (including *Sauvagesia* and *Cespedesia*; BS = 99%) and tribe Ochneae (BS = 94%). Within Ochneae, *Lophira* and *Elvasia* are successive sisters to subtribe Ochninae including *Brackenridgea*, *Gomphia*, *Diporidium*, *Ochna*, and *Ouratea*. *Ouratea* is non-monophyletic.

Violaceae. The phylogeny of Violaceae resolved here differs from the current infrafamilial classification (e.g., Ballard et al., 2013). *Fusispermum* is resolved as sister to the rest of Violaceae (BS = 100%). *Rinorea* (which appears as two distinct clades here and is non-monophyletic; see also Wahlert et al., 2014), *Decorsella*, *Paypayrola*, and *Rinoreocarpus* are successive sisters to the remainder of the family, in agreement with Wahlert et al. (2014), but *Decorsella* and *Paypayrola* are sisters in their study. Among the remaining genera, four clades are recognized. Clade I (as Clade 3 in Wahlert et al., 2014) with 54% BS support includes *Leonia*, *Gloeospermum*, *Amphirrhox*, *Orthion*, and *Mayanaea*; *Orthion* and *Mayanaea* are sisters (BS = 100%), and *Leonia* is nested in *Gloeospermum*. Clade II (as Clade 1 in Wahlert et al., 2014) is well supported with *Allexis* + (*Noisettia* + *Viola*). Clade III, ((*Agatea* + *Agation*) + *Corynostylis*) + *Anchietea*, is also strongly supported. Within Clade IV, *Hybanthus* and *Isodendrion* are sisters (BS = 100%), and this clade is then sister to *Melicytus* and *Hymenanthera* (BS = 63%), with *Hymenanthera* nested in *Melicytus*.

Achariaceae. Achariaceae are resolved as three strongly supported clades, Erythrospermeae + Lindackerieae (BS = 99%), *Hydnocarpus*, and the rest (including Pangieae, Achardieae, *Ryparosa*, etc.; BS = 50%), in agreement with Groppo et al. (2010). The monophyly of Lindackerieae has 100% BS support, but within Lindackerieae, only *Kuhlmanniodendron* and *Caloncoba* are resolved as sister groups (BS = 59%). Erythrospermeae include *Dasylepis* and *Erythrospermum*, but the relationship between them is not resolved. *Hydnocarpus*, *Pangium*, *Ryparosa*, *Trichadenia*, *Gynocardia*, *Chiangiodendron*, and *Kiggelaria* are successive sisters to Achardieae. Achardieae are composed of *Ceratiosicyos*, *Guthriea*, and *Acharia*, which are weakly supported as a clade.

Salicaceae. Salicaceae are resolved as three clades corresponding to Samydoideae (BS = 94%), Scyphostegioideae (only *Scyphostegia*), and Salicoideae (BS = 100%; see Stevens, 2001 onwards; Chase et al., 2002). Samydoideae are sister to Scyphostegioideae + Salicoideae with 94% BS support. Within Samydoideae, *Tetrathylacium* and *Lunania* are successive sisters to an unresolved clade composed of *Samyda*, *Casearia*, *Zuelania*, and *Laetia*. *Casearia* is non-monophyletic with some species sister to *Samyda* (Samarakoon, 2015). Within Salicoideae, two well-resolved subclades, *Poliothyrsis* + (*Idesia* + (*Salix* + *Populus*)) and *Hasseltia* + (*Abatia* + (*Banara* + *Prockia*)), are successive sisters to the rest of Salicoideae with strong support. Considering the rest of this clade, *Homalium* is not monophyletic and has *Bembicia* embedded in it. *Dovyalis* and *Trimeria* are successive sisters to an unresolved clade of *Flacourtia*, *Oncoba*, *Scolopia*, and *Xylosma*.

Passifloraceae. Passifloraceae are well resolved into three clades representing three subfamilies: Malesherbioideae (including only *Malesherbia*), Turneroideae (BS = 99%), and Passifloroideae (BS = 99%) (see Tokuoka, 2012). Malesherbioideae are sister to Turneroideae + Passifloroideae with 99% BS support at each node. Within Turneroideae, *Turnera* + *Piriqueta* (BS = 100%), *Erblichia*, *Stapfiella*, *Mathurina*, and then *Tricliceras* are successive sisters to *Streptopetalum* and *Loewia* with ≥60% BS support at each node (see Thulin et al., 2012; Tokuoka, 2012). However, *Streptopetalum* is non-monophyletic, with *Loewia* embedded in it. Within Passifloroideae, both tribe Paropsieae and tribe Passifloreae are monophyletic, with 100% and 99% BS support, respectively (see Tokuoka, 2012). Within tribe Paropsieae, *Paropsia* + *Viridivia* are sister to the rest of the clade (BS = 100%), with *Androsiphonia* sister to an interdigitated group of *Barteria* and *Smeathmannia* (BS = 76%). *Adenia* is sister to the rest of the tribe (BS = 99%), and a clade of *Basananthe* + ((*Schlechterina* + *Crossostemma*) + (*Efulensia* + *Deidamia*)) is sister to *Ancistrothyrsus* + (*Mitostemma* + *Dilkea*) and an unresolved group of *Passiflora–Tetrapathea–Tetrastylis*, with strong support. *Mitostemma* is non-monophyletic.

Peraceae. Within Peraceae, *Pogonophora* is sister to the rest of the family (BS = 78%). *Clutia* and *Pera* are successive sisters to *Trigonopleura* + *Chaetocarpus*, all with strong support.

Euphorbiaceae. Euphorbiaceae are not monophyletic in our four-gene-combined tree, with *Neoscortechinia* + *Cheilosa* (BS = 97%) sister to Rafflesiaceae (BS < 50%), rather than sister to the rest of Euphorbiaceae (also see Tokuoka, 2007), and *Neoscortechinia* is non-monophyletic. The rest of Euphorbiaceae form a clade with 79% BS support (Figs. 2A, S5), and two of the three currently recognized subfamilies are not monophyletic. Only Euphorbioideae are recovered as a clade (BS = 95%). Euphorbiaceae (exluding *Neoscortechinia* and

Cheilosa) are composed of seven major clades, including Acalyphoideae s.s. (BS = 94%), Erismantheae (BS = 96%), Euphorbioideae (BS = 95%), the articulated and inaperturate crotonoids clade (BS = 100%, 97%), Gelonieae (BS = 73%), and Adenoclineae s.l. (BS = 61%). The relationships among these clades are unclear, consistent with Wurdack et al. (2005).

Within Acalyphoideae s.s. (Fig. S5), the basic phylogenetic structure is congruent with Wurdack et al. (2005). Within this clade, a strongly supported clade of alchorneoids is sister to the remaining Acalyphoideae (BS = 94%), which are composed of eight subclades, A1–A8, as designated in Wurdack et al. (2005) for convenience of discussion. The relationships among some of these eight subclades are poorly supported. Subclade A1 is sister to the rest with 93% BS support, A2 (BS = 91%) and A3 (BS = 62%) are sisters (BS = 58%), followed by A4 (BS = 68%) with weak support; A7 (BS = 98%) and A8 (BS = 88%) are sisters (BS = 92%). Within the alchorneoids clade, a well-resolved clade (*Pseudagrostistachys* + *Necepsia*) + *Paranecepsia* is strongly supported as sister to the rest of the clade (BS = 97%), then (*Amyrea* + *Discoglypremna*) + *Cyttaranthus*, *Mareyopsis* + *Alchorneopsis*, and *Aubletiana* subsequently follow the sister group of (*Conceveiba–Gavarretia* + *Aparisthmium*) + *Alchornea–Bocquillonia* with ≥73% BS support. Within A1, *Cleidion* is sister to the rest of the genera of this clade but BS support is <50%; *Agrostistachys* and *Blumeodendron* are successive sisters to a strongly supported group composed of *Macaranga* and two paraphyletic genera, *Mallotus* and *Cordemoya*. Within A2, *Mercurialis*, *Discoclaoxylon*, and *Claoxylon* are successive sisters to a weakly supported clade (including *Lobanilia*, *Erythrococca*, and *Micrococca*) whose internal relationships receive BS support <50%. Within A3, *Afrotrewia* is sister to the rest (BS = 62%). *Pycnocoma*, *Argomuellera*, and *Droceloncia* form a highly supported clade, although intergeneric relationships are unresolved because *Pycnocoma* is paraphyletic. This clade is sister to *Wetria–Chondrostylis*, although *Wetria* is not monophyletic. The remaining members are two sister groups, (*Acalypha* + *Mareya*) + *Crotonogynopsis* (BS ≥ 72%) and *Spathiostemon* + *Homonoia* (BS = 100%). Within A4, (*Monotaxis* + *Amperea*) + *Adriana* is sister to the rest of the clade (BS = 68%). Relationships among the rest of the genera are generally not full resolved, except (*Discocleidion* + *Ricinus*) + *Speranskia*, (*Thyrsanthera* + *Melanolepis*) + *Sumbaviopsis*, and *Cleidiocarpon* + ((*Cephalocroton* + *Adenochlaena*) + (*Cephalomappa* + *Koilodepas*)). Within A5, *Leidesia* and *Seidelia* are sisters (BS = 100%). Within A6, *Enriquebeltrania* is sister to the rest of the clade, but BS support is <50%, and the rest of this clade is split into two groups. In one group, *Argythamnia* is nested in *Ditaxis*, which is sister to *Chiropetalum* (BS = 61%), with this clade sister to *Caperonia* + *Dysopsis* with BS support <50%. In the second group, *Philyra* is sister to a clade of *Adelia*, *Lasiocroton*, and *Leucocroton*, in which both *Adelia* and *Lasiocroton* are non-monophyletic. Within A7, *Adenophaedra* and *Bernardia* are sisters (BS = 99%), followed by *Caryodendron* (BS = 98%). Within A8, *Romanoa* + *Plukenetia* (BS = 87%), *Astrococcus* + *Dalechampia* (BS = 66%), and then *Cnesmone* + *Tragiella* (BS = 99%) are successive sisters to (*Tragia* + *Gitara*) + *Acidoton* with 88%, 64%, and 82% BS support, respectively.

Within Erismantheae, *Syndyophyllum* and *Moultonianthus* are sisters (BS = 96%). Within Euphorbioideae, *Pimelodendron*, *Plagiostyles*, and *Nealchornea* form a clade with 98% BS support, strongly sister to the rest of Euphorbioideae (also see Wurdack et al., 2005), which is composed of three subclades. (*Dichostemma* + *Anthostema*) + ((*Neoguillauminia* + *Calycopeplus*) + *Euphorbia–Pedilanthus*) is the first clade with strong support. *Euphorbia* is not monophyletic with *Pedilanthus* nested within it. This clade is sister to the other two subclades with 98% BS support. *Hura* is sister to the remainder of the second clade with 91% BS support, and the rest is divided into three groups, but support for the relationships among these groups is low: ((*Excoecaria* + *Spirostachys*) + *Sebastiania*)–(*Pachystroma* + (*Ophthalmoblapton* + *Tetraplandra*)), *Hippomane* + (*Bonania* + (*Grimmeodendron* + *Sebastiania*)), and *Colliguaja* + ((*Stillingia* + *Sapium*) + ((*Stillingia* + *Spegazziniophytum*) + *Adenopeltis*)). *Sapium*, *Sebastiania*, *Excoecaria*, and *Stillingia* are non-monophyletic. Within the third subclade, *Actinostemon–Maprounea* (BS = 99%) and *Senefelderopsis* are successive sisters to the rest of the clade with 74% and 94% BS support, respectively. Among the remaining genera, *Sapium* and *Neoshirakia* are sisters, and *Sclerocroton* is strongly supported as sister to *Ditrysinia* + *Microstachys* and *Mabea* + (*Sebastiania* + *Gymnanthes*). *Actinostemon* is not monophyletic, with some species placed within *Maprounea* and others in *Pseudosenefeldera*.

Within the articulated crotonoids, *Elateriospermum* and *Manihot* + *Cnidoscolus* (BS = 100%) are successive sisters to the rest (including *Hevea*, *Glycydendron*, *Micrandropsis*, and *Micrandra*) with strong BS support. Within the inaperturate crotonoids, we recovered two subclades. In the first clade, *Jatropha* is sister to the rest (BS = 97%), and (*Sandwithia* + *Sagotia*) + (*Astraea* + (*Acidocroton* + *Ophellantha*)) and *Brasiliocroton* are successive sisters to an unresolved *Croton–Colobocarpos–Moacroton* group (BS = 99%). *Croton* is non-monophyletic. Within the second clade (BS = 96%), only a few groups are well resolved: *Grossera* + *Cavacoa*, *Schinziophyton* + (*Ricinodendron* + *Givotia*), (*Tannodia* + *Domohinea*) + (*Neoholstia* + *Neoboutonia*), (*Blachia* + *Strophioblachia*) + *Codiaeum*, and *Aleurites* + ((((*Cocconerion* + *Baloghia*) + *Ricinocarpos*) + *Fontainea*) + (*Beyeria* + *Bertya*)). *Aleurites* and *Vernicia* are closely related but non-monophyletic.

Within Gelonieae, *Cladogelonium* and *Suregada* are sisters (BS = 73%). The relationships within Adenoclineae resolved in our study are identical to those found by Wurdack et al. (2005), corroborating that *Omphalea* belongs to this tribe, and further confirming that *Tetrorchidium* is non-monophyletic, with *Ditta* embedded within it.

Rafflesiaceae. Within Rafflesiaceae, *Sapria* is sister to *Rhizanthes* + *Rafflesia* with strong support (see Davis et al., 2007).

Pandaceae. Within Pandaceae, *Microdesmis* is sister to *Galearia* + *Panda*, with strong support at each node (see Xi et al., 2012).

Dichapetalaceae. *Dichapetalum* is not monophyletic, with *Tapura* nested within it.

Chrysobalanaceae. The phylogeny of Chrysobalanaceae is unresolved; *Licania*, *Couepia*, and *Hirtella* are intertwined within each other, and none of them is monophyletic (also see Sothers & Prance, 2014).

Rhizophoraceae. Four strongly supported clades are recovered within Rhizophoraceae (Schwarzbach & Ricklefs, 2000):

Macarisieae (BS = 100%), *Paradrypetes*, Gynotrocheae (BS = 100%), and Rhizophoreae (BS = 100%). *Paradrypetes* is sister to Macarisieae (BS = 99%), and these together are sister to Gynotrocheae + Rhizophoreae (BS = 100%). Within Macarisieae, *Sterigmapetalum* and *Cassipourea* are sisters (BS = 72%), but the relationships among *Macarisia*, *Dactylopetalum*, and *Blepharistemma* are unclear. Within Gynotrocheae, *Carallia* and *Crossostylis* are successive sisters to *Pellacalyx* + *Gynotroches* with strong support. Within Rhizophoreae, *Ceriops* + *Kandelia* (BS = 90%) is sister to *Rhizophora* + *Bruguiera* (BS = 53%), with strong support.

Erythroxylaceae. *Aneulophus* is sister to *Nectaropetalum* + *Erythroxylum* (BS ≥ 95%).

Humiriaceae. *Schistostemon* and *Sacoglottis* form a clade with strong support, which is then sister to an unresolved clade comprising *Vantanea* and *Humiria. Sacoglottis* is not monophyletic, with *Schistostemon* embedded in it.

Putranjivaceae. Within Putranjivaceae, the relationships among *Putranjiva*, *Drypetes*, and *Sibangea* are not resolved, and *Drypetes* is non-monophyletic.

Zygophyllales

Zygophyllales are sister to the rest of *Fabidae* and include two morphologically similar (Sheahan & Chase, 2000; Judd & Olmstead, 2004) sister families (BS = 98%), Krameriaceae (BS = 100%) and Zygophyllaceae (BS = 89%).

Zygophyllaceae. Zygophyllales are still highly enigmatic and share few non-DNA traits with any other rosid lineage (Soltis et al., 2005). We recovered the monophyly of five subclades within Zygophyllaceae, supporting the division of Zygophyllaceae into five subfamilies (see Sheahan & Chase, 2000), and we also confirmed the polyphyly of *Zygophyllum* with *Augea*, *Fagonia*, and *Tetraena* embedded within it, in agreement with the treatment of Beier et al. (2003). However, our *matR* gene analysis indicates that Zygophyllaceae are closely related to Crossosomatales, as do other analyses of mitochondrial data (Fig. S1; Zhu et al., 2007; Qiu et al., 2010; Sun et al., 2015). Unfortunately, this order is either unsampled or its relationships are not resolved in nuclear gene studies (Soltis et al., 1997; Zhang et al., 2012; Xi et al., 2014).

Relationships within *Malvidae*

Although *Malvidae* (eurosids II) is consistently recovered with strong support, the relationships among some of the members of this clade are still uncertain (Jansen et al., 2007; Zhu et al., 2007; Wang et al., 2009; Qiu et al., 2010; Soltis et al., 2011; Zhang et al., 2012; Ruhfel et al., 2014; Wang et al., 2014; Xi et al., 2014; Zeng et al., 2014). Within *Malvidae*, we found strong support for the six orders, except Myrtales and Geraniales form a clade that is sister to *Fabidae* and the remaining *Mavidae*. Brassicales (BS = 97%) and Malvales (BS = 85%) are sister groups (BS = 85%), which are successively followed by Huerteales (BS = 98%), Sapindales (BS = 75%), Picramniales (BS = 100%), and Crossosomatales (BS = 92%) (Fig. 2B). Bootstrap support for the relationships among Picramniales, Crossosomatales, and the rest of *Malvidae* are <50% BS. As previously noted, the *matR* analysis supports placement of the COM clade with *Malvidae* (BS = 55%; Fig. S1).

Brassicales

Brassicales are monophyletic with strong support (BS = 97%; for families see Table 1) and are resolved into three well-supported clades. The first clade is Tropaeolaceae + Akaniaceae (BS = 100%), the second is Moringaceae + Caricaceae (BS = 100%), and these two clades are successive sisters to the third clade, which is composed of 15 families (see Hall et al., 2004; Su et al., 2012; Edger et al., 2015). Within the third clade, Limnanthaceae, Setchellanthaceae, and Koeberliniaceae + Bataceae–Salvadoraceae are successive sisters to the remainder. The monogeneric Bataceae are nested within Salvadoraceae, causing Salvadoraceae to be non-monophyletic. We find a strongly supported clade of Gyrostemonaceae + (Borthwickiaceae + (Resedaceae + Stixaceae)) sister toEmblingiaceae–Tovariaceae + (Pentadiplandraceae + (Capparaceae + (Cleomaceae + Brassicaceae))) (see Hall et al., 2004; Qiu et al., 2010; Soltis et al., 2011; Su et al., 2012). Note that an expanded Resedaceae are proposed by including Borthwickiaceae and Stixaceae (APG IV, 2016; see Fig. 2B). The support for relationship between Emblingiaceae and Tovariaceae is <50%, but BS support for relationships among the rest are all >61%. In our *matR* analysis, Bataceae are closer to Capparaceae and Brassicaceae than to Resedaceae.

Tropaeolaceae. *Tropaeolum* is paraphyletic, with *Magallana* and *Trophaeastrum* embedded within it, and we therefore concur with Andersson & Andersson (2000) in extending *Tropaeolum* by including both of these additional genera.

Caricaceae. Within Caricaceae, we recovered *Cylicomorpha* + ((*Vasconcellea* + *Jacaratia*) + *Carica*–(*Horovitzia* + *Jarilla*)) (see Carvalho, 2013 onwards).

Gyrostemonaceae. *Gyrostemon* is not monophyletic, with *Tersonia* nested within it (BS = 75%).

Resedaceae (Borthwickiaceae and Stixaceae). *Borthwickia* is sister to the rest of this family (BS = 95%).

Caylusea + *Sesamoides*–(*Reseda*–*Ochradenus*) forms a clade with 100% BS support (see Su et al., 2012; Zhao et al., 2015), but the phylogenetic position of *Sesamoides* is unclear. *Ochradenus* is nested in *Reseda*, causing *Reseda* to be paraphyletic.

Stixis and *Tirania* are sisters (BS = 98%), and together they are sister to *Forchhammeria* with 73% BS support.

Capparaceae. Within Capparaceae, *Crateva* is strongly supported as sister to the rest of the family (BS = 95%); however, it is non-monophyletic with *Euadenia* nested in it. The relationships among the remaining genera of the family are not well resolved, with *Capparis* species widely scattered across this clade (Fig. S3; see Hall, 2008).

Cleomaceae. Cleomaceae are resolved as two sister groups here (BS = 97%), but still require further attention, as neither *Cleomella* nor *Cleome* is monophyletic (see Hall, 2008; Feodorova et al., 2010; Patchell et al., 2014). The first group is resolved as ((*Oxystylis* + *Wislizenia*) + *Cleomella*) + *Cleome* (Fig. S3; BS = 82%), corresponding to the North American cleomoids (Hall, 2008; Feodorova et al., 2010; Riser et al., 2013; Patchell et al., 2014), which agrees with the taxonomic revision of Roalson et al. (2015) that reduced the whole clade to the single large genus, *Cleomella*. The second group (BS = 52%) is composed of the majority of *Cleome* with *Polanisia*, *Dactylaena*, *Podandrogyne*, and *Dipterygium*. The

relationships among them agree with Patchell et al. (2014).

Brassicaceae. The phylogeny of Brassicaceae is poorly resolved, although the monophyly of the family is strongly recovered (BS = 98%). *Aethionema* is sister to the rest of the family (Fig. S3; also see Zunk et al., 1999; Koch et al., 2001; Beilstein et al., 2006; Huang et al., 2015). The remaining genera form a clade with 98% BS support, but relationships among them are still unresolved in this study, except for two main clades.

Within the first clade, *Physaria* is sister to *Synthlipsis* + *Dimorphocarpa* with BS support ≥60% at each node. *Phoenicaulis* + *Sandbergia* (BS = 92%) and *Cusickiella* + *Polyctenium* (BS = 90%) are sisters (BS = 81%), and they are collectively sister to some species of *Halimolobos* (BS = 54%). *Alyssum* is nested in *Lepidium* (BS = 98%), which makes *Lepidium* non-monophyletic; however, *Alyssum* is also non-monophyletic, with some species of *Alyssum* sister to *Berteroa* in the second clade with strong support. *Smelowskia* is also non-monophyletic with *Hedinia* nested within it (BS = 86%). *Cardamine* is not monophyletic, with species closely related to *Nasturtium* and *Nasturtiopsis* (BS = 85%). *Halimolobos* is not monophyletic, with some species of this genus sister to *Pennellia* (BS = 79%). *Pachycladon* and *Rorippa* also are not monophyletic.

Within the second clade, *Lunaria* and *Biscutella* are sisters (BS = 97%). *Arabis* is not monophyletic, appearing as two separated clades. *Baimashania* and these two *Arabis* clades are successive sisters to *Erophila* + *Draba* (BS = 97%) with strong BS support. *Isatis* is not monophyletic either, forming an *Isatis*–(*Pachypterygium* + *Conringia*–*Isatis*) clade with 98% BS support. *Schouwia* and *Zilla* are sisters (BS = 99%), as are *Carrichtera* and *Vella* (BS = 96%). *Orychophragmus*, *Sisymbrium* + (*Streptanthus* + (*Streptanthus* + *Stanleya*)), *Schouwia* + *Zilla*, *Psychine*, and *Vella* + *Carrichtera* are the successive sisters to the remainder of the second clade, but neither *Streptanthus* nor *Orychophragmus* is monophyletic. Among the remaining genera of this clade, *Raffenaldia* and *Hemicrambe* are successive sisters to a complex of *Erucastrum*–*Hirschfeldia*, with 70% and 100% BS support, respectively. *Sinapis* is non-monophyletic, and *Kremeriella* is sister to part of *Sinapis* (BS = 84%; *Sinapidendron* is sister to *Erucastrum* + *Diplotaxis* with 99% BS support, and *Coincya* is sister to *Muricaria* + *Crambe* (BS = 58%). *Erucaria*, *Crambella*, *Cakile*, and *Didesmus* form a clade with 95% BS support, but relationships among these genera are unclear because species of these genera are nested within each other. *Moricandia* is not monophyletic; some of *Moricandia* is sister to *Rytidocarpus* (BS = 73%), while some are nested between *Raphanus* and *Rapistrum*. *Rapistrum* appears non-monophyletic as well; some species of *Rapistrum* are sister to *Raphanus*, and others are sister to *Enarthrocarpus* (BS = 94%). *Alliaria* and *Thlaspi* are sisters (BS = 50%), *Dontostemon* and *Clausia* are sisters (BS = 70%), and *Diptychocarpus* and *Parrya* are also sisters (BS = 100%). *Anchonium* is sister to *Tauscheria* + *Sterigmostemum* (BS = 54%). *Christolea* and *Phaeonychium* are sisters (BS = 89%). *Arabidopsis* is sister to *Neotoruria* (BS = 61%). *Heliophila* is sister to *Iberis* + *Noccaea* (BS = 63%). *Lobularia* + *Notoceras* (BS = 73%) and *Anastatica* + *Diceratella* (BS = 95%) are collectivel sister to *Maresia* + (*Eremobium* + *Ricotia*) with 80% BS support.

Malvales

The monophyly of Malvales is recovered with 85% BS support. The clade comprises 10 families (APG III, 2009). Cytinaceae (BS = 100%) are sister to other Malvales (BS = 85%), followed by Neuradaceae (BS = 99%) and Thymelaeaceae (BS = 94%). The remaining families, Malvaceae (BS = 94%), Muntingiaceae (BS = 100%), Cistaceae (BS = 100%), Sarcolaenaceae (BS = 97%), Dipterocarpaceae (BS = 64%), Sphaerosepalaceae (BS = 99%), and Bixaceae (BS = 89%), form a clade with 51% BS support (Fig. 2B). Within this clade, (Dipterocarpaceae + Sarcolaenaceae) + Cistaceae (Dipterocarpean lineage; see Le Péchon & Gigord, 2014) are sister to Bixaceae + Sphaerosepalaceae (BS = 64%; Fig. 2B).

Thymelaeaceae. Thymelaeaceae are divided into two clades here, but without strong support. Within the first clade, *Synandrodaphne*, *Octolepis*, *Gonystylus*, *Lethedon*, and *Solmsia* are successive sisters to *Arnhemia* + *Deltaria* with 53%, 50%, 79%, 54%, 68%, and 51% BS support, respectively. Within the second clade, *Gyrinops* + *Aquilaria*, *Dicranolepis* + (*Enkleia* + (*Craterosiphon* + *Synaptolepis*)), *Edgeworthia* + ((*Daphne* + *Thymelaea*) + (*Stellera* + *Wikstroemia*)), and *Dais* + *Phaleria* are successive sisters to a clade composed of the remaining members, with 68%, 58%, 75%, and 64% BS support, respectively (see van der Bank et al., 2002; Motsi et al., 2010). The relationships among the rest of the clade are unclear, except that *Struthiola* + *Diarthron* has 90% BS support. Moreover, *Pimelea* would be monophyletic with 93% BS support if it is expanded to include *Thecanthes*, as suggested by Motsi et al. (2010).

Malvaceae. The phylogeny of Malvaceae is poorly resolved. Only seven clades are recovered, whereas nine subfamilies are recognized (Bayer et al., 1999). The clades recovered here correspond to Byttnerioideae–Grewioideae (see discussion below), Helicteroideae (including Durioneae), Dombeyoideae, Tilioideae, Brownlowioideae, Sterculioideae, and Bombacoideae–Malvoideae. The relationships among these seven lineages are unclear.

Within the Byttnerioideae–Grewioideae clade, only the monophyly of Grewioideae is recovered (BS = 65%). This clade is composed of two sister groups: *Apeiba* + (*Sparrmannia* + (*Corchorus* + *Triumfetta*–*Heliocarpus*)), with 67%, 50%, 69%, and 84% BS support from the external to internal nodes, respectively, with *Triumfetta* non-monophyletic, and *Trichospermum*–*Grewia* + (*Luehea* + (*Goethalsia* + (*Colona* + *Microcos*))), with 71%, 55%, 63%, and 60% BS support from the external to internal nodes, respectively. Byttnerioideae (BS < 50%) are divided into three groups. *Leptonychia*, *Abroma*, *Kleinhovia*, *Byttneria*, *Rayleya*, and *Ayenia* form a clade corresponding to tribe Byttnerieae (Bayer & Kubitzki, 2003), but only the relationship of *Abroma* + (*Kleinhovia* + *Byttneria*–*Rayleya*–*Ayenia*) received ≥90% BS support, and *Byttneria* and *Ayenia* are non-monophyletic (see Whitlock & Hale, 2011). Similarly, *Hermannia*, *Theobroma*, *Herrania*, *Guazuma*, and *Abelmoschus* form a clade but BS support is <50%. The remaining genera form a clade, but only the relationships of *Guichenotia* + some species of *Rulingia*, *Commersonia* + other species of *Rulingia*, and *Thomasia* +

Lasiopetalum are well supported. Both *Melochia* and *Rulingia* are non-monophyletic. Helicteroideae are resolved as two sister clades. *Neesia* and *Durio* are sisters, forming a clade corresponding to Durioneae (BS = 100%), which is sister to a well-resolved clade of *Reevesia* + (*Helicteres* + *Triplochiton*) with 71% BS support (see Nyffeler & Baum, 2000).

Dombeyoideae are resolved as two clades. *Nesogordonia*–((*Dombeya* + *Trochetiopsis*) + (*Schoutenia* + *Pterospermum*)) is resolved as the first clade with 81% BS support. Within the second clade, *Excentrodendron* + *Burretiodendron* are sister to *Schoutenia* + (*Corchoropsis* + (*Ruizia* + ((*Paramelhania* + *Trochetiopsis*)–(*Eriolaena* + (*Dombeya* + *Helmiopsiella*))))) with 52% BS support. *Dombeya*, *Trochetiopsis*, and *Schoutenia* are not monophyletic (see Le Péchon et al., 2015).

A clade of *Mortoniodendron*, *Craigia*, and *Tilia* corresponds to Tilioideae, but without strong support. Moreover, neither *Mortoniodendron* nor *Craigia* is monophyletic.

The phylogeny of Brownlowioideae is resolved as the remainder of *Mortoniodendron* + ((*Brownlowia* + *Pentace*) + (*Berrya* + *Christiana*)). In contrast, Nyffeler et al. (2005) suggested that *Mortoniodendron* is included within Tilioideae, and some species are placed there in our analysis as well (see above). *Cola*, *Firmiana*, *Hildegardia*, *Scaphium*, *Pterocymbium*, *Sterculia*, *Heritiera*, *Franciscodendron*, and *Brachychiton* form a clade, corresponding to Sterculioideae, but without strong support. We recovered *Franciscodendron* + *Brachychiton* (BS = 53%) and *Firmiana* + *Hildegardia* (BS = 60%).

Within Bombacoideae–Malvoideae, the delimitation of Bombacoideae and Malvoideae remains questionable, even though past studies indicated they were distinct but closely related (e.g., Nyffeler & Baum, 2000; Nyffeler et al., 2005). However, Bombacoideae appear to be non-monophyletic, with many genera placed in this tribe scattered throughout Malvoideae (Fig. S3). These genera are *Pentaplaris*, *Camptostemon*, *Lagunaria*, *Uladendron*, *Howittia*, *Septotheca*, *Ochroma*, *Fremontodendron*, *Chiranthodendron*, *Patinoa*, and the *Matisia* alliance (*Phragmotheca*, *Matisia*, and *Quararibea*), and most of them have been referred to either subfamily (Bombacoideae and Malvoideae) in previous studies (e.g., Alverson et al. 1999; Bayer et al. 1999; Nyffeler & Baum, 2000). The rest of Bombacoideae form a weakly supported clade. Within that clade, we recovered *Huberodendron* + *Gyranthera* (BS = 85%), and *Pachira* and *Eriotheca* are non-monophyletic.

Within Malvoideae, three tribes have been recognized: Hibisceae, Gossypieae, and Malveae (Bayer & Kubitzki, 2003). Within Hibisceae, *Hibiscus* is highly polyphyletic with species appearing throughout this clade (see Pfeil & Crisp, 2005; Koopman & Baum, 2008). A few sister groups are recovered with BS > 50%: *Perrierophytum*–*Kosteletzkya* + *Megistostegium*–*Humbertiella* (BS = 62%), *Hibiscus*–*Macrostelia*–*Helicteropsis* complex (BS = 52%), *Pavonia*–*Peltaea* + (*Malvaviscus* + (*Malachra* + *Pavonia*)) (BS ≥ 61%), *Hibiscus* + *Kosteletzkya* (BS = 89%), and *Kydia* + (*Talipariti*–*Hibiscus* + (*Decaschistia* + (*Urena* + *Hibiscus*))) (BS ≥ 58%). The monophyly of Gossypieae is recovered with 83% BS support, and *Kokia* + *Gossypioides* (BS = 99%), are sister to *Gossypium* (BS = 71%) and *Azanza* + *Thespesia*–*Hampea* (BS ≥ 74%). The monophyly of *Thespesia* can be maintained if it includes *Hampea*. Seelanan et al. (1997) suggested that *Azanza* also be included in *Thespesia*. The Malveae are recovered with 79% BS support, but the internal relationships within this tribe are not fully resolved, except for a few sister groups, including *Nototriche* + *Sphaeralcea*–*Acaulimalva* (BS = 77%), (*Urocarpidium* + *Palaua*) + (*Callirhoe* + *Modiolastrum*) (BS ≥ 66%), ((*Robinsonella* + *Periptera*) + *Plagianthus*) + *Lecanophora* (BS ≥ 79%), *Eremalche* + *Malope* (BS = 59%), and (*Kearnemalvastrum* + *Fuertesimalva*) + *Andeimalva*–*Modiola* (BS ≥ 66%). *Malva* is polyphyletic (see García et al., 2009), as is *Sida* (see Donnell et al., 2012).

Cistaceae. *Fumana* and *Lechea* are successively sisters to the remainder of the family with 100% and 92% BS support, respectively (see Guzmán & Vargas, 2009). The remaining genera form a clade with 73% BS support, including *Cistus*, *Halimium*, *Tuberaria* (BS = 99%), *Crocanthemum* (BS = 72%), *Hudsonia*, and *Helianthemum*. Both *Halimium* and *Helianthemum* are non-monophyletic (see Civeyrel et al., 2011).

Dipterocarpaceae. Within Dipterocarpaceae, *Monotes* (Monotoideae) is sister to the rest of this family (BS = 64%; see Dayanandan et al., 1999). The remaining genera form a clade corresponding to Dipterocarpoideae, with 89% BS support. Within Dipterocarpoideae, *Vateriopsis* is sister to a group composed of *Anisoptera*, *Stemonoporus*, *Vateria*, *Vatica*, *Cotylelobium*, and *Upuna* (BS = 70%), but relationships among these genera are unclear. *Anisoptera* and *Vatica* are non-monophyletic, and *Vatica* is weakly supported as sister to *Cotylelobium* (BS = 50%). The remaining genera are *Dipterocarpus* (BS = 100%), *Dryobalanops* (BS = 100%), *Shorea*, *Hopea*, *Parashorea*, and *Neobalanocarpus*. *Shorea*, *Hopea*, *Parashorea*, and *Neobalanocarpus* form a clade with 60% BS support. *Shorea* is paraphyletic, falling into two separate groups (Fig. S3), one with *Parashorea* nested within it (BS = 50%), and the other with *Hopea* + *Neobalanocarpus* nested within it (79%).

Bixaceae. Within Bixaceae, *Diegodendron* is sister to *Bixa* (BS = 99%). This clade is then sister to *Amoreuxia* + *Cochlospermum* (BS = 100%) with 89% BS support.

Huerteales

Huerteales (BS = 98%) are sister to Brassicales + Malvales (BS = 85%). Huerteales contain four small families forming two sister groups: Petenaeaceae (Christenhusz et al., 2010; APG IV, 2016) + Gerrardinaceae (BS = 73%), and Tapisciaceae + Dipentodontaceae (BS = 98%). The monophyly of these four families is well supported (Table 1).

Sapindales

Sapindales are monophyletic (BS = 75%). The clade is represented by nine families: Biebersteiniaceae (BS = 100%), Nitrariaceae (BS = 99%), Anacardiaceae (BS = 57%), Burseraceae (BS = 84%), Kirkiaceae (BS = 99%), Sapindaceae (BS = 78%), Meliaceae (BS = 98%), Rutaceae (BS = 99%), and Simaroubaceae (BS = 69%). Of these nine families within Sapindales, only two sister groups are recovered with BS > 50%, Biebersteiniaceae + Nitrariaceae (BS = 73%) and Anacardiaceae + Burseraceae (BS = 67%). However, the general topology agrees with Muellner et al. (2007) and Soltis et al. (2011).

Nitrariaceae. Within Nitrariaceae, *Peganum* and *Malacocarpus* are successive sisters to the *Nitraria*–*Tetradiclis* complex

with 99% and 59% BS support, respectively. *Tetradiclis* is embedded in *Nitraria*, making *Nitraria* paraphyletic.

Burseraceae. *Beiselia* is sister to other Burseraceae (BS = 84%). The relationships among the rest of the genera of Burseraceae are not well resolved here, and most genera are non-monophyletic in our trees. The family requires further study with more taxa and genes.

Anacardiaceae. Relationships among genera of Anacardiaceae generally are not well resolved. *Pegia* + *Spondias* (BS = 80%) are recovered as sister to the remaining members of this family (BS = 57%), which form two clades that generally correspond to Spondioideae (BS = 69%; excluding *Pseudosmodingium*, *Pegia*, and *Spondias*) and Anacardioideae (68%), respectively (Pell et al., 2011). However, the relationships between and within these two subfamilies lack BS support >50%.

The relationships among all genera in Spondioideae are well resolved with *Buchanania* + (*Lannea* + (((*Antrocaryon* + *Sclerocarya*) + *Harpephyllum*) + *Cyrtocarpa* + ((*Pleiogynium* + *Dracontomelon*) + *Choerospondias*)–*Tapirira*)), except that *Pseudosmodingium*, *Pegia*, and *Spondias* appear in the Anacardioideae clade (Fig. S3). This suggests that Spondioideae may not be monophyletic (Pell et al., 2011; Weeks et al., 2014). Within Anacardioideae, only a small clade is recovered as ((*Anacardium* + *Fegimanra*) + *Gluta*) + *Mangifera*, with ≥67% BS support at each node, and four genera (*Rhus*, *Cotinus*, *Toxicodendron*, and *Schinopsis*) are non-monophyletic in our analyses.

Sapindaceae. Within Sapindaceae (Fig. S3; Acevedo-Rodríguez et al., 2011), *Xanthoceras* in Xanthoceroideae is resolved as sister to the rest of the family (BS = 78%; see Buerki et al., 2010) and then subsequently followed by Hippocastanoideae (BS = 99%), Dodonaeoideae (BS = 96%), and Sapindoideae (BS = 62%), with 70% and 68% BS support, respectively.

Within Hippocastanoideae, *Acer* is non-monophyletic and contains *Dipteronia*, forming a clade with 100% BS support, sister to *Aesculus* + (*Billia* + *Handeliodendron*), with 62% and 100% BS support, respectively.

The monophyly of Dodonaeoideae is well supported, except that *Eurycorymbus*, which is assigned to Dodonaeoideae, is nested within Sapindoideae. *Dodonaea* is closely related to *Diplopeltis* and *Distichostemon*, but is not monophyletic.

Within Sapindoideae, *Ungnadia* and *Stadmania* + (*Smelophyllum* + *Koelreuteria*) are successive sisters to the rest of Sapindoideae, with 62% and 68% BS support, respectively. The relationships among the remaining genera in Sapindoideae are generally not well resolved. The clade of ((((*Allophylus*–*Sapindus* + *Thouinia*) + *Bridgesia*) + ((*Cardiospermum*–*Paullinia* + *Serjania*) + *Urvillea*)) + (*Athyana* + *Diatenopteryx*)) + *Guindilia* is well supported (Fig. S3). The sister groups are *Schleichera* + *Paranephelium* (BS = 100%), *Beguea* + *Macphersonia* (BS = 66%), *Plagioscyphus* + *Pappea* (BS = 59%), *Alectryon* + *Podonephelium* (BS = 85%), *Lepisanthes* + *Hebecoccus* (BS = 60%), *Atalaya* + *Pseudima* (BS = 55%), *Dimocarpus* + *Euphoria* (BS = 87%), and *Pometia* + *Nephelium* (BS = 96%). *Blighia* is not monophyletic, with *Lepidopetalum* nested within it. Similarly, *Sapindus* is nested within *Allophylus*, making *Allophylus* non-monophyletic. *Talisia* and *Diploglottis* are also non-monophyletic.

Meliaceae. Within Meliaceae, two clades are recovered, corresponding to two subfamilies Cedreloideae (BS = 67%) and Melioideae (BS = 64%) (Mabberley, 2011).

Within Cedreloideae, *Neobeguea* + *Chukrasia* (BS = 71%) is sister to the rest of this subfamily, which form a clade with 51% BS support. Within this clade, *Lovoa* and *Capuronianthus* are sisters with 84% BS support. *Swietenia* is not monophyletic, with *Schmardaea* nested within it. Similarly, *Cedrela* is embedded in *Toona*.

Within Melioideae, a few well-resolved clades of (*Melia* + *Azadirachta*) + *Owenia*, *Cipadessa* + (*Trichilia* + *Nymania*), *Astrotrichilia*, *Sandoricum* + (*Ekebergia* + *Quivisianthe*), and *Walsura* are successively sisters to the rest, with 64%, 64%, 84%, 61%, and 98% BS support, respectively. The rest of Melioideae form a clade with 83% BS support, but without internal resolution, and some genera (*Turraea*, *Aglaia*, *Dysoxylum*, and *Guarea*) are non-monophyletic.

Rutaceae. The relationships among genera in Rutaceae differ in many respects with the taxonomic structure of this family (Kubitzki et al., 2011). Cneoroideae (BS = 96%) are sister to other Rutaceae. The remaining members have the topology of (*Harrisonia* + (*Cneorum* + (*Cedrelopsis*–*Bottegoa* + *Ptaeroxylon*))) + *Spathelia*–*Dictyoloma*, and these are sister to the rest (BS = 99%). The remaining genera form two sister clades (BS = 99%; see Fig. S3): (i) a small clade of (*Dictamnus* + (*Skimmia* + *Orixa*)) + *Casimiroa* (BS = 98%) is sister to the remainders of this clade (BS = 97%); and (ii) *Ruta* + (*Boenninghausenia* + *Thamnosma*), *Cneoridium*, and *Haplophyllum* are subsequent sisters to a clade (BS = 99%) of *Aegle*, *Aeglopsis*, *Afraegle*, *Atalantia*, *Balsamocitrus*, *Burkillanthus*, *Citropsis*, *Citrus*, *Clausena*, *Clymenia*, *Feroniella*, *Fortunella*, *Glycosmis*, *Limonia*, *Luvunga*, *Merope*, *Merrillia*, *Micromelum*, *Monanthocitrus*, *Murraya*, *Naringi*, *Oxanthera*, *Pamburus*, *Paramignya*, *Pleiospermium*, *Severinia*, *Swinglea*, *Triphasia*, and *Wenzelia*. However, the monophyly of *Atalantia*, *Cedrelopsis*, *Citrus*, *Citropsis*, *Murraya*, *Melicope*, *Spathelia*, *Vepris*, and *Zanthoxylum* is not supported.

Simaroubaceae. The generic-level phylogeny of Simaroubaceae generally agrees with Clayton et al. (2007) (Fig. S3). Three major clades are recovered here: (i) the *Picrasma* clade, with *Picrasma* + (*Castela* + *Holacantha*), sister to the rest of the family (BS = 69%); (ii) the *Soulamea* clade, with *Nothospondias* sister to ((*Amaroria* + *Soulamea*) + *Brucea*) + *Leitneria*, but with BS support <50%; and (iii) the *Simarouba* clade, with (((*Simarouba* + *Pierreodendron*) + *Simaba*)–(*Eurycoma* + *Odyendea*)) + ((*Odyendea* + *Hannoa*) + (*Perriera* + *Gymnostemon*)). *Samadera*, *Quassia*, and *Picrolemma* are successively sister to the *Simarouba* clade (Fig. S3) with 91%, 99%, and 100% BS support, respectively, and then all are sister to the *Soulamea* clade (BS = 69%), with *Ailanthus* sister to all of the above (BS = 69%). The relationship between (*Simarouba* + *Pierreodendron*) + *Simaba* and *Eurycoma* + *Odyendea* lacks strong support, and *Odyendyea* is resolved as paraphyletic, with some of its species sister to *Hannoa* and some sister to *Eurycoma*.

Picramniales and Crossosomatales

Picramniales (Picramniaceae) are strongly supported as monophyletic and sister to Sapindales (Fig. 2B; also see APG III, 2009; Soltis et al., 2011).

Crossosomatales are monophyletic with strong support. The monophyly of all the families in this order is recovered with 100% BS support (Table 1). Crossosomatales are split into two subclades: Aphloiaceae + (Geissolomataceae + Strasburgeriaceae) (BS ≥ 75%) and Staphyleaceae + (Guamatelaceae + (Crossosomataceae + Stachyuraceae)) (BS ≥ 75%).

Relationships of the Geraniales–Myrtales clade

Geraniales and Myrtales are members of *Malvidae* in most analyses (Jansen et al., 2007; APG III, 2009; Burleigh et al., 2009; Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014). However, in our study, Geraniales and Myrtales form a clade (BS = 58%) that is sister to *Malvidae* and *Fabidae* with weak support (BS = 56%; Fig. 2B). In our *matR* analysis, Geraniales and Myrtales are also not placed within *Malvidae*, although with <50% BS support (Fig. S1). Nevertheless, similar topologies have also been reported in other studies (Zhu et al., 2007; Qiu et al., 2010; Zhang et al., 2012; Wang et al., 2014; Xi et al., 2014; Zeng et al., 2014; Sun et al., 2015; Yang et al., 2015).

Geraniales

Geraniales are a small order including three strongly supported families (BS = 100%; Fig. 2B; Table 1), with Geraniaceae + (Vivianiaceae + Melianthaceae). The relationships among the three families are also strongly supported (BS ≥ 99%; Fig. 2B). Note that nomenclatural priority requires that Francoaceae be substituted for Melianthaceae (per APG IV, 2016). Moreover, Vivianaceae are also included in Francoaceae in APG IV (2016). Following APG IV (2016), Geraniales comprise only Geraniaceae and Francoaceae (see Fig. 2B).

Geraniaceae. Within Geraniaceae, *Hypseocharis* is sister to the rest of the family (BS = 100%), as in Price & Palmer (1993) and Bakker et al. (1998). Other relationships are *Pelargonium* + ((*Geranium* + (*Erodium* + *California*)) + *Monsonia–Sarcocaulon*) with strong support (Fig. S3), which agrees with previous studies (Fiz et al., 2008; Palazzesi et al., 2012). The monotypic *California* is sister to *Erodium*, although BS support is only 52%, whereas it was unresolved in Fiz et al. (2008) using a combined dataset of *rbcL* and *trnL-F*. *Sarcocaulon* is nested in *Monsonia*, forming a clade with 100% BS support, making *Monsonia* paraphyletic.

Francoaceae (Melianthaceae and Vivianiaceae). The topology is (*Melianthus* + *Bersama*) + *Francoa–Greyia* with the relationship between the latter two genera unclear. *Wendtia* and *Viviania* form a clade with strong support.

Myrtales

Myrtales (BS = 77%) comprise nine families: Alzateaceae (BS = 100%), Combretaceae (BS = 98%), Crypteroniaceae (BS = 100%), Lythraceae (BS = 97%), Melastomataceae (BS = 98%), Myrtaceae (BS = 79%), Onagraceae (BS = 99%), Penaeaceae (BS = 62%), and Vochysiaceae (BS = 96%). In the present study, Combretaceae are sister to the remaining families of this order, which form a weakly supported clade of two sister groups: Lythraceae + Onagraceae (BS = 100%) and a clade of (Vochysiaceae + Myrtaceae) + (Melastomataceae + (Crypteroniaceae + (Alzateaceae + Penaeaceae))), which has moderate to high BS support (Fig. 2B).

Combretaceae. In our study, *Conocarpus* is sister to the rest of Combretaceae with strong support (BS = 98%; Fig. S3), in contrast to Maurin et al. (2010), who suggested that *Strephonema* was sister to the rest of the family based on analyses of a combined plastid and nuclear ITS dataset. The rest of the genera (BS = 62%) form two clades. The first (BS = 66%) is a single genus, *Strephonema*, corresponding to subfamily Strephonematoideae. The second (BS = 71%) corresponds to Combretoideae and in turn consists of two groups: (i) *Guiera* + (*Laguncularia* + *Lumnitzera*); and (ii) ((*Anogeissus* + *Terminalia–*(*Bucida* + *Buchenavia*)) + *Terminalia–Pteleopsis*) + (*Combretum* + *Thiloa*). The topology of the second group indicates that *Terminalia* is paraphyletic, and *Quisqualis* and *Calopyxis* are embedded in *Combretum* (see Maurin et al., 2010).

Lythraceae. Lythraceae are resolved as two clades (Fig. S3). The first is ((*Pleurophora* + *Cuphea*) + (*Woodfordia* + (*Koehneria* + (*Adenaria* + *Pehria*)))) + (*Punica* + (*Pemphis* + (*Lafoensia* + *Capuronia*))), and the second comprises two unresolved complexes subsequently followed by *Lythrum–Peplis* + *Decodon* and (*Didiplis* + *Rotala*) + *Heimia*. *Lythrum*, *Sonneratia*, *Nesaea*, and *Ammannia* are non-monophyletic (see Shi et al., 2000; Huang & Shi, 2002; Graham et al., 2005).

Onagraceae. Within Onagraceae, *Ludwigia* (subfamily Jussiaeoideae) is sister to the rest of the family (BS = 99%), which agrees with recent studies (Levin et al., 2003, 2004; Wagner & Hoch, 2005 onwards; Ford & Gottlieb, 2007). The remaining genera form a clade (BS = 100%), recognized as Onagroideae, with (*Fuchsia* + *Circaea*) + (*Clarkia* + *Oenothera–Gaura* + (*Epilobeae* + *Chamerion*)), which is generally consistent with Wagner & Hoch (2005 onwards). Our study also indicates that *Oenothera* is paraphyletic, as suggested by Levin et al. (2004).

Vochysiaceae and Myrtaceae. *Vochysia* is non-monophyletic. Similarly, a number of genera in Myrtaceae (*Agonis*, *Astartea*, *Babingtonia*, *Backhousia*, *Baeckea*, *Chamelaucium*, *Darwinia*, *Eugenia*, *Kunzea*, *Leptospermum*, *Melaleuca*, *Metrosideros*, *Pimenta*, *Syzygium*, and *Verticordia*) are non-monophyletic or polyphyletic.

Crypteroniaceae. In Crypteroniaceae, *Dactylocladus* is sister to *Crypteronia* and *Axinandra* with strong support (BS = 100%), but BS support for *Crypteronia* and *Axinandra* is only 51%.

Penaeaceae. A clade of *Olinia* and *Rhynchocalyx* is sister to the rest of the family (BS = 62%). Both *Brachysiphon* and *Stylapterus* are non-monophyletic.

Melastomataceae. Two major groups are recovered in Melastomataceae, generally agreeing with Clausing & Renner (2001). The first is *Memecylon–Mouriri* (BS = 98%), corresponding to Olisbeoideae (Memecylaceae or Memecyloideae), and the second group corresponds to Melastomatoideae (BS = 98%). In Melastomatoideae, *Pternandra* is resolved as sister to the rest (BS = 98%).

Conclusions and Future Prospects

The *Rosidae* phylogeny constructed here with the chloroplast genes *atpB*, *matK*, and *rbcL* and the more slowly evolving

mitochondrial *matR* is largely in agreement with previous work (Zhu et al., 2007; Wang et al., 2009; Qiu et al., 2010; Soltis et al., 2011; Ruhfel et al., 2014; Sun et al., 2015), with some notable areas of incongruence. For example, the COM clade may be an example of ancient lineage sorting or hybridization (Sun et al., 2015). Zygophyllales are sister to the nitrogen-fixing clade and the COM clade in *Fabidae* using plastid sequence data (Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014) and also in our three-chloroplastid-gene and four-gene-combined analyses (Figs. S2 and S3). However, our *matR* gene analysis indicates that Zygophyllaceae are closely related to Crossosomatales as do other mitochondrial data analyses (Fig. S1; Zhu et al., 2007; Qiu et al., 2010; Sun et al., 2015). Other examples (e.g., Geraniales, Myrtales, Casuarinaceae and Myricaceae, and Rafflesiaceae and Euphorbiaceae) require further study with more nuclear gene and taxa sampling.

Extensive taxon sampling can help elucidate the phylogeny of *Rosidae* at the generic level and clarify infrafamilial relationships. The dense taxon sampling strategy used here can improve the accuracy of phylogenetic analyses in some cases (e.g., Wiens, 2003; Heath et al., 2008), and it is helpful for testing the monophyly of many genera. However, considering all of the missing data in the supermatrix analyses, potentially different evolutionary histories among chloroplast and mitochondrial loci, the lack of nuclear data, and possible errors in the GenBank data, our results should be verified with more detailed study.

In our study, the *Superrosidae* (Moore et al., 2010; Soltis et al., 2011) are strongly supported (BS = 94%; Fig. 2B). Within the *Rosidae*, we find support for *Malvidae* and *Fabidae*, but Geraniales and Myrtales form a clade, outside of *Malvidae* and *Fabidae* (Fig. 2B), albeit with low support. This result differs from many previous studies (Jansen et al., 2007; APG III, 2009; Burleigh et al., 2009; Wang et al., 2009; Soltis et al., 2011; Ruhfel et al., 2014), but this topology has also been reported in other studies (Zhu et al., 2007; Burleigh et al., 2009; Qiu et al., 2010; Zhang et al., 2012; Wang et al., 2014; Xi et al., 2014; Zeng et al., 2014; Sun et al., 2015; Yang et al., 2015).

The abundance of phylogenetic data and advances in bioinformatics and computational methods have enabled the assembly of mega-phylogenies representing angiosperm relationships (e.g., Zanne et al., 2014; Hinchliff et al., 2015). In this study, we use a careful, botanically informed approach to curate and assemble the sequence data and build and evaluate the resulting tree of *Rosidae*. Hence this tree not only has useful phylogenetic implications, but it can be a helpful resource for future evolutionary and ecological analyses of *Rosidae*.

Acknowledgements

This work was supported by the National Key Basic Research Program of China (Grant No. 2014CB954100), National Natural Science Foundation of China (Grant Nos. 31590822 and 31270268), the US National Science Foundation (Grant DEB-1208428), the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists (2011T1S24), State Key Laboratory of Systematic and Evolutionary Botany (LSEB2011-10), and National Science & Technology Infrastructure – National Specimen Information Infrastructure (Grant No. 2005DKA21400).

References

Acevedo-Rodríguez P, van-Welzen PC, Adema FR, van der Ham WJM. 2011. Sapindaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 10: 357–407.

Altshul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Anderson C. 2011. Revision of *Ryssopterys* and transfer to *Stigmaphyllon* (Malpighiaceae). *Blumea* 56: 73–104.

Anderson WR, Anderson C, Davis CC. 2006. Malpighiaceae [online]. Available from http://herbarium.lsa.umich.edu/malpigh/index.html [accessed 20 August 2015].

Andersson L, Andersson S. 2000. A molecular phylogeny of Tropaeolaceae and its systematic implications. *Taxon* 49: 721–736.

Alverson WS, Whitlock BA, Nyffeler R, Bayer C, Baum DA. 1999. Phylogeny of core Malvales: Evidence from *ndhF* sequence data. *American Journal of Botany* 86: 1474–1486.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161: 105–121.

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. 2016. *Botanical Journal of the Linnean Society* 181: 1–20.

Bakker FT, Vassiliades DD, Morton C. 1998. Phylogenetic relationships of *Biebersteinia* Stephan (Geraniaceae) inferred from *rbcL* and *atpB* sequence comparisons. *Botanical Journal of the Linnean Society* 127: 149–158.

Ballard HEJ, de Paula-Souza J, Wahlert GA. 2013. Violaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 11: 303–322.

Bayer C, Fay MF, de Bruijn AY, Savolainen V, Morton CM, Kubitzki K, Chase MW. 1999. Support for an expanded family concept of Malvaceae within a recircumscribed order Malvales: A combined analysis of plastid *atpB* and *rbcL* DNA sequences. *Botanical Journal of the Linnean Society* 129: 267–303.

Bayer C, Kubitzki K. 2003. Malvaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 5: 225–311.

Bazinet AL, Zwickl DJ, Cummings MP. 2014. A gateway for phylogenetic analysis powered by grid computing featuring GARLI 2.0. *Systematic Biology* 63: 812–818.

Beier BA, Chase MW, Thulin M. 2003. Phylogenetic relationships and taxonomy of subfamily Zygophylloideae (Zygophyllaceae) based on molecular and morphological data. *Plant Systematics and Evolution* 240: 11–39.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. 2006. Brassicaceae phylogeny and trichome evolution. *American Journal of Botany* 93: 607–619.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* 97: 1296–1303.

Bello MA, Bruneau A, Forest F, Hawkins JA. 2009. Elusive relationships within order Fabales: Phylogenetic analyses using *matK* and *rbcL* sequence data. *Systematic Biology* 34: 102–114.

Bello MA, Rudall PJ, Hawkins JA. 2012. Combined phylogenetic analyses reveal interfamilial relationships and patterns of floral evolution in the eudicot order Fabales. *Cladistics* 28: 393–421.

Bruneau A, Forest F, Herendeen PS, Klitgaard BB, Lewis GP. 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Systematic Botany* 26: 487–514.

Bruneau A, Mercure M, Lewis GP, Herendeen PS. 2008. Phylogenetic patterns and diversification in the caesalpinioid legumes. *Botany* 86: 697–718.

Buerki S, Lowry II, Porter P, Alvarez N, Razafimandimbison SG, Küpfer P, Callmander MW. 2010. Phylogeny and circumscription of Sapindaceae revisited: Molecular sequence data, morphology and biogeography support recognition of a new family, Xanthoceraceae. *Plant Ecology and Evolution* 143: 148–159.

Burleigh JG, Bansal MS, Eulenstein O, Hartmann S, Wehe A, Vision TJ. 2011. Genome-scale phylogenetics: Inferring the plant tree of life from 18,896 gene trees. *Systematic Biology* 60: 117–125.

Burleigh JG, Hilu KW, Soltis DE. 2009. Inferring phylogenies with incomplete data sets: A 5-gene, 567-taxon analysis of angiosperms. *BMC Evolutionary Biology* 9: 61.

Cantino PD, Doyle JA, Graham SW, Judd WS, Olmstead RG, Soltis DE, Soltis PS, Donoghue MJ. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56: 822–846.

Cardoso D, de-Queiroz LP, Pennington RT, de-Lima HC, Fonty É, Wojciechowski MF, Lavin M. 2012. Revisiting the phylogeny of papilionoid legumes: New insights from comprehensively sampled early-branching lineages. *American Journal of Botany* 99: 1991–2013.

Cardoso D, Pennington RT, de-Queiroz LP, Boatwright JS, Van Wyk BE, Wojciechowski MF, Lavin M. 2013. Reconstructing the deep-branching relationships of the papilionoid legumes. *South African Journal of Botany* 89: 58–75.

Carvalho FA. 2013 onwards. e-Monograph of Caricaceae. Version 1, November 2013. [online]. Available from http://herbaria.plants.ox.ac.uk/bol/caricaceae [accessed 13 August 2015].

Chase MW, Zmarzty S, Lledó MD, Wurdack KJ, Swensen SM, Fay MF. 2002. When in doubt, put it in Flacourtiaceae: A molecular phylogenetic analysis based on plastid *rbcL* DNA sequences. *Kew Bulletin* 57: 141–181.

Chen ZD, Yang T, Lin L, Lu LM, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, Cao ZY, Liu HM, Wang XM, Wang W, Zhang JB, Meng Z, Cao W, Li JH, Wu SD, Zhao HL, Liu ZJ, Du ZY, Wang QF, Guo J, Tan XX, Su JX, Zhang LJ, Yang LL, Liao YY, Li MH, Zhang GQ, Chung SW, Zhang J, Xiang KL, Li RQ, Soltis DE, Soltis PS, Zhou SL, Ran JH, Wang XQ, Jin XH, Chen YS, Gao TG, Li JH, Zhang SZ, Lu AM, China Phylogeny Consortium. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54: 277–306.

Chin SW, Shaw J, Haberle R, Wen J, Potter D. 2014. Diversification of almonds, peaches, plums and cherries – molecular systematics and biogeographic history of *Prunus* (Rosaceae). *Molecular Phylogenetics and Evolution* 76: 34–48.

Christenhusz, MJM, Fay MF, Clarkson JJ, Gasson P, Morales Can J, Jiménez Barrios JB, Chase MW. 2010. Petaeaceae, a new angiosperm family with a distant relationship to Gerradina (Gerrardinaceae). *Botanical Journal of the Linnean Society* 164: 16–25.

Civeyrel L, Leclercq J, Demoly JP, Agnan Y, Quèbre N, Pélissier C, Otto T. 2011. Molecular systematics, character evolution, and pollen morphology of *Cistus* and *Halimium* (Cistaceae). *Plant Systematics and Evolution* 295: 23–54.

Clausing G, Renner SS. 2001. Molecular phylogenetics of Melastomataceae and Memecylaceae: Implications for character evolution. *American Journal of Botany* 88: 486–498.

Clayton JW, Fernando ES, Soltis PS, Soltis DS. 2007. Molecular phylogeny of the tree-of-heaven family (Simaroubaceae) based on chloroplast and nuclear markers. *International Journal of Plant Sciences* 168: 1325–1339.

Cook CD, Rutishauser R. 2001. Name changes in Podostemaceae. *Taxon* 50: 1163–1167.

Coughenour JM, Simmons MP, Lombardi JA, Yakobson K, Archer RH. 2011. Phylogeny of Celastraceae subfamily Hippocrateoideae inferred from morphological characters and nuclear and plastid loci. *Molecular Phylogenetics and Evolution* 59: 320–330.

Davis CC, Anderson WR. 2010. A complete generic phylogeny of Malpighiaceae inferred from nucleotide sequence data and morphology. *American Journal of Botany* 97: 2031–2048.

Davis CC, Latvis M, Nickrent DL, Wurdack KJ, Baum DA. 2007. Floral gigantism in Rafflesiaceae. *Science* 315: 1812–1812.

Dayanandan S, Ashton PS, Williams SM, Primack RB. 1999. Phylogeny of the tropical tree family Dipterocarpaceae based on nucleotide sequences of the chloroplast *rbcL* gene. *American Journal of Botany* 86: 1182–1190.

de Queiroz A, Gatesy JG. 2007. The supermatrix approach to systematics. *Trends in Ecology and Evolution* 22: 34–41.

Dobeš C, Paule J. 2010. A comprehensive chloroplast DNA-based phylogeny of the genus *Potentilla* (Rosaceae): Implications for its geographic origin, phylogeography and generic circumscription. *Plant Systematics and Evolution* 56: 156–175.

Donnell AA, Ballard HE, Cantino-Jr PD. 2012. *Callianthe* (Malvaceae): A new genus of neotropical Malvaceae. *Systematic Botany* 37: 212–222.

Doyle JJ, Doyle JL, Ballenger JA, Dickson EE, Kajita T, Ohashi H. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: Taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541–554.

Dressler S, Repplinger M, Bayer C. 2014. Linaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 11: 237–246.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Schranz ME, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW. 2015. The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences USA* 112: 8362–8366.

Edgar RC. 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5: 113.

Eriksson T, Hibbs MS, Yoder AD, Delwiche CF, Donoghue MJ. 2003. The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the *trnL-F* region of chloroplast DNA. *International Journal of Plant Sciences* 164: 197–211.

Eriksson T, Lundberg M, Töpel M, Östensson P, Smedmark JE. 2015. *Sibbaldia*: A molecular phylogenetic study of a remarkably polyphyletic genus in Rosaceae. *Plant Systematics and Evolution* 301: 171–184.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH. 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Systematic Botany* 35: 811–826.

Fiz O, Vargas P, Alarcón ML, Aedo M, Garcia JL, Aldasoro JJ. 2008. Phylogeny and historical biogeography of Geraniaceae in relation to climate changes and pollination ecology. *Systematic Botany* 33: 326–342.

Ford VS, Gottlieb LD. 2007. Tribal relationships within Onagraceae inferred from PgiC sequences. *Systematic Botany* 32: 348–356.

Forest F, Chase MW, Persson C, Crane PR, Hawkins JA. 2007. The role of biotic and abiotic factors in evolution of ant dispersal in the milkwort family (Polygalaceae). *Evolution* 61: 1675–1694.

Forrest LL, Hollingsworth PM. 2003. A recircumscription of *Begonia* based on nuclear ribosomal sequences. *Plant Systematics and Evolution* 241: 193–211.

Fougère-Danezan M, Maumont S, Bruneau A. 2007. Relationships among resin-producing Detarieae s.l. (Leguminosae) as inferred by molecular data. *Systematic Botany* 32: 748–761.

García PE, Schönswetter P, Aguilar JF, Feliner GN, Schneeweiss GM. 2009. Five molecular markers reveal extensive morphological homoplasy and reticulate evolution in the *Malva* alliance (Malvaceae). *Molecular Phylogenetics and Evolution* 50: 226–239.

Gehrke B, Bräuchler C, Romoleroux K, Lundberg M, Heubl G, Eriksson T. 2008. Molecular phylogenetics of *Alchemilla*, *Aphanes* and *Lachemilla* (Rosaceae) inferred from plastid and nuclear intron and spacer DNA sequences, with comments on generic classification. *Molecular Phylogenetics and Evolution* 47: 1030–1044.

Goloboff PA, Catalano SA, Marcos-Mirande J, Szumik CA, Salvador AJ, Källersjö M, Farris JS. 2009. Phylogenetic analysis of 73 060 taxa corroborates major eukaryotic groups. *Cladistics* 25: 211–230.

Graham SA, Hall J, Sytsma K, Shi SH. 2005. Phylogenetic analysis of the Lythraceae based on four gene regions and morphology. *International Journal of Plant Sciences* 166: 995–1017.

Groppo M, Fiaschi P, Salatino MLF, Ceccantini GCT, dos-Santos FDAR, Verola CF, Antonelli A. 2010. Placement of *Kuhlmanniodendron* Fiaschi & Groppo in Lindackerieae (Achariaceae, Malpighiales) confirmed by analyses of *rbcL* sequences, with notes on pollen morphology and wood anatomy. *Plant Systematics and Evolution* 286: 27–37.

Guzmán B, Vargas P. 2009. Historical biogeography and character evolution of Cistaceae (Malvales) based on analysis of plastid *rbcL* and *trnL-trnF* sequences. *Organisms Diversity & Evolution* 9: 83–99.

Hadiah JT, Conn BJ, Quinn CJ. 2008. Infra-familial phylogeny of Urticaceae, using chloroplast sequence data. *Australian Systematic Botany* 21: 375–385.

Hadiah JT, Quinn CJ, Conn BJ. 2003. Phylogeny of *Elatostema* (Urticaceae) using chloroplast DNA data. *Telopea* 10: 235–246.

Hall JC. 2008. Systematics of Capparaceae and Cleomaceae: An evaluation of the generic delimitations of *Capparis* and *Cleome* using plastid DNA sequence data. *Botany* 86: 682–696.

Hall JC, Iltis HH, Sytsma KJ. 2004. Molecular phylogenetics of core Brassicales, placement of orphan genera *Emblingia*, *Forchhammeria*, *Tirania*, and character evolution. *Systematic Botany* 29: 654–669.

Heath TA, Hedtke SM, Hillis DM. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution* 46: 239–257.

Herbert J, Chase MW, Möller M, Abbott RJ. 2006. Nuclear and plastid DNA sequences confirm the placement of the enigmatic *Canacomyrica monticola* in Myricaceae. *Taxon* 55: 349–357.

Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R, Sauquet H, Neinhuis C, Slotta TAB, Rohwer JG, Campbell CS, Chatrou LW. 2003. Angiosperm phylogeny based on *matK* sequence information. *American Journal of Botany* 90: 1758–1776.

Hinchliff CE, Smith SA, Allman JF, Burleigh JG, Chaudhary R, Coghill LM, Crandall KA, Deng J, Drew BT, Gazis R, Gude K, Hibbett DS, Katz LA, Laughinghouse HD, McTavish EJ, Midford PE, Owen CL, Ree RH, Rees JA, Soltis DE, Williams T, Cranston KA. 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences USA* 112: 12764–12769.

Hoffmann P, Kathriarachchi H, Wurdack KJ. 2006. A phylogenetic classification of Phyllanthaceae (Malpighiales; Euphorbiaceae *sensu lato*). *Kew Bulletin* 61: 37–53.

Hu JM, Lavin M, Wojciechowski MF, Sanderson MJ. 2002. Phylogenetic analysis of nuclear ribosomal ITS/5.8 S sequences in the tribe Millettieae (Fabaceae): *Poecilanthe–Cyclolobium*, the core Millettieae, and the *Callerya* group. *Systematic Botany* 27: 722–733.

Huang CH, Sun RR, Hu Y, Zeng LP, Zhang N, Cai LM, Zhang Q, Koch MA, Al-Shehbaz I, Edger PP, Oires JC, Tan DY, Zhong Y, Ma H. 2015. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular Biology and Evolution* 33: 394–412.

Huang YL, Shi SH. 2002. Phylogenetics of Lythraceae *sensu lato*: A preliminary analysis based on chloroplast *rbcL* gene, *psaA-ycf3* spacer, and nuclear rDNA internal transcribed spacer (ITS) sequences. *International Journal of Plant Sciences* 163: 215–225.

Huson DH, Scornavacca C. 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology* 61: 1061–1067.

Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack MJ, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences USA* 104: 19369–19374.

Judd WS, Olmstead RG. 2004. A survey of tricolpate eudicot phylogenetic relationships. *American Journal of Botany* 91: 1627–1644.

Junier T, Zdobnov EM. 2010. The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26: 1669–1670.

Kato M, Kita Y, Koi S. 2003. Molecular phylogeny, taxonomy and biogeography of *Malaccotristicha australis* comb. nov. (syn. *Tristicha australis*) (Podostemaceae). *Australian Systematic Botany* 16: 177–183.

Kathriarachchi H, Hoffmann P, Samuel R, Wurdack KJ, Chase MW. 2005. Molecular phylogenetics of Phyllanthaceae inferred from five genes (plastid *atpB*, *matK*, *3'ndhF*, *rbcL*, and nuclear *PHYC*). *Molecular Phylogenetics and Evolution* 36: 112–134.

Kathriarachchi H, Samuel R, Hoffmann P, Mlinarec J, Wurdack KJ, Ralimanana H, Chase MW. 2006. Phylogenetics of tribe Phyllantheae (Phyllanthaceae; Euphorbiaceae sensu lato) based on nrITS and plastid *matK* DNA sequence data. *American Journal of Botany* 93: 637–655.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper, A, Markowitz, S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.

Kim C, Deng T, Chase M, Zhang DG, Nie ZL, Sun H. 2015. Generic phylogeny and character evolution in Urticeae (Urticaceae) inferred from nuclear and plastid DNA regions. *Taxon* 64: 65–79.

Koch MA, Haubold B, Mitchell-Olds T. 2001. Molecular systematics of the Brassicaceae: Evidence from coding plastidic *matK* and nuclear Chs sequences. *American Journal of Botany* 88: 534–544.

Koenen EJM, de Vos JM, Atchison GW, Simon MF, Schrire BD, de Souza ER, de Queiroz LP, Hughes CE. 2013. Exploring the tempo of species diversification in legumes. *South African Journal of Botany* 89: 19–30.

Koi S, Kato M. 2010. Developmental morphology of seedling and shoot and phylogenetic relationship of *Diplobryum koyamae* (Podostemaceae). *American Journal of Botany* 97: 373–387.

Koi S, Kita Y, Hirayama Y, Rutishauser R, Huber KA, Kato M. 2012. Molecular phylogenetic analysis of Podostemaceae: Implications for taxonomy of major groups. *Botanical Journal of the Linnean Society* 169: 461–492.

Koi S, Rutishauser R, Kato M. 2009. Phylogenetic relationship and morphology of *Dalzellia gracilis* (Podostemaceae, subfamily Tristichoideae) with proposal of a new genus. *International Journal of Plant Sciences* 170: 237–246.

Koopman MM, Baum DA. 2008. Phylogeny and biogeography of tribe Hibisceae (Malvaceae) on Madagascar. *Systematic Botany* 33: 364–374.

Kubitzki K, Kallunki JA, Duretto M, Wilson PG. 2011. Rutaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 10: 276–356.

Kück P, Meusemann K. 2010. FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution* 56: 1115–1118.

Le Péchon T, Gigord LDB. 2014. On the relevance of molecular tools for taxonomic revision in Malvales, Malvaceae s.l, and Dombeyoideae. In: Besse P ed. *Molecular plant taxonomy methods and protocols*. New York: Humana Press. 337–363.

Le Péchon T, Dai Q, Zhang LB, Gao XF, Sauquet H. 2015. Diversification of Dombeyoideae (Malvaceae) in the Mascarenes: Old taxa on Young Island. *International Journal of Plant Sciences* 176: 211–221.

Lee EK, Cibrian-Jaramillo A, Kolokotronis SO, Katari MS, Stamatakis A, Ott M, Chiu JC, Little DP, Stevenson DW, McCombie WR, Martienssen RA, Coruzzi G, DeSalle R. 2011. A functional phylogenomic view of the seed plants. *PLoS Genetics* 7: e1002411.

Lee S, Wen J. 2001. A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. *American Journal of Botany* 88: 150–160.

Legume Phylogeny Working Group. 2013. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon* 62: 217–248.

Levin RA, Wagner WL, Hoch PC, Hahn WJ, Rodriguez A, Baum DA, Katinas L, Zimmer EA, Sytsma KJ. 2004. Paraphyly in tribe Onagreae: Insights into phylogenetic relationships of Onagraceae based on nuclear and chloroplast sequence data. *Systematic Botany* 29: 147–164.

Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA, Sytsma KJ. 2003. Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. *American Journal of Botany* 90: 107–115.

Li HL, Wang W, Li RQ, Zhang JB, Sun M, Naeem R, Su JX, Xiang XG, Mortimer PE, Li DZ, Hyde KD, Xu JC, Soltis DE, Soltis PS, Li JH, Zhang SZ, Chen ZD, Lu AM. 2016. A well resolved global tree of the N-fixing clade based on supermatrix approach: In comparison with a local tree from China. *Journal of Systematics and Evolution* 54: 392–399.

Li HL, Wang W, Mortimer PE, Li RQ, Li DZ, Hyde KD, Xu JC, Soltis DE, Chen ZD. 2015. Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Scientific Reports* 5: 14023.

Li RQ, Chen ZD, Lu AM, Soltis DE, Soltis PS, Manos PS. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *International Journal of Plant Sciences* 165: 311–324.

Li YQ, Dressler S, Zhang D, Renner SS. 2009. More Miocene dispersal between Africa and Asia – the case of *Bridelia* (Phyllanthaceae). *Systematics Botany* 34: 521–529.

Lo EYY, Donoghue MJ. 2012. Expanded phylogenetic and dating analyses of the apples and their relatives. *Molecular Phylogenetics and Evolution* 63: 230–243.

Lu LM, Wang W, Chen ZD, Wen J. 2013. Phylogeny of the non-monophyletic *Cayratia* Juss. (Vitaceae) and implications for character evolution and biogeography. *Molecular Phylogenetics and Evolution* 68: 502–515.

Luckow M, Fortunato RH, Sede S, Livshultz T. 2005. The phylogenetic affinities of two mysterious monotypic mimosoids from southern South America. *Systematic Botany* 30: 585–602.

Mabberley DJ. 2011. Meliaceae. In: Kubitzki K ed. *The families and genera of vascular plants*. Berlin: Springer-Verlag. 10: 185–211.

MacKinder B. 2005. Tribe Detarieae. In: Lewis GP, Schrire BD, MacKinder B eds. *Legumes of the world*. Kew: Royal Botanic Gardens. 68–109.

Magallón S, Crane PR, Herendeen PS. 1999. Phylogenetic pattern, diversity, and diversification of eudicots. *Annals of the Missouri Botanical Garden* 86: 297–372.

Manos PS, Stone DE. 2001. Evolution, phylogeny and systematics of the Juglandaceae. *Annals of the Missouri Botanical Garden* 88: 231–269.

Manos PS, Zhou ZK, Cannon CH. 2001. Systematics of Fagaceae: Phylogenetic tests of reproductive trait evolution. *International Journal of Plant Sciences* 162: 1361–1379.

Matthews ML, Endress PK. 2006. Floral structure and systematics in four orders of rosids, including a broad survey of floral mucilage cells. *Plant Systematics and Evolution* 260: 199–221.

Maurin O, Chase MW, Jordaan M, van der Bank M. 2010. Phylogenetic relationships of Combretaceae inferred from nuclear and plastid DNA sequence data: Implications for generic classification. *Botanical Journal of the Linnean Society* 162: 453–476.

McDill J, Simpson BB. 2011. Molecular phylogenetics of Linaceae with complete generic sampling and data from two plastid genes. *Botanical Journal of the Linnean Society* 165: 64–83.

McKenna MJ, Simmons MP, Bacon CD, Lombardi JA. 2011. Delimitation of segregate genera of *Maytenus* s.l. (Celastraceae) based on molecular and morphological characters. *Systematic Botany* 36: 922–932.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Miller MA, Pfeiffer W, Schwartz T eds. *Gateway computing environments workshop (GCE)*. Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, 14 November 2010. New York: IEEE. 1–8.

Moore MJ, Hassan N, Gitzendanner MA, Bruenn RA, Croley M, Vandeventer A, Horn JW, Dhingra A, Brockington SF, Latvis M, Ramdial J, Alexandre R, Piedrahita A, Xi Z, Davis CC, Soltis PS, Soltis DE. 2011. Phylogenetic analysis of the plastid inverted repeat for 244 species: Insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. *International Journal of Plant Sciences* 172: 541–558.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolved the early diversification of eudicots. *Proceedings of the National Academy of Sciences USA* 107: 4623–4628.

Morton MC. 2011. Newly sequenced nuclear gene (Xdh) for inferring angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 98: 63–89.

Motsi MC, Moteetee AN, Beaumont AJ, Rye BL, Powell MP, Savolainen V, van-der Bank M. 2010. A phylogenetic study of *Pimelea* and *Thecanthes* (Thymelaeaceae): Evidence from plastid and nuclear ribosomal DNA sequence data. *Australian Systematics Botany* 23: 270–284.

Muellner AN, Vassiliades DD, Renner SS. 2007. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. *Plant Systematics and Evolution* 266: 233–252.

Nyffeler R, Baum DA. 2000. Phylogenetic relationships of the durians (Bombacaceae–Durioneae or Malvaceae/Helicteroideae/Durioneae) based on chloroplast and nuclear ribosomal DNA sequences. *Plant Systematics and Evolution* 224: 55–82.

Nyffeler R, Bayer C, Alverson WS, Yen A, Whitlock BA, Chase MW, Baum DA. 2005. Phylogenetic analysis of the *Malvadendrina* clade (Malvaceae s.l.) based on plastid DNA sequences. *Organisms Diversity & Evolution* 5: 109–123.

Palazzesi L, Gottschling M, Barreda V, Weigend M. 2012. First Miocene fossils of Vivianiaceae shed new light on phylogeny, divergence times, and historical biogeography of Geraniales. *Botanical Journal of the Linnean Society* 107: 67–85.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* 63: 315–328.

Pattengale ND, Swenson KM, Moret BME. 2010. Uncovering hidden phylogenetic consensus. In: Borodovsky M, Gogarten JPeter, Przytycka TM, Rajasekaran S eds. *Bioinformatics research and applications*. Berlin: Springer-Verlag. 128–139.

Pell SK, Mitchell JD, Miller AJ, Lobova TA. 2011. Anacardiaceae. In: Kubitzki K ed. *The families and genera of flowering plants*. Berlin: Springer-Verlag. 10: 7–50.

Pfeil BE, Crisp MD. 2005. What to do with *Hibiscus*? A proposed nomenclatural resolution for a large and well known genus of Malvaceae and comments on paraphyly. *Australian Systematic Botany* 18: 49–60.

Potter D. 2003. Molecular phylogenetic studies in Rosaceae. In: Sharma AK, Sharma A eds. *Plant genome: Biodiversity and evolution*. Enfield: NH Scientific Publications. 1 (Part A): 319–351.

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS. 2007a. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266: 5–43.

Potter D, Still SM, Grebenc T, Ballian D, Bozic G, Franjiæ J, Kraigher H. 2007b. Phylogenetic relationships in tribe Spiraeeae (Rosaceae) inferred from nucleotide sequence data. *Plant Systematics and Evolution* 266: 105–118.

Potter D, Gao F, Bortiri PE, Oh SH, Baggett S. 2002. Phylogenetic relationships in Rosaceae inferred from chloroplast *matK* and *trnL-trnF* nucleotide sequence data. *Plant Systematics and Evolution* 231: 77–89.

Price RA, Palmer JD. 1993. Phylogenetic relationships of the Geraniaceae and Geraniales from *rbcL* sequence comparisons. *Annals of the Missouri Botanical Garden* 80: 661–671.

Qiu, YL, Li LB, Wang B, Xue JY, Hendry TA, Li RQ, Brown JW, Liu Y, Hudson YH, Chen ZD. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* 48: 391–425.

Richardson JE, Fay MF, Cronk QCB, Chase MW. 2000. A revision of the tribal classification of Rhamnaceae. *Kew Bulletin* 55: 311–340.

Riser JP II, Cardinal-McTeague WM, Hall JC, Hahn WJ, Sytsma KJ, Roalson EH. 2013. Phylogenetic relationships among the North American cleomoids (Cleomaceae): A test of Iltis's reduction series. *American Journal of Botany* 100: 2102–2111.

Roalson EH, Hall JC, Riser JP II, Cardinal-McTeague WM, Cochrane TS, Sytsma KJ. 2015. A revision of generic boundaries and nomenclature in the North American cleomoid clade (Cleomaceae). *Phytotaxa* 205: 129–144.

Rodman JE, Soltis PS, Soltis DE, Sytsma K, Karol K. 1998. Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. *American Journal of Botany* 85: 997–997.

Rossetto M, Crayn DM, Jackes BR, Porter C. 2007. An updated estimate of intergeneric phylogenetic relationships in the Australian Vitaceae. *Canadian Journal of Botany* 85: 722–730.

Ruhfel BR, Bittrich V, Bove CP, Gustafsson MHG, Philbrick CT, Rutishauser R, Xi Z, Davis CC. 2011. Phylogeny of the Clusioid clade (Malpighiales): Evidence from plastid and mitochondrial genomes. *American Journal of Botany* 98: 306–325.

Ruhfel BR, Gitzendanner MA, Soltis DE, Soltis PS, Burleigh JG. 2014. From algae to angiosperms – inferring the phylogeny of green plants (*Viridiplantae*) from 360 plastid genomes. *BMC Evolutionary Biology* 14: 23.

Samarakoon T. 2015. *Phylogenetic relationships of Samydaceae and taxonomic revision of the species of* Casearia *in south-central Asia*. Ph.D. Dissertation. Hattiesburg: University of Southern Mississippi.

Samuel R, Kathriarachchi H, Hoffmann P, Barfuss MH, Wurdack KJ, Davis CC, Chase MW. 2005. Molecular phylogenetics of Phyllanthaceae: Evidence from plastid *matK* and nuclear PHYC sequences. *American Journal of Botany* 92: 132–141.

Sauquet H, Ho S, Gandolfo M, Jordan G, Wilf P, Cantrill D, Bayly M, Bromham L, Brown G, Carpenter R. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: The case of *Nothofagus* (Fagales). *Systematic Biology* 61: 298–313.

Schenk JJ, Herschlag R, Thomas DW. 2015. Describing a new species into a polyphyletic genus: Taxonomic novelty in *Ledermanniella* s.l. (Podostemaceae) from Cameroon. *Systematic Botany* 40: 539–552.

Schneider JV, Bissiengou P, Amaral Mdo C, Tahir A, Fay MF, Thines M, Sosef MS, Zizka G, Chatrou LW. 2014. Phylogenetics, ancestral state reconstruction, and a new infrafamilial classification of the pantropical Ochnaceae (Medusagynaceae, Ochnaceae s.str., Quiinaceae) based on five DNA regions. *Molecular Phylogenetics and Evolution* 78: 199–214.

Schrire BD, Lavin M, Barker NP, Forest F. 2009. Phylogeny of the tribe Indigofereae (Leguminosae–Papilionoideae): Geographically structured more in succulent-rich and temperate settings than in grass-rich environments. *American Journal of Botany* 96: 816–852.

Schutte AL, van Wyk BE. 1998. The tribal position of *Hypocalyptus* Thunb. (Fabaceae). *Novon* 8: 178–182.

Schwarzbach AE, Ricklefs RE. 2000. Systematic affinities of Rhizophoraceae and Anisophylleaceae, and intergeneric relationships within Rhizophoraceae, based on chloroplast DNA, nuclear ribosomal DNA, and morphology. *American Journal of Botany* 87: 547–564.

Seelanan T, Schnabel A, Wendel JF. 1997. Congruence and consensus in the cotton tribe. *Systematic Botany* 22: 259–290.

Sheahan MC, Chase MW. 2000. Phylogenetic relationships within Zygophyllaceae based on DNA sequences of three plastid regions, with special emphasis on Zygophylloideae. *Systematic Botany* 25: 371–384.

Shi SH, Huang YL, Tan FX, He XJ, Boufford DE. 2000. Phylogenetic analysis of the Sonneratiaceae and its relationship to Lythraceae based on ITS sequences of nrDNA. *Journal of Plant Research* 113: 253–258.

Simmons MP, Bacon CD, Cappa JJ, McKenna MJ. 2012. Phylogeny of Celastraceae subfamilies Cassinoideae and Tripterygioideae inferred from morphological characters and nuclear and plastid loci. *Systematic Botany* 37: 456–467.

Simmons MP, Cappa JJ. 2013. Wilczekra, a new genus of Celastraceae for a disjunct lineage of *Euonymus*. *Systematic Botany* 28: 148–153.

Sinou C, Forest F, Lewis GP, Bruneau A. 2009. The genus *Bauhinia* s.l. (Leguminosae): A phylogeny based on the plastid *trnL-F* region. *Botany* 87: 947–960.

Smith SA, Beaulieu JM, Donoghue MJ. 2009. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology* 9: 37.

Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany* 98: 404–414.

Soltis DE, Gitzendanner MA, Soltis PS. 2007. A 567-taxon data set for angiosperms: The challenges posed by Bayesian analyses of large data sets. *International Journal of Plant Sciences* 168: 137–157.

Soltis DE, Moore MJ, Burleigh JG, Bell CD, Soltis PS. 2010. Assembling the angiosperm tree of life: Progress and future prospects. *Annals of the Missouri Botanical Garden* 97: 514–526.

Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi ZX, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu YL, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donoghue MJ and Soltis PS. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis PS, Soltis DE. 2013. Angiosperm phylogeny: A framework for studies of genome evolution. In: Leitch IJ, Greilhuber J, Dolezel J, Wendel JF eds. *Plant genome diversity*. Vienna: Springer-Verlag Wien. 2: 1–11.

Soltis PS, Soltis DE, Chase MW, Endress PK, Crane PR. 2004. The diversification of flowering plants. In: Cracraft J, Donoghue M eds. *The tree of life*. New York: Oxford University Press. 154–167.

Soltis DE, Soltis PS, Chase MW. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolaninen-Hahn VWH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

Soltis DE, Soltis PS, Endress PK, Chase MW. 2005. *Phylogeny and evolution of angiosperms*. Sunderland: Sinauer Associates.

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences USA* 92: 2647–2651.

Soltis DE, Soltis PS, Nickrent DL, Johnson LA, Hahn WJ, Hoot SB, Sweere JA, Kuzoff RK, Kron KA, Chase MW, Swensen SM. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 1: 1–49.

Sothers CA, Prance GT. 2014. Resurrection of *Angelesia*, a Southeast Asian genus of Chrysobalanaceae. *Blumea* 59: 103–105.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Stevens PF. 2001 onwards. Angiosperm Phylogeny Website. Version 12, July 2012. [online]. Available from http://www.mobot.org/MOBOT/research/APweb/ [accessed 20 January 2015].

Su JX, Wang W, Zhang LB, Chen ZD. 2012. Phylogenetic placement of two enigmatic genera, *Borthwickia* and *Stixis*, based on molecular and pollen data, and the description of a new family of Brassicales, Borthwickiaceae. *Taxon* 61: 601–611.

Sun M, Soltis DE, Soltis PS, Zhu XY, Burleigh JG, Chen ZD. 2015. Deep phylogenetic incongruence in the angiosperm clade *Rosidae*. *Molecular Phylogenetics and Evolution* 83: 156–166.

Sytsma KJ, Morawetz J, Pires JC, Nepokroeff M, Conti E, Zjhra M, Hall JC, Chase MW. 2002. Urticalean rosids: Circumscription, rosid ancestry, and phylogenetics based on *rbcL*, *trnL-F*, and *ndhF* sequences. *American Journal of Botany* 89: 1531–1546.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30: 2725–2729.

Thiv M, Ghogue JP, Grob V, Huber K, Pfeifer E, Rutishauser R. 2009. How to get off the mismatch at the generic rank in African Podostemaceae. *Plant Systematics and Evolution* 283: 57–77.

Thulin M, Razafimandimbison SG, Chafe P, Heidari N, Kool A, Shore JK. 2012. Phylogeny of the Turneraceae clade (Passifloraceae s.l.): Trans-Atlantic disjunctions and two new genera in Africa. *Taxon* 61: 308–323.

Tippery NP, Philbrick CT, Bove CP, Les DH. 2011. Systematics and phylogeny of neotropical river weeds (Podostemaceae: Podostemoideae). *Systematic Botany* 36: 105–118.

Tokuoka T. 2007. Molecular phylogenetic analysis of Euphorbiaceae *sensu stricto* based on plastid and nuclear DNA sequences and ovule and seed character evolution. *Journal of Plant Research* 120: 511–522.

Tokuoka T. 2012. Molecular phylogenetic analysis of Passifloraceae *sensu lato* (Malpighiales) based on plastid and nuclear DNA sequences. *Journal of Plant Research* 125: 489–497.

Trias-Blasi A, Parnell JA, Hodkinson TR. 2012. Multi-gene region phylogenetic analysis of the grape family (Vitaceae). *Systematic Botany* 37: 941–950.

van der Bank M, Fay MF, Chase MW. 2002. Molecular phylogenetics of Thymelaeaceae with particular reference to African and Australian genera. *Taxon* 51: 329–339.

van Velzen R, Bakker FT, Sattarian A, van der M. 2006. Evolutionary relationships of Celtidaceae. In: Sattarian A ed. *Contribution to the biosystematics of* Celtis *L. (Celtidaceae) with special emphasis on the African species*. Ph.D. Dissertation. Wageningen: Wageningen University.

Vorontsova MS, Hoffmann P. 2008. A phylogenetic classification of tribe Porantheae (Phyllanthaceae, Euphorbiaceae *sensu lato*). *Kew Bulletin* 63: 41–59.

Wagner WL, Hoch PC. 2005. Onagraceae, the evening primrose family website. [online]. Available from http://botany.si.edu/onagraceae/index.cfm [accessed 31 July 2015].

Wahlert GA, Marcussen T, Paula-Souza JD, Feng M, Ballard HE. 2014. A phylogeny of the Violaceae (Malpighiales) inferred from plastid DNA sequences: Implications for generic diversity and intrafamilial classification. *Systematic Botany* 39: 239–252.

Wang BH, Zhang Y, Wei PP, Sun M, Ma XF, Zhu XY. 2014. Identification of nuclear low-copy genes and their phylogenetic utility in rosids. *Genome* 57: 547–554.

Wang HC, Sun H, Campton D, Yang JB. 2006. A phylogenetic analysis of Thermopsideae (Leguminosae: Papilionoideae) inferred from nuclear ribosomal ITS sequences. *Botanical Journal of the Linnean Society* 151: 365–373.

Wang HC, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE. 2009. Rosidae radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences USA* 106: 3853–3858.

Weeks A, Zapata F, Pell SK, Daly DC, Mitchell JD, Fine PV. 2014. To move or to evolve: Contrasting patterns of intercontinental connectivity and climatic niche evolution in "Terebinthaceae" (Anacardiaceae and Burseraceae). *Frontiers in Genetics* 5: 409.

Whitlock BA, Hale AM. 2011. The phylogeny of *Ayenia*, *Byttneria*, and *Rayleya* (Malvaceae s.l.) and its implications for the evolution of growth forms. *Systematic Botany* 36: 129–136.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Grahamm SW, Mathew S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, Pokorny L, Shaw AJ, DeGironimo L, Stevenson DW, Surek B, Villarreal JC, Roureu B, Philippe H, dePamphilis CW, Chen T, Deyholos MK, Baucom RS, Kutchan TM, Augustin MM, Wang J, Zhang Y, Tian Z, Yan Z, Wu X, Sun X, Wong GKS, Leebens-Mack J. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences USA* 111: E4859–E4868.

Wiens JJ. 2003. Incomplete taxa, incomplete characters, and phylogenetic accuracy: Is there a missing data problem? *Journal of Vertebrate Paleontology* 23: 297–310.

Wojciechowski MF, Lavin M, Sanderson MJ. 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. *American Journal of Botany* 91: 1846–1862.

Wojciechowski MF, Sanderson M, Steele KP, Liston A. 2000. Molecular phylogeny of the "Temperate Herbaceous Tribes" of Papilionoid legumes: A supertree approach. *Advances in Legume Systematics* 9: 277–298.

Wu ZY, Monro AK, Milne RI, Wang H, Yi TS, Liu J, Li DZ. 2013. Molecular phylogeny of the nettle family (Urticaceae) inferred from multiple loci of three genomes and extensive generic sampling. *Molecular Phylogenetics and Evolution* 69: 814–827.

Wurdack KJ, Davis CC. 2009. Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *American Journal of Botany* 96: 1551–1570.

Wurdack KJ, Hoffmann P, Chase MW. 2005. Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae *sensu stricto*) using plastid *rbcL* and *trnL-F* sequences. *American Journal of Botany* 92: 1397–1420.

Wurdack KJ, Hoffmann P, Samuel R, de-Bruijn A, van-der Bank M, Chase MW. 2004. Molecular phylogenetic analysis of Phyllanthaceae (Phyllanthoideae *pro parte*, Euphorbiaceae *sensu lato*) using plastid rbcL sequences. *American Journal of Botany* 91: 1882–1900.

Xi ZX, Ruhfel BR, Schaefer H, Amorim AM, Sugumaran M, Wurdack KJ, Endress PK, Matthews M, Stevens PF, Mathews S, Davis CC. 2012. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation in Malpighiales. *Proceedings of the National Academy of Sciences USA* 109: 17519–17524.

Xi ZX, Liu L, Rest JS, Davis CC. 2014. Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. *Systematic Biology* 63: 919–932.

Xiang XG, Wang W, Li RQ, Lin L, Liu Y, Zhou ZK, Li ZY, Chen ZD. 2014. Large-scale phylogenetic analyses reveal fagalean diversification promoted by the interplay of diaspores and environments in the Paleogene. *Perspectives in Plant Ecology, Evolution and Systematics* 16: 101–110.

Xu XM, Dunn K, Field C. 2015. A robust anova approach to estimating a phylogeny from multiple genes. *Molecular Biology and Evolution* 32: 2186–2194.

Yang MQ, van-Velzen R, Bakker FT, Sattarian A, Li DZ, Yi TS. 2013. Molecular phylogenetics and character evolution of Cannabaceae. *Taxon* 62: 473–485.

Yang Y, Moore MJ, Brockington SF, Soltis DE, Wong GKS, Carpenter EJ, Zhang Y, Chen L, Yan ZX, Xie YL, Sage RF, Covshoff S, Hibberd JM, Nelson MN, Smith SA. 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution* 32: 2001–2014.

Zanne AE, Tank DC, Cornwell WK, Eastman JM, Smith SA, FitzJohn RG, McGlinn DJ, O'Meara BC, Moles AT, Reich PB, Royer DL, Soltis DE, Stevens PF, Westoby M, Wright IJ, Aarssen L, Robert IB, Calaminus A, Govaerts R, Hemmings F, Leishman MR, Oleksyn J, Soltis PS, Swenson NG, Warman L, Beaulieu JM. 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature* 506: 89–92.

Zeng LP, Zhang Q, Sun RR, Kong HZ, Zhang N, Ma H. 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nature Communications* 5: 4956.

Zhang LB, Simmons MP. 2006. Phylogeny and delimitation of the Celastrales inferred from nuclear and plastid genes. *Systematic Botany* 31: 122–137.

Zhang N, Wen J, Zimmer EA. 2015. Congruent deep relationships in the Grape family (Vitaceae) based on sequences of chloroplast genomes and mitochondrial genes via genome skimming. *PloS ONE* 10: e0144701.

Zhang N, Zeng LP, Shan HY, Ma H. 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytologists* 195: 923–937.

Zhao HL, Lu LM, Su JX, Chen ZD. 2015. Phylogeny of Stixeae and Borthwickiaceae based on morphological and molecular data. *Chinese Bulletin of Botany* 50: 473–481. (in Chinese)

Zhu XY, Chase MW, Qiu YL, Kong HZ, Dilcher DL, Li JH, Chen ZD. 2007. MatR sequences help to resolve deep phylogenetic relationships in rosids. *BMC Evolutionary Biology* 7: 217.

Zunk K, Mummenhoff K, Hurka H. 1999. Phylogenetic relationships in tribe Lepidieae (Brassicaceae) based on chloroplast DNA restriction site variation. *Canadian Journal of Botany* 77: 1504–1512.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12211/suppinfo

Table S1. Accession numbers of sequences retrieved from GenBank (released in December 2012).

Table S2. List of anomalous sequences and rogue taxa that were removed from the dataset prior to analysis.

Fig. S1. Maximum likelihood tree of *Rosidae* based on *matR*. Numbers above branches are bootstrap support (BS) values; BS $<$ 50% are not shown.

Fig. S2. Maximum likelihood tree of *Rosidae* based on *atpB*, *matK*, and *rbcL*. Numbers above branches are bootstrap support (BS) values; BS $<$ 50% are not shown.

Fig. S3. Maximum likelihood tree of *Rosidae* based on analysis of the four-gene (*matR*, *atpB*, *matK*, and *rbcL*) supermatrix. Numbers above branches are bootstrap support (BS) values; BS $<$ 50% are not shown.

Fig. S4. Summary tree showing infrafamilial relationships of Fabaceae. Numbers above branches are bootstrap support (BS) values; BS $<$ 50% are not shown. Most clade names follow Cardoso et al. (2012, 2013) and the Legume Phylogeny Working Group (2013).

Fig. S5. Summary tree showing infrafamilial relationships of Euphorbiaceae. Numbers above branches are bootstrap support (BS) values; BS $<$ 50% are not shown. Most clade names follow Wurdack et al. (2005).

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12201

Research Article

Global versus Chinese perspectives on the phylogeny of the N-fixing clade

Hong-Lei Li[1,2,3], Wei Wang[1], Rui-Qi Li[1], Jing-Bo Zhang[1], Miao Sun[1], Rehan Naeem[1], Jun-Xia Su[4], Xiao-Guo Xiang[1], Peter E. Mortimer[5,6], De-Zhu Li[6,7], Kevin D. Hyde[5,6,8], Jian-Chu Xu[5,6], Douglas E. Soltis[9,10,11], Pamela S. Soltis[10,11], Jianhua Li[12], Shou-Zhou Zhang[2], Hong Wu[3], Zhi-Duan Chen[1]*, and An-Ming Lu[1]*

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[2]Shenzhen Key Laboratory of Southern Subtropical Plant Diversity, Fairylake Botanical Garden, Shenzhen and Chinese Academy of Sciences, Shenzhen 518004, Guangdong, China
[3]College of Life Sciences, South China Agricultural University, Guangzhou 510642, China
[4]College of Life Sciences, Shanxi Normal University, Linfen 041004, Shanxi, China
[5]World Agroforestry Centre, East and Central Asia, Kunming 650201, China
[6]Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
[7]Plant Germplasm and Genomics Center, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
[8]School of Science, Mae Fah Luang University, Chiang Rai 57100, Thailand
[9]Department of Biology, University of Florida, Gainesville, FL 32611, USA
[10]Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
[11]Genetics Institute, University of Florida, Gainesville, FL 32610, USA
[12]Biology Department, Hope College, Holland, MI 49423, USA
*Authors for correspondence. Z.-D. Chen. E-mail: zhiduan@ibcas.ac.cn. Tel.: 86-10-62836434. Fax: 86-10-62836095.
A.-M. Lu. E-mail: anmin@ibcas.ac.cn. Tel.: 86-10-62836448. Fax: 86-10-62836095.
Received 9 November 2015; Accepted 3 March 2016; Article first published online 22 July 2016

Abstract There has been increasing interest in integrating a regional tree of life with community assembly rules in the ecological research. This raises questions regarding the impacts of taxon sampling strategies at the regional versus global scales on the topology. To address this concern, we constructed two trees for the nitrogen-fixing clade: (i) a genus-level global tree including 1023 genera; and (ii) a regional tree comprising 303 genera, with taxon sampling limited to China. We used the supermatrix approach and performed maximum likelihood analyses on combined *matK*, *rbcL*, and *trnL-F* plastid sequences. We found that the topology of the global and the regional tree of the N-fixing clade were generally congruent. However, whereas relationships among the four orders obtained with the global tree agreed with the accepted topology obtained in focused analyses with more genes, the regional topology obtained different relationships, albeit weakly supported. At a finer scale, the phylogenetic position of the family Myricaceae was found to be sensitive to sampling density. We expect that internal support throughout the phylogeny could be improved with denser taxon sampling. The taxon sampling approach (global vs. regional) did not have a major impact on fine-level branching patterns of the N-fixing clade. Thus, a well-resolved phylogeny with relatively dense taxon sampling strategy at the regional scale appears, in this case, to be a good representation of the overall phylogenetic pattern and could be used in ecological research. Otherwise, the regional tree should be adjusted according to the correspondingly reliable global tree.

Key words: global tree of life, N-fixing clade, phylogeny, regional tree of life, supermatrix.

The tree of life has been widely applied as a useful tool in different areas of ecological research (Ma, 2013; Lu et al., 2014). An upsurge in ecological studies incorporating phylogenetic information with community dynamics has been seen in the last decade (Webb et al., 2002; Kraft et al., 2007; Cavender-Bares et al., 2009; Asner & Martin, 2011; Gregory et al., 2014; Robert, 2015). Nitrogen-fixing (N-fixing) plants are an important component of biological communities. The ability to fix and use atmospheric N through a process known as biological N-fixation complements the limited bioavailability of N and has received a lot of attention. In particular, there is hope that N-fixing genes can be transferred to non-N-fixing crop species (Soltis et al., 1995; Ferguson et al., 2010; Santi et al., 2013; Venkateshwaran et al., 2013).

The N-fixing clade was first recovered and proposed by Soltis et al. (1995) based on initial molecular phylogenetic analyses of angiosperms. The clade as now recognized contains 28 families (of which, 10 families contain N-fixing species), over 1300 genera, and approximately 30 000 species. Many species of the N-fixing clade have great economic importance as crop plants including legumes (Fabaceae), fruit crops (Rosaceae), and vegetables (Cucurbitaceae). Many other species play a crucial role in biological communities because of their ability to fix atmospheric N through a symbiosis with N-fixing bacteria in root nodules (Li et al., 2015). Additionally, the clade possesses many woody species that are distributed in temperate and tropical forests (e.g., Betulaceae, Fagaceae, Fabaceae, and Ulmaceae) (Croat, 1978; Elias, 1980; Lopez et al., 1987), forests of extreme importance to biological communities (Wang et al., 2009; Fang et al., 2012; Huang et al., 2012).

The N-fixing clade has long been recognized as monophyletic and inclusive of the four orders Cucurbitales, Fabales, Fagales, and Rosales (Soltis et al., 2000, 2008, 2011; Zhu et al., 2007; APG III, 2009; Wang et al., 2009; Moore et al., 2010, 2011). Relationships among these orders, as well as within these orders, are generally well resolved. For example, the relationships among families in Fagales are well resolved except for the position of Myricaceae (Li et al., 2004; Herbert et al., 2006; Bell et al., 2010; Soltis et al., 2011; Xiang et al., 2014). Within Cucurbitales, the relationships of Begoniaceae, Datiscaceae, and Tetramelaceae remain unclear (Swensen et al., 1994, 1998; Zhang et al., 2006; Schaefer & Renner, 2011). The four families within Fabales were strongly supported as monophyletic, but branching orders of the families have not yet been clarified (Bello et al., 2009; Wang et al., 2009; Soltis et al., 2011), and the main clades of Fabaceae are generally resolved, however, the relationships among some of them are still unclear (Wojciechowski et al., 2004; Cardoso et al., 2012a, 2012b; The Legume Phylogeny Work Group, 2013). The interfamilial relationships within Rosales were resolved with Rosaceae sister to the rest of the order (e.g., Savolainen et al., 2000a, 2000b; Hilu et al., 2003; Soltis et al., 2007, 2011), with the remaining families forming two distinct clades: one clade of Ulmaceae and relatives, and a second comprising Rhamnaceae and relatives (see Richardson et al., 2000; Savolainen et al., 2000a; Sytsma et al., 2002; Zhang et al., 2011). However, phylogenetic relationships among Barbeyaceae, Dirachmaceae, Elaeagnaceae, and Rhamnaceae within Rosales remain unclear. Moreover, within Rosaceae, three subfamilies are monophyletic with strong support, however, the position of Dryadoideae remains uncertain (Potter, 2003; Potter et al., 2007; Chin et al., 2014). Informally, Rosaceae may be composed of clades of rhamnoids, ziziphoids, and ampeloziziphoids; however, the relationships among these are unresolved (Richardson et al., 2000).

The combination of many genes and whole plastid genome sequence data has led to an improved understanding of the deep-level phylogeny of the N-fixing clade within the framework of all angiosperms (Leebens-Mack et al., 2005; Jansen et al., 2007; Moore et al., 2010, 2011; Soltis et al., 2011). However, fewer than 40 taxa of the N-fixing clade were included in the most taxonomically robust of these studies. Some studies recovered the basic phylogenetic framework of the orders within the N-fixing clade using broad taxonomic coverage, but the support values for some crown clades was low (Bello et al., 2009; Zhang et al., 2011). Crown clades may be resolved with denser taxonomic sampling, but denser sampling typically requires a prohibitive volume of sequence data for a many-gene phylogeny. As an alternative, deep-level phylogenies of families or orders may be resolved with a smaller number of markers from spacer regions (Richardson et al., 2000; Li et al., 2004; Zhang et al., 2006). Using spacer regions is a possible way to construct a well-resolved phylogeny with dense sampling.

A well-supported and resolved topology is crucial to carry out reliable, downstream ecological analysis (Cavender-Bares et al., 2009; Roquet et al., 2013). A regional tree of life is useful for studying phylogenetic diversity, community assembly rules, conservation biology, and niche evolution in a distinct area (Whitney et al., 2009; Asner & Martin, 2011; Schaefer et al., 2011). However, the impact of the taxon sampling strategy on the topology at the regional compared to a global scale has not been rigorously studied.

In this study, we selected three plastid regions, *matK*, *rbcL*, and *trnL-F* spacer and reconstructed the most comprehensive phylogeny (global tree) of the N-fixing clade to date, comprising 1023 species at the generic level using the supermatrix approach. We also reconstructed a regional tree for the N-fixing clade with 303 genera native to China. Another two global trees that include the remaining 726 operational taxonomic units (OTUs) and 303 randomly selected OTUs from this remaining group (including 6 outgroups) were also reconstructed. Our aims are to establish a tree of life of the N-fixing clade at the generic level and to test the impact of the taxon sampling density at the regional or global scale by comparing the two topologies.

Material and Methods

Taxon sampling

Through our sampling approach we tried to maximize the taxonomic coverage of each of the previously recognized genera (Stevens, 2001 onwards and references therein) within the N-fixing clade. DNA samples for some species used here were extracted from dried materials in silica gel. Sequences of most species were obtained from GenBank. We constructed a three-marker data matrix for 1023 species. Species names and GenBank accession numbers are presented in Table S1.

For newly generated sequences, we isolated genomic DNA from silica gel-dried materials using a Plant Genomic DNA Kit (Beijing Biomed, Beijing, China) or from herbarium samples following a modified CTAB procedure (Doyle & Doyle, 1987). DNA regions were amplified with polymerase chain reaction (PCR). We carried out PCR amplifications using the primers in Li et al. (2013) and 2× *Taq* PCR MasterMix (Beijing Biomed) in 25-μL reactions with the following thermocycler program: 2 min at 95 °C for denaturation, then 35 cycles of 30 s at 95 °C, 30–60 s at 53–57 °C for annealing, 2 min 30 s at 72 °C for primer extension, and a 10-min incubation at 72 °C following the cycles. The PCR products were purified using a GFX PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA) and then directly sequenced them. Sequencing reactions were carried out using an ABI Prism BigDye Terminator Cycle Sequencing Kit (Applied Biosystems,

Beijing, China). We then processed the sequences using ABI 3730xl DNA Analysis Systems and following the manufacturer's protocols.

For sequences from GenBank, all available nucleotide sequences were selected from the three plastid regions (*matK*, *rbcL*, and *trnL-F*) representing the N-fixing clade. For each taxon, we tried to use the same species and DNA sample across the three plastid markers, but some composite accessions were necessary to represent genera. The longest sequence was selected when multiple sequences were available, and randomly selected one sequence when there were multiples of the same length. Most of the DNA sequences have been used in previously published studies (e.g., Li et al., 2004; Wojciechowski et al., 2004; Zhang et al., 2011).

DNA alignment

The *rbcL* sequences were aligned directly in the program MUSCLE using the default settings at the high accuracy parameter (Edgar, 2004), and the resulting alignment was manually adjusted by eye, using BioEdit version 5.0.9 (Hall, 1999). A two-step strategy was used to align the fast-evolving *matK* and *trnL-F* regions. First, we divided the sequences into clusters according to sequence length and taxonomic unit. Each cluster was aligned in MUSCLE under default high accuracy parameters, and then manually adjusted the alignment. Then we aligned the clusters with the profile–profile alignment algorithm in MUSCLE. Final adjustments were made to the alignments for these two genes using the MUSCLE refinement algorithm and then manually, especially to trim for quality and maximum coverage. The aligned global matrix contains 1023 OTUs. To compare with the global tree of the N-fixing clade, we constructed a regional matrix with 303 OTUs representing the N-fixing clade. In the regional matrix, all the genera have representatives distributed in China. The remaining 726 OTUs and 303 randomly selected OTUs from this remaining group (including 6 outgroups) were also prepared for the maximum likelihood (ML) analyses.

Phylogenetic analyses

The program RAxML version 7.6.6 (Stamatakis, 2006) was used to carry out the initial phylogenetic analysis under the ML criterion for each marker. No significant bootstrap (BS) support for conflicting nodes was evident (taken here as exceeding 70%), so the data from different markers for subsequent analyses were combined. Phylogenetic analyses of the combined dataset of three DNA regions using ML methods were carried out. The ML analysis was performed using RAxML with the following options: three data partitions (*rbcL*, *matK*, and *trnL-F*), GTR + I + Γ nucleotide substitution model, and 1000 non-parametric BS replicates. The gaps were treated as missing data. The program was run on the CIPRES network (Miller et al., 2010).

Results

For most nodes, the three-marker global tree showed higher BS support than the individual marker and regional trees. Thus, only the global tree is described below (Fig. S1; and interconnected subtrees in Fig. S2 for clearer visualization). We examine the regional tree (Figs. S3, S4) in the Discussion section under "Comparison of global and regional trees of the N-fixing clade" (below). The topologies with 726 and 303 OTUs are shown in Figs. S5 and S6.

Based on the combined three-marker dataset, we generated a well-resolved phylogeny of the N-fixing clade. Each of the four orders is strongly supported as monophyletic with BS value >80% (Fig. S1). Fabales are sister to the other three orders (BS = 100%), and Rosales are sister to Fagales and Cucurbitales (BS = 84%). Relationships within the four orders are summarized as follows.

Within Fabales, the monophyly of the four families are well supported (BS ≥ 99%). Within Fabaceae, subfamily Caesalpinioideae are paraphyletic and at the base of the family, whereas subfamilies Mimosoideae and Papilionoideae are well supported as monophyletic. In subfamily Caesalpinioideae, eight monophyletic clades were recovered: Cercideae, Deterieae s.l., Dialiinae s.l., Umtiza clade, Cassia clade, Caesalpinia clade, Tachigali clade, and Peltophorum clade. Within Mimosoideae, resolution of the large, higher-level mimosoid clades (e.g., tribal or generic level) is problematic. In subfamily Papilionoideae, 15 monophyletic clades were recovered: Swartzioid, Dipterygeae clade, Amburana clade, Cladrastis clade, Andira clade, Lecointeoid clade, Vataireoid, Dalbergioid s.l., Genistoid, Baphioid, Mirbelioids, Robinioid clade, inverted repeat-lacking clade (IRLC), Indigofereae, and Millettioid clade (inverted-repeat-lacking clade).

Within Rosales, Rosaceae were resolved as sister to other members of Rosales. The remaining families comprise a well-supported clade (BS = 99%). Within Rosaceae, three subfamilies Spiraeoideae, Dryadoideae, and Rosoideae were retrieved. Informally, we identified three well-supported clades in Rhamnaceae: Ampeloziziphoids (BS = 100%), Rhamnoids (BS = 99%), and Ziziphoids (BS = 100%). Ulmaceae comprise two well supported clades, each with BS = 99%: *Ampelocera* + *Holoptelea* and *Hemiptelea* + (*Zelkova* + *Ulmus*). Within Cannabaceae, *Aphananthe* was well supported as sister to the rest of the family. *Gironniera* + *Lozanella* was sister to the remainder. Within Moraceae, well-supported monophyletic clades of Castilleae (BS = 99%) and Dorstenieae s.l. (Clement & Weiblen, 2009) (BS = 99%) were detected. Two additional well-supported clades were Moreae (minus *Streblus*) (BS = 100%) and Artocarpeae (excluding *Hullettia* and *Parartocarpus*) (BS = 90%). In Urticaceae, four strongly supported clades (clade I–IV, Fig. S2k) were recognized and the relationships among them were well resolved.

In Cucurbitales, strong support (BS = 100%) was found for a clade of Corynocarpaceae + Coriariaceae as sister to a moderately supported (BS = 63%) clade consisting of the remaining Cucurbitales. Cucurbitaceae were well represented at the genus level in the current study. We recovered most tribes, including Fevilleeae, Actinostemmateae, Telfairieae, Bryonieae, Sicyeae, Schizopeponeae, Coniandreae, Cucurbiteae, and Benincaseae, *sensu* Schaefer & Renner (2011) based on the analyses of 14 DNA regions from the three plant genomes.

All families within Fagales had BS = 100%. Nothofagaceae were sister to the remaining Fagales (BS = 100%), followed by Fagaceae, which are sister to the remainder of the Fagales, with strong support (BS = 100%). The rest of Fagales formed

two clades: Casuarinaceae + (Ticodendraceae + Betulaceae) (BS = 100%) were sister to Myricaceae + (Rhoipteleaceae + Juglandaceae) (BS = 59%). In Betulaceae, *Alnus* was the sister to the remainder of Betulaceae (BS = 100%), with subsequent divergence order of *Betula* as sister to two clades: *Corylus* + *Ostryopsis* (BS = 99%) and *Ostrya* + *Carpinus* (BS = 100%). Within the Myricaceae, *Canacomyrica* was resolved as sister to *Myrica* + *Comptonia*. In Juglandaceae, two major clades were recovered: (i) *Alfaroa* + (*Engelhardia* + *Alfaropsis*) with BS = 99%; and (ii) *Annamocarya*, subsequently followed by *Platycarya*, *Cyclocarya*, and *Pterocarya* as sister to *Juglans* + *Carya* with BS = 98%.

Discussion

New interfamilial and intrafamilial relationships

Within the Fabales, defining the relationships among the four families has been particularly problematic in the past (Wojciechowski et al., 2004; Bruneau et al., 2008; Bello et al., 2009, 2012; Bell et al., 2010; Soltis et al., 2011). The topology resolved here is Quillajaceae + Surianaceae as sister to a weakly supported clade of Polygalaceae + Fabaceae. Persson (2001) suggested the relationships Polygalaceae + (Surianaceae + (Quillajaceae + Fabaceae)), but there was little support. In Doyle et al. (2000), Quillajaceae are sister to the other three families. Qiu et al. (2010) supported the relationships of Quillajaceae + (Fabaceae + (Surianaceae + Polygalaceae)) with weak support. In other analyses, the topology Polygalaceae + (Leguminosae + (Quillajaceae + Surianaceae)) is considered as the most likely hypothesis of interfamilial relationships of the order (Wojciechowski et al., 2004; Bruneau et al., 2008; Bello et al., 2009, 2012). Soltis et al. (2011) recovered a topology (Polygalaceae + Quillajaceae) + (Leguminosae + Surianaceae) upon the analyses of 17 genes, however, the support was weak and the taxon sampling in Fabales was low.

In the Rosales, the monophyly of Rhamnaceae has not been resolved by our work. Nevertheless, we identified three well-supported clades in Rhamnaceae: Ampeloziziphoids, Rhamnoids, and Ziziphoids. *Ventilago* was sister to Rhamnoids *sensu* Richardson et al. (2000) with strong support and should be included in Rhamnoids. Within Rosaceae, in agreement with Chin et al. (2014), Spiraeoideae are sister to Dryadoideae + Rosoideae. This result differs from a prior study focused on the family (Potter et al., 2007). However, the sister relationship of Dryadoideae and Rosoideae was supported by the result of the independent gene trees of *rbcL* and *matK* in Potter et al. (2007).

Relationships in Cucurbitales are similar to other recent analyses (e.g., Zhang et al., 2006; Soltis et al., 2007, 2011; Schaefer & Renner, 2011). However, we found that there was strong support (BS = 100%) for a clade of Corynocarpaceae + Coriariaceae as sister to the remaining Cucurbitales. Begoniaceae were resolved as sister to a well-supported (BS = 86%) clade of Datiscaceae + Tetramelaceae. However, Begoniaceae are resolved as sister to Datiscaceae with only moderate support in some analyses (Zhang et al., 2006; Schaefer et al., 2009; Schaefer & Renner, 2011).

Within the Fagales, the position of Myricaceae we present here is in agreement with the results of previous analyses (Li et al., 2004; Bell et al., 2010; Soltis et al., 2011). In contrast, Xiang et al. (2014) shows a close relationship between Myricaceae and clade Casuarinaceae + (Ticodendraceae + Betulaceae). In Betulaceae, *Alnus* is the sister to the remainder of Betulaceae (BS = 100%), followed by *Betula* as sister to two clades: *Corylus* + *Ostryopsis* (BS = 99%) and *Ostrya* + *Carpinus* (BS = 100%). These results agree well with Li et al. (2004). In some prior analyses (e.g., Forest et al., 2005; Grimm & Renner, 2013) *Betula* was resolved as sister to *Alnus*, but the support value was low.

Comparison of global and regional trees of the N-fixing clade

To test whether regional taxon sampling results in a tree with different branching patterns compared to a global tree, we compared the differences in the phylogenetic relationships of the N-fixing clade among our global and Chinese regional trees. The global and regional trees showed congruence in general, but the regional tree showed weaker support for some relationships (Figs. S1–S4). Within the Fabales, Polygalaceae were sister to Fabaceae in the regional tree, as in the global tree, although the support was lower (BS = 39% regional; BS = 45% global). In Surianaceae of the Fabales, our global tree showed strong support for *Recchia* + *Lundellia* as sister to *Suriana* + (*Cadellia* + *Stylobasium*), as described in Crayn et al. (1995), Forest et al. (2007), and Bello et al. (2009). In Polygalaceae, we found four monophyletic tribes with Xanthophylleae sister to the remaining Polygalaceae and Moutabeae sister to Carpolobieae + Polygaleae. These results are in agreement with previous molecular studies, especially Forest et al. (2007) and Bello et al. (2012). We recovered major clades in Fabaceae that were in accordance with previous studies (Doyle et al., 1997; Bruneau et al., 2001, 2008; Wojciechowski et al., 2004; Cardoso et al., 2012a, 2012b; Manzanilla & Bruneau, 2012). Similar topologies were also recovered in the regional tree. However, in a small number of relationships within Fabaceae, the regional tree showed higher support than the global tree. For example, Cassia clade was sister to Caesalpinia clade with BS = 66% in the regional tree, but support was <50% in the global tree. The sister relationship between Millettioid and Indigofereae got higher BS support in the regional tree (BS = 90%) than in the global tree (BS = 77%). In Rosales, both the global and regional trees agree with other analyses in providing strong support (BS = 99% global; BS = 98% regional) for the placement of Rosaceae as sister to other members of Rosales (Wang et al. 2009; Soltis et al. 2011; Zhang et al., 2011). The clade Ulmaceae + (Cannabaceae + (Moraceae + Urticaceae)) was recognized with strong support (BS = 99% global; BS = 100% regional) and the relationships among these four families were well resolved as in Soltis et al. (2011) and Zhang et al. (2011). Within Urticaceae, the global and regional trees resolved clade I as sister to clade IV (BS = 96% global; BS = 70% regional), and clade II as sister to clade III (BS = 98% global; BS = 68% regional). Similarly, the topology of Cucurbitales in the regional tree was comparable to the global tree, but the support value of the clade Cucurbitaceae + Tetramelaceae + Begoniaceae was lower (BS = 63% global; BS < 50% regional). In Fagales, our global tree showed that Nothofagaceae were sister to the remaining Fagales, followed by Fagaceae as sister to the remainder of Fagales. These findings

are congruent with other published phylogenies (Li et al., 2004; Soltis et al., 2007, 2011; Bell et al., 2010; Xiang et al., 2014).

In some cases, a number of deep level relationships were sensitive to taxon sampling, but the support values of these internal nodes were lower than 80%. The relationships among the four orders in the global tree are congruent with those recovered using large numbers of genes, but with a lower density of taxon sampling (Wang et al., 2009; Moore et al., 2010; Soltis et al., 2011). However, in the regional tree the relationships among the four orders were different or poorly resolved. In the regional tree, Cucurbitales are sister to the other three orders. Fabales were resolved as sister to a moderately supported Fagales + Rosales (BS = 65%) with BS < 50%, which is different from our global tree (Fig. S1). The placement of Fabales within the 726 global tree (Fig. S5) and 303 global tree (Fig. S6) are congruent with that in the global tree. Within Rosaceae, subfamily Dryadoideae were sister to Spiraeoideae in the regional tree with BS < 50%, but different in the global tree with Dryadoideae as sister to Rosoideae (BS = 76%). The relationships among the three subfamilies of Rosaceae in both the 726 global tree and 303 global tree are the same as that in the global tree, other than that in the regional tree. This indicates taxon sampling strategy at the regional scale could lead to a different topology in some cases compared with the sampling strategy at a global scale, albeit with BS < 80%. In Rhamnaceae, the sister relationship of Rhamnoids and Ziziphoids was found in our global tree, and in a number of studies (Richardson et al., 2000; Bell et al., 2010; Soltis et al., 2011; Zhang et al., 2011), although our BS support for the relationship was only 55%. The regional tree recovered a similar relationship, however, *Ventilago* of Rhamnoids is sister to Ziziphoids with low support (BS < 50%). In particular, Myricaceae were sister to Casuarinaceae + Betulaceae with BS = 72% in the regional tree, rather than sister to Juglandaceae in the global tree with BS = 59%. The position of Myricaceae in the 726 global tree agrees with that in the global tree, whereas the position of Myricaceae in the 303 global tree is the same as that in the regional tree. This indicates that the different placement of Myricaceae is caused by the density of taxa sampling other than the regional scale sampling strategy.

Conclusions and Perspectives

The N-fixing clade, *sensu* APG III, contains 28 families (of which 10 are N-fixing), over 1300 genera, and approximately 30 000 species. We present the most comprehensive genus-level phylogenetic hypothesis to date for the N-fixing clade, developed after analysis of three plastid loci, *matK*, *rbcL*, and *trnL-F*. Furthermore, we tested the impacts of taxon sampling strategy at the regional or global scale on the topology by comparing the global and regional trees.

Based on the combined three-marker dataset, we generated a well-resolved phylogeny of the N-fixing clade composed of four plant orders. Each of the four orders was strongly supported as monophyletic (Fig. S1). The deep and crown clades of the global tree recovered in our analyses are largely congruent with those in previous studies, highlighting the utility of spacer regions with sufficient taxon coverage for phylogenetic resolution. Generally, no strong conflicts (BS > 80%) are found among the major clades of global and regional trees of life. Internal support throughout the phylogeny could be improved with denser taxon sampling. A well-resolved phylogeny with relatively dense taxon sampling strategy at the regional scale does not have a negative impact on deep-level branching patterns of the N-fixing clade. Thus, a well-resolved phylogeny (internal nodes with BS > 80%) with a taxon sampling strategy at the regional scale could be used in ecological research. Otherwise, the regional tree should be adjusted according to the correspondingly reliable global tree before being used in ecological research.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant Nos. 31590822, 31500175, 31270268 and 40830209), the National Basic Research Program of China (Grant No. 2014CB954100), Chinese Academy of Sciences International Institution Development Program (No. SAJC201315), Chinese Academy of Sciences External Cooperation Program of the Bureau of International Co-operation (No. GJHZ201321), Chinese Academy of Sciences Visiting Professorship to Douglas E. Soltis as a senior international scientist (Grant No. 2011T1S24), and Shenzhen Key Laboratory of Southern Subtropical Plant Diversity (Grant No. SSTLAB-2014-04).

References

APG III (Angiosperm Phylogeny Group III). 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161: 105–121.

Asner GP, Martin RE. 2011. Canopy phylogenetic, chemical and spectra1 assembly in a lowland Amazonian forest. *New Phytologist* 189: 999–1012.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* 97: 1296–1303.

Bello MA, Bruneau A, Forest F, Hawkins JA. 2009. Elusive relationships within order Fabales: Phylogenetic analyses using *matK* and *rbcL* sequence data. *Systematic Botany* 34: 102–114.

Bello MA, Rudall PJ, Hawkins JA. 2012. Combined phylogenetic analyses reveal interfamilial relationships and patterns of floral evolution in the eudicot order Fabales. *Cladistics* 28: 393–421.

Bruneau A, Forest F, Herendeen PS, Klitgaard BB, Lewis GP. 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Systematic Botany* 26: 487–514.

Bruneau A, Mercure M, Lewis GP, Herendeen PS. 2008. Phylogenetic patterns and diversification in the caesalpinioid legumes. *Botany* 86: 697–718.

Cardoso D, de Lima HC, Rodrigues RS, de Queiroz LP, Pennington RT, 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, de Queiroz LP, Pennington RT, de Lima HC, Fonty É, Wojciechowski MF, Lavin M. 2012b. Revisiting the phylogeny of papilionoid legumes: New insights from comprehensively

sampled early-branching lineages. *American Journal of Botany* 99: 1991–2013.

Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW. 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters* 12: 693–715.

Chin S-W, Shaw J, Haberle R, Wen J, Potter D. 2014. Diversification of almonds, peaches, plums and cherries – Molecular systematics and biogeographic history of *Prunus* (Rosaceae). *Molecular Phylogenetics and Evolution* 76: 34–48.

Clement WL, Weiblen GD. 2009. Morphological evolution in the mulberry family (Moraceae). *Systematic Botany* 34: 530–552.

Crayn D, Fernando E, Gadek P, Quinn C. 1995. A reassessment of the familial affinity of the Mexican genus *Recchia* Moçiño & Sessé ex DC. *Brittonia* 47: 397–402.

Croat TB. 1978. *Flora of Barro Colorado Island*. Palo Alto: Stanford University Press.

Doyle JJ, Chappill JA, Bailey DC, Kajita T. 2000. Towards a comprehensive phylogeny of legumes: Evidence from *rbcL* sequences and non-molecular data. In: Herendeen PS, Bruneau A eds. *Advances in legume systematics*. Richmond: Royal Botanic Gardens, Kew. 1–20.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Doyle JJ, Doyle JL, Ballenger JA, Dickson EE, Kajita T, Ohashi H. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: Taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541–554.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Elias T. 1980. *The complete trees of North America: Field guide and natural history*. New York: van Nostrand Reinhold.

Fang J, Shen Z, Tang Z, Wang X, Wang Z, Feng J, Liu Y, Qiao X, Wu X, Zheng C. 2012. Forest community survey and the structural characteristics of forests in China. *Ecography* 35: 1059–1071.

Ferguson BJ, Indrasumunar A, Hayashi S, Lin M-H, Lin Y-H, Reid DE, Gresshoff PM. 2010. Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology* 52: 61–76.

Forest F, Chase MW, Persson C, Crane PR, Hawkins JA. 2007. The role of biotic and abiotic factors in evolution of ant dispersal in the milkwort family (Polygalaceae). *Evolution* 61: 1675–1694.

Forest F, Savolainen V, Chase MW, Lupia R, Bruneau A, Crane PR. 2005. Teasing apart molecular- versus fossil-based error estimates when dating phylogenetic trees: A case study in the birch family (Betulaceae). *Systematic Botany* 30: 118–133.

Gregory PA, Roberta EM, Raul T, Christopher BA, Felipe S, Loreli C-J, Paola M. 2014. Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences USA* 111: 5604–5609.

Grimm GW, Renner SS. 2013. Harvesting Betulaceae sequences from GenBank to generate a new chronogram for the family. *Botanical Journal of the Linnean Society* 172: 465–477.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Herbert J, Chase MW, Möller M, Abbott RJ. 2006. Nuclear and plastid DNA sequences confirm the placement of the enigmatic *Canacomyrica monticola* in Myricaceae. *Taxon* 55: 349–357.

Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R. 2003. Angiosperm phylogeny based on *matK* sequence information. *American Journal of Botany* 90: 1758–1766.

Huang J, Chen B, Liu C, Lai J, Zhang J, Ma K. 2012. Identifying hotspots of endemic woody seed plant diversity in China. *Diversity and Distributions* 18: 673–688.

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences USA* 104: 19369–19374.

Kraft NJB, Cornwell WK, Webb CO, Ackerly DD. 2007. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist* 170: 271–283.

Leebens-Mack J, Raubeson LA, Cui L, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, de Pamphilis CW. 2005. Identifying the basal angiosperm node in chloroplast genome phylogenies: Sampling one's way out of the Felsenstein zone. *Molecular Biology and Evolution* 22: 1948–1963.

Li H-L, Wang W, Lin L, Zhu XY, Li JH, Zhu XY, Chen ZD. 2013. Diversification of the phaseoloid legumes: Effects of climate change, range expansion and habit shift. *Frontiers in Plant Science* 4: 386.

Li H-L, Wang W, Mortimer PE, Li R-Q, Li D-Z, Hyde KD, Xu J-C, Soltis D-E, Chen Z-D. 2015. Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Scientific Reports* 5: 14023.

Li RQ, Chen ZD, Lu AM, Soltis DE, Soltis PS, Manos PS. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *International Journal of Plant Sciences* 165: 311–324.

Lopez AJ, Little EL, Ritz GF, Rombold JS, Hahn WJ. 1987. *Arboles comunes del Paraguay: Ñande yvyra mata kuera*. Washington: Cuerpo de Paz. (in Spanish)

Lu LM, Sun M, Zhang JB, Li HL, Lin L, Yang T, Chen M, Chen ZD. 2014. Tree of life and its applications. *Biodiversity Science* 22: 3–20.

Ma KP. 2013. A mini review on the advancement of biodiversity research in China in 2012. *Biodiversity Science* 21: 1–2.

Manzanilla V, Bruneau A. 2012. Phylogeny reconstruction in the Caesalpinieae grade (Leguminosae) based on duplicated copies of the sucrose synthase gene and plastid markers. *Molecular Phylogenetics and Evolution* 65: 149–162.

Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T. 2010. The CIPRES Portals [online]. Available from www.phylo.org [accessed 5 January 2015].

Moore MJ, Hassan N, Gitzendanner MA, Bruenn RA, Croley M, Vandeventer A, Horn JW, Dhingra A, Brockington SF, Latvis M, Ramdial J, Alexandre R, Piedrahita A, Xi Z, Davis CC, Soltis PS, Soltis DE. 2011. Phylogenetic analysis of the plastid inverted repeat for 244 species: Insights into deeper-level angiosperm relationships from a long, slowly evolving sequence region. *International Journal of Plant Sciences* 172: 541–558.

Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences USA* 107: 4623–4628.

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

Potter D. 2003. Molecular phylogenetic studies in Rosaceae. In: Sharma AK, Sharma A eds. *Plant genome: Biodiversity and evolution, Vol. 1, Part A: Phanerogams*. Enfield: Science Publishers. 319–351.

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS. 2007. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266: 5–43.

Qiu YL, Li L, Wang B, Xue JY, Hendry TA, Li RQ, Brown JW, Liu Y, Hudson GT, Chen ZD. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *Journal of Systematics and Evolution* 48: 391–425.

Richardson JE, Fay MF, Cronk QCB, Bowman D, Chase MW. 2000. A phylogenetic analysis of Rhamnaceae using *rbcL* and *trnL-F* plastid DNA sequences. *American Journal of Botany* 87: 1309–1324.

Robert ER. 2015. Intrinsic dynamics of the regional community. *Ecology Letters* 18: 497–503.

Roquet C, Thuiller W, Lavergne S. 2013. Building megaphylogenies for macroecology: Taking up the challenge. *Ecography* 36: 13–26.

Santi C, Bogusz D, Franche C. 2013. Biological nitrogen fixation in non-legume plants. *Annals of Botany* 111: 743–767.

Savolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, Fay MF, De Bruijn AY, Sullivan S, Qiu YL. 2000a. Phylogenetics of flowering plants based on combined analysis of plastid *atpB* and *rbcL* gene sequences. *Systematic Biology* 49: 306–362.

Savolainen V, Fay MF, Albach DC, Backlund A, van der Bank M, Cameron KM, Johnson S, Lledó MD, Pintaud JC, Powell M. 2000b. Phylogeny of the eudicots: A nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bulletin* 55: 257–309.

Schaefer H, Hardy OJ, Silva L, Barraclough TG, Savolainen V. 2011. Testing Darwin's naturalization hypothesis in the Azores. *Ecology Letters* 14: 389–396.

Schaefer H, Heibl C, Renner SS. 2009. Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proceedings of the Royal Society B: Biological Sciences* 276: 843–851.

Schaefer H, Renner SS. 2011. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). *Taxon* 60: 122–138.

Soltis DE, Bell CD, Kim S, Soltis PS. 2008. Origin and early evolution of Angiosperms. *Annals of the New York Academy of Sciences* 1133: 3–25.

Soltis DE, Gitzendanner MA, Soltis PS. 2007. A 567-taxon data set for angiosperms: The challenges posed by Bayesian analyses of large data sets. *International Journal of Plant Sciences* 168: 137–157.

Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu YL, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donoghue MJ, Soltis PS. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences USA* 92: 2647–2651.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stevens PF. 2001 onwards. Angiosperm Phylogeny Website. Version 12, July 2012 (and more or less continuously updated since) [online]. Available from www.mobot.org/MOBOT/research/APweb [accessed 5 January 2015].

Swensen SM, Luthi JN, Rieseberg LH. 1998. Datiscaceae revisited: Monophyly and the sequence of breeding system evolution. *Systematic Botany* 23: 157–169.

Swensen SM, Mullin BC, Chase MW. 1994. Phylogenetic affinities of Datiscaceae based on an analysis of nucleotide sequences from the plastid *rbcL* gene. *Systematic Botany* 19: 157–168.

Sytsma KJ, Morawetz J, Pires JC, Nepokroeff M, Conti E, Zjhra M, Hall JC, Chase MW. 2002. Urticalean rosids: Circumscription, rosid ancestry, and phylogenetics based on *rbcL*, *trnL–trnF*, and *ndhF* sequences. *American Journal of Botany* 89: 1531–1546.

The Legume Phylogeny Work Group. 2013. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon* 62: 217–248.

Venkateshwaran M, Volkening JD, Sussman MR, Ané J-M. 2013. Symbiosis and the social network of higher plants. *Current Opinion in Plant Biology* 16: 118–127.

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences USA* 106: 3853–3858.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

Whitney KD, Ahern JR, Campbell LG. 2009. Hybridization-prone plant families do not generate more invasive species. *Biological Invasions* 11: 1205–1215.

Wojciechowski MF, Lavin M, Sanderson MJ. 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. *American Journal of Botany* 91: 1846–1862.

Xiang X-G, Wang W, Li R-Q, Lin L, Liu Y, Zhou Z-K, Li ZY, Chen ZD. 2014. Large-scale phylogenetic analyses reveal fagalean diversification promoted by the interplay of diaspores and environments in the Paleogene. *Perspectives in Plant Ecology, Evolution and Systematics* 16: 101–110.

Zhang LB, Simmons MP, Kocyan A, Renner SS. 2006. Phylogeny of the Cucurbitales based on DNA sequences of nine loci from three genomes: Implications for morphological and sexual system evolution. *Molecular Phylogenetics and Evolution* 39: 305–322.

Zhang SD, Soltis DE, Yang Y, Li DZ, Yi TS. 2011. Multi-gene analysis provides a well-supported phylogeny of Rosales. *Molecular Phylogenetics and Evolution* 60: 21–28.

Zhu XY, Chase M, Qiu YL, Kong HZ, Dilcher D, Li JH, Chen ZD. 2007. Mitochondrial *matR* sequences help to resolve deep phylogenetic relationships in rosids. *BMC Evolutionary Biology* 7: 217.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12201/suppinfo:

Fig. S1. Summary tree resulting from maximum likelihood analysis of three genes (5564 bp; *matK*, *rbcL*, and *trnL-F*) for 1023 genera in the N-fixing clade, with tips representing families and major clades based on APG III (2009) and references therein.

Fig. S2. Majority-rule consensus of maximum likelihood trees containing 1023 genera in the N-fixing clade.
Fig. S3. Summary tree of the 303-species tree from maximum likelihood analysis of the N-fixing clade, with tips representing families and major clades based on APG III (2009) and references therein.
Fig. S4. Large-scale maximum likelihood majority-rule consensus containing 303 species of the N-fixing clade.
Fig. S5. Maximum likelihood majority-rule consensus containing 726 species that represent genera of the N-fixing clade with distributions outside China.
Fig. S6. Maximum likelihood majority-rule consensus containing 303 species randomly selected from the 726-species matrix of the N-fixing clade.
Table S1. Taxa used in this study of the N-fixing clade with GenBank accession numbers.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12192

Research Article

A supermatrix approach provides a comprehensive genus-level phylogeny for Gentianales

Lei-Lei Yang[1,2†], Hong-Lei Li[1,2†], Lei Wei[4], Tuo Yang[3], Dai-Yong Kuang[5], Ming-Hong Li[5], Yi-Ying Liao[2], Zhi-Duan Chen[3*], Hong Wu[1*], and Shou-Zhou Zhang[2*]

[1]State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China

[2]Shenzhen Key Laboratory of Southern Subtropical Plant Diversity, Fairylake Botanical Garden, Shenzhen and Chinese Academy of Sciences, Shenzhen 518004, Guangdong, China

[3]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

[4]Marine Environmental Monitor Center of Shanwei, State Oceanic Administration, Shanwei 516600, Guangdong, China

[5]Hengshan National Nature Reserve Administration, Hengshan 421900, Hunan, China

[†]These authors contributed equally to this work.

*Author for correspondence. Z.D. Chen. E-mail: zhiduan@ibcas.ac.cn. Tel.: 86-10-62836434. Fax: 86-10-62590843. H. Wu. E-mail: wh@scau.edu.cn. Tel./Fax: 86-20-85285280. S.Z. Zhang. E-mail: shouzhouz@126.com. Tel.: 86-755-25732477. Fax: 86-755-25734474.

Received 9 November 2015; Accepted 18 December 2015; Article first published online 23 February 2016

Abstract Gentianales consist of Apocynaceae, Gelsemiaceae, Gentianaceae, Loganiaceae, and Rubiaceae, of which the majority are woody plants in tropical and subtropical areas. Despite extensive efforts in reconstructing the phylogeny of Gentianales based on molecular data, some interfamily and intrafamily relationships remain uncertain. We reconstructed the genus-level phylogeny of Gentianales based on the supermatrix of eight plastid markers (*rbcL, matK, atpB, ndhF, rpl16, rps16*, the *trnL-trnF* region, and *atpB-rbcL* spacer) and one mitochondrial gene (*matR*) using maximum likelihood. The major clades and their relationships retrieved in the present study concur with those of previous studies. All of the five families of Gentianales are monophyletic with strong support. We resolved Rubiaceae as sister to the remaining families in Gentianales and showed support for the sister relationship between Loganiaceae and Apocynaceae. Our results provide new insights into relationships among intrafamilial clades. For example, within Rubiaceae we found that Craterispermeae were sister to Morindeae + (Palicoureeae + Psychotrieae) and that Theligoneae were sister to Putorieae. Within Gentianaceae, our phylogeny revealed that Gentianeae were sister to Helieae and Potalieae, and subtribe Lisianthiinae were sister to Potaliinae and Faroinae. Within Loganiaceae, we found *Neuburgia* as sister to Spigelieae. Within Apocynaceae, our results supported Amsonieae as sister to Melodineae, and Hunterieae as sister to a clade comprising Plumerieae + (Carisseae + APSA). We also confirmed the monophyly of Perplocoideae and the relationships among Baisseeae + (Secamonoideae + Asclepiadoideae).

Key words: Gentianales, maximum likelihood, phylogeny, supermatrix approach.

The order Gentianales (*sensu* APG III, 2009) includes Apocynaceae, Gelsemiaceae, Gentianaceae, Loganiaceae, and Rubiaceae and contains approximately 1200 genera and 20 000 species. Species in Gentianales are united by their opposite, entire leaves joined by a line across the stem (Bremer & Struwe, 1992; Nicholas & Baijnath, 1994; Struwe et al., 1994), regular and pentamerous flowers, and nuclear endosperm formation (Jensen, 1992). Gentianales have a nearly worldwide distribution, mostly in tropical and subtropical regions. Gentianales contain many ornamentals and economically important plants, such as oleander (*Nerium*), periwinkle (*Vinca*), gentian (*Gentiana*), and coffee (*Coffea*) (Simpson, 2010).

The order Gentianales has been recognized in many traditional classification systems (e.g., Bartling, 1830; Lindley, 1833; Wagenitz, 1959; Takhtajan, 1980, 1997; Cronquist, 1981, 1983; Thorne, 1992). The delimitation of the Gentianales was further improved by molecular phylogenetics over the past decade, which supported Gentianales including five families: Apocynaceae, Gelsemiaceae, Gentianaceae, Loganiaceae, and Rubiaceae (Downie & Palmer, 1992; Chase et al., 1993; Bremer et al., 1994; Endress et al., 1996; Sennblad & Bremer, 1996; Sennblad, 1997; Backlund et al., 2000; Olmstead et al., 2000; Oxelman & Bremer, 2000; Soltis et al., 2000, 2011; Bremer et al., 2002; Frasier, 2008; APG III, 2009; Refulio-Rodriguez & Olmstead, 2014). According to recent molecular phylogenetic studies, Rubiaceae represent the first diverged lineage of Gentianales, but the relationships among the other four families were uncertain. For instance, Backlund et al. (2000) placed Gentianaceae as sister to Apocynaceae, Gelsemiaceae,

and Loganiaceae based on analyses of 62 taxa with *rbcL* and *ndhF* genes, however, the relationships among the latter three families were poorly resolved. In Soltis et al. (2000), Apocynaceae were sister to Gelsemiaceae and Gentianaceae were sister to Loganiaceae based on three gene analysis. Frasier (2008) placed Loganiaceae as sister to Gelsemiaceae + (Apocynaceae + Gentianaceae) by analyzing four plastid DNA regions, which was also supported by Soltis et al. (2011) based on the analyses of 10 DNA regions. Loganiaceae were resolved as sister to Gelsemiaceae, and Apocynaceae were sister to Gentianaceae by Refulio-Rodriguez & Olmstead (2014).

Intrafamilial relationships of Rubiaceae have been the subject of several studies. Within the family, three monophyletic subfamilies referred to Rubioideae, Ixoroideae, and Cinchonoideae were recognized based on molecular phylogenetic analyses (Bremer, 1996a; Bremer et al., 1999; Andersson & Antonelli, 2005; Bremer & Eriksson, 2009; Rydin et al., 2009; Manns et al., 2012). However, the relationships among these subfamilies and other lineages, Coptosapelteae (*Coptosapelta* and *Acranthera*) and Luculieae (*Luculia*), continue to be disputed. Some studies have shown an unresolved position for Coptosapelteae and Luculieae (Bremer et al., 1999; Andersson & Antonelli, 2005; Bremer & Eriksson, 2009). In contrast, Bremer (1996a) placed Luculieae as sister to the remaining Rubiaceae based on *rbcL* data, whereas Bremer (1996b) placed Luculieae in the subfamily Cinchonoideae s.s. according to the analyses of a combined dataset of morphological and chloroplast DNA characters. The sister relationship between Coptosapelteae and Luculieae was recovered by Rydin et al. (2009), however, the placement of this clade within Rubiaceae was uncertain. In Manns et al. (2012), Luculieae were resolved as sister to Coptosapelteae and Rubioideae. Robbrecht & Manen (2006) placed Coptosapelteae as the sister to the rest of the Rubiaceae based on analyses of 300 genera using *rbcL*, *rps16*, *trnL-trnF*, and *atpB-rbcL* regions.

In Gentianaceae, molecular phylogenetic studies support six interfamilial clades: Saccifolieae, Exaceae, Chironieae, Gentianeae, Helieae, and Potalieae (Struwe et al., 2002; Yuan et al., 2005; Merckx et al., 2013; Rybczynski et al., 2014). Potalieae were sister to Gentianeae and Helieae in some analyses (Struwe et al., 2009; Merckx et al., 2013). However, the relationships among these three tribes were not resolved (Struwe et al., 2002; Yuan et al., 2005; Rybczynski et al., 2014). Within Potalieae, Potaliinae were sister to Lisianthiinae and Faroinae based on analyses of the secondary structure of the internal transcribed spacer (ITS) (Molina & Struwe, 2009). Nevertheless, Lisianthiinae were resolved as sister to Potaliinae and Faroinae based on the analysis of combined micromorphology and molecular data (Struwe et al., 2009).

Intrafamilial phylogenetic relationships within Apocynaceae and Loganiaceae have also been the subject of several studies. In Apocynaceae, a total of five subfamilies (Asclepiadoideae, Apocynoideae, Rauvolfioideae, Periplocoideae, and Secamonoideae) and 25 tribes were recognized according to the latest suprageneric classification with morphological and molecular evidence (Endress et al., 2014). Both Rauvolfioideae and Apocynoideae are paraphyletic, and the relationships among tribes within Apocynaceae remain largely unclear (Sennblad & Bremer, 1996; Sennblad et al., 1998; Livshultz et al., 2007; Simões et al., 2007; Lens et al., 2008; Livshultz, 2010). The placement of Asclepiadoideae, Periplocoideae, and Secamonoideae was conflicted among some studies (Livshultz et al., 2007; Simões et al., 2007; Lens et al., 2008; Straub et al., 2014). Within Rauvolfioideae, the relationships among Alyxieae, Amsonieae, Melodineae, and Hunterieae were still unclear (Simões et al., 2007; Livshultz, 2010). Within Apocynoideae, the placement of Nerieae was uncertain (Livshultz et al., 2007; Simões et al., 2007; Straub et al., 2014). Within Loganiaceae, the monophyly of Strychneae and the position of Loganieae were controversial in previous studies (Backlund et al., 2000; Popovkin et al., 2011).

Families of Gentianales show striking heterogeneity in species number. The species-rich family Rubiaceae have 13 150 species (611 genera); Apocynaceae 4555 (415), Gentianaceae 1675 (88), and Loganiaceae 420 (13). Gelsemiaceae only contain 11 species (2 genera). Of the ca. 1200 genera in Gentianales, less than 50 have been sampled in previous studies (Bremer, 1996a; Backlund et al., 2000; Olmstead et al., 2000; Soltis et al., 2000, 2011; Bremer et al., 2001; Backlund, 2005; Jiao & Li, 2007; Frasier, 2008; Refulio-Rodriguez & Olmstead, 2014). Thus, dense sampling at generic level is necessary to estimate the relationships among the families of Gentianales.

In the present study, we used the chloroplast genes *rbcL, matK, atpB, ndhF, rpl16*, and *rps16*, the *trnL-trnF* spacer, the *atpB-rbcL* spacer, and one mitochondrial gene (*matR*), to reconstruct the phylogeny of Gentianales. The objectives of this study were to resolve the interfamilial relationships within the order and intrafamilial relationships within each family.

Material and Methods

Taxon sampling

In total, we sampled 649 accessions including 221 genera (53.3%) of Apocynaceae, two genera (100%) of Gelsemiaceae, 57 genera (64.8%) of Gentianaceae, 11 genera (84.6%) of Loganiaceae, and 358 genera (58.6%) of Rubiaceae in this study. Our sampling represented all the families and most of the tribes within Gentianales *sensu* Backlund (2005). To reduce the numbers of missing data, we created composite samples using different accessions of the same species or genus (e.g., *Calycophyllum*, *Cephalanthus*) (Kim et al., 2004; Wang et al., 2009). We selected *Plocosperma buxifolium* Benth., *Syringa vulgaris* L., *Buddleja yunnanensis* L. F. Gagnep., and *Peltanthera floribunda* Benth. as outgroups according to Angiosperm Phylogeny Group (APG III).

We collected sequence data from three chloroplast genes, *rbcL*, *matK*, and *atpB*, and one mitochondrial gene (*matR*), for all families. We generated 394 new sequences for this study and collected others from GenBank. Of these newly sequenced sequences, *matKs* of 24 genera, *matRs* of 45 genera, *atpBs* of 99 genera, and *rbcLs* of 12 genera were included in molecular analyses for the first time. To increase the informative sites, we collected *rpl16*, *rps16*, and *trnL-trnF* for Apocynaceae, and *ndhF*, *rps16*, and *atpB-rbcL* for Rubiaceae from GenBank according to previous studies (Robbrecht & Manen, 2006; Livshultz et al., 2007; Karehed et al., 2008; Bremer & Eriksson, 2009). All voucher information and GenBank accession numbers are presented in Table S1.

Molecular methods

We extracted total genomic DNAs from fresh or silica gel-dried leaves (Chase & Hills, 1991) according to the methods of Doyle & Doyle (1987) or using a Plant Genomic DNA Kit (Biomed, Beijing, China). We used the standard polymerase chain reaction to amplify target regions.

We carried out polymerase chain reaction in a 50-μL volume containing 40 μL DNA, 5 μL 10 × Taq buffer, 5 μL dNTP, and 2 U Taq. The cycling program for all primers consisted of initial denaturation 4 min at 94 °C followed by 35 cycles of amplification at 94 °C for 30 s, 50–55 °C for 30–60 s, and 72 °C for 1 min, and ended by a final extension at 72 °C for 10 min. We carried out the sequencing reactions using an ABI Prism BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Beijing, China). Following the manufacturer's protocols, sequences were analyzed using ABI 3730xl DNA Analysis Systems.

Alignment and phylogenetic analyses

We aligned sequences using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/), and then manually adjusted them in BioEdit version 5.0.9 (Hall, 1999). We treated gaps as missing data.

We inferred the phylogenetic relationships within Gentianales using maximum likelihood (Felsenstein, 1973) in the program RAxML version 7.6.6 (Stamatakis, 2006) implemented on the CIPRES web cluster (Miller et al., 2010). We applied the GTR + I + Γ substitution model to each gene independently. We partitioned all genes within concatenated data, and applied unlinked substitution models.

Results

Dataset

Our dataset comprised 653 genera, and the final alignment included 11 951 characters from eight chloroplast regions, including *rbcL*, 1280 characters and 523 genera (19.9% missing data); *matK*, 1674 characters and 409 genera (37.4% missing data); *atpB*, 1404 characters and 131 genera (79.9% missing data); *ndhF*, 2081 characters and 225 genera (65.5% missing data); *rpl16*, 1045 characters and 118 genera (81.9% missing data); *rps16*, 1052 characters and 380 genera (41.8% missing data); *trnL-trnF*, 1052 characters and 64 genera (90.2% missing data); *atpB-rbcL*, 639 characters and 185 genera (71.7% missing data), and *matR* with 1724 characters and 69 genera (89.4% missing data).

Phylogenetic analyses

The topologies based on individual DNA data were largely congruent except in some of the terminal branches. The phylogeny based on the combined data gave higher bootstrap (BS) support than those based on individual markers. Hence, we present only the results from combined DNA data analyses below. The best tree from RAxML analyses of the 653 genera (Fig. 1, summary tree) was divided into separate, interconnected subtrees (Fig. 2). Gentianales were divided into five major lineages that are consistent with the five families: Rubiaceae (BS = 100%), Gentianaceae (BS = 100%), Gelsemiaceae (BS = 100%), Loganiaceae (BS = 99%), and Apocynaceae (BS = 100%). Rubiaceae were sister to the rest of the order with strong support (BS = 100%). Loganiaceae were resolved as sister to Apocynaceae with BS = 56%. The relationships among Gelsemiaceae, Gentianaceae, and Apocynaceae + Loganiaceae were not well resolved.

Within Rubiaceae, we applied tribes from Robbrecht & Manen (2006), and found that Coptosapelteae were sister to all other taxa of the family. The remaining Rubiaceae were split into four major clades: Ixoroideae, Cinchonoideae, Luculieae, and Rubioideae. Within Cinchonoideae (Fig. 2a), four monophyletic clades were strongly supported: (i) Hymenodictyeae were sister to Naucleeae (BS = 100%); (ii) Guettardeae were sister to Rondeletieae (BS = 100%); (iii) Isertieae were sister to Cin. C (subtribe Cinchoninae of Cinchoneae) (BS = 100%); and (iv) a subclade containing paraphyletic Hamelieae with respect to Hillieae were sister to the CCE complex (Catesbaeeae/Chiococceae alliance). Within Rubioideae (Figs. 2b, 2c), Ophiorrhizeae (BS = 100%), Urophylleae (BS = 68%), Lasiantheae (BS = 100%), and Coussareeae (BS = 100%) were monophyletic. The remaining Rubioideae were split into two major subclades: (i) a Psychotriidinae alliance *sensu* Razafimandimbison et al. (2008); and (ii) a Spermacoceae alliance *sensu* Bremer & Manen (2000). Within the Psychotriidinae alliance, Schizocolea was the first diverged lineage. Gaertnereae were paraphyletic with respect to Schradereae. Craterispermeae, followed by Morindeae (BS = 54%) were subsequently sister to Palicoureeae (BS = 100%) and Psychotrieae (BS = 100%). Within the Spermacoceae alliance, nine monophyletic tribes were recovered: Danaideae (BS = 94%), Anthospermeae (BS = 63%), Argostemmateae (BS = 100%), Paederieae (BS = 98%), Putorieae (BS = 100%), Theligoneae, Rubieae (BS = 100%), Knoxieae (BS = 100%), and Spermacoceae (BS = 100%). Paederieae, followed by Rubieae, were subsequently sister to Theligoneae and Putorieae. Within Ixoroideae (Figs. 2d, 2e), eight monophyletic tribes were recognized: Posoquerieae, Sipaneeae (BS = 100%), Condamineeae (BS = 99%), Sabiceeae (BS = 100%), Ixoreae (BS = 85%), Vanguerieae (BS = 77%), Pavetteae (BS = 70%), and Octoropideae (BS = 60%). Condamineeae were sister to (Posoquerieae + Spipaneeae) (BS = 100%). Sabiceeae were sister to Mussaendeae (BS = 74%).

Within Gentianaceae (Fig. 2f), Saccifolieae (BS = 100%) were the first diverged lineage. Exaceae (BS = 97%) were sister to the rest of the family (BS = 98%). Within the remaining of Gentianaceae, Chironieae (BS = 100%), followed by Gentianeae (BS = 99%), were subsequently sister to Helieae (BS = 77%) and Potalieae (BS = 88%). Within Chironieae, subtribe Chironiinae were well-supported as monophyletic (BS = 100%) and sister to Coutoubeinae and Canscorinae (BS = 94%). Within Potalieae, subtribe Lisianthiinae (BS = 97%) were sister to Potaliinae (BS = 93%) and Faroinae (BS = 88%).

Within Loganiaceae (Fig. 2f), Antonieae (BS = 100%) were sister to the rest of the family (BS = 92%). In the rest of Loganiaceae, Loganieae were sister to a clade (BS = 86%) comprising paraphyletic Strychneae with Spigelieae embedded in.

In Apocynaceae (Figs. 1, 2g–2i), the names of subfamilies and tribes were referred to Endress et al. (2014), and we found that Rauvolifioideae and Apocynoideae were paraphyletic. Asclepiadoideae (BS = 99%) and Secamonoideae (BS = 88%) were resolved as monophyletic with strong support.

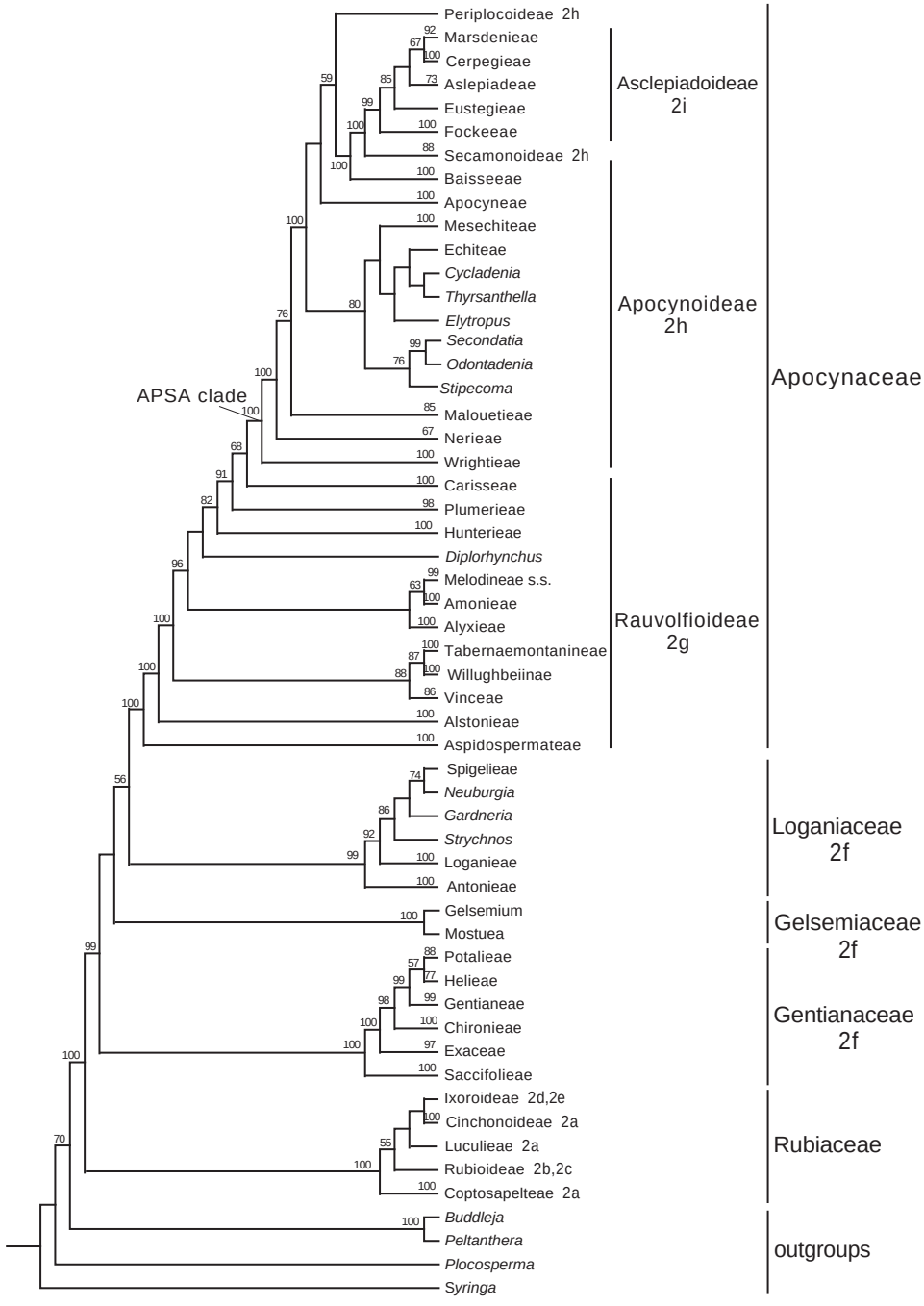


Fig. 1. Majority rule consensus of maximum likelihood trees resulting from the analysis of 653 taxa and the concatenated dataset of nine genes. Names of the families follow Angiosperm Phylogeny Group (APG) III. Numbers are bootstrap percentages (>50%). The number and letter following clade names (e.g., 2a–2i) refer to clade designations that are used to depict the separate portions of the complete tree in Fig. 2.

Aspidospermateae (BS = 100%) were the first diverged lineage of Apocynaceae. Alstonieae (BS = 100%), followed by a clade (BS = 88%) of Vinceae + (Willughbeiinae + Tabernaemontanineae) were subsequently sister to the rest of Apocynaceae. Within the rest of Apocynaceae, we resolved one major clade (BS = 82%) comprising Hunterieae (BS = 100%), Plumerieae (BS = 98%), Carisseae (BS = 100%), and the APSA clade (comprising Apocynoideae, Periplocoideae, Secamonoideae, and Asclepiadoideae) (BS = 100%). Hunterieae, followed by Plumerieae, were subsequently sister to Carisseae and the APSA clade. Within the APSA clade, Wrightieae, Nerieae, and Malouetieae were in turn sister to the rest (Fig. 2h). Within the rest of the APSA clade, we recovered three subclades: (i) Odontadenieae with Echiteae and Mesechiteae nested in; (ii) Apocyneae; and (iii) (Periplocoideae + (Baisseeae + (Secamonoideae + Asclepiadoideae))).

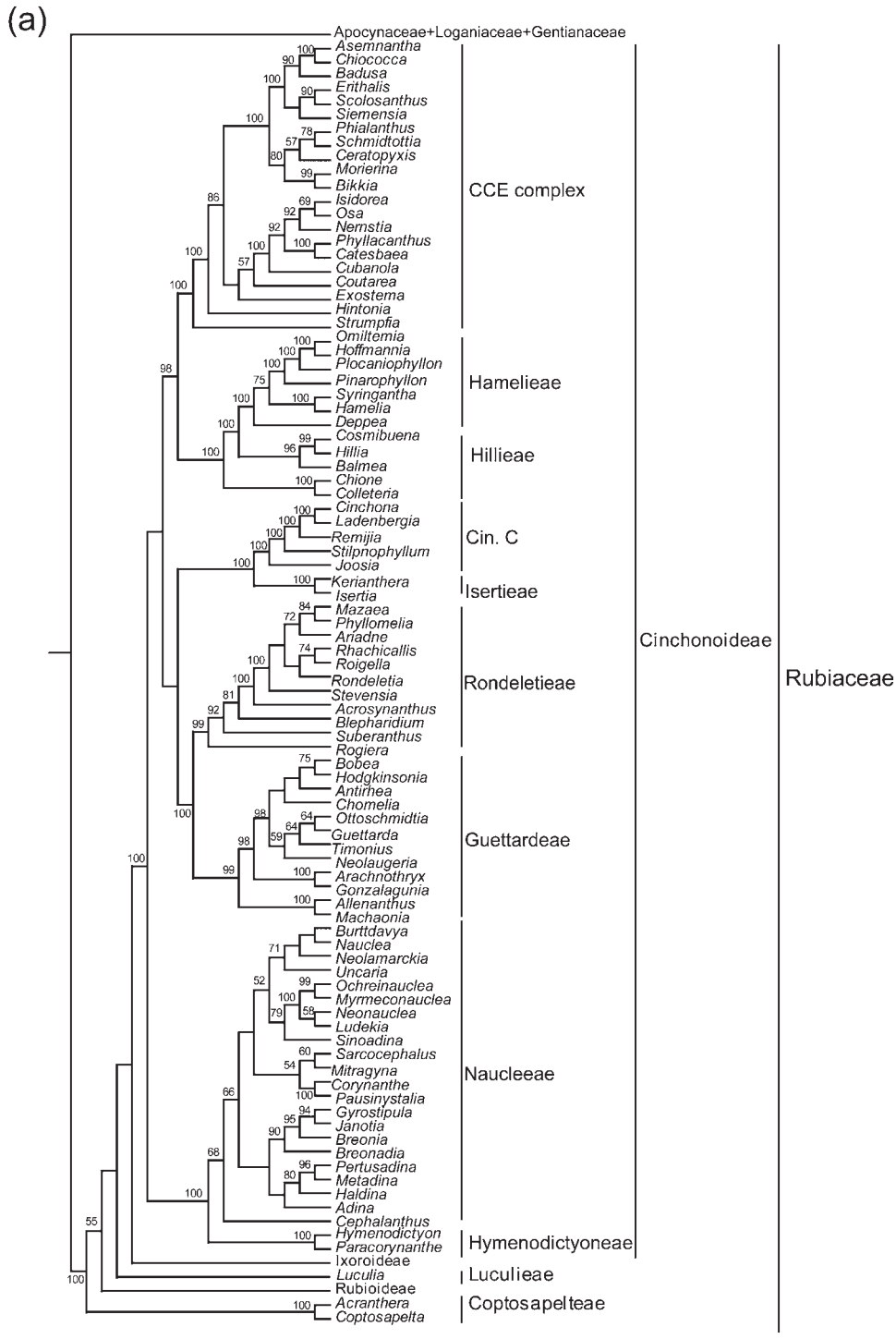


Fig. 2. Supertree of maximum likelihood trees resulting from the analysis of 653 taxa and the concatenated dataset of nine genes. APSA, Apocynoideae, Periplocoideae, Secamonoideae, and Asclepiadoideae; CCE, Catesbaeeae/Chiococceae alliance; Cin. C, subtribe Cinchoninae of Cinchoneae.

Discussion

Interfamilial phylogenetic relationships within Gentianales

Recent molecular phylogenetic studies have largely resolved relationships within most families of Gentianales (Yuan et al., 2005; Simões et al., 2007; Lens et al., 2008; Bremer & Eriksson, 2009; Rydin et al., 2009; Livshultz, 2010; Manns et al., 2012; Merckx et al., 2013; Rybczynski et al., 2014). However, the results from our study represent the most comprehensive genus-level phylogenetic hypothesis to date for Gentianales. Our sampling spanned the phylogenetic diversity of most families, thus, out results are unlikely constrained by limited

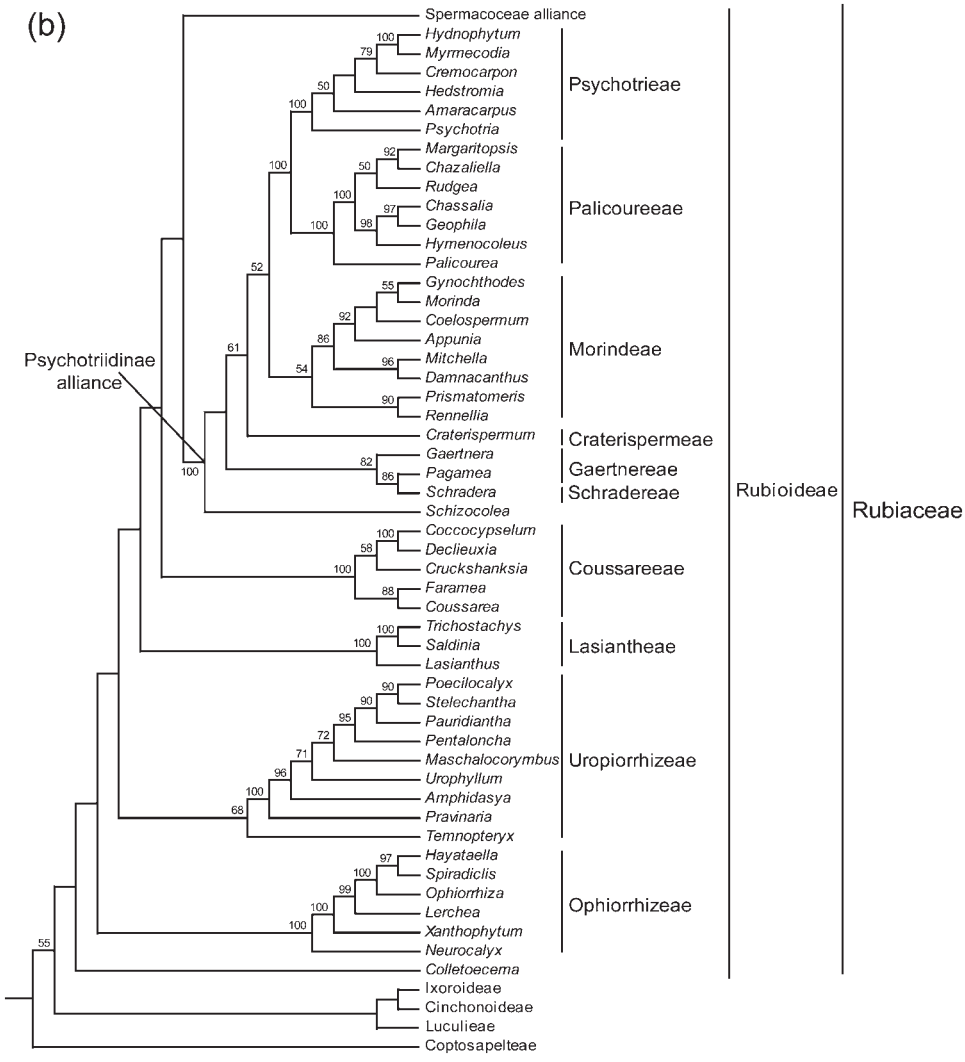


Fig. 2. Continued

taxon sampling. According to the concatenated dataset with 649 genera, we found strong support for the monophyly of Gentianales, Apocynaceae, Loganiaceae, Gelsemiaceae, Gentianaceae, and Rubiaceae which were resolved as sister to the rest of the order. Our results are consistent with a number of previous analyses (Bremer, 1996a; Backlund et al., 2000; Olmstead et al., 2000; Soltis et al., 2000, 2011; Bremer et al., 2001; Backlund, 2005; Jiao & Li, 2007; Frasier, 2008; Refulio-Rodriguez & Olmstead, 2014).

The exact relationships among Apocynaceae, Loganiaceae, Gelsemiaceae, and Gentianaceae have long been an open question. An investigation based on *ndhF* and *rbcL* supported Gentianaceae and Loganiaceae as subsequent sisters to Apocynaceae and Gelsemiaceae (Backlund et al., 2000). In contrast, a study based on *RPB2* DNA data showed that Gentianaceae were sister to all other Gentianales, and Loganiaceae were sister to Apocynaceae + Gelsemiaceae (Oxelman & Bremer, 2000). Other studies using one to three genes supported Gentianaceae as sister to Loganiaceae, and Apocynaceae as sister to Gelsemiaceae were weakly supported or unsolved (Olmstead et al., 2000; Soltis et al., 2000; Bremer et al., 2001). A study by Refulio-Rodriguez & Olmstead (2014) supported Apocynaceae + Gentianaceae as sister to Gelsemiaceae + Loganiaceae, while Jiao & Li (2007) resolved Gentianaceae and Gelsemiaceae as subsequent sisters to Apocynaceae + Loganiaceae. Other studies have supported Loganiaceae as sister to Gelsemiaceae + (Apocynaceae + Gentianaceae) (Frasier, 2008; Soltis et al., 2011). Our results are consistent with Jiao & Li (2007), by showing Apocynaceae as sister to Loganiaceae. The Apocynaceae + Loganiaceae clade was more strongly supported in our results (BS = 56%) than in Jiao & Li (2007), but the support for the relationship in both studies was low. The sister relationship between Apocynaceae and Loganiaceae was also supported by characteristics of pretext leaf (Li, 1982).

Intrafamilial phylogenetic relationships of the Gentianales

Rubiaceae

Previous molecular studies generally recovered three major lineages within Rubiaceae (Bremer, 1996a, 1996b; Bremer et al., 1999; Rova et al. 2002; Andersson & Antonelli, 2005; Robbrecht & Manen, 2006; Bremer & Eriksson, 2009; Rydin et al., 2009; Manns et al., 2012); the subfamilies Rubioideae, Ixoroideae, and Cinchonoideae *sensu* Bremer et al. (1999), and

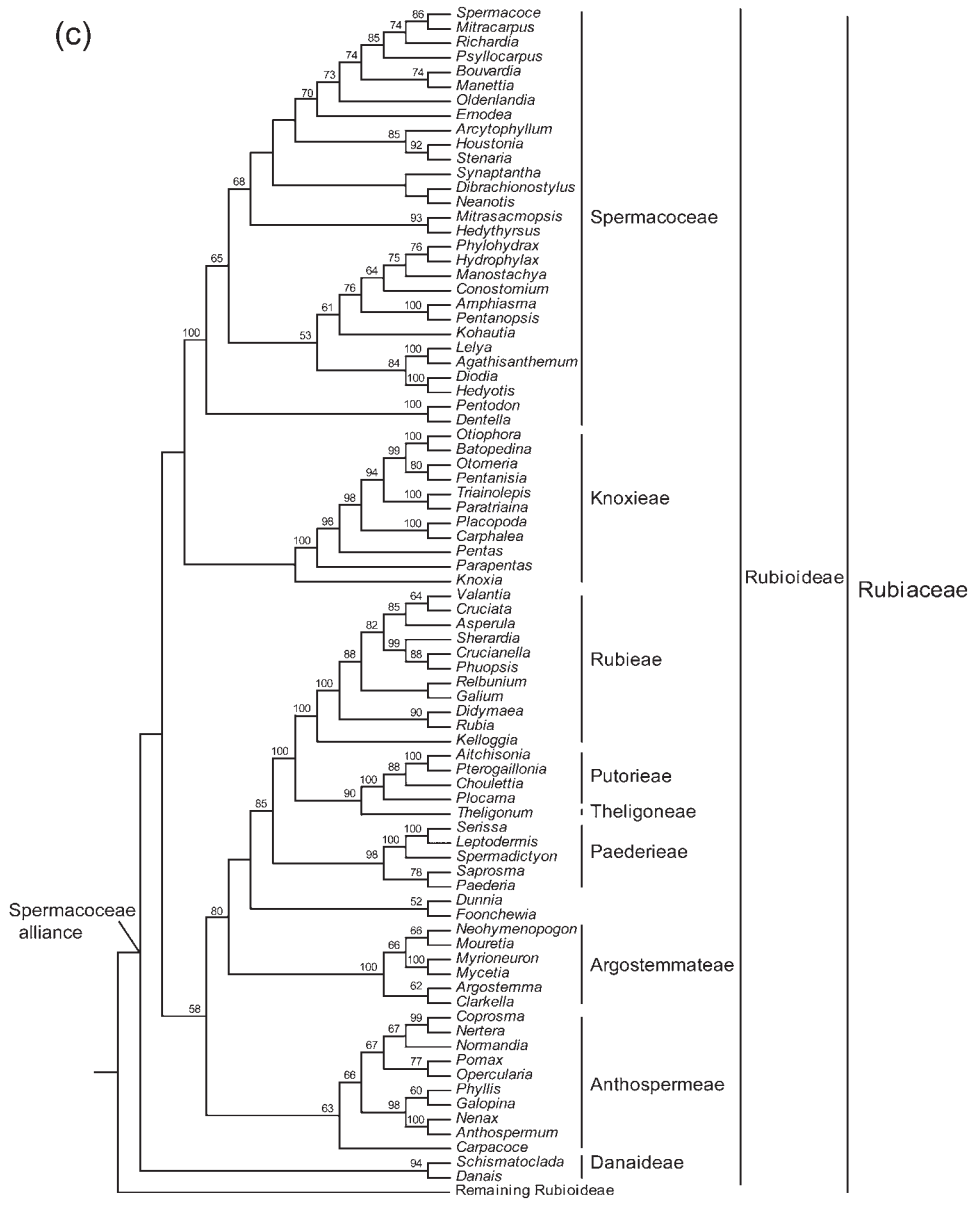


Fig. 2. Continued

two additional tribes, namely Luculieae and Coptosapelteae, which have phylogenetic positions outside the three subfamilies. Our result supported the monophyly of the three subfamilies and two tribes, but the subfamilies received low support (BS < 50%). We resolved Coptosapelteae as sister to the remaining Rubioideae, and this is consistent with several prior studies (Bremer et al., 1999; Robbrecht & Manen, 2006). However, other prior studies differ in the placement of Coptosapelteae. For example, Manns et al. (2012) showed Coptosapelteae as sister to Rubioideae, and Rydin et al. (2009) presented it as sister to Luculieae. The relationships among Rubioideae, Ixoroideae, and Cinchonoideae were poorly resolved by our study.

Within Cinchonoideae, our results recognized four clades, but the relationships among them were not well resolved. One clade included Hymenodictyeae and Naucleeae. Another included Guettardeae and Rondeletieae. An additional clade comprised Isertieae and Cin. C, and the last clade was the CCE complex + (Hamelieae + Hillieae). These results are consistent with several previous studies (Robbrecht & Manen, 2006; Manns et al., 2012). Bremer & Eriksson (2009) and Rydin et al. (2009) recovered these clades, except for Isertieae + Cin. C, and the CCE complex + (Hamelieae + Hillieae), respectively.

Within Rubioideae, our analyses recovered five basal lineages (*Colletoecema*, Ophiorrhizeae, Urophylleae, Lasiantheae, and Coussareeae), and two crown lineages (Spermacoceae alliance and Psychotrieae alliance). Our results are consistent with previous studies (Robbrecht & Manen, 2006; Bremer & Eriksson, 2009; Rydin et al., 2009; Manns et al., 2012), except that our BS support for the internal nodes within these clades and the monophyly of the Spermacoceae alliance were lower. In the Psychotrieae alliance, we found that Schradereae were nested within Gaertnereae, whereas Bremer & Eriksson (2009) resolved Schradereae as sister to

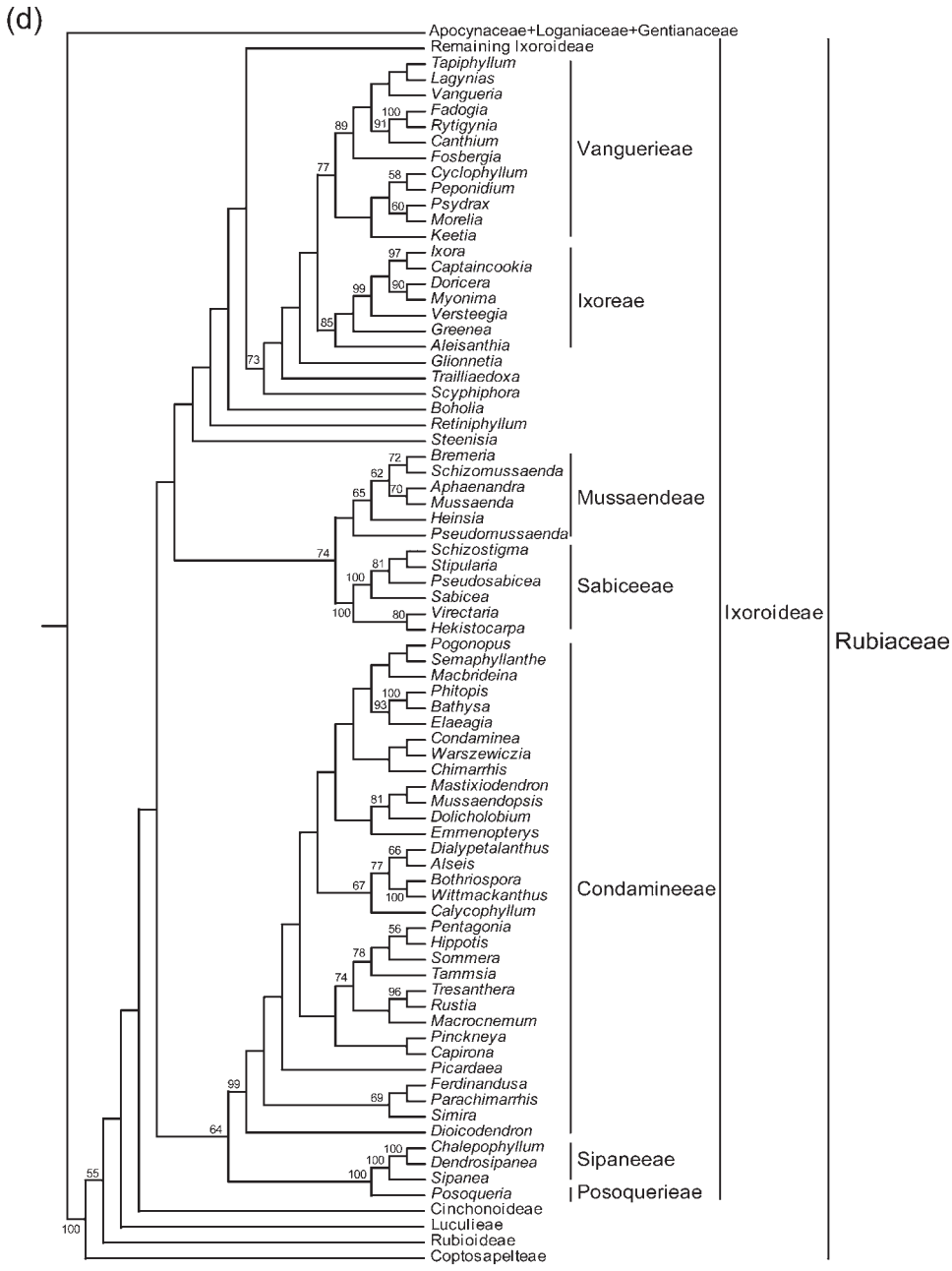


Fig. 2. Continued

Gaertnereae. The close relationship between Schradereae and Gaertnereae was not recovered by Robbrecht & Manen (2006). We showed that Morindeae were sister to Palicoureeae and Psychotrieae, and this is consistent with previous studies (Rydin et al., 2009; Manns et al., 2012), which had more sparse taxon sampling at generic level. Our results further supported Craterispermeae as sister to Morindeae + (Palicoureeae + Psychotrieae), which is a novel grouping, but the support is low. In the Spermacoceae alliance, we recovered one monophyletic subclade of Anthospermeae + (Argostemmateae + (Paederieae + (Rubieae + (Putorieae + Theligoneae)))), which has been resolved in prior studies (Robbrecht & Manen, 2006; Bremer & Eriksson, 2009; Rydin et al., 2009), differing only in the position of Theligoneae, in which they found the tribe a sister of Rubieae. Our results supported Theligoneae as sister to Putorieae (BS = 90%), and this relationship has not been previously reported.

Within Ixoroideae, we recovered nine monophyletic tribes, including Posoquerieae, Sipaneeae, Condamineeae, Sabiceeae, Mussaendeae, Ixoreae, Vanuerieae, Octoropideae, and Pavetteae, which have been recovered in previous studies (Robbrecht & Manen, 2006; Bremer & Eriksson, 2009; Rydin et al., 2009; Manns et al., 2012). We found that Condamineeae were sister to Posoquerieae and Sipaneeae. Sabiceeae were sister to Mussaendeae. These results agree with previous molecular phylogenetic studies (Robbrecht & Manen, 2006; Bremer & Eriksson, 2009; Rydin et al., 2009; Manns et al., 2012).

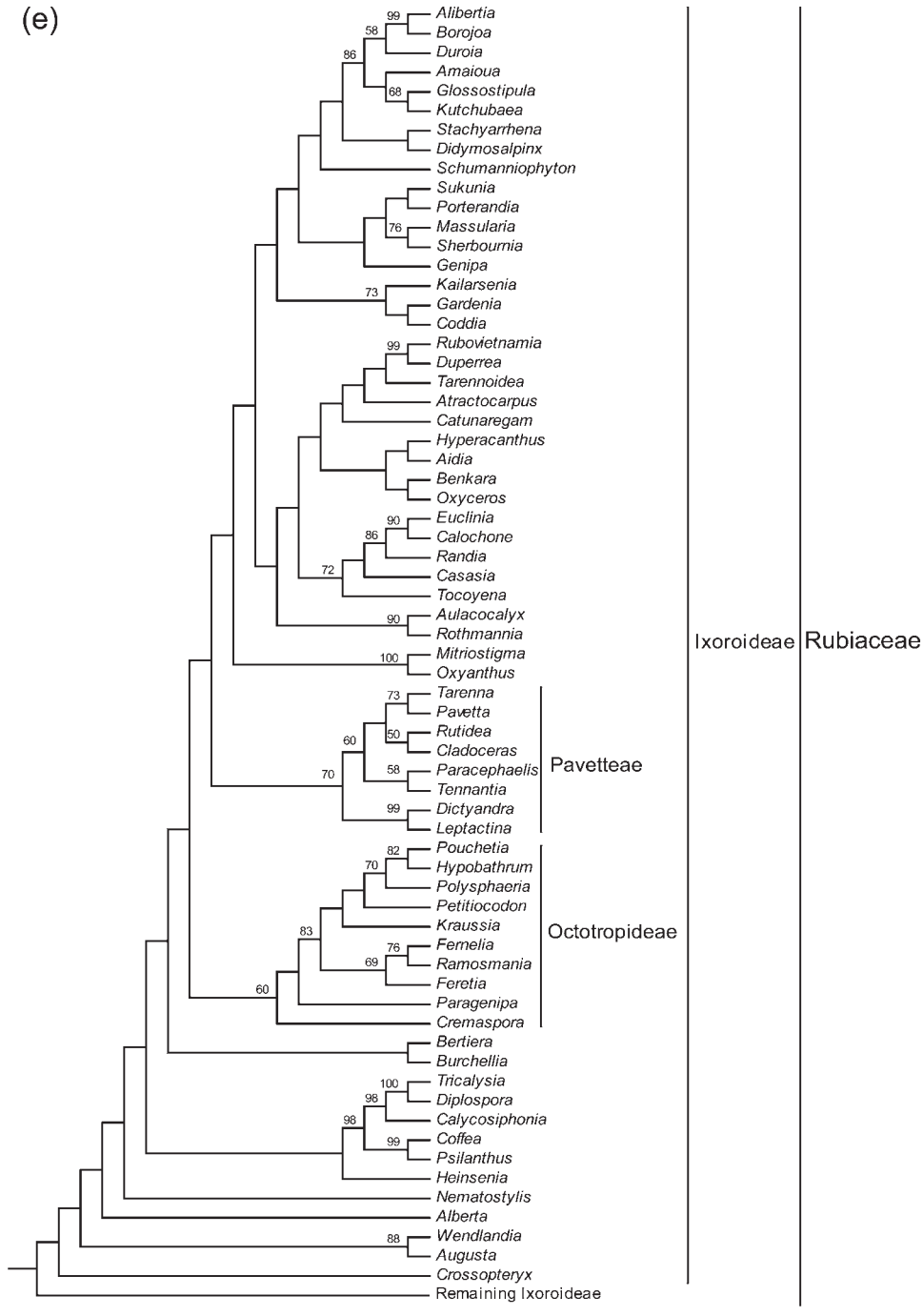


Fig. 2. Continued

Gentianaceae

Our deep-level topological relationships within Gentianaceae are largely in agreement with Rybczynski et al. (2014), who circumscribed seven tribes including Gentianeae, Helieae, Potalieae, Chironieae, Exaceae, Saccifolieae, and *incertae sedis* (*Voyria*). The genus *Voyria* was not included in our taxon sampling, but our work supported all the other six tribes as monophyletic (Fig. 2d). In a prior study based on the secondary structure of ITS, the monophyly of Helieae was not supported (Molina & Struwe, 2009). Our results showed that Saccifolieae were the first diverged lineage of Gentianaceae, and that Exaceae were sister to the reminder of family. Chironieae were sister to a clade containing the remaining three tribes. These results are consistent with previous studies (Yuan et al., 2005; Struwe et al., 2009; Merckx et al., 2013). Within Chironieae, we resolved three subtribes as monophyletic. Chironiinae were sister to Coutoubeinae and Canscorinae with strong support. The same relationship has been recovered by Merckx et al. (2013) based on sparse sampling, but the BS support values in that study were lower (BS $<$ 50%). A study by Yuan et al. (2005) supported Coutoubeinae as sister to Chironiinae, but was based on a

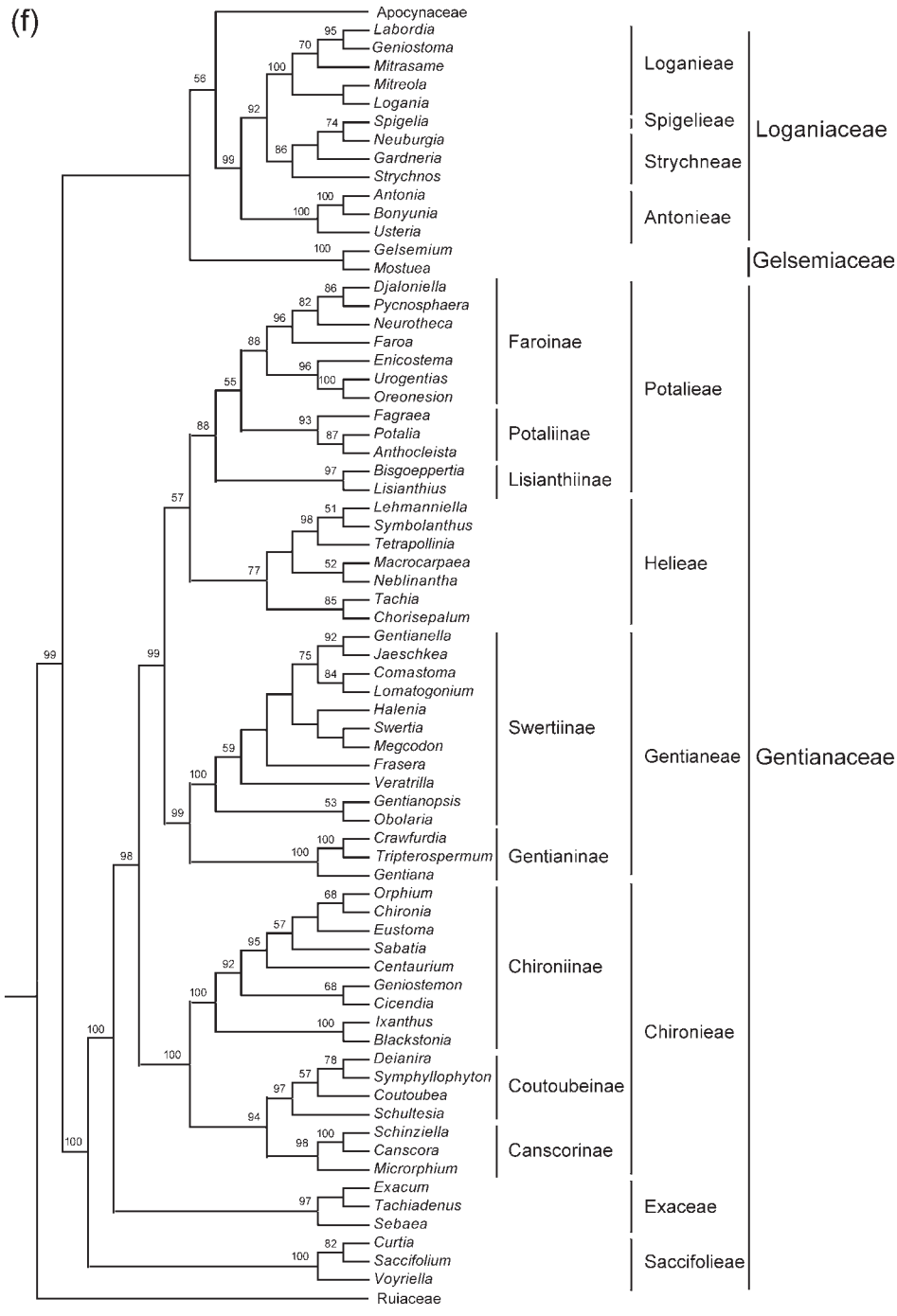


Fig. 2. Continued

less dense taxon sampling at the generic level and the BS support value was only 60%. Our results supported Gentianeae as sister to Helieae and Potalieae, but the BS support value for the sister relationship between Helieae and Potalieae was only 57%. The sister relationship between Helieae and Potalieae is inconsistent with some prior studies based on nuclear and mitochondrial gene sequences (Merckx et al., 2013) or combined molecular and morphological characters (Struwe et al., 2009), in which Helieae were sister to Gentianeae. However, the support value was lower in Struwe et al. (2009) than in Merckx et al. (2013) who included more taxa. The close affinities between Helieae and Gentianeae were also recovered by Molina & Struwe (2009), but Helieae were paraphyletic. In some studies, the relationships among these three tribes were not resolved (Yuan et al., 2005; Struwe et al., 2009). Within Potalieae, our combined analyses resolved Lisianthiinae as sister to Potaliinae and Faroinae, and this is in disagreement with Molina & Struwe (2009), who placed Potaliinae as sister to Faroinae and Lisianthiinae based on the analyses of the secondary structure of ITS.

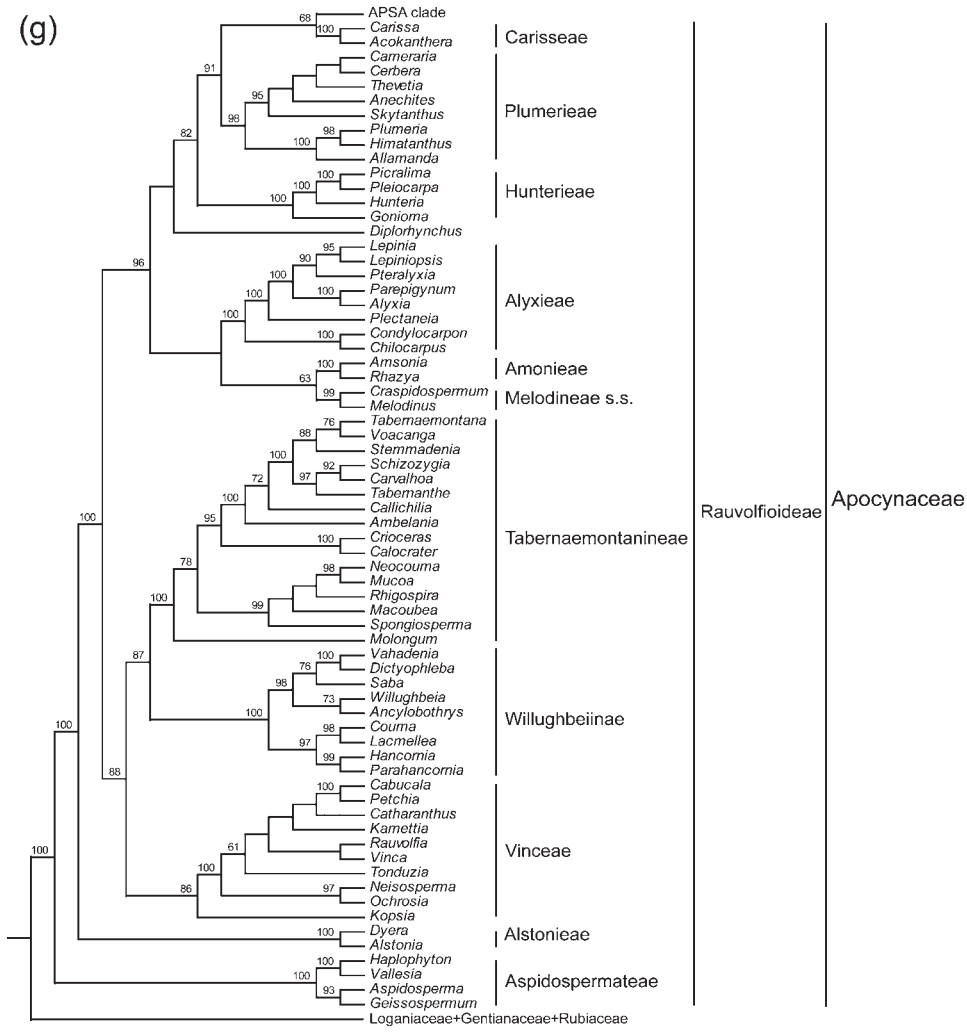


Fig. 2. Continued

Loganiaceae

According to the current interfamilial classification, Loganiaceae include 15 genera in four tribes (Antonieae, Loganieae, Spigelieae, and Strychneae) (Frasier, 2008). Our results strongly supported a division of Loganiaceae into three clades. The first comprises Antonieae, which were sister to the reminder of Loganiaceae. This is consistent with previous studies (Backlund et al., 2000; Oxelman & Bremer, 2000; Frasier, 2008; Popovkin et al., 2011). However, our results for the relationships among Loganieae, Spigelieae, and Strychneae are in conflict with some previous studies. For example, Oxelman & Bremer (2000) found that Loganieae were sister to Spigelieae and embedded in Strychneae, whereas Frasier (2008) resolved the sister of Loganieae as Spigelieae + *Gardneria*, which were sister to the other Strychneae (not including *Gardneria*). In contrast, Popovkin et al. (2011) used phylogenetic analyses of ITS and showed that Loganieae were paraphyletic and included Spigelieae, which were sister to *Neuburgi* + *Strychnos* of Strychneae (Popovkin et al., 2011). Our results are generally consistent with Backlund et al. (2000), except that they found Spigelieae as sister to *Strychnos*.

Apocynaceae

According to Endress et al. (2014), Apocynaceae were divided into 25 tribes within five subfamilies (Asclepiadoideae, Perplocoideae, Secamonoideae, Apocynoideae and Rauvolfioideae). Our 649-taxon analyses (Figs. 1, 2g–2i) are in agreement with prior studies (e.g., Endress et al., 1996; Sennblad & Bremer, 1996, 2000, 2002; Livshultz et al., 2007; Simões et al., 2007; Lens et al., 2008) in that: (i) Rauvolfioideae and Apocynoideae were paraphyletic; (ii) APSA were monophyletic and strongly supported; (iii) Asclepiadoideae and Secamonoideae were monophyletic and strongly supported; (iv) Perplocoideae were monophyletic albeit with low support; and (v) Asclepiadoideae were sister to Secamonoideae. Moreover, our phylogeny resolved Perplocoideae as sister to a clade containing Baisseeae, Secamonoideae, and Asclepiadoideae with moderate to strong support, and this is congruent with several previous studies (Simões et al., 2007; Lens et al., 2008). However, other studies, which included a smaller number of genera, showed an unresolved position for Perplocoideae or resolved it as sister to members of Echiteae (Sennblad & Bremer, 1996; Livshultz et al., 2007; Straub et al., 2014).

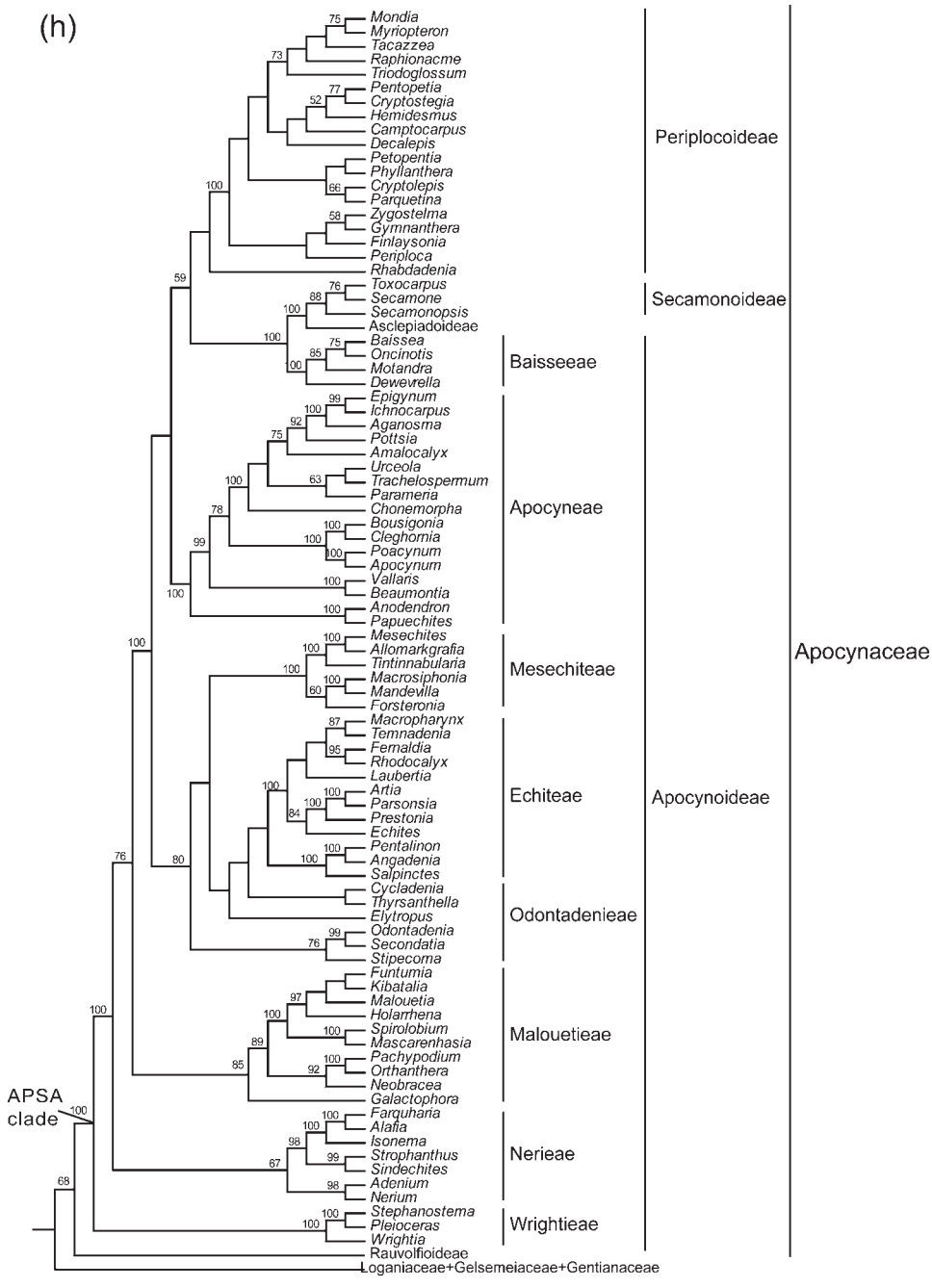


Fig. 2. Continued

Within Rauvolfioideae (Fig. 2g), most of the terminal clades recovered in our analyses corresponded to prior work based on five chloroplast DNA regions (Simões et al., 2007). Aspidospermateae and Alstonieae were subsequently sisters to the rest of the family. In particular, we found a strongly supported clade including Vinceae as sister to Willughbeiinae + Tabernaemontaneae as sister to the remaining Rauvolfioideae and the APSA clade. The sister relationship between Willughbeiinae and Tabernaemontaneae received higher BS in our study than in prior studies. Our work also supported the monophyly of Alyxieae, and this is in agreement with Simões et al., (2007). However, Alyxieae were paraphyletic according to Livshultz et al., (2007). The relationships among clades Alyxieae, Amsonieae, Melodineae, and Hunterieae were not resolved in previous studies (Simões et al., 2007; Livshultz, 2010). In the present study, Amsonieae were resolved as sister to Melodineae with moderate support (BS = 63%), and Hunterieae as sister to a clade comprising Plumerieae + (Carisseae + APSA) with strong support (BS = 82%). We report these results for the first time. Plumerieae, followed by Carisseae, were subsequently sister to the APSA clade with strong support, and this is congruent with previous studies (Livshultz et al., 2007; Simões et al., 2007; Livshultz, 2010).

Within the APSA clade (Figs. 2h, 2i), the Wrightieae clade was the first diverged lineage, and this is congruent with

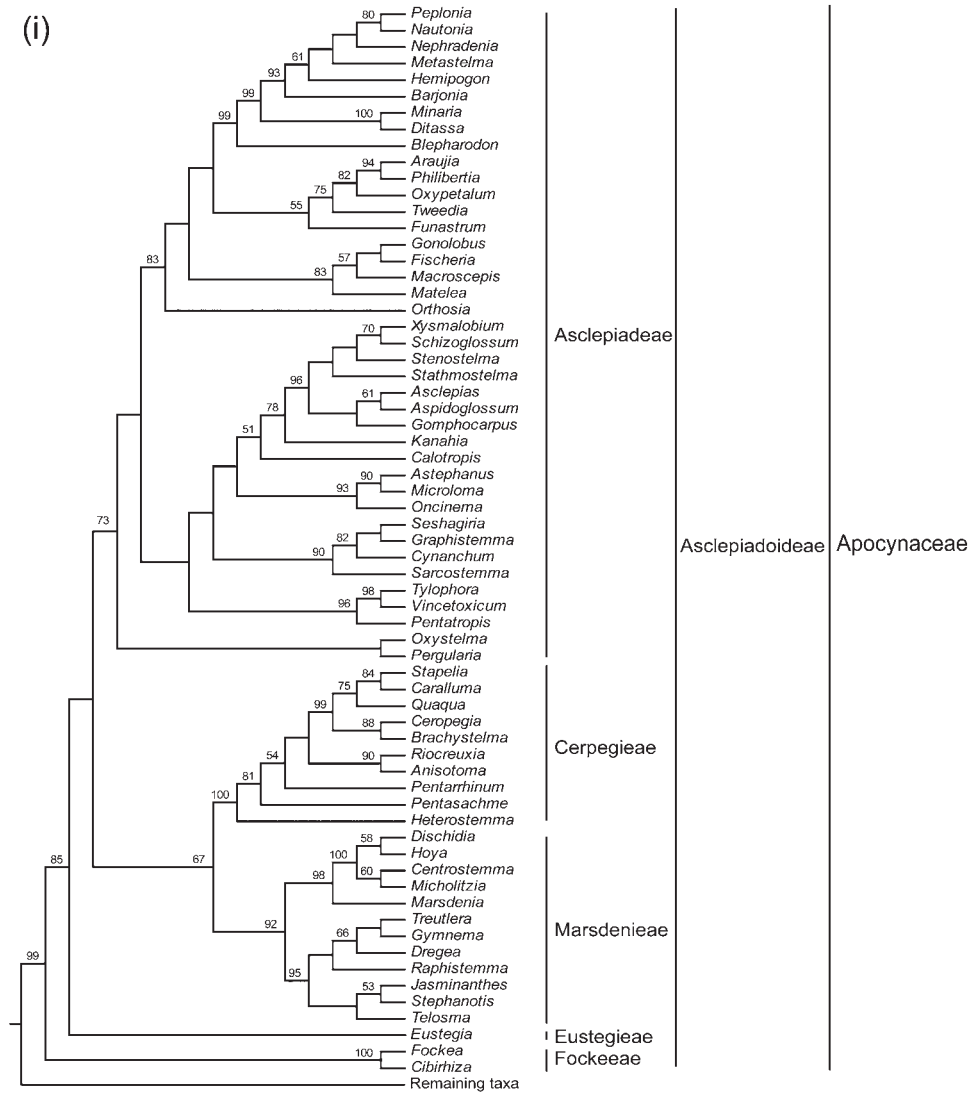


Fig. 2. Continued

previous studies (Livshultz et al., 2007; Simões et al., 2007; Lens et al., 2008; Straub et al., 2014). Nerieae and Malouetieae were mutually monophyletic subsequent sisters to the rest of the APSA clade. The monophyly of the Nerieae clade was not resolved in the molecular phylogenetic study of Simões et al. (2007). However, our results for Nerieae are consistent with a number of previous studies (Livshultz et al., 2007; Straub et al., 2014). We found that Malouetieae were sister to the remainder of the APSA clade, which comprised three subclades: (i) Odontadenieae with Echiteae and Mesechiteae nested in; (ii) Apocyneae; and (iii) Periplocoideae and Baisseeae as subsequent sisters to Secamonoideae and Asclepiadoideae. Our results showed that Perplocoideae were monophyletic, albeit with low support (BS $<$ 50%), and the relationships of (Baisseeae + (Secamonoideae + Asclepiadoideae)) were in agreement with previous studies (Simões et al., 2007; Lens et al., 2008). In some prior studies, the monophyly of Perplocoideae was not recovered and the clade (Perplocoideae + (Baisseeae + (Secamonoideae + Asclepiadoideae))) was nested within Apocynoideae (Livshultz et al., 2007; Straub et al., 2014). Our phylogeny resolved five tribes of Asclepiadoideae as monophyletic including Asclepiadeae, Ceropegieae, Marsdenieae, Eustegieae, and Fockeeae, and showed Fockeeae as sister to the other tribes. Additionally, we recovered Ceropegieae as sister to Marsdenieae, and this is consistent with Livshultz et al. (2007).

Conclusions and Perspectives

The order Gentianales, *sensu* APG III, consists of Apocynaceae, Gelsemiaceae, Gentianaceae, Loganiaceae, and Rubiaceae, and includes approximately 1200 genera and 20 000 species. In this paper, we provide a phylogeny for the genera of Gentianales, developed after analysis of the supermatrix of eight plastid markers (*rbcL, matK, atpB, ndhF, rpl16, rps16*, the *trnL-trnF* region, and the *atpB-rbcL* spacer) and one mitochondrial gene (*matR*) using maximum likelihood. The major clades and their relationships retrieved in the present study concur with those of previous studies. Our results provide new insights into relationships among intrafamilial clades. For example, within Rubiaceae we found that Craterispermeae

were sister to Morindeae + (Palicoureeae + Psychotrieae) and within Apocynaceae, Hunterieae were sister to a clade comprising Plumerieae + (Carisseae + APSA).

In the present study, we sampled 649 genera of Gentianales. Approximately 46% of the genera of Gentianales still have not been included in molecular phylogenetic analyses. Although deep and crown clades of the Gentianales recovered in our analyses are largely congruent with those in previous studies, in some cases, the support for some lineages was low, for example, Perplocoideae of Apocynaceae and Ixoroideae of Rubiaceae. In other cases, the markers showed different resolutions on different lineages, for example, in Rubiaceae, relationships among tribes of Ixoroideae were not resolved by comparison with that of Cinchonoideae. Thus, we advise the future studies to highlight the following two aspects: (i) sufficient taxon sampling including as many genera as possible; and (ii) more appropriate markers that could provide more informative sites for the phylogenetic analyses of Gentianales.

Acknowledgements

This research was supported by The Scientific Research Fund of Fairylake Botanical Garden (the study of deep relationships in Apocynaceae phylogeny), Shenzhen Key Laboratory of Southern Subtropical Plant Diversity (Grant No. SSTLAB-2014-04), National Key Basic Research Program of China (Grant No. 2014CB954100), Chinese Academy of Sciences International Institution Development Program (Grant No. SAJC201315), the Shenzhen Science and Technology Innovation Council Fund (Grant No. KQC 201105310009A), Chinese Academy of Sciences External Cooperation Program of BIC (Grant No. GJHZ201321), and the National Natural Science Foundation of China (Grant No. 31500175).

References

Andersson L, Antonelli A. 2005. Phylogeny of the tribe Cinchoneae (Rubiaceae), its position in Cinchonoideae, and description of a new genus, *Ciliosemina*. *Taxon* 54: 17–28.

APG III (Angiosperm Phylogeny Group III). 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161: 105–121.

Backlund M. 2005. *Phylogenetic studies in the Gentianales – approaches at different taxonomic levels*. Ph.D. Dissertation. Uppsala: Uppsala University.

Backlund M, Oxelman B, Bremer B. 2000. Phylogenetic relationships within the Gentianales based on *ndhF* and *rbcL* sequences, with particular reference to the Loganiaceae. *American Journal of Botany* 87: 1029–1043.

Bartling FG. 1830. *Ordines naturales plantarum eorumque characters et affinitates*. Göttingen: Dieterichianus.

Bremer B. 1996a. Phylogenetic studies within Rubiaceae and relationships to other families based on molecular data. *Opera Botanica Belgica* 7: 33–50.

Bremer B. 1996b. Combined and separate analyses of morphological and molecular data in the plant family Rubiaceae. *Cladistics* 12: 21–40.

Bremer B, Manen J-F. 2000. Phylogeny and classification of the subfamily Rubioideae (Rubiaceae). *Plant Systematics and Evolution* 225: 43–72.

Bremer B, Bremer K, Heidari N, Olmstead RG, Anderberg AA, Källersjö M, Barkhordarian E. 2002. Phylogenetics of asterids based on 3 coding and 3 noncoding chloroplast DNA markers and the utility of noncoding DNA at higher taxonomic levels. *Molecular Phylogenetics and Evolution* 24: 274–301.

Bremer B, Eriksson T. 2009. Time tree of Rubiaceae: Phylogeny and dating the family, subfamilies, and tribes. *International Journal of Plant Sciences* 170: 766–793.

Bremer B, Jansen RK, Oxelman B, Backlund M, Lantz H, Kim KJ. 1999. More characters or more taxa for a robust phylogeny in a case study from the coffee family (Rubiaceae). *Systematic Biology* 48: 413–435.

Bremer B, Olmstead RG, Struwe L, Bremer B. 1994. *Rbcl* sequences support exclusion of *Retzia*, *Desfontmnia*, and *Nicodemia* from the gentianales. *Plant Systematics and Evolution* 190: 213–230.

Bremer B, Struwe L. 1992. Phylogeny of the Rubiaceae and the Loganiaceae: Congruence of conflict between morphological and molecular data? *American Journal of Botany* 79: 1171–1184.

Bremer K, Backlund A, Sennblad B, Swenson U, Andreasen K, Hjertson M, Lundberg J, Backlund M, Bremer B. 2001. A phylogenetic analysis of 100 + genera and 50 + families of euasterids based on morphological and molecular data. *Plant Systematics and Evolution* 229: 137–169.

Chase MW, Hills HH. 1991. Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon* 40: 215–220.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Heden M, Gaut BS, Jansen RK, Kim K-J, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q, Plunkett GM, Soltis PS, Swensen S, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Goldenberg E, Learn GH, Graham SW, Barrett SCH, Dayanandan S, Albert VA. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–581.

Cronquist A. 1981. *An integrated system of classification of flowering plants*. New York: Columbia University Press.

Cronquist A. 1983. Some realignments in the dicotyledons. *Nordic Journal of Botany* 3: 75–83.

Downie SR, Palmer JD. 1992. Restriction site mapping of the chloroplast DNA inverted repeat: A molecular phylogeny of the Asteridae. *Annals of the Missouri Botanical Garden* 79: 266–283.

Doyle JJ, Doyle JI. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 19: 11–15.

Endress ME, Liede-Schumann S, Meve U. 2014. An updated classification for Apocynaceae. *Phytotaxa* 159: 175–194.

Endress ME, Sennblad B, Nilsson S, Civeyrel L, Chase MW, Huysmans S, Grafstrom E, Bremer B. 1996. A phylogenetic analysis of Apocynaceae s. str. and some related taxa in Gentianales: A multidisciplinary approach. *Opera Botanica Belgica* 7: 59–102.

Felsenstein J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology* 22: 240–249.

Frasier CL. 2008. *Evolution and systematics of the angiosperm order Gentianales with an in-depth focus on Loganiaceae and its species rich and toxic genus Strychnos*. Ph.D. Dissertation. New Brunswick: Rutgers University.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT[C]. *Nucleic Acids Symposium Series* 41: 95–98.

Jensen SR. 1992. Systematic implications of the distribution of iridoids and other chemical compounds in the Loganiaceae and other families of the Asteridae. *Annals of the Missouri Botanical Garden* 79: 284–302.

Jiao Z, Li J. 2007. Phylogeny of intercontinental disjunct Gelsemiaceae inferred from chloroplast and nuclear DNA sequences. *Systematic Botany* 32: 617–627.

Karehed J, Groeninckx I, Dessein S, Motley TJ, Bremer B. 2008. The phylogenetic utility of chloroplast and nuclear DNA markers and the phylogeny of the Rubiaceae tribe Spermacoceae. *Molecular Phylogenetics and Evolution* 49: 843–866.

Kim S, Soltis DE, Soltis PS, Zanis MJ, Suh Y. 2004. Phylogenetic relationships among early-diverging eudicots based on four genes: Were the eudicots ancestrally woody? *Molecular Phylogenetics and Evolution* 31: 16–30.

Lens F, Endress ME, Baas P, Jansen S, Smets E. 2008. Wood anatomy of Rauvolfioideae (Apocynaceae): A search for meaningful non-DNA characters at the tribal level. *American Journal of Botany* 95: 1199–1215.

Li P-T. 1982. On the trends of morphological differentiation and evolutionary system in Chinese apocynales. *Guihaia* 2: 165–170.

Lindley J. 1833. *Nixus Plantarum*. London: Ridgway & Sons.

Livshultz T. 2010. The phylogenetic position of milkweeds (Apocynaceae subfamilies Secamonoideae and Asclepiadoideae): Evidence from the nucleus and chloroplast. *Taxon* 59: 1016–1030.

Livshultz T, Middleton DJ, Endress ME, Williams J. 2007. Phylogeny of Apocynoideae and the APSA clade. *Annals of the Missouri Botanical Garden* 94: 323–361.

Manns U, Wikström N, Taylor CM, Bremer B. 2012. Historical biogeography of the predominantly neotropical subfamily Cinchonoideae (Rubiaceae): Into or out of America? *International Journal of Plant Sciences* 173: 261–286.

Merckx VS, Kissling J, Hentrich H, Janssens SB, Mennes CB, Specht CD, Smets EF. 2013. Phylogenetic relationships of the mycoheterotrophic genus *Voyria* and the implications for the biogeographic history of Gentianaceae. *American Journal of Botany* 100: 712–721.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Gateway Computing Environments Workshop (GCE), 2010. Piscataway: IEEE. http://dx.doi.org/10.1109/GCE2010.5676129.

Molina J, Struwe L. 2009. Utility of secondary structure in phylogenetic reconstructions using nrDNA ITS sequences – an example from Potalieae (Gentianaceae: Asteridae). *Systematic Biology* 34: 414–428.

Nicholas A, Baijnath H. 1994. A consensus classification for the order Gentianales with additional details on the suborder Apocynineae. *The Botanical Review* 60: 440–482.

Olmstead RG, Kim KJ, Jansen RK, Wagstaff SJ. 2000. The phylogeny of the Asteridae sensu lato based on chloroplast *ndhF* gene sequences. *Molecular Phylogenetics and Evolution* 16: 96–112.

Oxelman B, Bremer B. 2000. Discovery of paralogous nuclear gene sequences coding for the second-largest subunit of RNA polymerase II (RPB2) and their phylogenetic utility in Gentianales of the asterids. *Molecular Biology Evolution* 17: 1131–1145.

Popovkin AV, Mathews KG, Santos JCM, Molina MC, Struwe L. 2011. *Spigelia genuflexa* (Loganiaceae), a new geocarpic species from the Atlantic forest of northeastern Bahia, Brazil. *PhytoKeys* 6: 47–65.

Razafimandimbison SG, Rydin C, Bremer B. 2008. Evolution and trends in the *Psychotrieae alliance* (Rubiaceae)—A rarely reported evolutionary change of many-seeded carpels from one-seeded carpels. *Molecular Phylogenetics and Evolution* 48: 207–223.

Refulio-Rodriguez NF, Olmstead RG. 2014. Phylogeny of Lamiidae. *American Journal of Botany* 101: 287–299.

Robbrecht E, Manen CF. 2006. The major evolutionary lineages of the coffee family (Rubiaceae, angiosperms). *Systematics and Geography of Plants* 76: 85–146.

Rova JH, Delprete PG, Andersson L, Albert VA. 2002. A *trnL-F* cpDNA sequence study of the Condamineeae-Rondeletieae-Sipaneeae complex with implications on the phylogeny of the Rubiaceae. *American Journal of Botany* 89: 145–159.

Rybczynski JJ, Mikula A, Davey MR. 2014. *The Gentianaceae – Volume 1: Characterization and ecology*. New York: Springer Heidelberg.

Rydin C, Kainulainen K, Razafimandimbison SG, Smedmark JEE, Bremer B. 2009. Deep divergences in the coffee family and the systematic position of *Acranthera*. *Plant Systematics and Evolution* 278: 101–123.

Sennblad B. 1997. *Phylogeny of the Apocynaceae s.l.* Ph.D. Dissertation. Uppsala: Acta Universitatis Upsaliensis.

Sennblad B, Bremer B. 1996. The familial and subfamilial relationships of Apocynaceae and Asclepiadaceae evaluated with *rbcL* data. *Plant Systematics and Evolution* 202: 153–175.

Sennblad B, Bremer B. 2000. Is there a justification for differential a priori weighting in coding sequences? A case study from *rbcL* and Apocynaceae s.l. *Systematic Biology* 49: 101–113.

Sennblad B, Bremer B. 2002. Classification of Apocynaceae s.l. according to a new approach combining Linnaean and phylogenetic taxonomy. *Systematic Biology* 51: 389–409.

Sennblad B, Endress ME, Bremer B. 1998. Morphology and molecular data in phylogenetic fraternity: The tribe Wrightieae (Apocynaceae) based on plastid and nuclear DNA. *Systematic Botany* 37: 795–806.

Simões AO, Livshultz T, Conti E, Endress ME. 2007. Phylogeny and systematics of the Rauvolfioideae (Apocynaceae) based on molecular and morphological evidence. *Annals of the Missouri Botanical Garden* 94: 268–297.

Simpson MG. 2010. *Plant systematics*. San Diego: Elsevier-Academic Press.

Soltis DE, Smith SA, Cellinese N, Kenneth JW, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu YL, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donghue MJ, Soltis PS. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from 18s rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Straub SC, Moore MJ, Soltis PS, Soltis DE, Liston A, Livshultz T. 2014. Phylogenetic signal detection from an ancient rapid radiation: Effects of noise reduction, long-branch attraction, and model

selection in crown clade Apocynaceae. *Molecular Phylogenetics and Evolution* 80: 169–185.

Struwe L, Albert VA, Bremer B. 1994. Cladistics and family level classification of the Gentianales. *Cladistics* 10: 175–206.

Struwe L, Albert VA, Calió FM, Frasier C, Lepis KB, Mathews KG, Grant JR. 2009. Evolutionary patterns in neotropical Helieae (Gentianaceae): Evidence from morphology, chloroplast and nuclear DNA sequences. *Taxon* 58: 479–499.

Struwe L, Kadereit JW, Klackenberg J, Nilsson S, Thiv M, Von Hagen KB, Albert VA. 2002. *Systematics, character evolution, and biogeography of Gentianaceae, including a new tribal and subtribal classification. Gentianaceae: Systematics and natural history*. Cambridge: Cambridge University Press.

Takhtajan AL. 1980. Outline of the classification of flowering plants (Magnoliophyta). *The Botanical Review* 46: 225–359.

Takhtajan AL. 1997. *Diversity and classification of flowering plants*. New York: Columbia University Press.

Thorne RF. 1992. Classification and geography of the flowering plants. *The Botanical Review* 58: 225–327.

Wagenitz G. 1959. Die systematische Stellung der Rubiaceae. Ein Beitrag zum System der Sympetalen. *Botanische Jahrbücher für Systematik* 79: 17–35. (In German)

Wang W, Lu AM, Ren Y, Endress ME, Chen ZD. 2009. Phylogeny and classification of Ranunculales: Evidence from four molecular loci and morphological data. *Perspectives in Plant Ecology, Evolution and Systematics* 11: 81–110.

Yuan YM, Wohlhauser S, Möller M, Klackenberg J, Callmander MW, Küpfer P. 2005. Phylogeny and biogeography of *Exacum* (Gentianaceae): A disjunctive distribution in the Indian Ocean basin resulted from long distance dispersal and extensive radiation. *Systematic Biology* 54: 21–34.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12192/suppinfo:

Table S1. List of taxa used in this study with GenBank accession numbers.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12216

Research Article

A comprehensive generic-level phylogeny of the sunflower family: Implications for the systematics of Chinese Asteraceae

Zhi-Xi Fu[1,2], Bo-Han Jiao[1,2], Bao Nie[1,2], Guo-Jin Zhang[1,2], Tian-Gang Gao[1*], and China Phylogeny Consortium [†]

[1]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 10093, China
[2]University of the Chinese Academy of Sciences, Beijing 100049, China
[†]Members of the China Phylogeny Consortium are listed in the Appendix.
*Author for correspondence. E-mail: gaotg@ibcas.ac.cn. Tel.: 86-10-62836447. Fax: 86-10-62590843.
Received 16 April 2016; Accepted 17 June 2016; Article first published online 22 July 2016

Abstract The sunflower family (Asteraceae) is the largest and the most diverse flowering plant family, comprising 24 000–30 000 species and 1600–1700 genera. In China, Asteraceae are also the largest family, with approximately 2336 indigenous species in 248 genera. In the past two decades, molecular phylogenetic analyses has contributed greatly to our understanding of the systematics of Asteraceae. Nevertheless, the large-scale analyses and knowledge about the relationships of Chinese Asteraceae at the generic level as a whole are far from complete due to difficulties in sampling. In this study, we presented a three-marker (*rbcL*, *ndhF*, and *matK*) phylogeny of Asteraceae, including 506 genera (i.e., approximately one-third of Asteraceae genera). The study sampled 200 Chinese genera (i.e., approximately 80% of Chinese Asteraceae genera). The backbones of the new phylogeny were largely congruent with earlier studies, with 13 subfamilies and 45 tribes recognized. Chinese Asteraceae were distributed in 7 subfamilies (Mutisioideae, Wunderlichioideae, Carduoideae, Pertyoideae, Gymnarrhenoideae, Cichorioideae, and Asteroideae) and 22 tribes (Mutiseae, Hyalideae, Cardueae, Pertyeae, Gymnarrheneae, Vernonieae, Cichorieae, Doroniceae, Senecioneae, Astereae, Anthemideae, Gnaphalieae, Calenduleae, Inuleae, Athroismeae, Helenieae, Coreopsideae, Neurolaeneae, Tageteae, Millieae, Eupatorieae, and Heliantheae). Chinese Asteraceae lacked 6 basal subfamilies and 23 tribes. Several previously ambiguous relationships were clarified. Our analyses also resolved some unplaced genera within Chinese Asteraceae. Finally, our phylogenetic tree was used to revise the classification for all genera of Chinese Asteraceae. In total, 255 genera, 22 tribes, and 7 subfamilies in China are recognized.

Key words: Asteraceae, China, classification, phylogeny, supermatrix.

Asteraceae are the largest family of flowering plants in the world with over 1600 genera including 23 000 species (Anderberg et al., 2007). The members of the family are distributed in every continent but Antarctica (Funk et al., 2005). The family is placed in Eudicots–Asterids–Campanulids–Asterales (APG IV, 2016).

Historically, Asteraceae were classified into two subfamilies (Asteroideae and Cichorioideae) and 13 tribes (Bentham, 1873). This classification was used in some floras and handbooks (e.g., Ling et al., 1985a in *Flora Reipublicae Popularis Sinicae*). However, there have been major changes in the classification of Asteraceae in recent decades with a better phylogenetic framework (Jansen & Palmer, 1987; Kim et al., 1992; Kim & Jansen, 1995; Bayer & Starr, 1998; Kim et al., 2002; Panero & Funk, 2002, 2008; Goertzen et al., 2003; Panero, 2005; Funk et al., 2005, 2009a, 2009b, 2009c; Funk & Specht, 2007; Smith et al., 2009; Torices, 2010; Panero et al., 2014; Mandel et al., 2015). Based on 10 or 14 chloroplast DNA (cpDNA) markers, Panero & Funk (2002, 2008) and Panero et al. (2014) reconstructed the robust "backbone" of Asteraceae with 12–13 major clades (subfamilies) identified. Chinese Asteraceae comprise approximately 2336 indigenous species (ca. 1145 endemic) and 248 genera (nearly 15% of the world genera, Shih et al., 2011). During the last two decades, several molecular phylogenetic studies sampled Chinese Asteraceae. But these studies largely focused on either generic- or species-level relationships (e.g., *Nannoglottis* of Qinghai–Tibet Plateau (QTP), Liu et al., 2002; *Ligularia–Cremanthodium–Parasenecio* (LCP) complex of QTP, Liu et al., 2006; *Saussurea* of QTP, Wang & Liu, 2004, Wang et al., 2009b; Himalayan endemic *Dolomiaea*, *Diplazoptilon* Ling, and *Xanthopappus*, Wang et al., 2007; *Nemosenecio*, *Sinosenecio*, and *Tephroseris*, Wang et al., 2009a; *Parasyncalathium*, *Soroseris*, *Stebbinsia*, and *Syncalathium* of QTP, Zhang et al., 2011a, 2011b; *Ajania* and *Chrysanthemum*, Zhao et al., 2010, Liu et al., 2012; *Aster*, Li et al., 2012; *Anaphalis*, Nie et al., 2013, 2015; *Lactuca* alliance, Wang et al., 2013b; *Crepidiastrum*, Peng et al., 2014; *Faberia*, Liu et al., 2013, Wang

et al., 2014; *Youngia*, Deng et al., 2014; *Diplazoptilon* and *Saussurea*, Yuan et al., 2015). These studies provided new insights into the phylogeny of Asteraceae and led to a number of taxonomic changes regarding the circumscription of genera. However, knowledge of relationships at the generic level of the whole Chinese Asteraceae remained poorly understood due to difficulties in sampling. In addition, the placements of some genera of Chinese Asteraceae (e.g., *Ainsliaea*, *Myripnois*, *Pertya*, *Cavea*, *Echinops*, *Atractylodes*, *Carlina*, *Tugarinovia*, *Formania*, *Centipeda*, and *Doronicum*) into tribes or subfamilies were still disputed or even completely unknown.

Given the large number of available plastid sequences in Asteraceae and the fact that no robust large-scale phylogenies existed for Chinese Asteraceae at the generic level, this study includes ca. 80% genera (200), all tribes (22), and all subfamilies (7) of Chinese Asteraceae. The main objectives of this study were to: (i) produce a most comprehensive generic-level phylogeny of Chinese Asteraceae; (ii) elucidate phylogenetic relationships of Chinese Asteraceae at the generic level and resolve phylogenetic placements of some genera with uncertain or unknown affinities; and (iii) evaluate the current classification (Shih et al., 2011) and provide an updated generic classification of Chinese Asteraceae.

Materials and Methods

Taxon sampling

A supermatrix of 512 genera, 805 species (including outgroup species), and 1840 sequences was constructed, including representatives of all (13) subfamilies, all (45) tribes, and 33% (506 of 1600) genera (according to recent molecular studies, Panero, 2005; Panero & Funk, 2002, 2008; Funk et al., 2009a, 2009c; Panero et al., 2014). Chinese Asteraceae (Asteraceae distributed in China, both native and introduced) were broadly sampled, including 313 species in 200 genera. Six genera and seven species of two closely related families (*Goodenia varia* R. Br. from Goodeniaceae; *Acicarpha tribuloides* Juss., *Acicarpha spathulata* R. Br., *Boopis anthemoides* Juss., *Calycera crassifoliav* (Miers) Hicken, *Nastanthus spathulatus* (Phil.) Miers, and *Scaevola aemula* R. Br. from Calyceraceae) were selected as outgroup species according to recent studies (Funk et al., 2005; Lundberg, 2009; Winkworth et al., 2008; APG IV, 2016). A total of 51 species representing 38 Chinese genera were newly sequenced in the Tree of Life for the genera of Chinese vascular plants project (Chen et al., 2016).

DNA extraction, polymerase chain reaction amplification, and sequencing

Three markers (*rbcL*, *matK*, and *ndhF*) from the plastid genome were used in the phylogenetic analyses. Total genomic DNA was extracted from silica gel-dried leaf material using a modified CTAB protocol (Doyle & Doyle, 1987) and Plant Genomic DNA Kit (Tiangen Biotech, Beijing, China).

All primers are provided in Table 1. The reaction volume was 25 μL, containing 7.5–8.5 μL ddH_2O, 12.5 μL Mix (0.05 U/μL Taq polymerase, 4 mol/L $MgCl_2$, and 0.4 mol/L each dNTP; TransGen Biotech, Beijing, China), 1.5 μL each primer (10 pmol/μL), and 50–100 ng template DNA. Polymerase chain reaction products were purified using an agarose gel purification kit (Qiagen, Hilden, Germany) following the recommended protocols. The PCR conditions were: 1 cycle of 5 min at 94 °C for denaturation, 40 cycles of 1 min at 94 °C for denaturation (for *rbcL*, 35 cycles; for *ndhF*, 30 cycles), 1.5 min of annealing at 50 °C (for *rbcL*, 30 s), and 1.5 min at 72 °C for extension (for *matK*, 2 min; for *trnK*, 3 min), with a final 10 min extension at 72 °C.

Sequencing reactions were carried out using an ABI Prism BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Sequences were analyzed on an ABI 3730xl DNA Analysis System (Applied Biosystems) following the manufacturer's protocols. GenBank accession numbers of the 120 sequences newly generated from the 38 Chinese genera were deposited in GenBank (Table S1; see Chen et al., 2016).

Molecular markers and DNA alignment

Three genes (plastid *rbcL*, *matK*, and *ndhF*) of the Asteraceae were obtained from GenBank (National Center for Biotechnology Information (NCBI), http://www.ncbi.nlm.nih.gov) using a Perl script. Three datasets were available by 1 April 2014 (except for *Famatinanthus decussatus* (Hieron.) Ariza & S. E. Freire, Panero et al., 2014).

A three-step strategy was used for each region to generate high-quality alignments. First, the profile alignments of three markers were carried out using MAFFT version 7.0 (http://mafft.cbrc.jp/alignment/software/; Katoh & Standley, 2013). Then, the alignments were checked and adjusted manually with BioEdit version 7.1.3 (Hall, 1999). Gaps were treated as

Table 1 List of primers used in polymerase chain reaction amplification and cycle sequencing

Gene	Primer name	Sequence	Reference
ndhF	*ndhF*-5F	ATGGAACAGACATATCAATATTAAT	Olmstead & Palmer (1994)
	ndhF-1318R	CGAAACATATAAAATGC(AG)GTTAATCC	Olmstead & Sweere (1994)
	ndhF-972F	GTCTCAATTGGGTTATATGATG	Olmstead & Sweere (1994)
	ndhF-2110R	CCCCCTA(CT)ATATTTGATACCTTCTCC	Olmstead & Sweere (1994)
rbcL	*rbcL*-1F	ATGTCACCACAAACAGAAACTAAAGC	Fay et al. (1997)
	rbcL-1460R	CTTTTAGTAAAAGATTGGGCCGAG	Chase et al. (1993)
matK	*matK*-AF	CTATATCCACTTATCTTTCAGGAGT	Kato et al. (1998)
	matK-8R	AAAGTTCTAGCACAAGAAAGTCGA	Kato et al. (1998)
	trnK-3914F	GGGGTTGCTAACTCAACGG	Johnson & Soltis (1994)
	trnK-2R	AACTAGTCGGATGGAGTAG	Johnson & Soltis (1994)

missing data and no gap coding was applied. All characters were treated as equally weighted. In the second step, several preliminary maximum likelihood (ML) trees were constructed using RAxML (Stamatakis et al., 2008) to identify any obviously problematic taxa. Duplicate sequences were eliminated from the same taxon, keeping the sequence with longest length. In the final step, a generic balance sampling strategy (i.e., each genus included 1–4 samplings) was adopted by the present supermatrix.

In the final supermatrix, names of tribes, genera, and species were checked based on *Flora of China* (Shih et al., 2011), *Species Catalogue of China* (Gao & Zhang, 2016), Tropicos (http://www.tropicos.org/Home.aspx), and The Plant List (http://www.theplantlist.org).

Phylogenetic analyses

The best-fitting nucleotide substitution model for each gene was evaluated using the program Modeltest 3.7 (Posada & Crandall, 1998) according to the Akaike Information Criterion. It was the general time reversible model incorporating sites and a gamma distribution (GTR + I + G) for three genes.

Phylogenetic analyses were undertaken using ML and Bayesian inference (BI). Maximum likelihood analyses were generated by RAxML version 7.2.8. (Stamatakis et al., 2008). All parameter values for the tree search were calculated with 1000 non-parametric inferences to assess nodal support. Bootstrap values (BS) of 80%–100% were interpreted as strong support, 60%–80% as moderate. The BI analysis was carried out in MrBayes version 3.2.2 (Ronquist et al., 2012). Four Markov chain Monte Carlo chains were run, sampling one tree every 1000 generations for 8 000 000 generations, starting with a random tree. Bayesian posterior probabilities (PP) were calculated for the majority consensus tree of all sampled trees after discarding 25% of trees sampled. Posterior probabilities of 0.88–1.00 were considered to be strong support, 0.70–0.87 to be moderate. The ML and BI analyses were both undertaken in the CIPRES science gateway portal (https://www.phylo.org/portal2/; Miller et al., 2010). Finally, the trees were visualized by FigTree version 1.4.0 (Rambaut, 2012).

Results

Characteristics of sequence data

The complete data matrix contained 805 species and three markers for a total of 120 newly determined sequences and 1720 previously published sequences (see Table S1). The total length of the three regions of cpDNA was 5125 bp. Sequence characteristics by genes are summarized in Table 2.

Resolution and backbone of major clades within Asteraceae

A summary of the ML tree based on the rapid bootstrapping analysis from RAxML (final optimized $-\ln L = 104\,470.14$) is shown in Figs. 1 and 2. The phylogeny estimated using BI analysis of three markers shared the same topology with the ML tree. An overview of inferred topologies and bootstrap values is given in Fig. 1 and 13 major clades were recognized (as shown in Figs. 1, 2 in different colors). Recent synonyms of species from NCBI sequences are listed in brackets in Fig. 2.

In both BI and ML analyses, the monophyly of Asteraceae was strongly supported (PP = 1.00; BS = 92). Twelve of 13 major clades were well supported along the backbone of the Asteraceae (PP = 0.88–1.00; Fig. 1) (the only exception, subfamily Wuderlichioideae were weakly supported: PP = 0.64; BS = 49). The Barnadesioideae were resolved as a monophyletic clade with strong support (PP = 1.00; BS = 100). They were sister to the remaining major clades (Famatinanthoideae, Mutisioideae, Gochnatioideae + Stifftioideae + Wunderlichioideae, Hecastocleidoideae, Carduoideae, Pertyoideae, Gymnarrhenoideae, Cichorioideae, Corymbiodeae, and Asteroideae, PP = 1.00; BS = 96), however, the interrelationship within Gochnatioideae + Stifftioideae + Wunderlichioideae was unresolved in the BI and ML analyses.

Phylogenetic relationships of Chinese Asteraceae

The detailed topology of Chinese Asteraceae was investigated using the 200 genera (nearly 80% of genera in China) and 313 species (Fig. 2, taxon names in black). The ML and BI analyses recognized seven well-distinguished clades of Chinese Asteraceae: Mutisioideae, Wunderlichioideae, Carduoideae, Pertyoideae, Gymnarrhenoideae, Cichorioideae, and Asteroideae.

Subfamily Mutisioideae

Within Mutisioideae (Figs. 1, 2A), Nassauvieae + Mutisieae and Onoserideae were supported as monophyletic, albeit with incongruent levels of support between the two inference methods (PP = 0.94; BS = 48). Each tribe was recovered with strong support (PP = 1.00; BS > 99). Within Mutisieae, The Chinese *Gerbera* (including *Piloselloides*) was strongly supported as monophyletic (PP = 1.00; BS = 97). The relationship between *Adenocaulon* and the remaining groups was not resolved.

Subfamily Wunderlichioideae

Support for the monophyly of Wunderlichioideae was weak (PP = 0.64; Figs. 1, 2A). Within tribe Hyalideae, Chinese genera *Nouelia* and *Leucomeris* were supported as the monophyletic group with strong support (PP = 1.00). They were supported as sister to the monophyletic South American *Ianthopappus* + *Hyalis* (PP = 1.00; BS = 100).

Table 2 Statistics from analyses of the chloroplast datasets of Asteraceae used in this study

Data	Aligned length	Taxa	Newly produced sequences/GenBank	Variable sites	Parsimony informative sites	Missing data in matrix, %
ndhF	2716	439	31/408	1478	1036	45.1
matK	1113	702	41/661	773	584	26.3
rbcL	1296	699	48/651	542	371	28
Combined	5125	805	120/1720	2793	1991	33.1

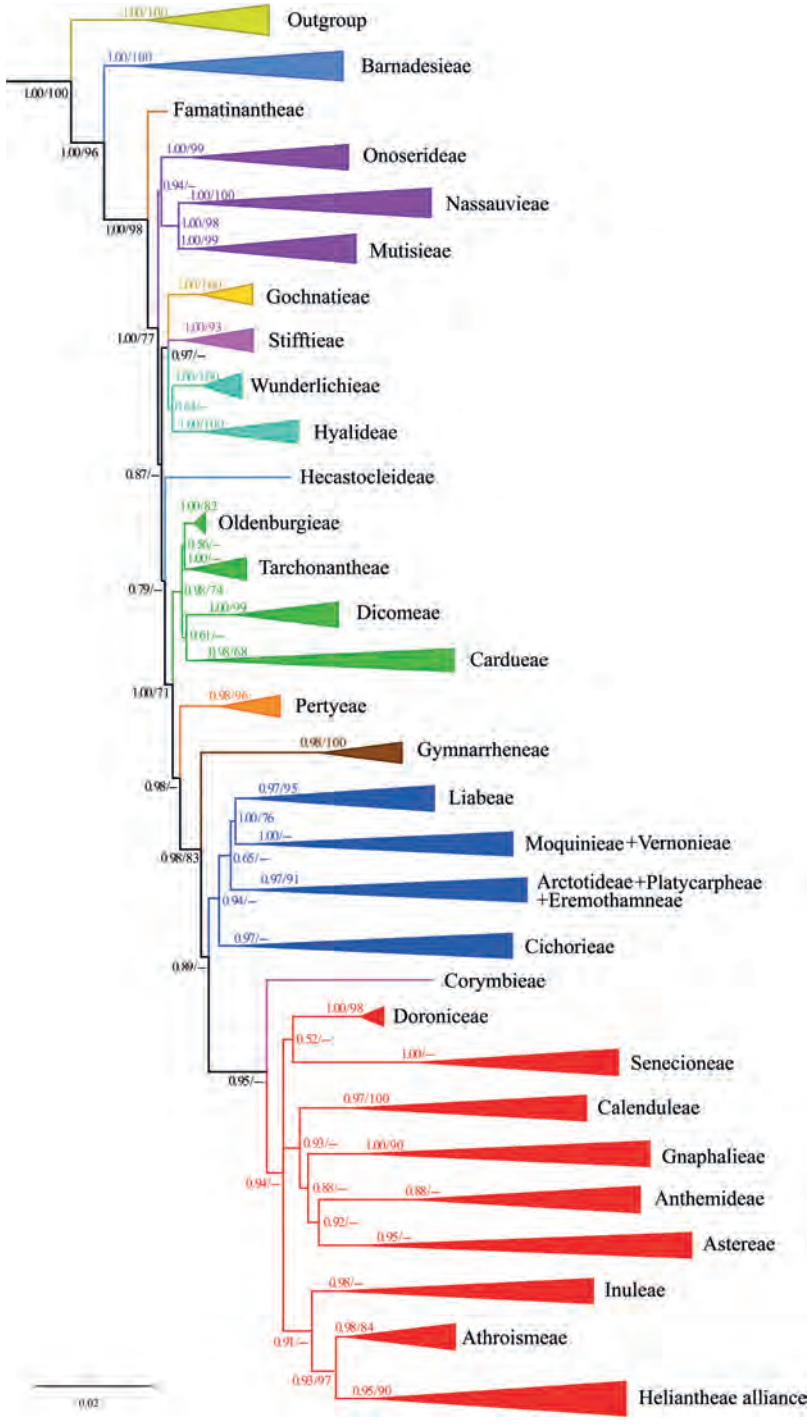


Fig. 1. Skeletal representation of the 805 species tree from Bayesian inference (BI) and maximum likelihood (ML) analyses, with tips representing tribes of Asteraceae based on the taxonomic arrangement of Funk et al. (2009c) and Panero et al. (2014). Branches and terminals are color-coded by the subfamilies of Asteraceae: aqua, Wunderlichioideae; blue, Cichorioideae; brown, Gymnarrhenoideae; dark orange, Famatinantheae; gold, Gochnatioideae; green, Carduoideae; light blue, Barnadesieae; light green, Carduoideae; magenta, Corymbiodeae; mid blue, Hecastocleidoideae; orange, Pertyoideae; pink, Stifftioideae; purple, Mutisioideae; red, Asteroideae; Support values are provided for each node (BI/ML). –, Values <0.50 (BI) or <50% (bootstrap support).

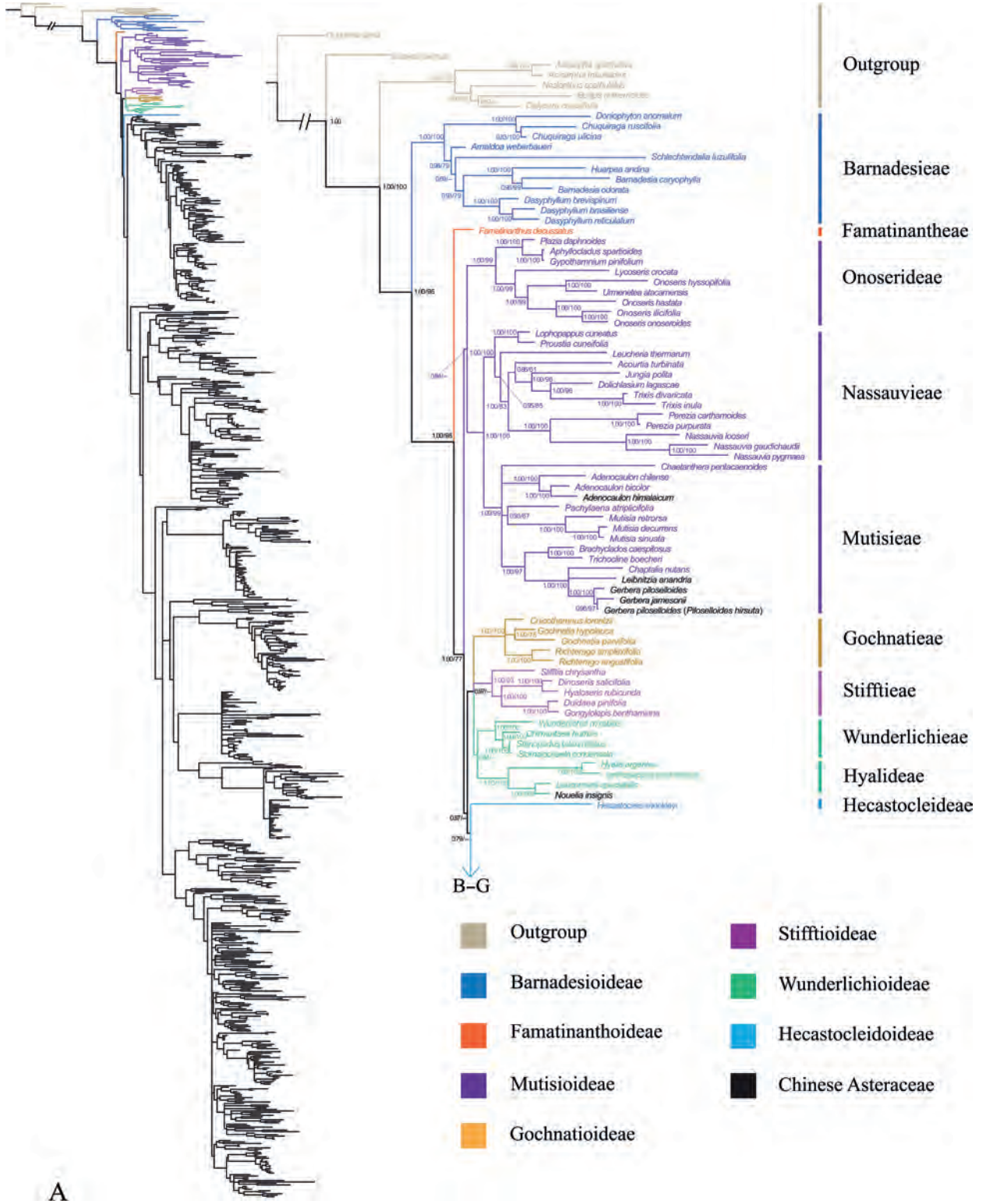


Fig. 2. Large-scale Bayesian and maximum likelihood estimate (values >0.50 or 50% are shown) of the Asteraceae phylogeny. **A**, Barnadesioideae, Famatinanthoideae, Mutisioideae, Gochnatioideae, Stifftioideae, Wunderlichioideae, and Hecastocleidoideae. **B**, Carduoideae. **C**, Pertyoideae, Gymnarrhenoideae, and Cichorioideae. **D**, Asteroideae and Corymbiodeae. **E–G**, Asteroideae; taxon names of Chinese Asteraceae in black. The tree contains 805 species represented by up to 5125 bp of sequence data from *ndhF*, *matK*, and *rbcL*.

Continued

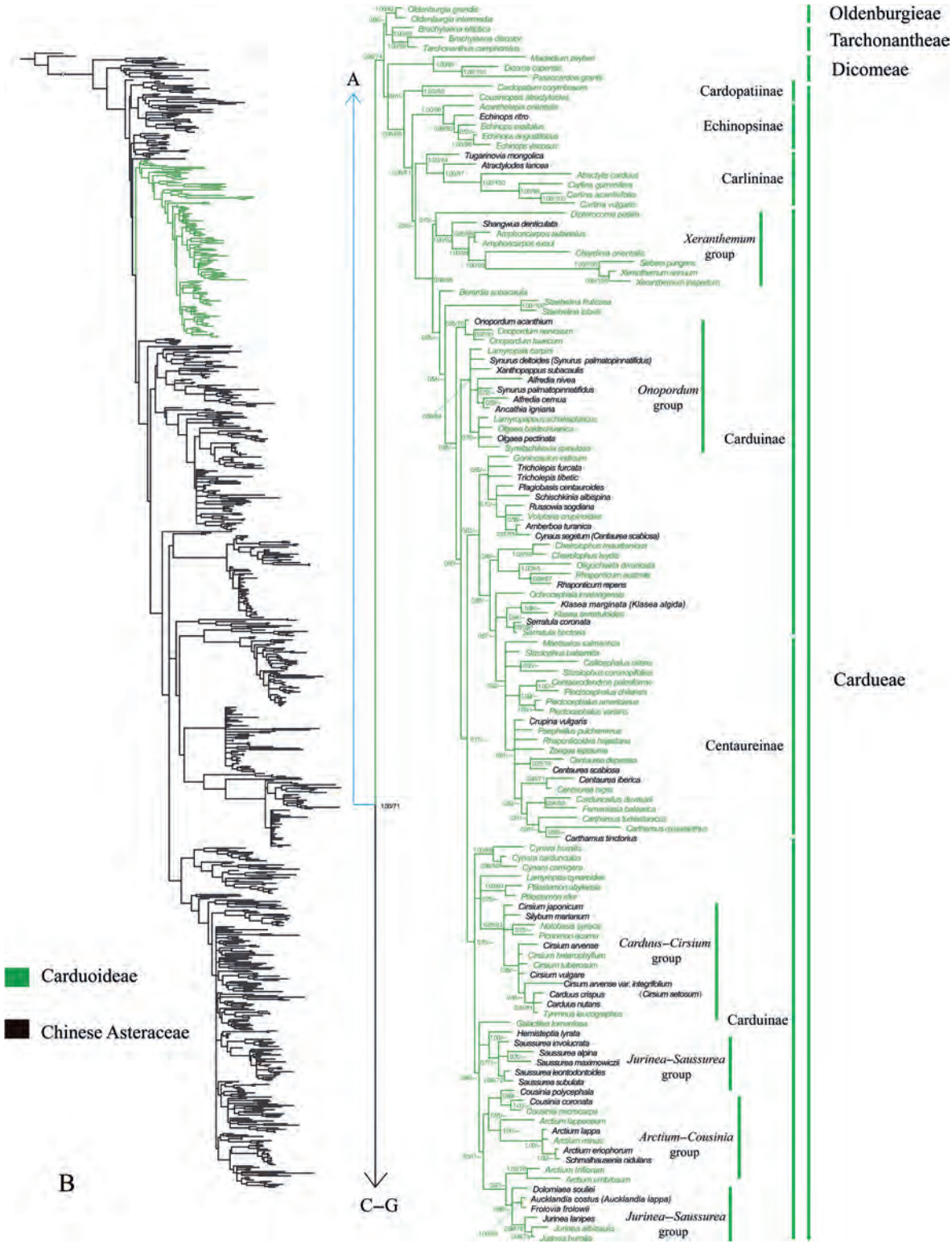


Fig. 2. Continued

Subfamily Carduoideae

The Carduoideae were supported as monophyletic with strong support (PP = 0.98; BS = 74; Figs. 1, 2B). Within Carduoideae, a subclade composed of tribes Oldenburgieae + Tarchonantheae was found sister to a subclade containing tribes Dicomea + Cardueae (PP = 0.98; BS = 74), but interrelationships in the two subclades were poorly resolved (PP = 0.56/0.61). Within Cardueae, Cardopatiinae, Echinopinae, and Carlininae were resolved as successive sisters to Carduinae and Centaureinae (PP > 0.75). Within Carlininae, the BI and ML

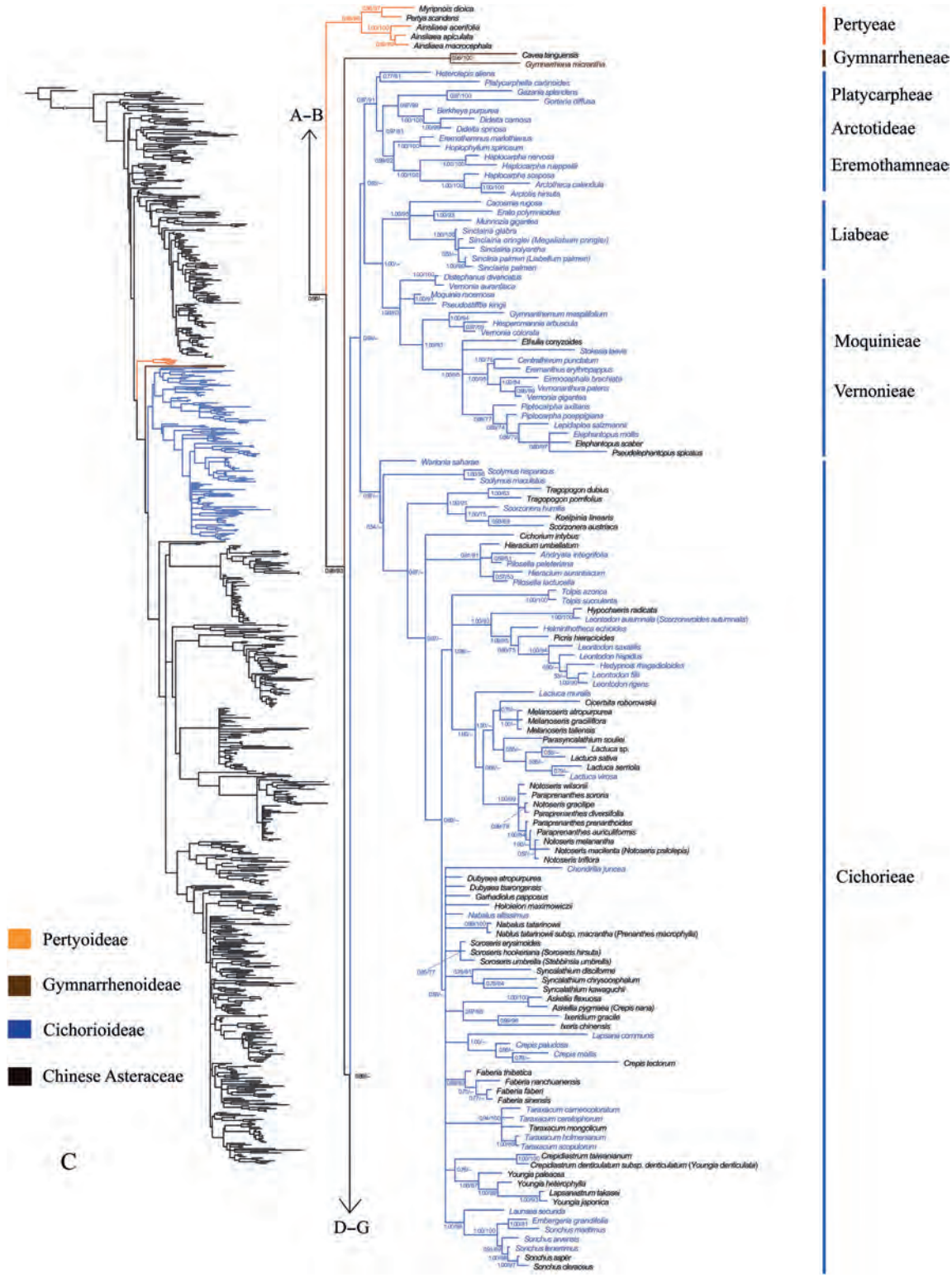


Fig. 2. Continued

hypotheses supported Chinese genera *Tugarinovia* and *Atractylodes* as successive sister group to *Atractylis* + *Carlina* (PP = 1.00). Within Carduinae and Centaureinae, the evidence presented here provided strong support for close relationships between *Aucklandia* + *Frolovia* (PP = 1.00), *Hemisteptia* + *Saussurea* (PP = 1.00), and *Klasea* + *Serratula* (PP = 0.84). The *Arctium–Cousinia* group comprising the representatives of genera *Arctium* and *Schmalhausenia* was supported as

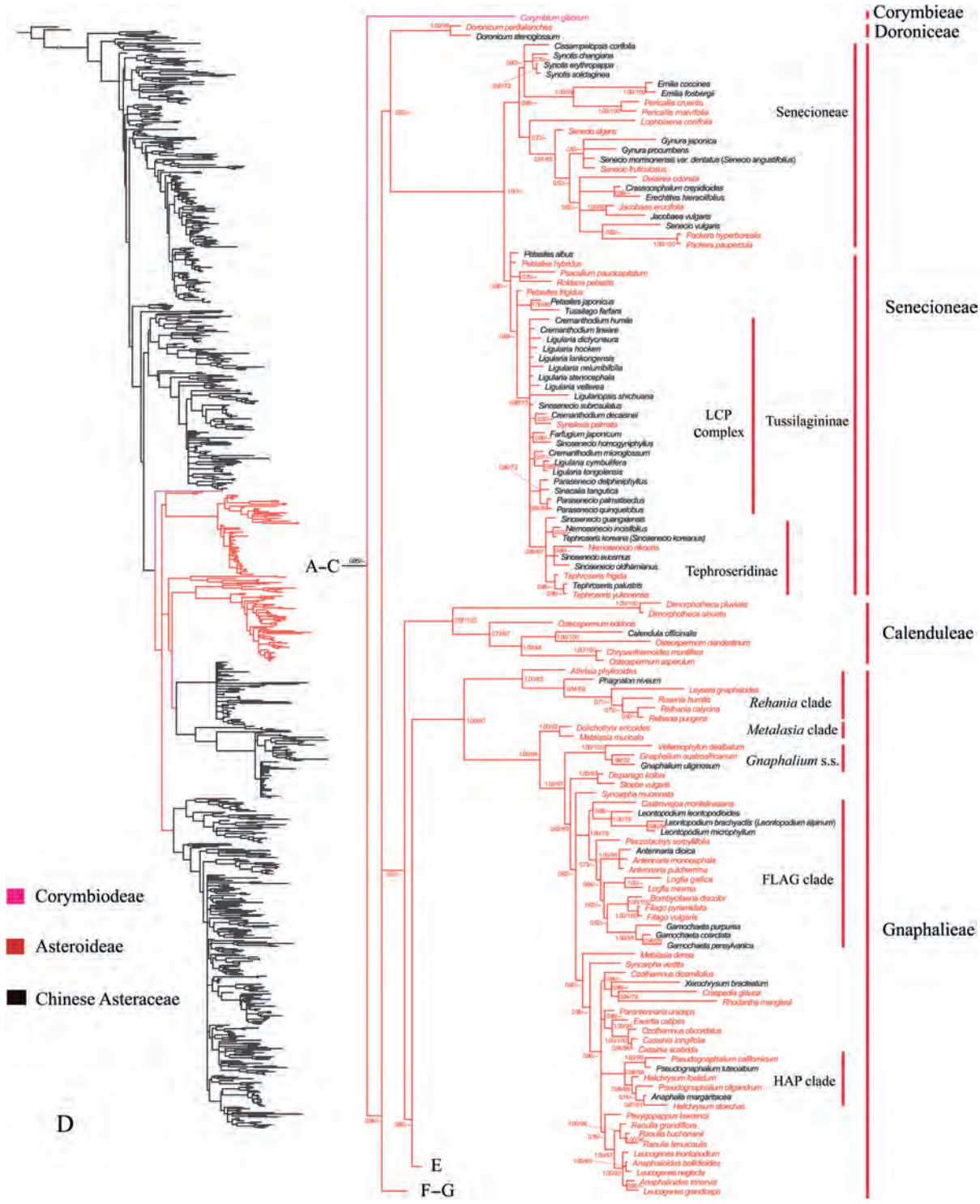


Fig. 2. Continued

monophyletic (PP = 0.98; BS = 74). This group was allied with *Cousinia* with strong support value (PP = 0.95).

Subfamily Pertyoideae

The clade of *Myripnois* and *Pertya* (PP = 0.89; BS = 98; Fig. 2C) was supported as sister to *Ainsliaeae* with strong support (PP = 0.98; BS = 96).

Subfamily Gymnarrhenoideae

The monotypic genus *Cavea* was closely allied to *Gymnarrhena* with strong support (PP = 1.00; BS = 96).

Subfamily Cichorioideae

The Cichorioideae were rendered monophyletic with high support in both analyses (PP = 0.97; BS = 95; Figs., 1, 2C). The

Fig. 2. Continued

subfamily consists of two subclades. Within the first subclade (Figs. 1, 2C), the sister relationship of the tribe Liabeae (PP = 0.97; BS = 95) and Moquinieae + Vernonieae (PP = 1.00) was well supported (PP = 1.00; BS = 76); however, the well supported tribes containing Eremothamneae + Arctotideae + Platycarpheae (PP = 0.97; BS = 91) were sister to former tribes with weak support (PP = 0.65). Within Vernonieae (Fig. 2C), a moderately supported group (PP = 0.66; BS = 70) comprising genera *Pseudelephantopus* + *Elephantopus* + *Piptocarpha* was supported as sister to *Lepidaploa* with moderate support (PP = 0.69; BS = 74). The second subclade was the strongly supported tribe Cichorieae (PP = 0.97; Fig. 2C). *Askellia* and *Ixeris* + *Ixeridium*, *Lapsanastrum*, and *Youngia* were recovered with strong support (PP > 0.97;

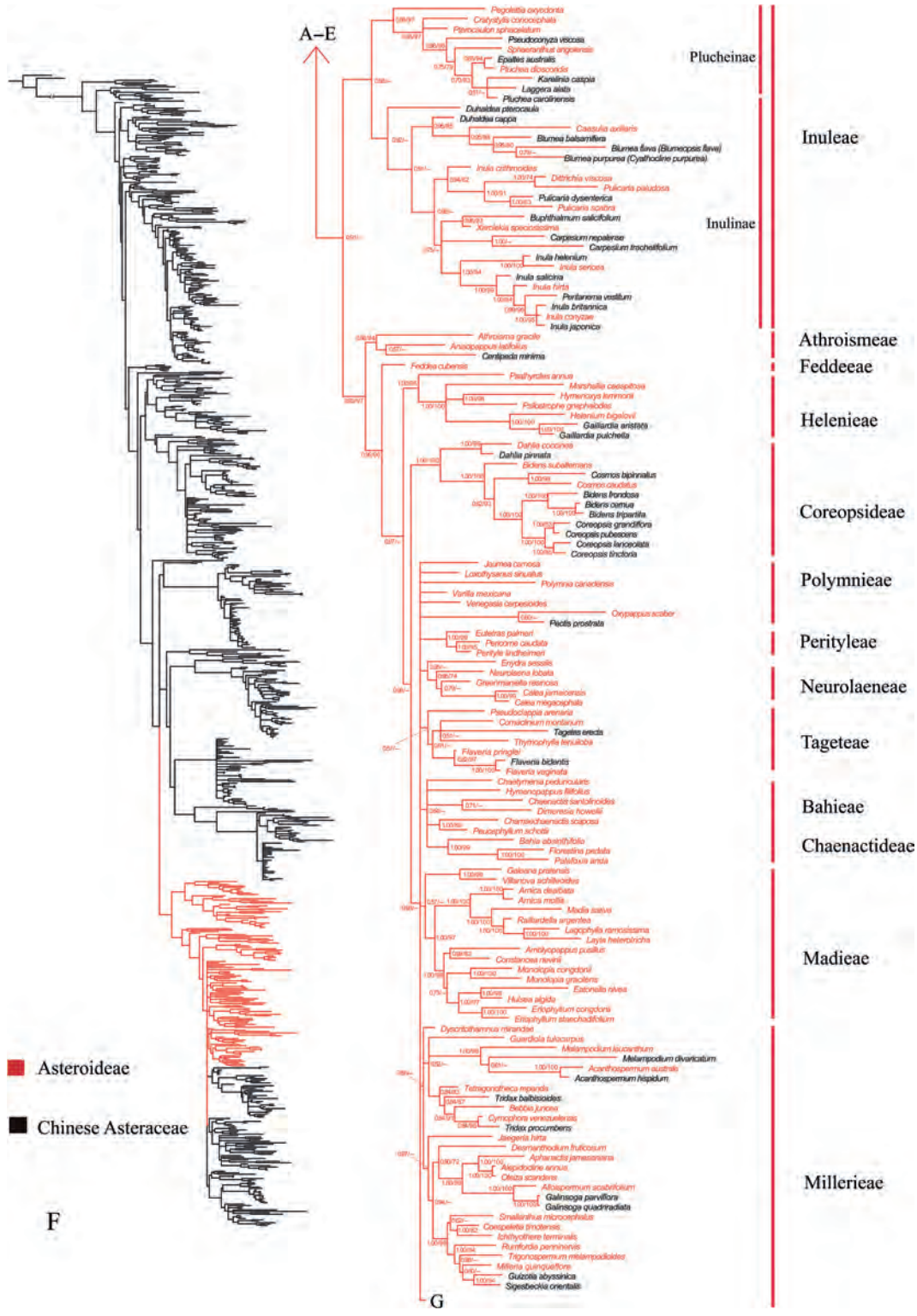


Fig. 2. Continued

BS > 65). *Cicerbita* was sister to *Melanoseris* in the BI analysis with moderate support (PP = 0.78). The monophyly of *Parasyncalathium* and *Lactuca* was poorly supported (PP = 0.65). *Paraprenanthes* and *Notoseris* were supported as monophyletic (PP = 1.00; BS = 79), but neither genus was itself recovered as monophyletic.

Subfamily Asteroideae

Asteroideae were further subdivided into three subclades. However, interrelationships among the three subclades were not resolved (Fig. 1).

Within the first subclade, one additional lineage (Doroniceae) was supported, but the relationships between

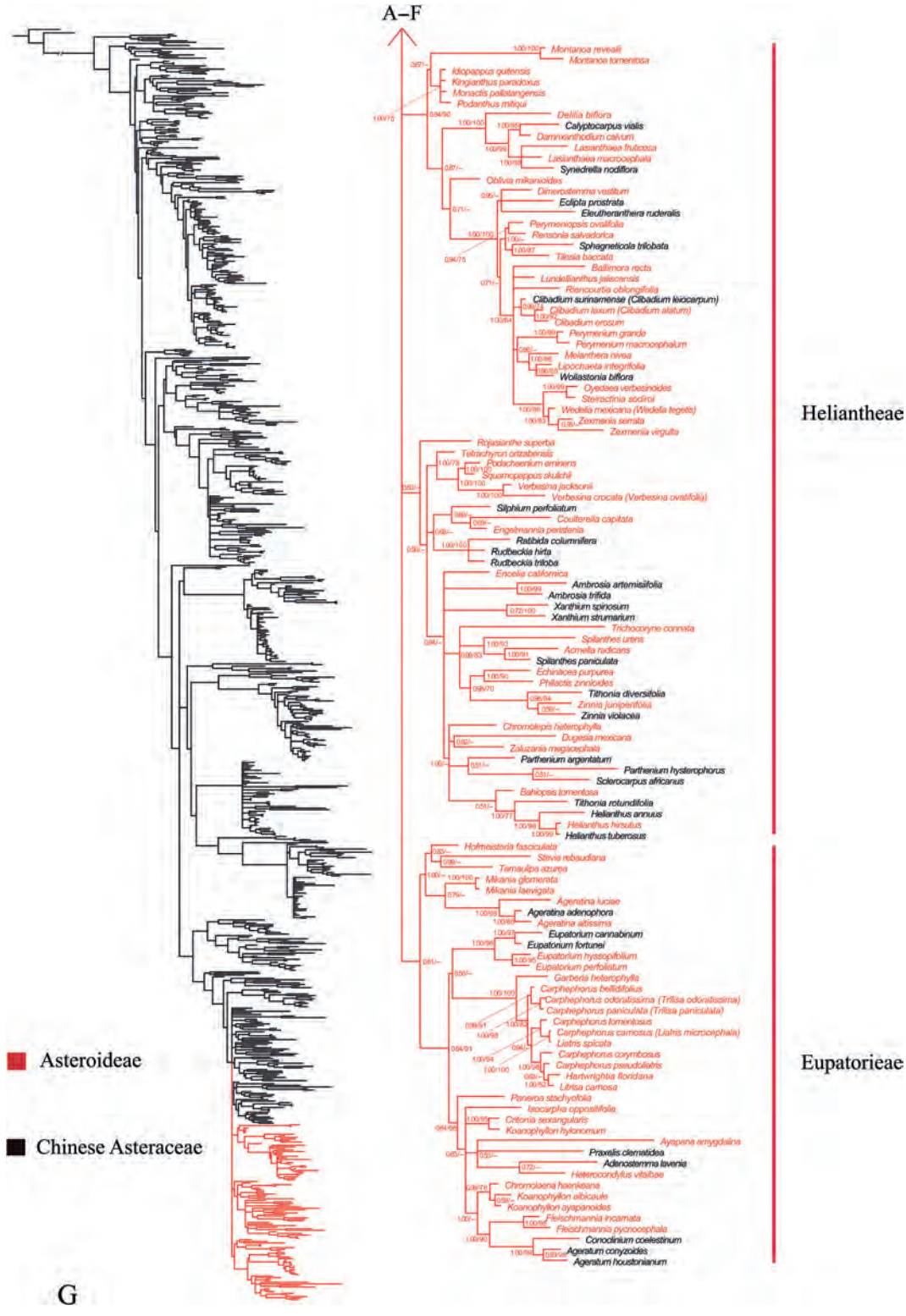


Fig. 2. Continued

Doroniceae and Senecioneae were poorly resolved (PP = 0.52; Figs., 1, 2D). The monophyletic Senecioneae (PP = 1.00) consists of two strongly supported subtribes, Senecioninae (PP = 0.96) and Tussilagininae (PP = 0.90; BS = 60). The sister relationships within Senecioninae were supported as follows, *Emilia* + *Pericallis* (PP = 1.00, BS = 100) and *Crassocephalum* + *Erechtites* (PP = 0.99, BS = 78). However, the phylogenetic relationship of Chinese genera *Synotis* (PP = 0.76) and *Cissampelopsis* + *Emilia* + *Pericallis* received moderate resolution (PP = 0.60). *Senecio* (*sensu* Chen, 1999)

were rendered polyphyletic (Fig. 2D). Within Tussilagininae, the LCP complex (*sensu* Liu et al., 2006) comprising representatives of the genera, such as *Ligularia*, *Cremanthodium*, *Parasenecio*, and *Sinosenecio*, were rendered monophyletic with high support (PP = 0.99; BS = 73). However, each genus was not supported as monophyletic group. Within the LCP complex, *Sinosenecio* + *Nemosenecio* + *Tephroseris* were grouped together with high support (PP = 0.98; BS = 67). Additionally, the polyphyly of *Sinosenecio* was supported.

Within the second subclade, tribes Calenduleae, Gnaphalieae, and Anthemideae were subsequent sisters to Astereae. These placements all received strong support (PP > 0.88). Within Calenduleae, a group of *Calendula*, *Osteospermum*, and *Chrysanthemoides* was resolved with strong support values (PP = 1.00; BS = 98; Fig. 2D). The relationships within Gnaphalieae (Fig. 2D) were resolved with high support indices, *Athrixia* + (*Phagnalon* + *Relhania* + *Leysera*) (PP=1.00; BS = 85), *Xerochrysum* + (*Craspedia* + *Rhodanthe*) (PP = 0.99), *Gnaphalium* and *Vellereophyton* (PP = 1.00), *Helichrysum–Anaphalis–Pseudognaphalium* (HAP clade, PP = 0.99; BS = 88). Within tribe Anthemideae (Fig. 2E), *Ajania*, *Stilpnolepis*, and *Artemisia* were confirmed to be a monophyletic group with weak support (PP = 0.65), although the relationship within the group was collapsed. The phylogenetic relationships of some genera were identified as follows, *Glebionis* + *Argyranthemum*, *Chamaemelum* + *Santolina*, *Leucanthemum* + *Mauranthemum* (PP > 0.95). The tribe Astereae was supported as a monophyletic group (PP = 0.88) (Fig. 2E). However, within the tribe, all samples formed a large polytomy without further resolution (Fig. 2E). *Formania* and *Thespis* were supported as the members of Astereae (Fig. 2E).

Our results indicated that a third subclade (Figs. 1, 2F, 2G) could be divided into two well-supported sister groups, Inuleae (PP = 0.98) and Athroismeae + Heliantheae alliance (PP = 0.93; BS = 97). The Inuleae consists of Plucheinae and Inulinae. Within Plucheinae, *Pseudoconyza* was found sister to the group containing *Sphaeranthus*, *Laggera*, *Pluchea*, *Karelinia*, and *Epaltes* with strong support (PP = 0.98; BS = 96). However, the latter relationships were unambiguously supported. Within Inulinae, *Inula* was retrieved as polyphyletic. The close relationship of *Blumea* + *Caesulia* was strongly supported (PP = 0.95; BS = 88). Within Athroismeae, the close relationship of *Centipeda*, *Anisopappus*, and *Athroisma* received strong support (PP = 0.98; BS = 84). Within the Heliantheae alliance, there were moderate support values for division of the alliance into 13 tribes (including Feddeeae, Helenieae, Coreopsideae, Polymnieae, Perityleae, Neurolaeneae, Tageteae, Bahieae, Chaenactideae, Madieae, Millieae, Eupatorieae, and Heliantheae) (Fig. 2F, 2G; Panero, 2007; Funk et al., 2009c). Within Millerieae and Eupatorieae, the sister relationships of *Melampodium* + *Acanthospermum*, *Galinsoga* + *Alloispermum*, *Sigesbeckia* + *Guizotia*, and *Ageratum* + *Conoclinium* were resolved with high support indices (PP = 1.00). The following sister relationships of the tribe Heliantheae (Fig. 2G) were strongly recovered, *Eleutheranthera* + (*Dimerostemma* + *Eclipta*), *Synedrella* + *Lasianthaea*, *Calyptocarpus* + *Damnxanthodium*, *Sphagneticola* + *Tilesia*, *Wollastonia* + *Lipochae*, *Ratibida* + *Rudbeckia*, and *Spilanthes* + *Acmella* (PP > 0.94).

Discussion

Phylogenetic relationships within Asteraceae

Based on the comprehensive generic-level sampling, the backbone of Asteraceae using three chloroplast markers corresponded well to those recovered by recent studies based on 10 or 14 cpDNA markers (Panero & Funk, 2002, 2008; Panero et al., 2014). In our phylogenetic tree (Figs. 1, 2), 13 clades (subfamilies) were identified and 12 of them were statistically supported in our BI analysis (PP > 0.94), with the exception of Wuderlichioideae (PP = 0.64; BS = 49, the same result with Panero & Funk, 2008). Our analyses provided new insights into some previously ambiguous relationships. Within Carduoideae, several investigators had reported various lineages with uncertain relationships, for example: two sister lineages, Dicomeae + (Oldenburgieae + Tarchonantheae + Cardueae) (Funk et al., 2005); three unresolved lineages, Dicomeae, Cardueae, and (Oldenburgieae + Tarchonantheae) (Panero & Funk, 2008; Funk et al., 2009c); and four unresolved distant lineages, Dicomeae, Cardueae, Oldenburgieae, and Tarchonantheae (Ortiz et al., 2013). Our results support the tribes Oldenburgieae + Tarchonantheae as the sister to the tribes Dicomeae + Cardueae (PP = 0.98; BS = 74; Figs. 1, 2B).

There were still some uncertainties in our results. The monophyly of Gochnatioideae + Stifftioideae + Wunderlichioideae received high support in our BI analysis (PP = 0.97), but their interrelationship was not resolved in our ML analysis (Figs. 1, 2A). The recent molecular phylogenetic study (Panero et al., 2014), in fact, clarified the close sister relationship of Stifftioideae and Wunderlichioideae + Gochnatioideae in the BI analysis. Within Cichorioideae, the sister relationship between Cichorieae and the remaining tribes was well supported by our analysis and recent studies (Funk et al., 2004; Funk & Chan., 2009). However, our analyses indicated that the placements of remaining tribes, such as Arctotideae, Liabeae, Eremothamneae, and *Heterolepis*, were still in doubt (Fig. 2C).

Systematics of Chinese Asteraceae

The analyses supported 13 clades (13 subfamilies, including 45 tribes) (Fig. 1, shown in different colors). Chinese Asteraceae were not monophyletic and were placed into seven major robust clades (subfamilies), Mutisioideae, Wunderlichioideae, Carduoideae, Pertyoideae, Gymnarrhenoideae, Cichorioideae, and Asteroideae, and 22 tribes, Mutiseae, Hyalideae, Cardueae, Pertyeae, Gymnarrheneae, Vernonieae, Cichorieae, Doroniceae, Senecioneae, Astereae, Anthemideae, Gnaphalieae, Calenduleae, Inuleae, Athroismeae, Helenieae, Coreopsideae, Neurolaeneae, Tageteae, Millieae, Eupatorieaea, and Heliantheae (Fig. 2). Chinese Asteraceae lacked 6 subfamilies Barnadesioideae, Famatinanthoideae, Gochnatioideae, Stifftioideae, Corymbiodeae, and Hecastocleidoideae and 23 tribes, Barnadesieae, Famatinantheae, Onoserideae, Nassauvieae, Gochnatieae, Stifftieae, Wunderlichieae, Hecastocleideae, Oldenburgieae, Tarchonantheae, Dicomeae, Moquinieae, Liabeae, Arctotideae, Eremothamneae, Platycarpheae, Corymbieae, Feddeeae, Polymnieae, Perityleae, Madieae, Chaenactideae, and Bahieae (names of subfamilies and tribes following previous studies, Panero & Funk, 2002, 2008; Panero, 2005; Funk et al., 2009c; Panero et al., 2014). Details of the phylogenetic relationships within major Chinese clades were discussed below.

Subfamily Mutisioideae

The Mutisioideae include 630 species and ca. 44 genera in three tribes (Onoserideae, Nassauvieae, and Mutisieae; Figs. 1, 2A). The tribe Mutisieae contains ca. 14 genera and over 200 species.

Tribe Mutiseae

Mutisioideae are poorly represented in China with only 1 tribe Mutisieae (PP = 1.00; BS = 99; Figs. 1, 2A), 3 genera, and 13 indigenous species (three endemic spp.). The monophyly of *Gerbera* received strong support as sister to the East Asian–North American disjunct *Leibnitzia* (PP = 1.00; BS = 100; Fig. 2A), confirming previous hypotheses (Baird et al., 2010; see also Wen et al., 2010). The placement of *Adenocaulon* was unresolved within Mutiseae, as shown by Kim et al. (2002).

Subfamily Wunderlichioideae

The Wunderlichioideae consist of two tribes, Wunderlichieae and Hyalideae (Figs. 1, 2A). The family (8 genera and 42 species) is disjunctly distributed in northeastern South America and southwestern China.

Tribe Hyalideae

Wunderlichioideae are poorly represented in China with one tribe Hyalideae (Fig. 2A), two genera (*Nouelia* and *Leucomeris*), and two indigenous species. Only *Nouelia insignis* Franch. is endemic to China. Within Hyalideae, a group comprising the sister genera *Nouelia* and *Leucomeris* was supported as sister to the South American genera *Ianthopappus* + *Hyalis* (PP = 1.00; BS = 100; Fig. 2A). The relationship was also supported by recent analyses (e.g., Kim et al., 2002; Panero & Funk, 2008). Therefore, the treatment of placing *Nouelia* and *Leucomeris* into Mutisieae (e.g., Hind, 2007; Gao & Hind, 2011) needs to be revised.

Subfamily Carduoideae

The Carduoideae consist of ca. 2850 species and 85 genera in 4 tribes (Garcia-Jacas et al., 2002; Funk et al., 2005; Susanna & Garcia-Jacas, 2007, 2009; Ortiz et al., 2013; Figs. 1, 2B).

Tribe Cardueae

The Carduoideae are represented in China with only one tribe Cardueae (Fig. 2B), four subtribes (Echinopsinae, Carliniaea, Carduinae, and Centaureinae), except for Cardopatiinae (Fig. 2B, Table 3), 41 genera, and ca. 464 species (244 endemic spp.). They are also the most morphologically diverse and species-rich tribe in China. Within Carlininae, Chinese genera *Tugarinovia* and *Atractylodes* were supported as successive sisters to *Atractylis* + *Carlina* (PP = 1.00; BS = 84; Fig. 2B), which was congruent with the findings reported by Susanna et al. (2006) and Barres et al. (2013). The Carduinae includes seven groups (*sensu* Susanna & Garcia-Jacas, 2009). Five of seven groups are represented in China (Table 3). Within Carduinae, the close relationship of *Shangwua* and *Xeranthemum* was supported by our analysis and recent studies (Wang et al., 2009b; Wang et al., 2013a). The close relationships of *Cousinia* + (*Arctium* + *Schmalhausenia*) and *Aucklandia* + *Frolovia*, previously inferred by López-Vinyallonga et al. (2009) and Wang et al. (2007), separately, were also supported in our phylogenetic inferences (PP > 0.95; Fig. 2B). Our results also showed that samplings of the morphological diversity of *Saussurea* (ca. 300 spp. in China and 400 spp. in the world) were still far from complete. The intrarelationships and interrelationships between *Saussurea* and related genera (e.g., the Chinese monotypic genus *Bolocephalus*) remained unresolved (Raab-Straube, 2003; Kita et al., 2004; Wang & Liu, 2004; Wang et al., 2009b). Therefore, the full classification of the *Saussurea* is in need of revision. Within Centaureinae, *Serratula* and *Klasea* were supported as monophyletic (PP = 0.84; Fig. 2B), as stated by Barres et al. (2013).

Subfamily Pertyoideae

The Pertyoideae consist of one tribe (Pertyeae), four genera, and ca. 80 species distributed only in Asia.

Tribe Pertyeae

The Pertyeae is a well-represented tribe in China with three genera and ca. 58 species (45 spp. endemic). Based on incomplete morphological studies (Cabrera, 1977; Hind, 2007; Katinas et al., 2008; Gao et al., 2011), the genera *Ainsliaea*, *Pertya*, and *Myripnois* were previously treated as members of tribe Mutisieae. However, our results (Figs. 1, 2C) and recent molecular studies (Kim et al., 2002; Panero & Funk, 2002, 2008; Mitsui et al., 2008) showed that these genera form a distinct clade (Pertyoideae and Pertyeae, recognized by Panero & Funk, 2002) nested above the Carduoideae and the monophyly of *Ainsliaea* and *Pertya* + *Myripnois* received strong support (PP = 0.98; BS = 96; Fig. 2C). Furthermore, *Myripnois*, a genus endemic to North China, was embedded within the genus *Pertya*. These two genera were very similar in gross morphology (e.g., shrub, dioecious, capitula solitary, terminal on branchlets, subsessile or with short peduncle, Gao et al., 2011). Further sampling of more species will certainly contribute to the redefinition of the two genera.

Subfamily Gymnarrhenoideae

The Gymnarrhenoideae include only one tribe and two monotypic genera, *Gymnarrhena* and *Cavea* (Fig. 2C).

Tribe Gymnarrheneae

The present result (PP = 0.98; BS = 96; Fig. 2C) and Anderberg & Ohlson (2012) strongly supported the monophyly of *Cavea* and *Gymnarrhena*. *Gymnarrhena* is a rosulate and dwarf desert annual herb, which is mainly distributed in North Africa and the Middle East. *Cavea* is a perennial herb with branched stems that grows on gravelly ground near streams and glaciers of high mountains in the Himalaya area. There were no obvious habitat or morphological characters between *Cavea* and *Gymnarrhena* to support the monophyly of Gymnarrhenoideae, although the two genera might share an important synapomorphy, that is, two types of flowers (capitula) with a tendency towards dioecism (Anderberg & Ohlson, 2012).

Subfamily Cichorioideae

The Cichorioideae include ca. 2900 species and ca. 250 genera in seven described tribes (Cichorieae, Vernonieae, Arctotideae, Liabeae, Platycarpheae, Eremothamneae, and Moquinieae) and one unplaced genus *Heterolepis* (Figs. 1, 2C; Funk & Chan, 2009). Two tribes (Vernonieae and Cichorieae), 41 genera, and 426 species are indigenous to China (199 endemic spp.).

Tribe Vernonieae

Six genera and 39 species are indigenous to China (10 spp. endemic). Both chloroplast and nuclear DNA datasets strongly supported the sister relationship between *Elephantopus* and

Table 3 Revised taxonomy of the Chinese Asteraceae at the generic level

Subfamily 1. Mutisioideae (Cass.) Lindl. [1829]. (3 genera).
Tribe 1. Mutisieae Cass. [1819]. (3 genera).
Adenocaulon Hook., *Leibnitzia* Cass., *Gerbera* L.
Subfamily 2. Wunderlichioideae Panero & V. A. Funk [2007]. (2 genera).
Tribe 2. Hyalideae Panero [2007]. (2 genera).
Leucomeris D. Don, *Nouelia* Franch.
Subfamily 3. Carduoideae Cass. ex Sweet [1829]. (41 genera).
Tribe 3. Cardueae Cass. [1819]. (41 genera).
Subtribe 3.1 Echinopsinae (Cass.) Dumort.
Echinops L.
Subtribe 3.2 Carlininae Dumort.
Atractylodes DC.,*Carlina* L., *Tugarinovia* Iljin
Subtribe 3.3 Carduinae (Cass.) Dumort.
Jurinea–Saussurea group
Aucklandia Falc.,† *Bolocephalus* Hand.-Mazz., *Dolomiaea* DC., *Frolovia* (DC.) Lipsch., *Hemisteptia* Bunge ex Fischer & C. A. Meyer, *Himalaiella* Raab-Straube, *Jurinea* Cass., *Saussurea* DC.
Arctium–Cousinia group
Arctium L., *Cousinia* Cass., *Schmalhausenia* C. Winkl.
Onopordum group
Alfredia Cass., *Ancathia* DC., *Synurus* Iljin, *Syreitschikovia* Pavlov, *Olgaea* Iljin, *Onopordum* L., *Xanthopappus* C. Winkl.
Carduus–Cirsium group
Carduus L., *Cirsium* Mill.
Xeranthemum group
Shangwua Yu J. Wang, Raab-Straube, Susanna & J. Quan Liu
Subtribe 3.4 Centaureinae (Cass.) Dumort.
Amberboa Vaill., *Archiserratula* L. Martins, *Carthamus* L.,† *Centaurea* L.,‡ *Crupina* (Pers.) DC., *Cyanus* Mill.,† *Klasea* Cass., *Oligochaeta* (DC.) K. Koch, *Plagiobasis* Schrenk, *Psephellus* Cass., *Rhaponticoides* Vaill., *Rhaponticum* Vaill., *Russowia* C. Winkl., *Schischkinia* Iljin, *Serratula* L., *Tricholepis* DC.
Subfamily 4. Pertyoideae Panero & V. A. Funk [2002]. (3 genera).
Tribe 4. Pertyeae Panero & V. A. Funk [2002]. (3 genera).
Ainsliaea DC., *Myripnois* Bunge, *Pertya* Sch.-Bip.
Subfamily 5. Gymnarrhenoideae Panero & V. A. Funk [2002]. (1 genus).
Tribe 5. Gymnarrheneae Panero & V. A. Funk [2002]. (1 genus).
Cavea W. W. Smith & J. Small
Subfamily 6. Cichorioideae (Juss.) Chev. [1828]. (42 genera).
Tribe 6. Vernonieae Cass. [1819]. (6 genera).
Camchaya Gagnep., *Distephanus* Cass., *Elephantopus* L., *Ethulia* L.f., *Pseudelephantopus* Rohr,† *Vernonia* Schreb.‡
Tribe 7. Cichorieae Lam. & DC. [1806]. (36 genera).
Askellia W. A. Weber, *Cicerbita* Wallr., *Crepidiastrum* Nakai, *Crepis* L., *Dubyaea* DC., *Epilasia* (Bunge) Benth., *Faberia* Hemsl., *Garhadiolus* Jaub. & Spach, *Chondrilla* L., *Cichorium* L.,† *Heteracia* Fisch. & C. A. Mey., *Hieracium* L., *Hololeion* Kitam., *Hypochaeris* L.,‡ *Ixeridium* (A. Gray) Tzvelev, *Ixeris* (Cass.) Cass., *Koelpinia* Pall., *Lactuca* L.,‡ *Lapsanastrum* J. H. Pak & K. Bremer, *Launaea* Cass., *Melanoseris* Decne., *Nabalus* Cass., *Notoseris* C. Shih, *Paraprenanthes* C. C. Chang ex C. Shih, *Parasyncalathium* J. W. Zhang, Boufford & H. Sun, *Picris* L., *Pilosella* Vaill., *Podospermum* DC., *Scorzonera* L., *Sonchella* Sennikov, *Sonchus* L.,‡ *Soroseris* Stebbins, *Syncalathium* Lipsch., *Taraxacum* F. H. Wiggers,‡ *Tragopogon* L.,‡ *Youngia* Cass.
Subfamily 7. Asteroideae (Cass.) Lindl. [1829]. (163 genera).
Tribe 8. Doroniceae Panero [2005]. (1 genus).
Doronicum L.
Tribe 9. Senecioneae Cass. [1819]. (23 genera).
Subtribe 9.1 Tussilagininae s. str. clade
Cremanthodium Benth., *Dicercoclados* C. Jeffrey & Y. L. Chen, *Farfugium* Lindl., *Ligularia* Cass., *Ligulariopsis* Y. L. Chen, *Nemosenecio* (Kitam.) B. Nord., *Parasenecio* W. W. Smith & J. Small, *Petasites* Mill., *Sinacalia* H. Rob. & Brettell, *Sinosenecio* B. Nord., *Syneilesis* Maxim., *Tephroseris* (Reichenb.) Reichenb., *Tussilago* L.
Subtribe 9.2 Senecioninae
Cissampelopsis (DC.) Miq., *Crassocephalum* (DC.) Miq.,† *Emilia* Cass.,‡ *Erechtites* Raf.,† *Gynura* Cass., *Hainanecio* Y. Liu & Q. E. Yang, *Jacobaea* Mill., *Pericallis* D. Don,† *Senecio* L., *Synotis* (C. B. Clarke) C. Jeffrey & Y. L. Chen
Tribe 10. Calenduleae Cass. [1819]. (1 genus).
Calendula L.†
Tribe 11. Gnaphalieae (Cass.) Lecoq & Juillet [1831]. (11 genera).
Anaphalis DC., *Antennaria* Gaertn., Filago L., *Gamochaeta* Wedd.,‡ *Gnaphalium* L., *Gnomophalium* Greuter, *Helichrysum* Mill., *Leontopodium* R. Br. ex Cass., *Phagnalon* Cass., *Pseudognaphalium* Kirp., *Xerochrysum* Tzvelev†
Tribe 12. Astereae Cass. [1819]. (29 genera).

Continued

Table 3 Continued

African lineages
Bellis L.,† *Crinitina* Soják, *Galatella* Cass., *Grangea* Adans., *Tripolium* Nees, *Nannoglottis* Maxim.
Australasian lineages
Aster L., *Asterothamnus* Novopokr., *Arctogeron* DC., *Callistephus* Cass., *Calotis* R. Br., *Eschenbachia* Moench, *Formania* W.W. Smith & J. Small, *Heteroplexis* C.C. Chang, *Lagenophora* Cass., *Myriactis* Less., *Neobrachyactis* Brouillet, *Psychrogeton* Boiss., *Sheareria* S. Moore, *Thespis* DC., *Turczaninovia* DC., *Rhinactinidia* Novopokr.
North American lineages
Erigeron L.,‡ *Eurybia* (Cass.) Cass., *Grindelia* Willd.,† *Solidago* L.,‡ *Symphyotrichum* Nees‡
Unplaced genera
Dichrocephala L'Hér. ex DC., *Microglossa* DC.
Tribe 13. Anthemideae Cass. [1819]. (29 genera).
Southern Hemisphere grade
Subtribe 13.1 Cotulinae Kitt.
Cotula L., *Soliva* Ruiz & Pav.†
Asian–South African grade
Subtribe 13.2 Artemisiinae Less.
Ajania Poljakov, *Artemisia* L., *Brachanthemum* DC., *Chrysanthemum* L., *Crossostephium* Less., *Elachanthemum* Y. Ling & Y. R. Ling, *Filifolium* Kitam., *Hippolytia* Poljakov, *Kaschgaria* Poljakov, *Leucanthemella* Tzvelev, *Microcephala* Pobed., *Neopallasia* Poljakov, *Stilpnolepis* Krasch.
Genera of the Asian–South African grade unassigned to a subtribe
Ajaniopsis C. Shih, *Cancrinia* Kar. & Kir., *Opisthopappus* C. Shih
Subtribe 13.3 Handeliinae Bremer & Humphries
Allardia Decne, *Handelia* Heimerl, *Pseudohandelia* Tzvelev, *Richteria* Kar. & Kir.
Eurasian grade
Subtribe 13.4 Matricariinae Willk.
Achillea L.,‡ *Matricaria* L.
Subtribe 13.5 Anthemidinae (Cass.) Dumort.
Anthemis L.,† *Tanacetum* L.,‡ *Tripleurospermum* Sch.-Bip.
Mediterranean clade
Subtribe 13.6 Leucantheminae Bremer & Humphries
Leucanthemum Mill.†
Subtribe 13.7 Glebionidinae Oberprieler & Vogt
Glebionis Cass.†
Tribe 14. Inuleae Cass. [1819]. (14 genera).
Subtribe 14.1 Inulinae Dumort.
Blumea DC., *Buphthalmum* L.,† *Carpesium* L., *Duhaldea* DC., *Inula* L., *Pentanema* Cass., *Pulicaria* Gaertn.‡
Subtribe 14.2 Plucheinae Dumort.
Epaltes Cass., *Karelinia* Less., *Laggera* Sch.-Bip. ex Benth. & J. D. Hook., *Pluchea* Cass.,‡ *Pseudoconyza* Cuatr., *Pterocaulon* Ell., *Sphaeranthus* L.
Tribe 15. Athroismeae Panero [2002]. (3 genera).
Anisopappus Hook. & Arnott, *Centipeda* Lour., *Symphyllocarpus* Maxim.
Heliantheae alliance (tribes 16–22)
Tribe 16. Helenieae (Cass.) Lindl. [1826]. (1 genus).
Helenium L.†
Tribe 17. Coreopsideae Lindl. [1829]. (4 genera).
Bidens L.,‡ *Cosmos* Cav.,† *Coreopsis* L.,† *Dahlia* Cav.†
Tribe 18. Neurolaeneae Rydb. [1927]. (1 genus).
Enydra Lour.
Tribe 19. Tageteae Cass. [1819]. (5 genera).
Flaveria Juss.,† *Glossocardia* Cass., *Pectis* L.,† *Tagetes* L.,† *Dyssodia* Cav.†
Tribe 20. Millieae Lindl. [1829]. (8 genera).
Acanthospermum Schrank,† *Blainvillea* Cass., *Galinsoga* Ruiz & Pav.,† *Guizotia* Cass.,‡ *Sigesbeckia* L., *Smallanthus* Mack.,† *Tridax* L.,† *Melampodium* L.†
Tribe 21. Eupatorieae Cass. [1819]. (10 genera).
Adenostemma J. R. Forst. & G. Forst., *Ageratina* Spach,† *Ageratum* L.,† *Austroeupatorium* R. M. King & H. Rob.,† *Chromolaena* DC.,† *Conoclinium* DC.,† *Eupatorium* L., *Gymnocoronis* DC.,† *Mikania* Willd.,‡ *Praxelis* Cass.†
Tribe 22. Heliantheae Cass. [1819]. (23 genera).
Acmella Pers.,‡ *Ambrosia* L.,† *Calyptocarpus* Less.,† *Clibadium* F. Allam. ex L.,† *Eleutheranthera* Poit. ex Bosc.,† *Eclipta* L.,† *Gaillardia* Foug.,† *Helianthus* L.,† *Lagascea* Cav.,† *Melanthera* Rohr, *Parthenium* L.,† *Rudbeckia* L.,† *SanvitaliaLam.*,† *Sclerocarpus* Jacq.,† *Silphium* L.,† *Sphagneticola* O. Hoffm.,‡ *Synedrella* Gaertn.,† *Tithonia* Desf. ex Juss.,† *Wollastonia* DC. ex Decaisne, *Xanthium* L.,† *Heliopsis* Pers.,† *Ratibida* Raf.,† *Zinnia* L.†

Names of subfamilies, tribes, subtribes, clades, grades, lineages, and groups are in bold. Publication dates are in brackets. †Genera containing only introduced species. ‡Indigenous and introduced species.

Chrysolaena + *Lepidaploa* + *Lessingianthus* (Keeley et al., 2007). A moderate support lineage consisting of *Lepidaploa* + (*Elephantopus* + *Pseudelephantopus*) was also observed in our study (PP = 0.66; BS = 70; Fig. 2C).

Tribe Cichorieae

With ca. 95 genera and 2500 species, the Cichorieae is mainly distributed in the Northern Hemisphere (Shih et al., 2011). Based on a recent analysis, Kilian et al. (2009) recognized 11 subtribes. It is the third most species-rich tribe of Asteraceae in China. Eight subtribes, 35 genera, and 387 species are indigenous to China (the great majority, ca. 189 spp., endemic). The monophyly of *Soroseris* and *Stebbinsia* was supported in our analysis (PP = 0.95; BS = 77; Fig. 2C) and previous study (Zhang et al., 2011b). Inclusion of *Lapsanastrum* in *Youngia* observed in our result (PP = 1.00; BS = 87) were in accord with the recently published results from nuclear ribosomal internal transcribed space (nrITS) analyses (Deng et al., 2014). In addition, *Askellia* + (*Ixeris* + *Ixeridium*) were recovered as a monophyletic group (PP = 0.97; BS = 66; Fig. 2C). A monophyletic group including *Cicerbita* and *Melanoseris* was recovered with moderate support (PP = 0.78). *Parasyncalathium* was loosely allied with *Lactuca* in a weakly supported clade (PP = 0.65). More studies are needed to determine the taxonomic status of its generic affiliation. *Paraprenanthes* and *Notoseris* were nested within an unresolved trichotomy from our analysis. Based on the analysis of other multiple five cpDNA and nrITS data, Wang et al. (2013b) speculated that the monophyly of *Paraprenanthes* and *Notoseris* might be the result of introgressive hybridization.

Subfamily Asteroideae

Asteroideae are the largest subfamily of Asteraceae, comprising 22 tribes (Fig. 1), ca. 1150 genera, and 16 000 species (data from http://angio.bergianska.se/). Within Asteroideae, ca. 163 genera and 1400 species are indigenous to China (emend from Shih et al., 2011). Fifteen tribes (Doroniceae, Senecioneae, Astereae, Anthemideae, Gnaphalieae, Calenduleae, Inuleae, Athroismeae, Helenieae, Coreopsideae, Neurolaeneae, Tageteae, Millieae, Eupatorieae, and Heliantheae) are represented in China (ca. 66% tribes and 60% genera of Chinese Asteraceae; Fig. 2D–2G).

Tribe Doroniceae

The Doroniceae (Panero, 2005) includes only one genus with ca. 40 species in Eurasia and Northern Africa. It is represented in China by seven species (four endemic spp.). The position of *Doronicum* varied. It was treated either as the members of tribe Senecioneae (e.g., Jeffrey & Chen, 1984; Álvarez Fernández et al., 2001) or as a separate and uncertain clade of Asteroideae (Goertzen et al., 2003; Pelser et al., 2007; Nordenstam et al., 2009). The present BI analysis strongly supported the second alternative (Fig. 2D; PP = 1.00). It was also suggested here to reinstate *Doronicum* as an independent tribe (*sensu* Panero, 2005), because *Doronicum* might occupy a basal position of Senecionodae and Asterodae (C. F. Zhang et al., Fudan University, Shanghai, pers. comm.).

Tribe Senecioneae

The Senecioneae, with an estimated 150–170 genera and 3500 species, is the largest tribe of Asteraceae. In China, the Senecioneae is the second most species-rich tribe, consisting of two subtribes, 23 genera, and 457 species (the great majority, ca. 311 spp., endemic). The subtribe Senecioninae was supported as sister to Tussilagininae (PP = 1.00; Fig. 2D), as in previous analyses (Pelser et al., 2007; Nordenstam et al., 2009). Within Senecioninae, the close relationships of *Emilia* + *Pericallis* (PP = 1.00, BS = 100) and *Crassocephalum* + *Erechtites* (PP = 0.99, BS = 78) supported in our analyses were in agreement with previous results reported by Pelser et al. (2007, 2010). The Chinese *Synotis*, *Cissampelopsis*, *Emilia*, and *Pericallis* formed a moderately supported monophyletic group (PP = 0.60; Fig. 2D). The present result, as in Pelser et al. (2007, 2010), showed strong phylogenetic divergence within the polyphyletic *Senecio* (*sensu* Chen, 1999).

The Tussilagininae includes species with an almost exclusively East Asian distribution. The monophyly of the LCP complex (*sensu* Liu et al., 2006, ca. 12 genera and 400 species) was supported in the present analysis (PP = 0.99; BS = 73; Fig. 2D), as indicated by some recent analyses (Golden et al., 2001; Liu et al., 2006; Pelser et al., 2007; Wang et al., 2009a). However, in our analysis (Fig. 2D), some genera of the LCP complex, that is, *Ligularia* (ca. 140 spp.), *Cremanthodium* (ca. 70 spp.), *Parasenecio* (ca. 60 spp.), and *Sinosenecio* (ca. 41 spp.), as described by Jeffrey & Chen (1984) and Chen (1999), were not monophyletic (Fig. 2D). These findings indicated that a number of generic problems exist in the current classification of Tussilagininae. An enhanced sampling of the LCP complex is needed to resolve their generic affiliation. Tephroseridinae (*Sinosenecio–Nemosenecio–Tephroseris*) was nested within the Tussilagininae and the species of *Sinosenecio* had different positions in the present trees (Fig. 2D), consistent with results of recent analyses (Nordenstam et al., 2009; Wang et al., 2009a).

Tribe Calenduleae

The Calenduleae consists of 12 genera and ca. 120 species, which are mainly distributed in southern Africa (80% spp.). The tribe is poorly represented in China with one introduced genus (one species, *Calendula officinalis* L.; Fig. 2D). The close relationship among *Calendula*, *Osteospermum*, and *Chrysanthemoides* was resolved with strong support values (PP =n1.00; BS = 98; Fig. 2D), congruent with the finding reported by Nordenstam & Kallersjo (2009).

Tribe Gnaphalieae

The Gnaphalieae is a moderately large tribe with ca. 185 genera and 1240 species. There are only a few taxa of the Gnaphalieae in the Northern Hemisphere (Anderberg, 1991; Bayer et al., 2007). Therefore, Gnaphalieae is poorly represented in China with 11 genera and 120 species (ca. 62 endemic spp.). Our result (PP = 1.00; BS = 85; Fig. 2D) and recent molecular phylogenies (e.g., Ward et al., 2009; Nie et al., 2015) consistently supported the *Relhania* clade as the basal group, which included a representative group of genera *Athrixia*, *Leysera*, *Relhania*, and Chinese *Phagnalon*. Within the crown radiation group, the monophyly of the HAP clade was strongly supported by our analysis (PP = 0.99; BS = 88; Fig. 2D). Given that *Anaphalis* and *Pseudoganaphalium* rendered *Helichrysum* paraphyletic, Nie et al. (2013) and Galbany-Casals et al. (2014) suggested that the traditional generic concept of *Helichrysum* was not supported. Furthermore, the monophyly of *Anaphalis* was weakly supported (Nie et al., 2013). Within the *Filago*, *Leontopodium*, *Antennaria*,

Gamochaeta (FLAG) clade (Fig. 2D), the monophyly of *Leontopodium* (ca. 58 spp., 37 in China), hypothesized by Blöch et al. (2010), was supported in our analysis. Safer et al. (2011) identified 10 groups of *Leontopodium*, however, the infrageneric relationships were not fully resolved. Their (*Anaphalis* and *Leontopodium*) taxonomic status merits further study. The sister relationship of *Gnaphalium* and *Vellereophyton* was strongly supported in our analysis (PP = 1.00; BS = 100; Fig. 2D), in accordance with the recent analysis (Smissen et al., 2011). The monophyly of *Xerochrysum* and *Craspedia* + *Rhodanthe* was recovered with high support (PP = 0.99).

Tribe Astereae

With approximately 225 genera and 3100 species, the tribe Astereae is the second largest tribe of Asteraceae. Twenty-nine genera and 237 species are indigenous to China (112 spp. endemic) (emend from Ling et al., 1985b). Due to the limited value of the three cpDNA markers and samplings, support was not sufficient to separate different genetic clusters within the tribe (Figs. 1, 2E). However, we found that the enigmatic *Formania* was deeply nested within the tribe Astereae. The systematic position of monotypic and endemic genus *Formania* has puzzled taxonomists for a long time. Our analysis indicates that it should be a member of the tribe Astereae (Chen & Brouillet, 2011) (Fig. 2E), not the tribe Anthemideae as suggested by Shih & Fu (1983). Additionally, based on cpDNA analysis, *Thespis* was also imbedded in Astereae, as proposed from nrDNA data by Zhong et al. (2014). Recently, Brouillet et al. (2009) hypothesized that Asian *Aster* (ca. 152 species worldwide, with ca. 123 spp. distributed in China) and allies were nested in the Australasian lineages (Table 3). Li et al. (2012) showed that *Aster* was paraphyletic and several allies (e.g., genera *Kalimeris*, *Miyamayomena*, *Turczaninowia*, and *Heteropappus*) from Asia should be merged with *Aster*. Furthermore, many well-recognized species of *Aster* s.s. (e.g., some shrubby taxa such as *Aster* ser. *Albescentes*, *Aster* ser. *Hersileoides* (*sensu* Ling et al., 1985b), and some alpine taxa) were not closely related to the *Aster* clade (including *A. amellus* L., the type species). A thorough taxonomic revision of *Aster* and its allies combining morphological and molecular analyses is warranted.

Tribe Anthemideae

Based on molecular phylogenetic analyses, Oberprieler et al. (2007, 2009) proposed a classification consisting of 14 subtribes, ca. 110 genera, and 1750 species. They form the fourth most species-rich group in Chinese Asteraceae (including 7 subtribes, ca. 29 genera, 364 species, and 138 endemic spp.; Table 3). The Chinese genera were mainly found in the Artemisiinae of the Asia–South African grade (Oberprieler et al., 2009). Within Artemisiinae, a group composed of *Ajania*, *Chrysanthemum*, *Stilpnolepis*, and *Artemisia* was supported as monophyletic with moderate statistical support in our analysis (PP = 0.65; Fig. 2E). The results partly corroborated the molecular studies of Zhao et al. (2010) and Liu et al. (2012), which also suggested a close relationship among *Elachanthemum*, *Ajania*, and *Chrysanthemum*. A detailed examination and taxonomic treatment of these genera are needed. The genus *Artemisia* includes approximately 180 species in China (ca. 400 in the world). However, the interspecific relationships within the genus and among subgenera were still in doubt (e.g., Watson et al., 2002; Sanz et al., 2008; Pellicer et al., 2011). More sampling and taxonomic work are necessary. The close relationships of *Glebionis* + *Argyranthemum*, *Santolina* + *Chamaemelum*, and *Leucanthemum* + *Mauranthemum* were supported with strong bootstrap values in the present result (PP > 0.95; Fig. 2E), which corroborated the study of Oberprieler et al. (2009).

Tribe Inuleae

The Inuleae includes two subtribes, with ca. 60 genera and 600 species worldwide. There are 2 subtribes, 14 genera, and 92 species indigenous to China (16 spp. endemic). The subtribe Plucheinae was supported as a sister group to Inulinae (PP = 0.98; Fig. 2F), as in most previous studies (Anderberg, 2007, 2009; Englund et al., 2009; Nylinder & Anderberg, 2015). Within Plucheinae, the Chinese *Pseudoconyza* was resolved as sister to the rest of *Sphaeranthus*, *Laggera*, *Pluchea*, *Karelinia*, and *Epaltes* with strong support (PP = 0.98; BS = 96; Fig. 2F). Within Inulinae, the close relationship of *Blumea* + *Caesulia* and polyphyletic *Inula* observed in this study (Fig. 2F) were largely consistent with previous molecular analyses (e.g., Englund et al., 2009; Nylinder & Anderberg, 2015). The present study (Fig. 2F) and a recent study (Li et al., 2014) both identified *Cyathocline purpurea* (Buch.-Ham. ex D. Don) Kuntze (former members of Astereae, *sensu* Ling et al., 1985b) as congeneric with *Blumea*.

Tribe Athroismeae

The Athroismeae was a small tribe with ca. 7 genera and 60 species. They are poorly represented in China by three genera (*Anisopappus*, *Centipeda*, and *Symphyllocarpus*) and three species (no endemic spp.). Shih & Gilbert (2011) once mentioned the difficulty in determining the position of the *Centipeda*. Strong support was observed in the present analysis for the close relationship of *Centipeda*, *Anisopappus*, and *Athroisma* (Fig. 2F). The sister relationship between *Centipeda* and *Anisopappus* + *Athroisma* + *Blepharispermum* was also supported in an earlier nrITS study, as shown by Wagstaff & Breitwieser (2002).

Heliantheae alliance

The Heliantheae alliance (recognized by Panero, 2007) appears to be the most derived in the third subclade of Asteroideae. A large putative monophyletic assemblage had been identified, including 13 tribes, ca. 460 genera, and 5500 species. They were mostly distributed in the New World (Baldwin et al., 2002; Panero & Funk, 2002; Panero, 2007; Baldwin, 2009; Funk et al. 2009c), and poorly represented in China with ca. 100 species from 7 tribes (Helenieae, Coreopsideae, Neurolaeneae, Tageteae, Millieae, Eupatorieae, and Heliantheae; Fig. 2F, 2G) and 52 genera (39 genera containing only introduced species, and 13 genera including both introduced and indigenous species; see Table 3). Within the Coreopsideae, the present study supported the polyphyly of *Bidens* (Fig. 2F), which is consistent with the studies of Kim et al. (1999), Kimball & Crawford (2004), and Crawford et al. (2009) based on nrITS analyses. Within Eupatorieae, the close relationship between *Ageratum* and *Conoclinium* was strongly supported (our results and Robinson et al., 2009; Fig. 2G). Within Heliantheae, the monophyly of *Eleutheranthera* + (*Dimerostemma* + *Eclipta*) was recovered by our result (PP = 0.95; Fig. 2G). According to the data from J. L. Panero

(released from NCBI), we recovered some clades including some Chinese introduced genera (PP > 0.94; Fig. 2G), *Synedrella* + *Lasianthaea*, *Calyptocarpus* + *Damnxanthodium*, *Sphagneticola* + *Tilesia*, *Wollastonia* + *Lipochae*, *Ratibida* + *Rudbeckia*, and *Spilanthes* + *Acmella*.

Revised taxonomy of Chinese Asteraceae at the generic level

Asteraceae are the largest angiosperm family in China in terms of species number (Wang et al., 2015). In this study, the molecular phylogeny largely resolved the relationships of Chinese Asteraceae at the generic level, although 55 Chinese genera remain to be sampled. It also provided a framework for revising the recent classification of Asteraceae in *Flora of China* (Shih et al., 2011). The following rearrangements on the classification of Chinese Asteraceae were suggested: transferring the genera *Leucomeris* and *Nouelia* from the tribe Mutisieae to the tribe Hyalideae (Wunderlichioideae); *Ainsliaea*, *Myripnois*, and *Pertya* from the tribe Mutisieae to the tribe Pertyeae (Pertyoideae); *Echinops* from the tribe Echinopeae and *Atractylodes*, *Carlina*, and *Tugarinovia* from the tribe Carlineae to the tribe Cardueae (Carduoideae); *Cavea* from the genera incertae sedis to the tribe Gymnarrhenoideae (Gymnarrhenoideae), *Centipeda* from the genera incertae sedis to the tribe Athroismeae (Asteroideae); and *Doronicum* from the tribe Senecioneae to the tribe Doroniceae (Asteroideae). Other systematic studies related to Chinese Asteraceae (e.g., Liu, 2005; Anderberg et al., 2007; Gao & Liu, 2007; Funk et al. 2009b; Blöch et al., 2010; Fan et al. 2011; Liu & Yang, 2011; Zhang et al., 2011a; Wang et al., 2013a, 2013b; Li et al., 2014; Yuan et al., 2015) and The Phylogeny of Angiosperms (http://angio.bergianska.se/) were also considered here. Finally, we herein proposed an updated classification of Chinese Asteraceae at the generic level to reflect the recent phylogenetic and taxonomic changes (see Table 3, containing their placements in subfamilies, tribes, subtribes, and groups). The updated classification accounted for 7 subfamilies, 22 tribes, and 255 genera (48 introduced). A new classification of Chinese Asteraceae based on broader sampling, more markers, and detailed morphological and cytological evidence is still needed in the near future.

Acknowledgements

This study was financially supported by grants from the National Natural Science Foundation of China (Grant Nos. 31270237, 31570204, 31070167, 30670148, and J1310002), S & T Basic Work (Grant Nos. 2013FY112100, 2014FY210300), the National Key Basic Research Program of China (Grant No. 2014CB954100), the Chinese Academy of Sciences International Institution Development Program (Grant No. SAJC201315), the Chinese Academy of Sciences External Cooperation Program of BIC (Grant No. GJHZ201321), and the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists (Grant No. 2011T1S24), awarded to Prof. Zhi-Duan Chen et al. We would like to thank Drs. Jun Wen, Zhi-Duan Chen, Wei Wang, and Cai-Fei Zhang for critical and valuable comments on the improvement of our manuscript, and Yan-Chao Bi and Yu Han for data collection.

References

Álvarez Fernández I, Fuertes Aquilar J, Panero JL, Nieto Feliner G. 2001. A phylogenetic analysis of *Doronicum* (Asteraceae, Senecioneae) based on morphological, nuclear ribosomal (ITS), and chloroplast (*trnL-F*) evidence. *Molecular Phylogenetics and Evolution* 20: 41–64.

Anderberg AA. 1991. Taxonomy and phylogeny of the tribe Gnaphalieae (Asteraceae). *Opera Botanica* 104: 1–195.

Anderberg AA. 2007. Tribe Inuleae Cass. In: Kadereit JW, Jeffrey C eds. *The families and genera of vascular plants*, Vol. 8, *Flowering plants. Eudicots. Asterales*. Berlin: Springer. 374–391.

Anderberg AA. 2009. Inuleae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 667–680.

Anderberg AA, Baldwin BG, Bayer RG, Breitwieser J, Jeffrey C, Dillon MO, Eldenäs P, Funk VA, Garcia-Jacas N, Hind DJN, Karis PO, Lack HW, Nesom G, Nordenstam B, Oberprieler C, Panero JL, Puttock C, Robinson H, Stuessy TF, Susanna A, Urtubey E, Vogt R, Ward J, Watson LE. 2007. Compositae. In: Kadereit JW, Jeffrey C eds. *The families and genera of vascular plants*, Vol. 8, *Flowering plants. Eudicots. Asterales*. Berlin: Springer. 61–588.

Anderberg AA, Ohlson JL. 2012. The genus *Cavea*, an addition to the tribe Gymnarrheneae (Asteraceae–Gymnarrhenoideae). *Compositae Newsletter* 50: 46–55.

APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.

Baird KE, Funk VA, Wen J, Weeks A. 2010. Molecular phylogenetic analysis of *Leibnitzia* Cass. (Asteraceae: Mutisieae: *Gerbera*-complex), an Asian–North American disjunct genus. *Journal of Systematics and Evolution* 48: 161–174.

Baldwin BG. 2009. Heliantheae alliance In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 689–711.

Baldwin BG, Wessa BL, Panero JL. 2002. Nuclear rDNA evidence for major lineages of helenioid Heliantheae (Compositae). *Systematic Botany* 27: 161–198.

Barres L, Sanmartín I, Anderson CL, Susanna A, Buerki S, Galbany-Casals M, Vilatersana R. 2013. Reconstructing the evolution and biogeographic history of tribe Cardueae (Compositae). *American Journal of Botany* 100: 867–882.

Bayer MB, Breitweiser I, Ward J, Puttock CF. 2007. Tribe Gnaphalieae. In: Kadereit JW, Jeffrey C eds. *The families and genera of vascular plants*, Vol. 8, *Flowering plants. Eudicots. Asterales*. Berlin: Springer. 246–283.

Bayer RJ, Starr JR. 1998. Tribal phylogeny of the Asteraceae based on two non-coding chloroplast sequences, the *trnL* intron and *trnL/trnF* intergenic spacer. *Annals of the Missouri Botanical Garden* 85: 242–256.

Bentham G. 1873. Compositae. In: Bentham G, Hooker JD eds. *Genera Plantarum* Vol. 2 (1). London: Reeve. 163–533.

Blöch C, Dickoré WB, Samuel R, Stuessy TF. 2010. Molecular phylogeny of the edelweiss (*Leontopodium*, Asteraceae–Gnaphalieae). *Edinburgh Journal of Botany* 67: 235–264.

Brouillet L, Lowrey TK, Urbatsch L, Karaman-Castro V, Sancho G, Wagstaff S, Semple JC. 2009. Astereae. In: Funk VA, Susanna A, Stuessy T, Bayer R eds. *Systematics, evolution, and biogeography of the Compositae*. Vienna: IAPT. 449–490.

Cabrera A. 1977. Mutisieae—systematic review. In: Heywood VH, Harborne JE, Turner B eds. *The biology and chemistry of the compositae*, Vol. 2. London: Academic Press. 1039–1066.

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y, Kron KA, Rettig JH, Conti E, Palmer JD, Manhart JR, Sytsma KJ, Michaels HJ, Kress WJ, Karol KG, Clark WD, Hedren M, Gaut BS, Jansen RK, Kim KJ, Wimpee CF, Smith JF, Furnier GR, Strauss SH, Xiang Q, Plunkett GM, Soltis PS, Swensen SM, Williams SE, Gadek PA, Quinn CJ, Eguiarte LE, Golenberg E, Learn GH, Jr, Graham SW, Barrett SCH, Dayanandan S, Albert VA. 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–548, 550–580.

Chen YL. 1999. Senecioneae. In: *Flora Reipublicae Popularis Sinica*. Beijing: Science Press. 77(1): 5–326.

Chen YL, Brouillet L. 2011. *Formania*. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20–21: 569.

Chen ZD, Yang T, Lin L, Lu LM, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, Cao ZY, Liu HM, Wang XM, Wang W, Zhang JB, Meng Z, Cao W, Li JH, Wu SD, Zhao HL, Liu ZJ, Du ZY, Wang QF, Guo J, Tan XX, Su JX, Zhang LJ, Yang LL, Liao YY, Li MH, Zhang GQ, Chung SW, Zhang J, Xiang KL, Li RQ, Soltis DE, Soltis PS, Zhou SL, Ran JH, Wang XQ, Jin XH, Chen YS, Gao TG, Li JH, Zhang SZ, Lu AM, China Phylogeny Consortium. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54: 277–306.

Crawford DJ, Tadesse M, Mort ME, Kimball RT, Randle CP. 2009. Coreopsideae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 713–730.

Deng T, Zhang JW, Zhu XX, Zhang DG, Nie ZL, Sun H. 2014. *Youngia zhengyiana* (Asteraceae, Crepidinae), a new species from south China, with notes on the systematics of *Youngia* inferred from morphology and nrITS phylogeny. *Phytotaxa* 170: 259–268.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Englund M, Pornpongrungrueng P, Gustafsson M, Anderberg AA. 2009. Phylogenetic relationships and generic delimitation in Inuleae subtribe Inulinae (Asteraceae) based on ITS and cpDNA sequence data. *Cladistics* 25: 319–352.

Fay MF, Swensen SM, Chase MW. 1997. Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae). *Kew Bulletin* 52: 111–120.

Fan XY, Zhang SM, Gao TG. 2011. *Grindelia*, a new naturalized genus of the tribe Astereae, Asteraceae in China. *Plant Diversity and Resources* 33: 171–173.

Funk VA, Anderberg AA, Baldwin BG, Bayer RJ, Bonifacino JM, Breitwieser I, Brouillet L, Carbajal R, Chan R, Coutinho AXP, Crawford DJ, Crisci JV, Dillon MO, Freire SE, Galbany-Casals M, Garcia-Jacas N, Gemeinholzer B, Gruenstaeudl M, Hansen HV, Himmelreich S, Kadereit JW, Kallersjo M, Karaman-Castro V, Karis PO, Katinas L, Keeley SC, Kilian N, Kimball RT, Lowrey TK, Lundberg J, McKenzie RJ, Tadesse M, Mort ME, Nordenstam B, Oberprieler C, Ortiz S, Pelser PB, Randle CP, Robinson H, Roque N, Sancho G, Semple JC, Serrano M, Stuessy TF, Susanna A, Unwin M, Urbatsch L, Urtubey E, Valles J, Vogt R, Wagstaff S, Ward J, Watson LE. 2009a. Compositae metatrees: The next generation. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 747–777.

Funk VA, Bayer RJ, Keeley S, Chan R, Watson L, Gemeinholzer B, Schilling EE, Panero JL, Baldwin BG, Garcia-Jacas NT, Susanna A, Jansen RK. 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. *Biologiske Skrifter* 55: 343–374.

Funk VA, Chan R. 2009. Introduction to Cichorioideae. In: Funk VA, Susana A, Stuessy TF, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 335–342.

Funk VA, Chan R, Keeley SC. 2004. Insights into the evolution of the tribe Arctoteae (Compositae: subfamily Cichorioideae s.s.) using *trnL-F, ndhF*, and ITS. *Taxon* 53: 637–655.

Funk VA, Specht C. 2007. Meta-trees: Grafting for a global perspective. *Proceedings of the Biological Society of Washington* 120: 232–240.

Funk VA, Susanna A, Stuessy T, Bayer RJ eds. 2009b. *Systematics, evolution, and biogeography of the Compositae*. Vienna: IAPT.

Funk VA, Susanna A, Stuessy TF, Robinson H. 2009c. Classification of Compositae. In: Funk VA, Susana A, Stuessy TF, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 171–189.

Galbany-Casals M, Unwin M, Garcia-Jacas N, Smissen RD, Susanna A, Bayer RJ. 2014. Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon* 63: 608–624.

Gao TG, Freire SE, Hind DJN. 2011. *Ainsliaea, Pertya and Myripnois*. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20–21: 15–32.

Gao TG, Hind DJN. 2011. *Nouelia, Leucomeris*. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20–21: 10.

Gao TG, Liu Y. 2007. *Gymnocoronis*, a new naturalized genus of the tribe Eupatorieae, Asteraceae in China. *Acta Phytotaxonomica Sinica* 45: 329–332.

Gao TG, Zhang GJ. 2016. *Species Catalogue of China. Volume I: Spermatophytes X*. Beijing: Science Press. (in press)

Garcia-Jacas N, Garnatje T, Susanna A, Vilatersama R. 2002. Tribal and subtribal delimitation and phylogeny of the Cardueae (Asteraceae): A combined nuclear and chloroplast DNA analysis. *Molecular Phylogenetics and Evolution* 22: 51–64.

Goertzen LR, Cannone JJ, Gutell RR, Jansen RK. 2003. ITS secondary structure derived from comparative analysis: Implications for sequence alignment and phylogeny of Asteraceae. *Molecular Phylogenetics and Evolution* 29: 216–234.

Golden JL, Kim YD, Bain JF. 2001. A re-evaluation of North American *Tephroseris* and *Sinosenecio* (Asteraceae: Senecioneae) based on molecular and micromorphological data. *Canadian Journal of Botany* 79: 1195–1201.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Hind DJN. 2007. Mutisieae. In: Kadereirt J, Jeffrey C eds. *The families and genera of vascular plants*, Vol. 8. Berlin and Heidelberg: Springer. 90–123.

Jansen RK, Palmer JD. 1987. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proceedings of the National Academy of Sciences USA* 84: 5818–5822.

Jeffrey C, Chen YL. 1984. Taxonomic studies on the tribe Senecioneae (Compositae) of Eastern Asia. *Kew Bulletin* 39: 205–446.

Johnson LA, Soltis DE. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143–156.

Katinas L, Pruski J, Sancho G, Tellería MC. 2008. The subfamily Mutisioideae (Asteraceae). *Botanical Review* 74: 469–716.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Kato HK, Oginuma Z, Gu B, Tobe H. 1998. Phylogenetic relationships of Betulaceae based on *matK* sequences with particular reference to

the positions of *Ostryopsis*. *Acta Phytotaxonomica et Geobotanica* 49: 89–97.

Keeley SC, Forsman ZH, Chan R. 2007. A phylogeny of the "evil tribe" (Vernonieae: Compositae) reveals Old/New World long distance dispersal: Support from separate and combined congruent datasets (*trnL-F*, *ndhF*, ITS). *Molecular Phylogenetics and Evolution* 44: 89–103.

Kilian N, Gemeinholzer B, Lack HW. 2009. Cichorieae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of the Compositae*. Vienna: IAPT. 343–383.

Kim HG, Loockerman DJ, Jansen RK. 2002. Systematic implications of *ndhF* sequence variation in the Mutisieae (Asteraceae). *Systematic Botany* 27: 598–609.

Kim KJ, Jansen RK. 1995. *ndhF* sequence evolution and the major clades in the sunflower family. *Proceedings of the National Academy of Sciences USA* 92: 10379–10383.

Kim KJ, Jansen RK, Wallace RS, Michaels HJ, Palmer JD. 1992. Phylogenetic implications of *rbcL* sequence variation in the Asteraceae. *Annals of the Missouri Botanical Garden* 79: 428–445.

Kim SC, Crawford DJ, Tadesse M, Berbee M, Ganders FR, Pirseyedi M, Esselman EJ. 1999. ITS sequences and phylogenetic relationships in *Bidens* and *Coreopsis* (Asteraceae). *Systematic Botany* 24: 480–493.

Kimball RT, Crawford DJ. 2004. Phylogeny of Coreopsideae (Asteraceae) using ITS sequences suggests lability in reproductive characters. *Molecular Phylogenetics and Evolution* 33: 127–139.

Kita Y, Fujikawa K, Ito M, Ohba H, Kato M. 2004. Molecular phylogenetic analyses and systematics of the genus *Saussurea* and related genera (Asteraceae, Cardueae). *Taxon* 53: 679–690.

Li WP, Qian FM, Yang XL, Chen SM. 2014. Systematic position of *Cyathocline* Cass. (Asteraceae): Evidences from molecular, cytological and morphological data. *Plant Systematics and Evolution* 300: 595–606.

Li WP, Yang FS, Jivkova T, Yin GS. 2012. Phylogenetic relationships and generic delimitation of Eurasian *Aster* (Asteraceae: Astereae) inferred from ITS, ETS and *trnL-F* sequence data. *Annals of Botany* 109: 1341–1357.

Ling Y, Chen YL, Shih C. 1985a. Compositae. *Flora Reipublicae Popularis Sinicae*, Beijing: Science Press 74: 1–2.

Ling Y, Chen YL, Shih C. 1985b. Astereae. *Flora Reipublicae Popularis Sinicae*, Beijing: Science Press 74: 70–353.

Liu JQ, Gao TG, Chen ZD, Lu AM. 2002. Molecular phylogeny and biogeography of the Qinghai–Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics and Evolution* 23: 307–325.

Liu JQ, Wang Y, Wang AL, Hideaki O, Abbott RJ. 2006. Radiation and diversification within the *Ligularia–Cremanthodium–Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai–Tibetan Plateau. *Molecular Phylogenetics and Evolution* 38: 31–49.

Liu PL, Wan Q, Guo YP, Yang J, Rao GY. 2012. Phylogeny of the genus *Chrysanthemum* L.: Evidence from single-copy nuclear gene and chloroplast DNA sequences. *Plos One* 7(11): e48970.

Liu QR. 2005. *Flaveria* Juss. (Compositae), a newly naturalized genus in China. *Acta Phytotaxonomica Sinica* 43: 178–180.

Liu Y, Chen YS, Yang QE. 2013. Generic status, circumscription, and allopolyploid origin of *Faberia* (Asteraceae: Cichorieae) as revealed by ITS and chloroplast DNA sequence data. *Taxon* 62: 1235–1247.

Liu Y, Yang QE. 2011. *Hainanecio*, a new genus of the Senecioneae, Asteraceae from China. *Botanical Studies* 52: 115–120.

López-Vinyallonga S, Mehregan I, Garcia-Jacas N, Tscherneva O, Susanna A, Kadereit JW. 2009. Phylogeny and evolution of the *Arctium–Cousinia* complex (Compositae, Cardueae–Carduinae). *Taxon* 58: 153–171.

Lundberg J. 2009. Asteraceae and relationships within Asterales. In: Funk VA, Susana A, Stuessy TF, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 157–169.

Mandel JR, Dikow RB, Funk VA. 2015. Using phylogenomics to resolve mega-families: An example from Compositae. *Journal of Systematics and Evolution* 53: 391–402.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: *Gateway Computing Environments Workshop (GCE 2010)*. Proceedings of a meeting held New Orleans, LA, 14 November, 2010. New York: IEEE. 1–8.

Mitsui Y, Chen ST, Peng CI, Deng YF, Setoguchi H. 2008. Phylogeny and biogeography of the genus *Ainsliaea* (Asteraceae) in the Sino-Japanese region based on nuclear rDNA and plastid DNA sequence data. *Annals of Botany* 101: 111–124.

Nie ZL, Funk VA, Meng Y, Deng T, Sun H, Wen J. 2015. Recent assembly of the global herbaceous flora: Evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytologist* 209: 1795–1806.

Nie ZL, Funk VA, Sun H, Deng T, Meng Y, Wen J. 2013. Molecular phylogeny of *Anaphalis* (Asteraceae, Gnaphalieae) with biogeographic implications in the Northern Hemisphere. *Journal of Plant Research* 126: 17–32.

Nordenstam B, Kallersjo M. 2009. Calenduleae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 527–538.

Nordenstam B, Pelser PB, Kadereit JW, Watson LE. 2009. Senecioneae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 503–525.

Nylinder S, Anderberg AA. 2015. Phylogeny of the Inuleae (Asteraceae) with special emphasis on the Inuleae–Plucheinae. *Taxon* 64: 110–130.

Oberprieler C, Himmelreich S, Kallersjo M, Valles J, Watson LE, Vogt R. 2009. Anthemideae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 631–666.

Oberprieler C, Himmelreich S, Vogt R. 2007. New subtribal classification of the tribe Anthemideae (Compositae). *Willdenowia* 37: 89–114.

Olmstead RG, Palmer JD. 1994. Chloroplast DNA systematics: A review of methods and data analysis. *American Journal of Botany* 81: 1205–1224.

Olmstead RG, Sweere JA. 1994. Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.

Ortiz S, Carbajal R, Serrano M, Rodríguez-Oubiña J, Iglesias I. 2013. Phylogeny of the African Mutisieae s.l. (Asteraceae) based on ndhF and *trnL-F* sequences (cpDNA). *Taxon* 62: 525–536.

Panero JL. 2005. New combinations and infrafamilial taxa in the Asteraceae. *Phytologia* 87: 1–14.

Panero JL. 2007. Key to the tribes of the Heliantheae alliance. In: Kadereit JW, Jeffrey C eds. *The families and genera of vascular plants, Vol. 8, Flowering plants. Eudicots. Asterales*. Berlin: Springer. 391–395.

Panero JL, Funk VA. 2002. Toward a phylogenetic subfamilial classification for the Compositae (Asteraceae). *Proceedings of the Biological Society of Washington* 115: 909–922.

Panero JL, Funk VA. 2008. The value of sampling anomalous taxa in phylogenetic studies: Major clades of the Asteraceae revealed. *Molecular Phylogenetics and Evolution* 47: 757–782.

Panero JL, Susana EF, Espinar LA, Crozier BS, Barboza GE, Cantero JJ. 2014. Resolution of deep nodes yields an improved backbone phylogeny and a new basal lineage to study early evolution of Asteraceae. *Molecular Phylogenetics and Evolution* 80: 43–53.

Pellicer J, Vallès J, Korobkov AA, Garnatje T. 2011. Phylogenetic relationships of *Artemisia* subg. *Dracunculus* (Asteraceae) based on ribosomal and chloroplast DNA sequences. *Taxon* 60: 691–704.

Pelser PB, Kennedy AH, Tepe EJ, Shidler JB, Nordenstam B, Kadereit JW, Watson LE. 2010. Patterns and causes of incongruence between plastid and nuclear senecioneae (Asteraceae) phylogenies. *American Journal of Botany* 97: 856–873.

Pelser PB, Nordenstam B, Kadereit JW, Watson LE. 2007. An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of *Senecio* L. *Taxon* 56: 1077–1104.

Peng YL, Zhang Y, Gao XF, Tong LJ, Li L, Li RY, Zhu ZM, Xian JR. 2014. A phylogenetic analysis and new delimitation of *Crepidiastrum* (Asteraceae, tribe Cichorieae). *Phytotaxa* 159: 241–255.

Posada D, Crandall KA. 1998. Modeltest: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Raab-Straube EV. 2003. Phylogenetic relationships in *Saussurea* (Compositae, Cardueae) *sensu lato*, inferred from morphological, ITS and *trnL-trnF* sequence data, with a synopsis of *Himalaiella* gen. nov., *Lipschitziella* and *Frolovia*. *Willdenowia* 33: 379–402.

Rambaut A. 2012. FigTree v 1.4.0. [online]. Available from tree.bio.ed.ac.uk/Software/figtree [accessed 16 April 2016].

Robinson H, Schilling E, Panero JL. 2009. Eupatorieae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 731–744.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Safer S, Tremetsberger K, Guo YP, Kohl G, Samuel MR, Stuessy TF, Stuppner H. 2011. Phylogenetic relationships in the genus *Leontopodium* (Asteraceae: Gnaphalieae) based on AFLP data. *Botanical Journal of the Linnean Society* 165: 364–377.

Sanz M, Vilatersana R, Hidalgo O, Garcia-Jacas N, Susanna A, Schneeweiss GM, Vallès J. 2008. Molecular phylogeny and evolution of floral characters of *Artemisia* and allies (Anthemideae, Asteraceae): Evidence from nrDNA ETS and ITS sequences. *Taxon* 57: 66–78.

Shih C, Chen YL, Chen YS, Lin YR, Liu SW, Ge XJ, Gao TG, Zhu SX, Liu Y, Yang QE, Humphries CJ, Raab-Straube EV, Gilbert MG, Nordenstam B, Kilian N, Brouillet L, Illarionova ID, Hind DJN, Jeffrey C, Bayer RJ, Kirschner J, Greuter W, Anderberg AA, Semple JC, Stěpánek J, Freire SE, Martins L, Koyama H, Kawahara T, Vincent L, Sukhorukov AP, Mavrodiev EV, Gottschlich G. 2011. Asteraceae (Compositae). In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20–21: 1–894.

Shih C, Fu GX. 1983. *Formania*. In: *Flora Reipublicae Popularis Sinica*. Beijing: Science Press. 76(1): 81.

Shih C, Gilbert MG. 2011. *Centipeda*. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20–21: 892–893.

Smissen RD, Galbany-Casals M, Breitwieser I. 2011. Ancient allopolyploidy in the everlasting daisies (Asteraceae: Gnaphalieae): Complex relationships among extant clades. *Taxon* 60: 649–662.

Smith SA, Beaulieu JM, Donoghue MJ. 2009. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology* 9: 37.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic Biology* 57: 758–771.

Susanna A, Garcia-Jacas N. 2007. Tribe Cardueae Cass. In: Kadereit JW, Jeffrey C eds. *The families and genera of vascular plants*, Vol. 8, *Flowering plants. Eudicots. Asterales*. Berlin: Springer. 123–146.

Susanna A, Garcia-Jacas N. 2009. Cardueae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 293–313.

Susanna A, Garcia-Jacas N, Hidalgo O, Vilatersana R, Garnatje T. 2006. The Cardueae (Compositae) revisited: Insights from ITS, *trnL-trnF*, and *matK* nuclear and chloroplast DNA analysis. *Annals of the Missouri Botanical Garden* 93: 150–171.

Torices R. 2010. Adding time-calibrated branch lengths to the Asteraceae supertree. *Journal of Systematics and Evolution* 48: 271–278.

Wagstaff SJ, Breitwieser I. 2002. Phylogenetic relationships of New Zealand Asteraceae inferred from ITS sequences. *Plant Systematics and Evolution* 231: 203–224.

Wang GY, Meng Y, Deng T, Yang YP. 2014. Molecular phylogeny of *Faberia* (Asteraceae: Cichorieae) based on nuclear and chloroplast sequences. *Phytotaxa* 167: 223–234.

Wang LS, Jia Y, Zhang XC, Qin HN. 2015. Overview of higher plant diversity in China. *Biodiversity Science* 23: 217–224.

Wang LY, Pelser PB, Nordenstam B, Liu JQ. 2009a. Strong incongruence between the ITS phylogeny and generic delimitation in the *Nemosenecio–Sinosenecio–Tephroseris* assemblage (Asteraceae: Senecioneae). *Botanical Studies* 50: 435–442.

Wang YJ, Liu JQ. 2004. Phylogenetic analyses of *Saussurea* sect. *Pseudoeriocoryne* (Asteraceae: Cardueae) based on chloroplast DNA *trnL-F* sequences. *Biochemical Systematics and Ecology* 32: 1009–1023.

Wang YJ, Liu JQ, Miehe G. 2007. Phylogenetic origins of the Himalayan endemic *Dolomiaea*, *Diplazoptilon* and *Xanthopappus* (Asteraceae: Cardueae) based on three DNA fragments. *Annals of Botany* 99: 311–322.

Wang YJ, Raab-Straube EV, Susanna A, Liu JQ. 2013a. *Shangwua* (Compositae), a new genus from the Qinghai–Tibetan Plateau and Himalayas. *Taxon* 62: 984–996.

Wang YJ, Susanna A, Raab-Straube EV, Milne R, Liu JQ. 2009b. Island-like radiation of *Saussurea* (Asteraceae: Cardueae) triggered by uplifts of the Qinghai–Tibetan Plateau. *Biological Journal of the Linnean Society* 97: 893–903.

Wang ZH, Peng H, Kilian N. 2013b. Molecular phylogeny of the *Lactuca* alliance (Cichorieae Subtribe Lactucinae, Asteraceae) with focus on their Chinese centre of diversity detects potential events of reticulation and chloroplast capture. *Plos One* 8(12): e82692.

Ward J, Bayer R, Breitwieser I, Smissen R, Galbany M, Unwin M. 2009. Gnaphalieae. In: Funk VA, Susanna A, Stuessy T, Bayer RJ eds. *Systematics, evolution, and biogeography of Compositae*. Vienna: IAPT. 539–588.

Watson LE, Bates PL, Evans TM, Unwin MM, Estes JR. 2002. Molecular phylogeny of subtribe Artemisiinae (Asteraceae), including *Artemisia* and its segregate genera. *BMC Evolutionary Biology* 2: 17.

Wen J, Ickert-Bond, SM, Nie ZL, Li R. 2010. Timing and modes of evolution of eastern Asian–North American biogeographic disjunctions in seed plants. In: Long M, Gu H, Zhou Z eds. *Darwin's heritage today: Proceedings of the Darwin 200 Beijing International Conference*. Beijing: Higher Education Press. 252–269.

Winkworth RC, Lundberg J, Donoghue MJ. 2008. Toward a resolution of Campanulid phylogeny, with special reference to the placement of Dipascales. *Taxon* 57: 53–65.

Yuan Q, Bi YC, Chen YS. 2015. *Diplazoptilon* (Asteraceae) is merged with *Saussurea* based on evidence from morphology and molecular systematics. *Phytotaxa* 236: 53–61.

Zhang JW, Boufford DE, Sun H. 2011a. *Parasyncalathium* JW Zhang, Boufford, H. Sun (Asteraceae, Cichorieae): A new genus endemic to the Himalaya–Hengduan Mountains. *Taxon* 60: 1678–1684.

Zhang JW, Nie ZL, Wen J, Sun H. 2011b. Molecular phylogeny and biogeography of three closely related genera, *Soroseris*, *Stebbinsia*, and *Syncalathium* (Asteraceae, Cichorieae), endemic to the Tibetan Plateau, SW China. *Taxon* 60: 15–26.

Zhao HB, Chen FD, Chen SM, Wu GS, Guo WM. 2010. Molecular phylogeny of *Chrysanthemum*, *Ajania* and its allies (Anthemideae, Asteraceae) as inferred from nuclear ribosomal ITS and chloroplast *trnL-F* IGS sequences. *Plant Systematics and Evolution* 284: 153–169.

Zhong CX, Li WP, Yang XL, Tang M, Liao W, Chen SM. 2014. Molecular phylogeny of Chinese *Conyza*, *Microglossa* and *Thespis* (Asteraceae: Astereae) based on two nuclear ribosomal DANN regions. *Plant Science Journal* 32: 216–227.

Appendix

Members of the China Phylogeny Consortium

Zhi-Duan CHEN, An-Ming LU, Hong-Zhi KONG, Xiao-Quan WANG, Yin-Zheng WANG, Shi-Liang ZHOU
Institute of Botany, Chinese Academy of Sciences, Beijing, China
Shou-Zhou ZHANG, Xiao-Ming WANG
Fairylake Botanical Garden, Shenzhen, China
Zhong-Jian LIU
The Orchid Conservation and Research Center of Shenzhen, Shenzhen, China
Qing-Feng WANG
Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China
Jian-Hui LI
Computer Network Information Center, Chinese Academy of Sciences, Beijing, China
De-Zhu LI, Ting-Shuang YI
Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China
Hong MA
Fudan University, Shanghai, China
Douglas E. SOLTIS, Pamela S. SOLTIS
University of Florida, Gainesville, USA
Jian-Hua LI
Hope College, Holland, USA
Cheng-Xin FU
Zhejiang University, Hangzhou, China
Qi-Xin LIU
Nanjing Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12216/suppinfo

Table S1. Taxa and GenBank accession numbers for DNA sequences used in this study.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12204

Research Article

Using nuclear genes to reconstruct angiosperm phylogeny at the species level: A case study with Brassicaceae species

Liming Cai[1,3] and Hong Ma[1,2]*

[1]Ministry of Education Key Laboratory of Biodiversity Sciences and Ecological Engineering, Institute of Plant Biology, Institute of Biodiversity Sciences, Center for Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai 200438, China
[2]Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China
[3]Present address: Harvard University Herbaria, Cambridge, MA 02138, USA
*Author for correspondence. E-mail: hongma@fudan.edu.cn. Tel.: 86-21-51630500. Fax: 86-21-51630504.
Received 14 January 2016; Accepted 15 March 2016; Article first published online 8 June 2016

Abstract Angiosperm phylogeny has been investigated extensively using organellar sequences; recent efforts using nuclear genes have also been successful in reconstructing angiosperm phylogenies at family or deeper levels. However, it is not clear whether nuclear genes are also effective in understanding relationships between species in a genus. Here we present a case study of phylogeny at generic and specific levels with nuclear genes, using Brassicaceae taxa as examples. Brassicaceae includes various crops and the model plant *Arabidopsis thaliana*. A recent study showed that nuclear genes can provide well-resolved relationships between tribes and larger lineages in Brassicaceae, but few species were included in any given genus. We present a phylogeny with multiple species in each of five genera within Brassicaceae for a total of 65 taxa, using three protein-coding nuclear genes, *MLH1*, *SMC2*, and *MCM5*, with up to approximately 10 200 base pairs (in both exons and introns). Maximum likelihood and Bayesian analyses of the separate gene regions and combined data reveal high resolution at various phylogenetic depths. The relationships between genera here were largely congruent with previous results, with further resolution at the species level. Also, we report for the first time the affinity of *Cardamine rockii* with tribe Camelineae instead of other *Cardamine* members. In addition, we report sequence divergence at three levels: across angiosperms, among Brassicaceae species, and between *Arabidopsis* ecotypes. Our results provide a robust species-level phylogeny for a number of Brassicaceae members and support an optimistic perspective on the phylogenetic utility of conserved nuclear data for relatively recent clades.

Key words: *Brassica*, Brassicaceae, *Cardamine*, *Lepidium*, *Rorippa*, nuclear gene, species-level phylogeny.

The cabbage family (Brassicaceae) is one of the most important and well-known plant groups, with various vegetables (such as cabbage and broccoli) and oilseed crop species cultivated throughout the world. In addition, a well-known member of the family is the model organism *Arabidopsis thaliana* (L.) Heynh., the studies of which have revolutionized our knowledge in almost every field of modern plant biology. Several studies have been carried out on molecular phylogeny across this family in recent years (Bailey et al., 2006; Beilstein et al., 2006, 2008; Warwick et al., 2010; Al-Shehbaz, 2012). In these studies, phylogenetic inference provided the basics for understanding the patterns of evolutionary history. Transcribed nuclear ribosomal spacer (ITS), plastid DNA (cpDNA), and ribosomal DNA (rDNA) markers are most commonly used because of the highly conserved sequences and high copy numbers, making them easily accessible without cloning (Baldwin et al., 1995; Baldwin & Markos, 1998; Álvarez & Wendel, 2003). However, the entire plastid genome is a single linkage group (Birky, 1995), whereas rDNAs are sometimes not completely identical within a genome (Buckler et al., 1997). As a result, these markers are not capable of dealing with the cases involving hybridization and polyploidization. More importantly, building phylogenies below genus level requires rapidly evolving markers that vary among closely related species. Organellar and ribosomal genes are sometimes too conserved in this context (Small et al., 1998; Sang, 2002). The limitation of traditional cpDNA and rDNA markers and the growing availability of large genomic datasets from multiple taxa have prompted us to mine the plant genomes for appropriate low-copy nuclear protein-coding genes as phylogenetic markers.

Compared to organellar genes, protein-coding nuclear genes are biparentally inherited, recording genetic information from both male and female inherence. Additionally, their third positions of codons, introns, and untranslated 3′/5′-untranslated regions have relatively high evolutionary rates, presenting better resolution among closely related species (Zimmer & Wen, 2013). Therefore, we hypothesized that better resolved relationships, especially at the species level,

would be uncovered using suitable nuclear genes. At the same time, events of hybridization and polyploidization might also be uncovered. Despite these advantages, building nuclear phylogeny is still challenging in several aspects. The highly complex nuclear genome, particularly with the high frequency of polyploidy events in angiosperms (Soltis et al., 2009), makes it a serious problem to identify low-copy orthologs. For example, the ancestors of extent angiosperms (Jiao et al., 2011), of core eudicots (Blanc & Wolfe, 2004; De Bodt et al., 2005; Soltis et al., 2009), and of Brassicaceae (β and α) (Blanc & Wolfe, 2004; Schranz & Mitchell-Olds, 2006) all underwent one or more rounds of whole genome duplication. Repeated genome duplication and gene losses have made it difficult to distinguish orthologs from paralogs. In some cases, the loss of different duplicates in separate lineages can result in "hidden paralogs", yielding conflicting gene trees that exacerbate the uncertainty in phylogenetic reconstruction (Maddison, 1997). Thus it is crucial to identify genes that tend to return to single copy status quickly after genome duplication, thus behaving as ortholog among a wide range of plants.

In recent years, nuclear genes have been used to resolve relatively deep relationships among major land plant groups (Wickett et al., 2014), across angiosperms (Zhang et al., 2012; Zeng et al., 2014), and at the levels of order and family (Ding et al., 2012; Yang et al., 2015; Huang et al., 2016). Several of these studies used phylogenomics and/or phylotranscriptomics, with many gene sequences from large-scale datasets (Wickett et al., 2014; Zeng et al., 2014; Yang et al., 2015; Huang et al., 2016), illustrating the great power of the vast quantity of nuclear genes. However, for many questions concerning closely related species, it might not be necessary to use many hundreds of genes, and the need to sample a relatively large number of taxa might make large-scale datasets too costly to obtain. Also, the use of a large number of genes for many taxa will increase the need for great computation power. Thus, it is desirable to be able to quickly obtain a small number of genes from a large number of taxa. However, isolation of nuclear marker sequences is often time consuming and laborious (Small et al., 2004). Furthermore, the need for sequence variation to provide phylogenetic signals makes it difficult to develop universal primers. Therefore, identification of appropriate nuclear genes that are easy to amplify and sequence is a major objective, before one can address phylogenetic questions using this approach.

The identification of suitable nuclear genes initially involves gene comparisons across genomes and transcriptomes of a wide range of taxa, followed by lineage-specific analysis to test the usability of primers for amplification. Previous efforts on nuclear gene phylogenies within plant groups were largely focused on single copy genes, for example, *LFY* (Oh & Potter, 2005; Nie et al., 2008; Kim et al., 2010), or easily distinguished paralogs, such as *Adh1* and *Adh2* (Gaut & Clegg, 1991; Fukuda et al., 2005). A recent study showed that five nuclear genes were sufficient to resolve many deep relationships in the angiosperm phylogeny (Zhang et al., 2012). These nuclear genes were selected from four different gene families, the *MCM, SMC, MLH*, and *MSH* gene families. They were verified for their extensive single-copy and orthology across most angiosperm groups (Gozuacik et al., 2003; Lin et al., 2007; Surcel et al., 2008). However, whether these nuclear markers are appropriate for phylogenetic study at or below the genus level remains unknown.

As mentioned above, Brassicaceae provide an excellent model group for phylogenetic study because of its moderately large size of approximately 3700 species, relatively rapid evolutionary rate, and easy access for many of the members (Al-Shehbaz, 2012; Huang et al., 2016). Furthermore, the wealth of nuclear sequence information from multiple Brassicaceae genomic datasets (*Arabidopsis thaliana*, *Arabidopsis lyrata* (L.) O'Kane & Al-Shehbaz, *Capsella rubella* (Almq.) Almq., *Capsella grandiflora* Bioss., *Boechera stricta* (Graham) Al-Shehbaz, *Brassica rapa* L., and *Eutrema salsuginea* O. E. Schulz) allows preliminary screening of candidate nuclear genes (https://phytozome.jgi.doe.gov/pz/portal.html). Previous systematic research placed this family sister to Cleomaceae and divided it into 51 tribes; in addition, phylogenetic analyses support a clade (core Brassicaceae) with three main lineages (Lineages I, II, and III) and *Aethionema* as the sister to core Brassicaceae (Al-Shehbaz, 2012; Huang et al., 2016). Many of the core Brassicaceae species are grouped into Lineages I, II, and III with moderate to low support. A preliminary study (Ding et al., 2012) used three low-copy nuclear genes, including two from Zhang et al. (2012), to reconstruct phylogenetic relationships among 13 Brassicaceae species. The results were encouraging and indicated that nuclear genes performed better than plastid genes and ITS segments in untangling species-level phylogeny. Moreover, recent work with 55 large-scale datasets has yielded a highly resolved phylogeny of major lineages (Huang et al., 2016), but few of the genera included had more than one species.

In this study, to test the effectiveness of nuclear genes in resolving species-level phylogeny, a total of 35 taxa were selected from three genera (*Brassica* L., *Lepidium* L., and *Cardamine* L.). These three genera are widely distributed and have typical morphological features as to be distinguished from other cruciferous plants. Seventeen available Brassicaceae genomes and 11 transcriptomes were also included to expand our sampling among both closely related species and more distant ones. To test genes that are conserved among all plants (even other eukaryotes) to facilitate future use in other families, we sampled three representatives of eukaryote-wide gene families (Gozuacik et al., 2003; Lin et al., 2007; Surcel et al., 2008): *MLH1*, *SMC2*, and *MCM5*, and investigated their utility in resolving relationships among low-rank taxonomic hierarchies. We also examined the sequence similarities of these three genes across three levels of evolutionary distances: among divergent angiosperm species, between members of Brassicaceae, and within the same species—*Arabidopsis thaliana*. Our results illustrate the effectiveness of conserved nuclear genes in resolving species phylogeny within a genus and provide useful information for future investigation of relatively close relationships in many other groups.

Material and Methods

Taxon sampling

A total of 63 accessions from across Brassicaceae were included, together with two out-group species *Cleome serrulata* Pursh and *Populus trichocarpa* Torr. & A. Gray ex

Hook. The taxon sampling included 35 accessions of extracted DNAs (Table 1) and 28 accessions from genome and transcriptome datasets (Table 2). To test for the effectiveness of nuclear genes in resolving species within a genus, sampled taxa were largely members of three genera, including 8 from *Lepidium*, 8 from *Cardamine*, and 19 from *Brassica*. In addition, 6 species from *Rorippa* Scop., 1 species from *Nasturtium* W. T. Aiton, 1 species from *Leavenworthia* Torr., and 4 accessions from *Raphanus* L. were included as to expand our sampling variety within tribe Cardamineae and Brassiceae. *Brassica* is the type genus of Brassiceae with tremendous diversity, in part owing to domestication. *Cardamine* and *Lepidium* are also type genera of tribes Cardamineae and Lepidiumeae, respectively. All three genera are widely distributed in China and easily accessible. Sampled materials included both wild species and domestic vegetables. Species belonging to

Table 1 Scientific names and collection information of DNA materials of Brassicaceae species used to reconstruct angiosperm phylogeny with nuclear genes *MLH1*, *MCM5*, and *SMC2*

Taxon	*MLH1*	*MCM5*	*SMC2*	Voucher	Collection locality	Collection year
Brassica campestris L. var. *chinensis*	+	+	+	L. Cai 020102	Shanghai, China	2014
Brassica campestris L. var. *narinosa*	+	−	+	L. Cai 020701	Shanghai, China	2014
Brassica juncea (L.) Czern. var. *multisecta*	−	+	−	H. Ma 020302	Pennsylvania, USA	2014
Brassica oleracea L. var. *botrytis* L.	+	+	+	L. Cai 020501	Shanghai, China	2014
Brassica oleracea L. var. *caulorapa* Metzg.	+	+	+	L. Cai 020502	Shanghai, China	2014
Brassica oleracea L. var. *gemmifera* DC.	−	+	+	H. Ma 020505	Pennsylvania, USA	2014
Brassica oleracea L. var. *italic* Plenck	+	+	−	L. Cai 020504	Shanghai, China	2014
Brassica oleracea L. var. *gongylodes* L.	+	+	+	H. Ma 020508	Pennsylvania, USA	2014
Brassica oleracea L. var. *capitata* L.	−	+	+	H. Ma 020507	Pennsylvania, USA	2014
Brassica oleracea L. var. *acephala* (DC.) Metzg.	−	+	+	H. Ma 020506	Pennsylvania, USA	2014
Brassica oleraceae L. var. *alboglabra* Beiley	+	+	−	L. Cai 020601	Shanghai, China	2014
Brassica rapa L. var. *pekinensis* (Lour.) Kitam.	+	+	+	L. Cai 020402	Shanghai, China	2014
Brassica rapa L. var. *rapa*	+	−	+	H. Ma 020403	Pennsylvania, USA	2014
Brassica rapa L. var. *ruvo*	+	+	+	H. Ma 020405	Pennsylvania, USA	2014
Brassica napus L. var. *napobrassica* (L.) Hanelt	+	+	−	H. Ma 020701	Pennsylvania, USA	2014
Cardamine flexuosa With.	+	+	+	L. Cai 440208	Shanghai, China	2013
Cardamine hirsute L.	+	+	+	L. Cai 440502	Shanghai, China	2013
Cardamine lyrata Bunge	+	−	+	N. Zhang 440101	Shanghai, China	2011
Cardamine macrophylla Willd.	+	+	+	ZY 440401	China	2012
Cardamine oligosperma Nutt.	+	+	+	L. Cai 440503	California, USA	2013
Cardamine rockii O. E. Schulz	+	−	+	CGBOWS 440601	Yunnan, China	2014
Cardamine tangutorum O. E. Schulz	+	+	+	ZWJ 440301	Gansu, China	2012
Lepidium apetalum Willd.	−	−	+	H. Ma 100402	Shandong, China	2012
Lepidium cuneiforme C. Y. Wu	+	+	+	CGBOWS 101001	Sichuan, China	2014
Lepidium ferganense Korsh.	−	−	+	CGBOWS 100901	Xinjiang, China	2014
Lepidium latifolium L.	+	+	+	L. Cai 100201	Xinjiang, China	2012
Lepidium perfoliatum L.	+	+	−	L. Cai 100601	Xinjiang, China	2012
Lepidium ruderale L.	−	−	+	L. Cai 100301	Shandong, China	2012
Nasturtium officinale W. T. Aiton	+	+	+	L. Cai 570102	Shanghai, China	2014
Raphanus sativus L. var. (1)	+	+	+	H. Ma 060105	Pennsylvania, USA	2014
Raphanus sativus L. var. *longipinnatus* L. H. Bailey	+	+	+	L. Cai 060104	Shanghai, China	2014
Raphanus sativus L. var. (2)	+	+	−	L. Cai 060106	Shanghai, China	2014
Rorippa cantoniensis (Lour.) Ohwi	+	+	+	L. Cai 560201	Shanghai, China	2013
Rorippa dubia (Pers.) Hara	−	+	+	L. Cai 560501	Shanghai, China	2013
Rorippa islandica (Oeder) Borbás	+	−	−	L. Cai 560301	Xingjiang, China	2012

+, DNA sequences obtained for certain genes in that taxa; −, DNA sequences not obtained for certain genes in that taxa.

Table 2 Source information for Brassicaceae genomes/transcriptomes used in this study

Species	Data type	Source
Aethionema subulatum Bioss.	Transcriptome	Huang et al., 2016
Barbarea vulgaris W. T. Aiton	Transcriptome	Huang et al., 2016
Brassica nigra (L.) W. D. J. Koch	Transcriptome	Huang et al., 2016
Cardamine pensylvanica Muhl. ex Willd.	Transcriptome	Huang et al., 2016
Erysimum cheiranthoides L.	Transcriptome	Huang et al., 2016
Erysimum cheiri Crantz	Transcriptome	Huang et al., 2016
Lepidium campestre (L.) W. T. Aiton	Transcriptome	Huang et al., 2016
Lepidium didymum L.	Transcriptome	Huang et al., 2016
Rorippa indica (L.) Hiern	Transcriptome	Huang et al., 2016
Rorippa sylvestris (L.) Besser	Transcriptome	Huang et al., 2016
Rorippa globosa (Turcz.) Vassilcz.	Transcriptome	Huang et al., 2016
Aethionema arabicum (L.) Andrz. ex DC.	Genome	NCBI
Arabidopsis halleri (L.) O'Kane & Al-Shehbaz subsp. *gemmifera* (Matsum.) O'Kane & Al-Shehbaz	Genome	NCBI
Arabidopsis lyrata (L.) O'Kane & Al-Shehbaz	Genome	Phytozome v10.0
Arabidopsis thaliana (L.) Heynh.	Genome	Phytozome v10.0
Arabis alpina L. cultivar Pajares	Genome	NCBI
Boechera stricta (Graham) Al-Shehbaz	Genome	NCBI
Brassica napus L. cultivar ZS11	Genome	NCBI
Brassica oleracea L. var. *oleracea* cultivar TO1000	Genome	NCBI
Brassica rapa L. cultivar FPsc	Genome	Phytozome v10.0
Camelina sativa (L.) Crantz	Genome	NCBI
Capsella grandiflora Bioss.	Genome	NCBI
Capsella rubella (Almq.) Almq.	Genome	Phytozome v10.0
Eutrema salsuginea O. E. Schulz	Genome	Phytozome v10.0
Leavenworthia alabamica Rollins	Genome	NCBI
Raphanus raphanistrum L. subsp. *raphanistrum*	Genome	NCBI
Schrenkiella parvula (Schrenk) D. A. German & Al-Shehbaz	Genome	thellungiella.org
Sisymbrium irio L.	Genome	NCBI

NCBI, National Center for Biotechnology Information; v, version.

Cardamine, *Lepidium*, and *Rorippa* are mostly wild plants that are genetically more distant from other species. Cultivars of the same *Brassica* species were also sampled to represent lower-level divergence from relatively recent domestication histories. As a result, such a sampling strategy provided an effective test for the utility of the DNA markers to recover phylogenies at distinct levels. Source information for these accessions was listed in Table 1. Total genomic DNAs were extracted from leaves using the CTAB method (Stewart & Via, 1993).

Apart from our DNA sampling, we also took advantage of genome and transcriptome datasets to facilitate candidate gene screening. Transcriptome datasets used here were selected from those reported by Huang et al. (2016), including two species from *Lepidium*, one from *Cardamine* and three from *Rorippa*. Within Brassicaceae Lineage I, the clades Camelineae and Erysimeae were represented by six and two species, respectively. Likewise, within Lineage II, tribe Brassiceae was represented by eight species, including five in *Brassica*. Other tribes were represented by species with sequenced genomes: *Sisymbrium irio* L., *Eutrema salsuginea*, *Schrenkiella parvula* (Schrenk) D. A. German & Al-Shehbaz, *Thlaspi arvense* L., and *Arabis alpina* L. Two species from the basal lineage, *Aethionema subulatum* Bioss. and *Aethionema arabicum* (L.) Andrz. ex DC., were specially selected to represent the deepest genetic divergence within Brassicaceae.

Preliminary screening of candidate genes

Previously, five low-copy nuclear genes (*SMC1*, *SMC2*, *MLH1*, *MSH1*, and *MCM5*) were used to reconstruct a highly supported angiosperm phylogeny (Zhang et al., 2012). These five genes remain orthologous across a wide range of angiosperms and have conserved exon sequences, which can facilitate primer design and global alignment. Yet resolving species-level phylogeny requires rapidly evolving markers with sufficient variable sites among closely related species. Thus introns, with more variable sequence, are expected to play a key role in untangling recent and rapid radiation of species rendering their high evolutionary rate. These five genes all contain both exons and introns and encode proteins with at least 300 amino acids. They are all housekeeping genes with conserved functions. We then examined copy numbers of these five and related genes using eight Brassicaceae species with sequenced genomes from Phytozome version 10 (https://phytozome.jgi.doe.gov/pz/portal.html) (Fig. 1).

We compared genomic sequences of the five genes and calculated the following average nucleotide sequence identities: *SMC1*, 78.58%; *SMC2*, 82.20%; *MLH1*, 74.49%; *MSH1*, 65.96%; and *MCM5*, 77.21%. *MSH1* was not tested further due to its

Organism	MLH1	MLH3	MLH4	MCM2	MCM3	MCM4	MCM5	MCM6	MCM7	MCM8	MCM9	SMC1	SMC2	SMC3	SMC4	SMC5	SMC6	Source
Arabidopsis thaliana	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	Phytozome v9.1
Arabidopsis lyrata	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	Phytozome v9.1
Brassica rapa	1	1	1	2	2	1	1	1	2	1	1	2	2	2	1	1	2	Phytozome v9.1
Capsella rubella	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	Phytozome v9.1
Capsella grandiflora	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	Phytozome v10
Boechera stricta	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	Phytozome v10
Eutrema salsuginea	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	Phytozome v9.1
Schrenkiella parvula	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	Thellungiel v2
Populus trichocarpa	1	1	1	2	1	1	2	1	1	1	1	1	2	1	1	1	1	Phytozome v9.1

Fig. 1. Gene copy number of *MLH*, *MCM*, and *SMC* family members in selected Brassicaceae species with sequenced genomes. The gene names are provided above and copy numbers are highlighted by different colors. Data source information is provided on the right. v, version.

relatively low sequence identity and possible difficulty in polymerase chain reaction (PCR) amplification. To further test the feasibility of using the remaining four genes in resolving species-level phylogeny, we examined the genomic sequences for both conserved regions for primer design flanking variable regions for phylogenetic signals and found *SMC2*, *MHL1*, and *MCM5* to be more favorable.

Primer design, amplification, and sequencing

For PCR reactions to amplify specific regions of the *SMC2*, *MHL1*, and *MCM5* genes, we designed primers based on comparison (Fig. 2) of genomic sequences of eight Brassicaceae species. We identified primers matching regions conserved among all these genome sequences, then used the leave-one-out approach (Berger et al., 2011) for assessing site-specific congruence as implemented in RAxML 8.0.2 (Stamatakis, 2006). This test prunes a single taxon at a time from a reference tree, carrying out site-specific computations for it, followed by reinserting into the original position. A sliding window of 100 sites in the alignment is used for calculation. The mean distance between the best placements for all sliding windows is calculated, assessing the phylogenetic variability of different areas of a gene. Targeted loci were located in the regions with the lowest average node distance. Degenerate sites were used when there were different sites among sampled species.

Several primer pairs for each of the three genes (*SMC2*, *MHL1*, and *MCM5*) were tested by PCR with multiple template DNAs and the one that was chosen had most reliably yielded amplified products across all samples; the primer sequences and their properties are listed in Table 3. All three primer pairs amplified loci with conserved exons flanking variable introns (Fig. 2). *SMC2* was typically amplified in a segment that contained two partial exons, three exons and four introns, making up 1174 base pairs (bp) in *A. thaliana*. The amplified *A. thaliana MLH1* segment contained two partial exons, five exons, and six introns, with 1462 bp. The *MCM5* segment from *A. thaliana* contained two partial exons, four exons, and five introns totaling 1313 bp. We used DNA polymerase from Takara (Otsu, Shiga, Japan) for PCR reactions. This DNA polymerase possesses high proofreading activity and can reduce mismatch in DNA amplification. Cycling conditions for *SMC2* started with initial denaturation at 94.0 °C for 5 min, followed by 30 cycles of amplification, 94.0 °C for 60 s,

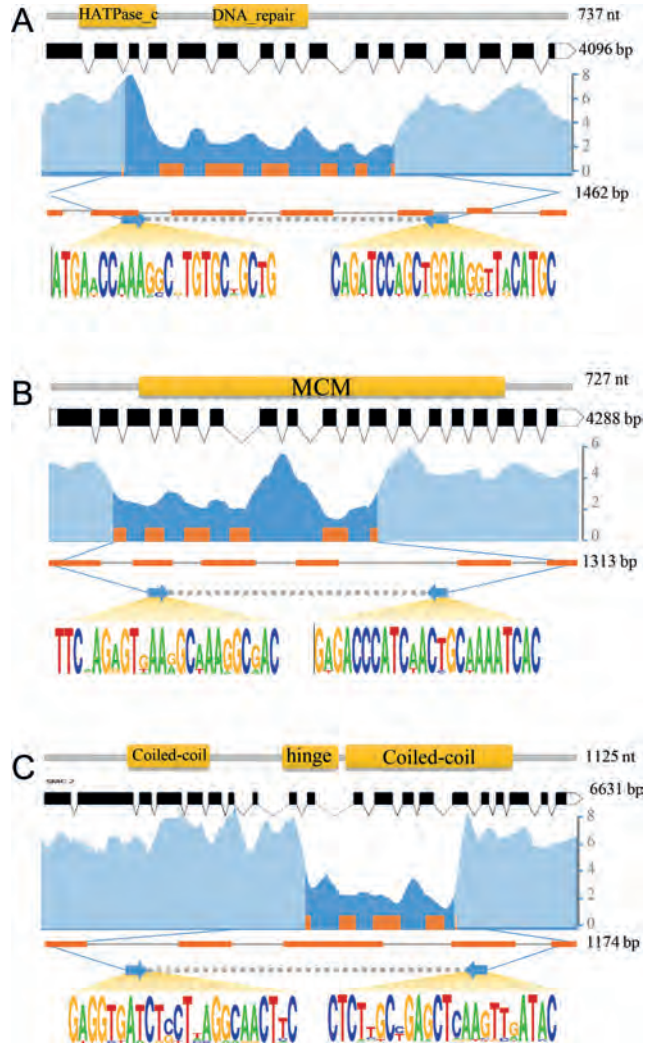


Fig. 2. Domains and nucleotide sequence conservation of *MLH1* (**A**), *MCM5* (**B**), and *SMC2* (**C**). Protein domains predicted by SMART are shown as yellow boxes at the top, with the number of translated amino acids in *Arabidopsis thaliana* shown on the right. Below the protein domain shown in black is the complete *A. thaliana* gene structure displayed by Exon-Intron Graphic Maker (http://wormweb.org/exonintron), with the number of nucleotides on the right. Below is a graph showing levels of site-specific congruence of phylogenetic signal identities among Brassicaceae species calculated by RAxML 8.0.2 (Stamatakis, 2006) with the window size of 100. Low value indicates high level of congruent phylogenetic signal. The region in dark blue represents the polymerase chain reaction (PCR)-amplified portion with the exons marked by orange, which give the most uniform phylogenetic signal. The gene structure of the PCR-amplified region is shown below with the exonic region in orange, intronic regions in gray, and the length of the *A. thaliana* PCR product shown on the right. Primers used in this study are marked as arrows, with the conserved and divergent sequences shown as generated using WebLogo (http://weblogo.berkeley.edu/logo.cgi). HATPase_c, histidine kinase-like ATPase, C-terminal domain (SMART accession number SM00387).

Table 3 Primer sequences and relative characters including length, Tm, GC%, and degeneracy

Gene	Primer	Sequence	Length, bp	Tm	GC%	Degeneracy
MLH1	MLH1_1F	5′-ATGAACCAAAGGCGTGTGCDGCTG-3′	24	66.3	54.2	3
MLH1	MLH1-2R	5′-ATTGATATCAACATGTTCVCGTGGCA-3′	25	63.9	52	1
MCM5	MCM5_1F	5′-TTCVAGAGTGAARGCAAAGGCGAC-3′	24	63.7	50	6
MCM5	MCM5_2R	5′-GTGATTTTGCAGTTGATGGGTCTC-3′	24	60.9	45.8	1
SMC2	SMC2_1F	5′-GAGGTGATCTCCTYAGGCAACTTC-3′	24	61	50	2
SMC2	SMC2_1R	5′-GTATYAACTTGAGCTCGGCAAGAG-3′	24	61.7	50	2

F, forward; R, reverse; Tm, annealing temperature; GC-content (GC%), percentage of either guanine or cytosine bases in a DNA molecule.

59.0 °C for 60 s, and 72.0 °C for 1.5 min, followed by a final extension at 72.0 °C for 5 min. Annealing temperatures for the other two genes were 60 °C for *MLH1* and 58 °C for *MCM5*. Additional cloning steps were used for some species within *Cardamine*, using vector pGEM-T and competent cell DH5 alpha. Three to five clones per amplification were sequenced on both forward and reverse strands using the same primer pair. Forward and reverse sequence strands were assembled with the ContigExpress program (http://www.contigexpress.com) and were then confirmed manually. Nucleotide sequences obtained from PCR and transcriptomes were submitted to GenBank, with accession numbers provided in Table S1.

Phylogenetic analysis

For sequence alignment and phylogenetic analysis, full-length *MLH1, MCM5*, and *SMC2* genes (with exons and introns) including 3051, 3005, and 4156 bp (for *A. thaliana*), respectively, were retrieved from genomic datasets, while the exon portion of their homologs were obtained from transcriptomic datasets. In addition, the PCR-amplified regions contained partial sequences with several exonic and intronic regions (mentioned above), whereas regions flanking the PCR-amplified loci were treated as missing data in the alignment for those taxa with only PCR-amplified sequences for the three genes. The sequences of each taxon were concatenated, forming a supermatrix with Seaview 4.4.2 (Gouy et al., 2010) and aligned using muscle 3.8.31 (Edgar, 2004) and subsequently adjusted manually. Maximum likelihood analyses were carried out using RAxML 8.0.2 (Stamatakis, 2006). Analysis was performed under the general time reversible model with the shape of the gamma distribution ($GTR + \Gamma$) as determined by Modeltest 3.7 (Posada & Crandall, 1998). Optimal tree searches were carried out with 100 random sequence addition replicates. Branch support was assessed using 100 rapid bootstrap replicates. Maximum likelihood bootstrap proportions (BP) $\geq$70% were considered strong support (Hillis & Bull, 1993).

Bayesian analyses were implemented with MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) under a time-free model. Posterior probability (PP) support values $\geq$0.95 were considered strong support for individual clades. MrBayes analyses were performed on the concatenated data. For consistency of results, two independent Markov chain Monte Carlo analyses were carried out for 200 000 generations to calculate PP. Prior probabilities for all trees were equal, starting trees were random, sampling every 1000 generations, and burn-in values were determined empirically from the likelihood values. The consistency of stationary-phase likelihood values and estimated parameter values was determined using Tracer 1.5 (Rambaut & Drummond, 2009). Bayesian PPs were determined by building a 50% majority-rule consensus tree from two Markov chain Monte Carlo analyses after discarding the 20% burn-in generations.

Results

Single-copy nuclear genes are excellent phylogenetic markers

As shown by Zhang et al. (2012), *MLH1*, *SMC2*, and *MCM5* are maintained as single-copy in most angiosperm species, consistent with an earlier report that genes engaged in DNA/RNA metabolisms tend to lose duplicate and remain orthologous after duplication (Blanc & Wolfe, 2004). Extensive phylogenetic studies have shown that members of *SMC*, *MCM*, and *MLH* gene families are maintained as one copy in most species (Forsburg, 2004; Lin et al., 2007; Surcel et al., 2008). These nuclear genes provide excellent markers to trace the evolution history of plants. We inspected their copy numbers in eight Brassicaceae species with fully sequenced genomes and found that most of them had only one copy for these species, except for *Brassica rapa*, which has undergone a recent whole genome duplication (Lysak et al., 2005; Yang et al., 2006). Two members from the *SMC* family (*SMC2* and *SMC6*) each have two copies across Brassicaceae. The *SMC2* duplication event occurred before the divergence of Brassicaceae (Fig. S1). There have been enough variations accumulated between the two copies, so that we could easily distinguish them, which is also seen in the application of two paralogous copies of *Adh1* and *Adh2*. In *Arabidopsis thaliana*, one copy of *SMC2* functions as subunit E in chromosome condensation complex, condensin. It is located on chromosome 5 (*AT5G62410*). Another copy functions similarly as structural maintenance of chromosome protein 2 and is mapped to chromosome 3 (*AT3G47460*). Both copies are maintained in all of the sequenced Brassicaceae genomes (Fig. 1). In our study, the homologs of *AT5G62410* were used as representative of *SMC2* for all sampled species.

The percentages of successful PCR for *SMC2*, *MCM5*, and *MLH1* were 93.7%, 87%, and 90.1%, respectively. These, along with gene sequences retrieved from public databases, include in total 53 *SMC2*, 55 *MCM5*, and 55 *MLH1* gene sequences from 63 Brassicaceae species. The lengths of each gene's initial alignment range from 3005 bp (*MCM5*) to 4156 bp (*SMC2*), with over one-third of the aligned regions covered by

Table 4 Characters of selected nuclear genes

Gene	Length, bp	Length of exon, bp	PI characters, bp (%)	Variable sites	GC%	α Parameter	Function anotation
MLH1	4096	2322	908 (39.1)	1265	41.49	0.4655	MUTL-homologue 1, DNA mismatch repair protein
MCM5	4288	2208	753 (34.1)	1066	42.69	0.4936	Minichromosome maintenance family protein, DNA replication licensing factor
SMC2	6364	3528	1333 (37.7)	1920	41.49	0.4655	Structural maintenance of chromosomes (SMC) family protein

GC-content (GC%), percentage of either guanine or cytosine bases in a DNA molecule.
Sequence lengths of selected genes, including length of exons, came from *Arabidopsis thaliana* (L.) Heynh. Parsimony-informative (PI) sites, variable sites, GC%, and gamma parameter for site rates (α Parameter) were calculated by mega. Function annotations were cited from Phytozome version 10.0 (http://phytozome.jgi.doe.gov/pz/portal.html).

the amplified loci, which contain both highly conserved and divergent sites. Gene length, length of exons, species coverage, parsimony-informative (PI) sites, variable sites, Guanine-cytosine (GC) frequencies, and α parameter (gamma parameter for site rates) are summarized in Table 4. The lengths of coding regions among different taxa are generally conserved. But lengths of introns could sometimes vary dramatically across taxa. We deleted such highly variable intronic regions in our alignment to reduce phylogenetic noise and the proportion of missing data. These highly variable regions mostly occurred in the middle of the intron with length varying from 6 bp to 25 bp in the alignment. Their corresponding positions in *A. thaliana* can be found in Table S2. The nucleotide sequences are highly conserved in exons (>92% global identities of all three genes). The conservation in terms of length, copy number, and exon sequences could be attributed to their functions in DNA/RNA metabolism and is important for primer design and sequence alignment. In the cases of our taxon sampling, the gamma parameters for site rates range from 0.466 for both *SMC2* and *MLH1*, to 0.494 for *MCM5*; thus, the three genes showed similar patterns of variability. Further analyses indicated that these genes are phylogenetically informative, with average frequency of PI sites greater than 30% (Table 4). The PI site proportion does not show significant heterogeneity between genes, but was especially high at third codon positions and in introns.

Further analyses of the nucleotide substitution model shows that $GTR+I+\Gamma$ is the fittest model for all three genes (Table S3). This result indicates these genes may have evolved under essentially very similar evolutionary patterns. A more detailed examination of site-specific placement bias in each gene reveals that the PCR-amplified region within each gene has the most stable phylogenetic signal (Fig. 2). Although the average node placement distance can be high in other regions of the genes, with possibly incongruent phylogenetic signals, this would not have a strong impact on our phylogenetic reconstructions in that only few genome or transcriptome data are available for those regions. For example, the PCR-amplified region from 540 to 1900 bp of the *MCM5* coding DNA sequence region (Fig. 2B) has a relatively low average node distance, whereas the intron region has higher phylogenetic divergence. Consequently, the exon regions with uniform phylogenetic signal are useful for resolving the deep nodes of early diversification. The divergent intron regions can provide crucial information to discern the subtle differences between closely related species.

Single-gene phylogenies were reconstructed for each of the three genes (Figs. S2–S4). They were largely consistent with well-established organismal relationships, suggesting that these genes were orthologous and phylogenetically informative. Although multiple sequences were occasionally obtained in one species, they always formed adjacent terminal branches in phylogenetic trees, suggesting that they resulted from recent polyploidization and should not affect phylogenetic relationships of more distantly related groups in this study.

Strongly supported phylogenies within *Cardamine*, *Lepidium*, and *Brassica*

Using a concatenated supermatrix of the sequences isolated from PCR amplification and retrieved from public databases, we built phylogenetic trees including 65 taxa. Our results showed that a combination of three nuclear loci resulted in a well-resolved phylogeny at various phylogenetic depths (Fig. 3). In the topology here with *Populus trichocarpa* as an outgroup, *Cleome serrulata* (Cleomaceae) was sister to a maximally supported clade of Brassicaceae; furthermore, *Aethionema* (including *Aethionema subulatum* and *Aethionema arabicum*) was the sister to all other Brassicaceae species (members of the core Brassicaceae), in agreement with previous studies (Al-Shehbaz et al., 2002; Huang et al., 2016). Within the core Brassicaceae clade, the phylogeny grouped all but one (*Arabis alpina*) of the remaining taxa into two clades. One included species from the previously defined Lineage I, and the other has those of Lineage II, again in agreement with recent results on Brassicaceae phylogeny (Al-Shehbaz et al., 2002; Huang et al., 2016). Eight species of the genus *Lepidium* formed a maximally supported group as the first divergent branch of Lineage I. Within the *Lepidium* clade, *L. campestre* (L.) W. T. Aiton and *L. perfoliatum* L. clustered as the sister to a large clade with the other *Lepidium* species (100 BP/1.0 PP). In addition, *L. didymum* L., which was formerly identified as *Coronopus didymus* (L.) Sm., is nested within *Lepidium* (100 BP/1.0 PP). This result was in agreement with a previous report on the systematics of *Lepidium* (Al-Shehbaz et al., 2002).

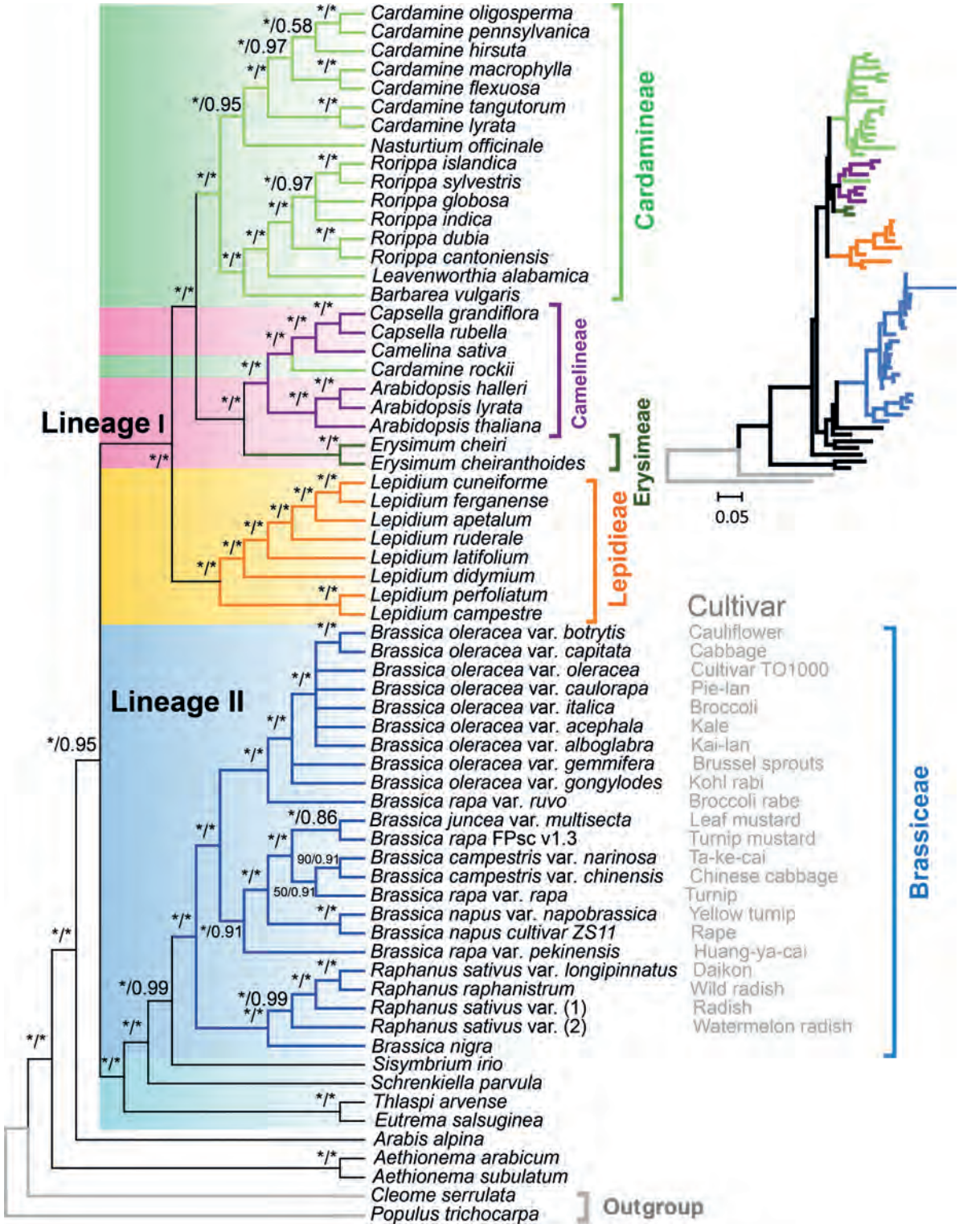


Fig. 3. Maximum likelihood majority-rule consensus tree of Brassicaceae based on the concatenated *MLH1*, *MCM5*, and *SMC2* datasets. Values above branches are maximum likelihood bootstrap values (left) and Bayesian inference posterior probabilities (right). Star indicates either a bootstrap proportion of 100 or posterior probability of 1.0. The phylogram of the reconstructed phylogeny is displayed on the upper right. Within clade Brassiceae, common names of cultivated vegetables are shown on the right.

A

	Amborella trichopoda	*Oryza sativa*	*Zea mays*	*Elaeis guineensis*	*Aquilegia coerulea*	*Solanum lycopersicum*	*Populus trichocarpa*	*Fragaria vesca*	*Arabidopsis thaliana*	*Medicago truncatula*
Amborella trichopoda	100%	69.64%	68.90%	66.94%	71.93%	69.29%	71.43%	70.19%	67.61%	68.05%
Oryza sativa	71%	100%	86.09%	75.84%	66.80%	67.32%	68.02%	67.37%	64.97%	66.80%
Zea mays	69.37%	85.15%	100%	77.11%	66.30%	66.68%	67.34%	67.76%	64.94%	66.05%
Elaeis guineensis	75%	73.75%	72.50%	100%	64.40%	69.49%	66.10%	72.45%	66.94%	60.59%
Aquilegia coerulea	71.89%	66.15%	66.10%	71.25%	100%	69.93%	72.96%	72.22%	69.11%	72.86%
Solanum lycopersicum	70.26%	66.71%	65.12%	72.50%	69.97%	100%	73.63%	73.16%	70.68%	71.44%
Populus trichocarpa	72.18%	67.27%	66.24%	72.50%	72.40%	73.88%	100%	75.93%	73.53%	73.66%
Fragaria vesca	71.59%	66.29%	66.52%	77.50%	73.54%	74.44%	77.70%	100%	71.76%	74.98%
Arabidopsis thaliana	68.49%	65.45%	65.54%	73.75%	69.69%	70.67%	74.10%	72.71%	100%	71.77%
Medicago truncatula	68.49%	65.39%	63.70%	67.50%	72.17%	70.33%	74.01%	74.01%	69.91%	100%

B

	Amborella trichopoda	*Oryza sativa*	*Zea mays*	*Elaeis guineensis*	*Aquilegia coerulea*	*Solanum lycopersicum*	*Populus trichocarpa*	*Fragaria vesca*	*Arabidopsis thaliana*	*Medicago truncatula*
Amborella trichopoda	100%	71.70%	71.28%	75.53%	72.87%	70.86%	70.54%	70.15%	69.88%	72.40%
Oryza sativa	80.08%	100%	88.03%	80.29%	70.68%	70.41%	70.05%	69.92%	69.64%	70.41%
Zea mays	79.66%	96.71%	100%	80.06%	70.59%	69.95%	69.69%	69.82%	69.45%	70.18%
Elaeis guineensis	81.69%	90.84%	90%	100%	77.12%	74.91%	75.36%	72.81%	73.66%	75.42%
Aquilegia coerulea	76.18%	75.93%	75.38%	81.52%	100%	73.74%	75.12%	73.15%	71.15%	73.92%
Solanum lycopersicum	76.04%	76.47%	75.92%	81.69%	79.25%	100%	74.04%	73.15%	72.93%	74.68%
Populus trichocarpa	73.39%	74.93%	74.24%	78.47%	76.90%	78.08%	100%	76.02%	75.37%	77.85%
Fragaria vesca	72.43%	74.68%	74.12%	77.62%	75.24%	77.35%	77.07%	100%	72.87%	76.67%
Arabidopsis thaliana	72.56%	74.07%	74.34%	79.32%	74.13%	78.46%	78.18%	76.79%	100%	74.22%
Medicago truncatula	74.79%	75.92%	75.78%	80.16%	78.14%	79.15%	79.58%	80.87%	78.87%	100%

C

	Amborella trichopoda	*Oryza sativa*	*Zea mays*	*Elaeis guineensis*	*Aquilegia coerulea*	*Solanum lycopersicum*	*Populus trichocarpa*	*Fragaria vesca*	*Arabidopsis thaliana*	*Medicago truncatula*
Amborella trichopoda	100%	70.36%	68.86%	73.07%	73.53%	70.83%	74.25%	72.44%	70.24%	71.93%
Oryza sativa	73.71%	100%	83.69%	77.40%	70.16%	68.10%	70.63%	74.57%	68.91%	69.33%
Zea mays	65.59%	82.07%	100%	NA	71.06%	NA	69.41%	NA	67.58%	69.47%
Elaeis guineensis	76.89%	81.61%	NA	100%	74.42%	71.89%	73.97%	72.44%	70.37%	72.11%
Aquilegia coerulea	73.53%	71.72%	64.33%	77.74%	100%	72.58%	76.01%	72.30%	71.41%	74.85%
Solanum lycopersicum	71.49%	70.10%	NA	75.54%	71.80%	100%	73.91%	71.59%	70.88%	73.04%
Populus trichocarpa	75.17%	73.63%	67.56%	78.24%	77.51%	75.80%	100%	77.41%	75.54%	78.06%
Fragaria vesca	65.57%	69.26%	NA	67.62%	66.80%	67.21%	72.13%	100%	74.57%	75.85%
Arabidopsis thaliana	71.30%	69.39%	62.72%	72.51%	71.18%	70.72%	76.04%	65.98%	100%	73.20%
Medicago truncatula	73.19%	73.38%	67.56%	76.22%	75.80%	75.50%	81.64%	71.31%	75.19%	100%

Fig. 4. Nucleotide and amino acid sequence identity of *MLH1* (**A**), *MCM5* (**B**), and *SMC2* (**C**) among 10 representative angiosperm species. Pairwise sequence identities are calculated using the SIAS webserver (http://imed.med.ucm.es/Tools/sias.html). Nucleotide and amino acid sequence identities are shown at upper right and bottom left, respectively. NA, data not available.

The other major branch of Lineage I was comprised of species belonging to the tribes Cardamineae, Camelineae, and Erysimeae. The Camelineae and Erysimeae members were more closely related (100 BP/1.0 PP) than they are to Cardamineae. The tribe Camelineae was represented by six species with sequenced genomes: *A. thaliana*, *A. lyrata*, *A. halleri* O'Kane & Al-Shehbaz, *C. grandiflora*, *C. rubella*, and *Camelina sativa* (L.) Crantz. In addition, *Cardamine rockii* O. E.

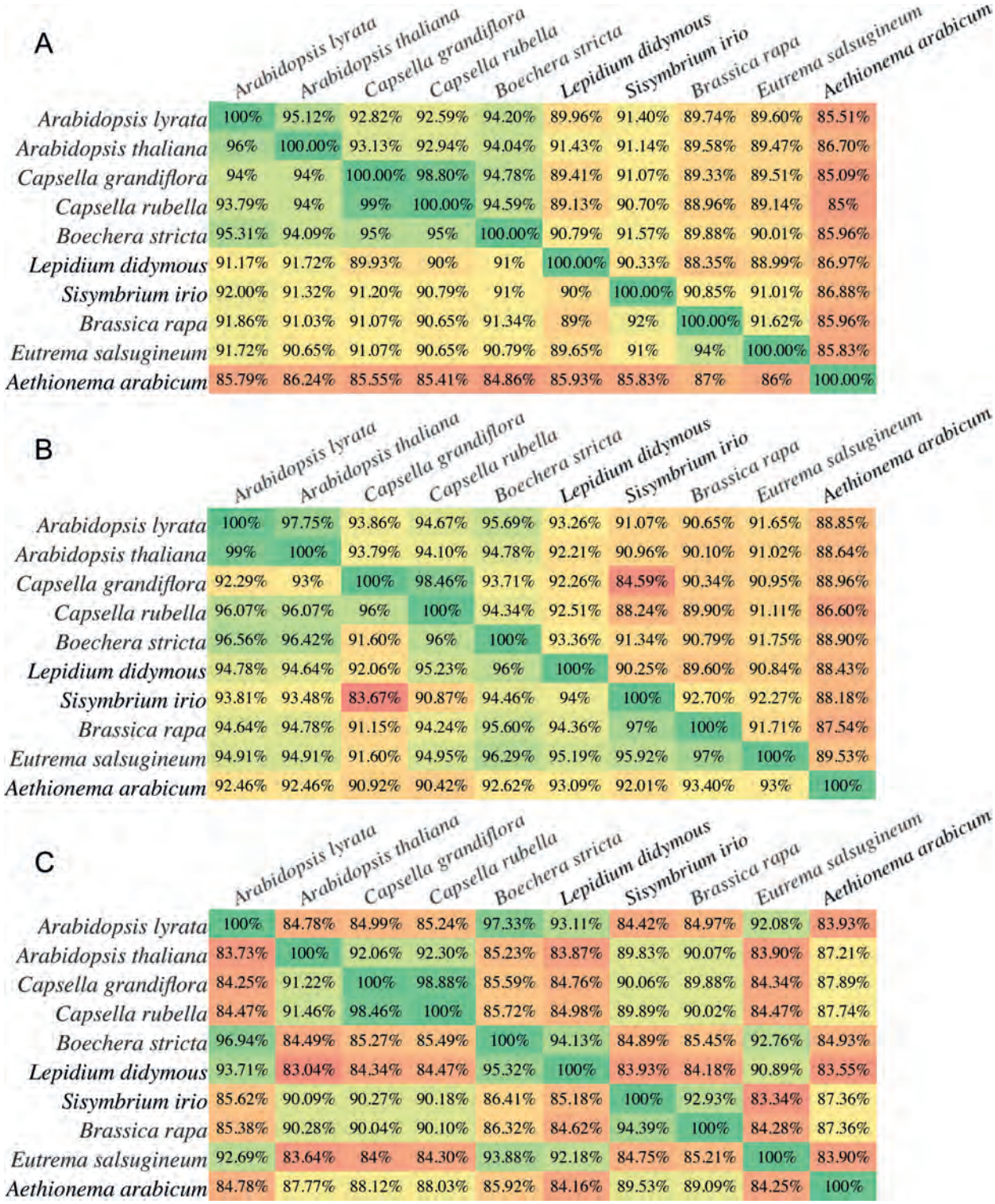


Fig. 5. Nucleotide and amino acid sequence identity of *MLH1* (**A**), *MCM5* (**B**), and *SMC2* (**C**) among 10 representative Brassicaceae species. Pairwise sequence identities are calculated using the SIAS webserver (http://imed.med.ucm.es/Tools/sias.html). Nucleotide and amino acid sequences identities are displayed at upper right and bottom left, respectively.

Schulz clustered with a clade composed of *Camelina* Crantz and *Capsella* Medik. (100 BP/1.0 PP), rather than being close to members of Cardamineae. *Erysimum cheiri* Crantz and *Erysimum cheiranthoides* L., with sequences from transcriptomics, are sisters with maximal support, representing the tribe Erysimeae. Our sampling of the tribe Cardamineae included 16 species from five genera. Cardamineae were divided into two clades (100 BP/1.0 PP),

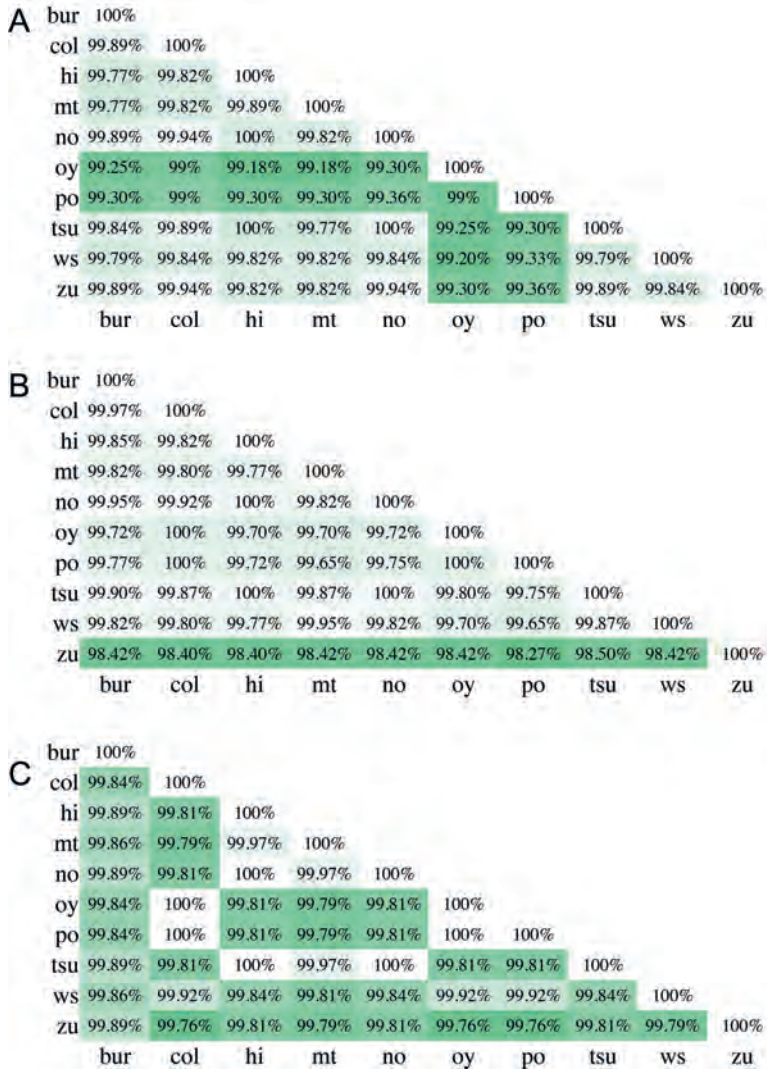


Fig. 6. Pairwise nucleotide sequence identity of *MLH1* (**A**), *MCM5* (**B**), and *SMC2* (**C**) among 10 representative *Arabidopsis thaliana* ecotypes, calculated using the SIAS webserver (http://imed.med.ucm.es/Tools/sias.html). Data were collected from the 19 genomes of the *A. thaliana* project (http://mus.well.ox.ac.uk/19genomes/). Abbreviations for ecotype names follow Gan et al. (2011): bur, Burren; col, Columbia; hi, Hilversum; mt, Martuba; no, Nossen; oy, Oystese; po, Poppelsdorf; tsu, Tsu; ws, Wassilewskija; zu, Zurich.

one containing genera *Barbarea* W. T. Aiton, *Leavenworthia*, and *Rorippa* (100 BP/1.0 PP), the other containing *Nasturtium* and *Cardamine* (100 BP/0.95 PP). Within *Rorippa*, six species were divided into two clades, containing two and four species, respectively. *Rorippa dubia* (Pers.) Hara and *Rorippa cantoniensis* (Lour.) Ohwi formed one of the clades (100 BP/1.0 PP). Within the other clade, *R. islandica* (Oeder) Borbás and *R. sylvestris* (L.) Besser share the closest affinity (100 BP/1.0 PP), with their relationship to the other two species unclear. Similarly, *Nasturtium officinale* W. T. Aiton was sister to a well-supported clade of seven *Cardamine* species. Relationships within genus *Cardamine* were more complicated. A basal clade was composed of two species *C. lyrata* Bunge and *C. tangutorum* O. E. Schulz (100 BP/1.0 PP). *Cardamine oligosperma* Nutt. and *C. pensylvanica* Muhl. ex Willd. were sisters with maximal support, as were *C. flexuosa* With. and *C. macrophylla* Willd. (both 100 BP/1.0 PP), and the position of *C. hirsuta* L. as sister to the clade of *C. oligosperma* and *C. pensylvanica* was moderately supported (100 BP/0.58 PP).

In Lineage II, four species belonging to the previously defined Expanded Lineage II formed successive sisters to the tribe Brassiceae, with maximal support. These species included *Sisymbrium irio*, *Thlaspi arvense*, *Eutrema salsuginea*, and *Schrenkiella parvula*. Our sampled taxa of the tribe Brassiceae included two genera, *Brassica* and *Raphanus* and they form a maximally supported clade. Within Brassiceae, all four *Raphanus sativus* L. cultivars form a well-supported clade (100 BP/1.0 PP), and most *Brassica* taxa form another maximally supported clade. Our results indicated that *Brassica nigra* (L.) W. D. J. Koch was more closely related to *Raphanus* than to other *Brassica* species (100 BP/1.0 PP), which was also

consistent with previous findings that *Brassica* is not monophyletic (Yang et al., 2002; Arias & Pires, 2012). Within the large *Brassica* clade, the genetic similarity and complicated domestication histories of various cultivars brought challenge for phylogenetic reconstruction. However, the introns and third codon position of the nuclear markers provided crucial information for good resolution even within species. Four *Brassica rapa* accessions held four different positions. Cultivars belonging to *Brassica oleracea* L. composed the majority of our sampling diversity within *Brassica*, forming a maximally supported clade with internal resolutions of either high or low support values. The relationships among *B. oleracea*, *B. rapa*, *B. napus* L., *B. juncea* (L.) Czern., and *B. campestris* L. was not clear due to the limitation of sequence data and the non-monophyletic results. Among these vegetables, *B. napus* L. var. *napobrassica* (L.) Hanelt, or yellow turnip, is a cross between the cabbage and the turnip. Its placement on the tree suggested that it was more similar to cabbage (*B. oleracea*) than others. *Arabis alpina* was the basal-most core Brassicaceae species (100 BP/1.0 PP) among taxa sampled here.

Analysis of sequence similarity for three marker genes among angiosperms

The above phylogenetic analysis suggested that the *MCM5*, *MHL1*, and *SMC2* genes could be effective markers for revolving relationships between members of a family or even a genus. To further explore the sequence similarities of these genes for their potential applications in the phylogenetic studies of plants of various evolutionary diversities, we obtained their homologs in representative angiosperms, Brassicaceae species, as well as *A. thaliana* ecotypes (Gan et al., 2011) and compared their pairwise sequence identity (Figs. 4–6). Ten angiosperm species were selected to be phylogenetically representative. This included the basal-most angiosperm *Amborella trichopoda* Baill., three monocot species, and representatives of major eudicot clades. Within this angiosperm sampling, the pairwise sequence identity of *MCM5* ranged between 72.43%–96.71% for amino acid sequences and 69.45%–88.03% for nucleotide acid sequences (Fig. 4). Similar results can also be seen for *SMC2* and *MLH1*. When we compared sequence identities within the Brassicaceae family, higher pairwise identities could be obtained. Generally, the nucleotide sequence identity ranges from 83.34% to 98.88% and the amino acid sequence identity ranges from 83.04% to 99.00%. Likewise, results from the 10 *A. thaliana* accessions revealed that, although both nucleotide and amino acid sequences are highly similar among different ecotypes, we can still find SNP and indel sites within the sampled loci. Because these genes are unusually long with 3000 bp or more, even a low percentage of differences can provide dozens or more sites for comparison. As shown in Fig. 1, other members of these three gene families are also stably maintained as single copy or low copy, providing additional markers if more information is needed.

Discussion

Newly identified nuclear genes are suitable for phylogenetic reconstructions

Extensive phylogenetic analyses have been undertaken using mainly organellar or rDNA markers. Although they are easy to obtain, they are often too conserved to provide sufficient signals for resolving relatively close relationships. Nuclear genes are both numerous and rich in phylogenetic signals, but many nuclear genes have paralogs and should be used with care to avoid misleading signals. Previous work on angiosperm phylogeny revealed that nuclear markers were also highly informative for low-rank taxonomic groups (Zhu & Ge, 2005; Yuan et al., 2009; Salas-Leiva et al., 2013). These genes are primarily single-copy and conserved. The nuclear markers used here were previously described for use in a study of angiosperm-wide phylogeny (Zhang et al., 2012) and they were comparable to other markers that were screened from ~1000 low-copy putative orthologous genes by comparing genomes of representative angiosperms (Zeng et al., 2014). They were reported to be suitable for angiosperm phylogenetic reconstruction when conserved exon sequences were used. Here, to provide more divergent sequences with signals for within-genus relationships, we took advantage of both exonic and intronic regions for phylogenetic reconstruction. As illustrated by our result, the highly conserved exonic regions could facilitate the design of primers with high amplification efficiency for a wide range of organisms. First, these loci are easy to amplify by PCR reaction. Cloning steps are not necessary unless there are occasional recent gene duplication events. Second, conserved exons make it easy to align across distantly related species. Finally, together with intron sequences, the nuclear gene markers provide information at various phylogenetic depths. They are especially powerful for resolving relationships involving recent and rapid radiation.

With the development of sequencing technology, more and more genome and transcriptome data will be available. The growing genome datasets will facilitate the identification of low-copy nuclear genes as phylogenetic markers for more and more plant groups. At the same time, the PCR-based approach presented here provides a complementary means for obtaining a small number of genes from a large number of taxa, without the expense of transcriptomics and avoiding the need for great computation capability that is associated with the analysis of many genes. Furthermore, the information on the sequence similarity indicate that there are many variable sites from the comparison of different Brassicaceae species; there are even variations between different *Arabidopsis thaliana* ecotypes. This information and the phylogenetic results (see below) together provide strong evidence that these genes can serve as effective markers for investigation of relationships between species in the same genus.

Potential newly defined species relationships in Brassicaceae

One of the advantages of nuclear gene markers is their usability for phylogenies at various depths. For example, with three nuclear loci, we obtained placement of all sampled genera congruent to the latest and most comprehensive analysis (Al-Shehbaz, 2012) with strong support (100 BP/>0.95 PP). Within each genus, internal nodes are well resolved, providing important information to investigate their recent evolutionary history. At the same time, greater resolution for some of the relationships among species would probably benefit from some additional

sequences, either from these three genes, or from other similarly conserved genes. In addition, very difficult relationships might also need more genes, as shown recently by the phylogenetic study of a greater number of tribes in Brassicaceae (Huang et al., 2016).

Cardamineae is a tribe containing 14 genera and 352 species. We sampled 16 species from five genera, which were grouped into two clades: one clade with *Cardamine* and *Nasturtium*, the other containing *Rorippa*, *Leavenworthia*, and *Barbarea*. Previous studies found a close relationship between *Cardamine* and *Rorippa* (Yang et al., 1999), and between *Cardamine* and *Nasturtium* (Beilstein et al., 2006), but they did not include all these genera. Our placement of the five genera within tribe Cardamineae has received the highest support so far. This is in good agreement with the initial hypothesis that *Nasturtium* is more closely related to *Cardamine* than to other genera in tribe Cardamineae (Al-Shehbaz & Price, 1998). Our results contribute to the understanding of the early divergence events within Cardamineae. Additionally, we found that one *Cardamine* species, *Cardamine rockii*, did not cluster with other *Cardamine* species, but rather it was grouped with members of Camelineae with strong support (100 BP/1.0 PP). Such phylogenetic placement has not been reported before. Thus this result suggests that this species might be misclassified and further investigation with more *Cardamine* species and Camelineae members will be needed to test this idea.

Within *Lepidium*, all relationships were strongly supported in the phylogeny here (100 BP/1.0 PP). *Lepidium perfoliatum* and *L. campestre* (100 BP/1.0 PP) formed the basal clade of *Lepidium*, consistent with previous results (Mummenhoff et al., 2001; Lee et al., 2002). Both of these studies placed the two species as the basal lineage with a larger sampling size in *Lepidium*. Three species, *L. apetalum* Willd., *L. ferganense* Korsh., and *L. cuneiforme* C. Y. Wu, form a highly supported group, allowing the placement of *L. cuneiforme*, which is endemic to China, for the first time. In addition, the placement of *L. apetalum* is similar to the results based on ITS sequences (Mummenhoff et al., 2001). The markers used here might be able to resolve the relationships in *Lepidium* when more species can be analyzed in the future.

Previous studies have shown that *Brassica* is not a monophyletic group (Yang et al., 2002; Arias & Pires, 2012). For example, a phylogenetic study with *B. rapa*, *B. nigra*, and *R. sativus* (Yang et al., 2002) indicated that *B. nigra* is sister to the clade of *B. rapa* and *R. sativus*. Human domestication of taxa within tribe Brassiceae also caused difficulty for tracing their ancient origins. In particular, multiple hybridization events between different cultivars along with whole genome duplication make it a major challenge to build phylogenetic trees across *Brassica* using nuclear genes. Our attempt revealed that the three nuclear markers used, *MLH1*, *SMC2*, and *MCM5*, are sufficiently variable and useful for detecting subtle differences between accessions within a species. Human domestication of *Brassica* is often explained by the U triangle theory (Nagaharu, 1935). The theory states the evolutionary relationships between modern vegetables and three ancestral species of *Brassica* by comparing their chromosome number. However, the long-discussed U triangle theory has not been rigorously tested by phylogenetic analysis. Our sampling of U triangle species included assumed ancestor species (*B. nigra*, *B. oleracea*, and *B. rapa*), as well as modern vegetables and oil seed crops (*B. juncea* and *B. napus*). This provides a unique opportunity to test the U triangle theory and further investigate the impact of hybridization on phylogenetic reconstruction. The problem lies in that our four accessions of species *B. rapa* hold four different places on the tree. This could be a result either from incorrect classification of vegetables, or other problems. More marker genes and more taxa are needed to resolve the proposed hybridization events. Nevertheless, the preliminary analysis here suggests that these nuclear genes contain variable sequences with phylogenetic signals that could be used to address such difficult questions.

Acknowledgements

We thank Professors Dunyan Tan, Hongzhi Kong, Guangyuan Rao, and Ji Yang for their help in plant sample collection and species identification, and Yi Hu for RNA extraction of some taxa. We are also grateful to Ning Zhang, Liping Zeng, and Qi Li for their help in computational analysis. In addition, assistance from Yue Jin, Engao Zhu, Chenjiang You, and Bailong Zhang in laboratory work was of great importance and we wish to express our gratitude. This work was supported by a grant from the National Natural Science Foundation of China (Grant No. 91531301) to H.M.

References

Al-Shehbaz IA. 2012. A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon* 61: 931–954.

Al-Shehbaz IA, Mummenhoff K, Appel O. 2002. *Cardaria*, *Coronopus*, and *Stroganowia* are united with *Lepidium* (Brassicaceae). *Novon* 12: 5–11.

Al-Shehbaz IA, Price RA. 1998. Delimitation of the genus *Nasturtium* (Brassicaceae). *Novon* 8: 124–126.

Álvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417–434.

Arias T, Pires JC. 2012. A fully resolved chloroplast phylogeny of the *Brassica* crops and wild relatives (Brassicaceae: Brassiceae): Novel clades and potential taxonomic implications. *Taxon* 61: 980–988.

Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL, Warwick SI, Windham MD, Al-Shehbaz IA. 2006. Toward a global phylogeny of the Brassicaceae. *Molecular Biology and Evolution* 23: 2142–2160.

Baldwin BG, Markos S. 1998. Phylogenetic utility of the External Transcribed Spacer (ETS) of 18S–26S rDNA: Congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution* 10: 449–463.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247–277.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. 2006. Brassicaceae phylogeny and trichome evolution. *American Journal of Botany* 93: 607–619.

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA. 2008. Brassicaceae phylogeny inferred from phytochrome A and *ndhF*

sequence data: Tribes and trichomes revisited. *American Journal of Botany* 95: 1307–1327.

Berger SA, Krompass D, Stamatakis A. 2011. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Systematic Biology* 60: 291–302.

Birky CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences USA* 92: 11 331–11 338.

Blanc G, Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16: 1667–1678.

Buckler ES, Ippolito A, Holtsford TP. 1997. The evolution of ribosomal DNA divergent paralogues and phylogenetic implications. *Genetics* 145: 821–832.

De Bodt S, Maere S, Van de Peer Y. 2005. Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution* 20: 591–597.

Ding M, Zeng L, Zhang N, Ma H. 2012. The use of low-copy nuclear genes for reconstructing the phylogeny of low-level taxonomic hierarchies: Evidence from Brassicaceae. *Plant Diversity and Resources* 34: 211–221.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Forsburg SL. 2004. Eukaryotic MCM proteins: Beyond replication initiation. *Microbiology and Molecular Biology Reviews* 68: 109–131.

Fukuda T, Yokoyama J, Nakamura T, Song I-J., Ito T, Ochiai T, Kanno A, Kameya T, Maki M. 2005. Molecular phylogeny and evolution of alcohol dehydrogenase (*Adh*) genes in legumes. *BMC Plant Biology* 5: 6.

Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Rätsch G, Mott R. 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature* 477: 419–423.

Gaut BS, Clegg MT. 1991. Molecular evolution of *Alcohol Dehydrogenase 1* in members of the grass family. *Proceedings of the National Academy of Sciences USA* 88: 2060–2064.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* 27: 221–224.

Gozuacik D, Chami M, Lagorce D, Faivre J, Murakami Y, Poch O, Biermann E, Knippers R, Bréchot C, Paterlini-Bréchot P. 2003. Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Research* 31: 570–579.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 42: 182–192.

Huang CH, Sun R, Hu Y, Zeng L, Zhang N, Cai L, Zhang Q, Koch MA, Al-Shehbaz I, Edger PP, Pires JC. 2016. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular Biology and Evolution* 33: 394–412.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.

Kim C, Shin H, Chang YT, Choi HK. 2010. Speciation pathway of *Isoëtes* (Isoëtaceae) in East Asia inferred from molecular phylogenetic relationships. *American Journal of Botany* 97: 958–969.

Lee JY, Mummenhoff K, Bowman JL. 2002. Allopolyploidization and evolution of species with reduced floral structures in *Lepidium* L. (Brassicaceae). *Proceedings of the National Academy of Sciences USA* 99: 16835–16840.

Lin Z, Nei M, Ma H. 2007. The origins and early evolution of DNA mismatch repair genes – multiple horizontal gene transfers and co-evolution. *Nucleic Acids Research* 35: 7591–7603.

Lysak MA, Koch MA, Pecinka A, Schubert I. 2005. Chromosome triplication found across the tribe Brassiceae. *Genome Research* 15: 516–525.

Maddison WP. 1997. Gene trees in species trees. *Systematic Biology* 46: 523–536.

Mummenhoff K, Brüggemann H, Bowman JL. 2001. Chloroplast DNA phylogeny and biogeography of *Lepidium* (Brassicaceae). *American Journal of Botany* 88: 2051–2063.

Nagaharu U. 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Japanese Journal of Botany* 7: 389–452.

Nie ZL, Wen J, Azuma H, Qiu YL, Sun H, Meng Y, Sun W-B, Zimmer EA. 2008. Phylogenetic and biogeographic complexity of Magnoliaceae in the Northern Hemisphere inferred from three nuclear data sets. *Molecular Phylogenetics and Evolution* 48: 1027–1040.

Oh SH, Potter D. 2005. Molecular phylogenetic systematics and biogeography of tribe Neillieae (Rosaceae) using DNA sequences of cpDNA, rDNA, and *LEAFY*. *American Journal of Botany* 92: 179–192.

Posada D, Crandall KA. 1998. Modeltest: Testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

Rambaut A, Drummond A. 2009. Tracer version 1.5.0. [online]. Available from http://tree.bio. ed.ac.uk/software/tracer/ [accessed 1 July 2013].

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Sang T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology* 37: 121–147.

Schranz ME, Mitchell-Olds T. 2006. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* 18: 1152–1165.

Salas-Leiva DE, Meerow AW, Calonje M, Griffith MP, Francisco-Ortega J, Nakamura K, Stevenson DW, Lewis CE, Namoff S. 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: Congruence of concatenated parsimony, likelihood and species tree inference methods. *Annals of Botany* 112: 1263–1278.

Small RL, Cronn RC, Wendel JF. 2004. LAS Johnson Review No. 2. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145–170.

Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF. 1998. The tortoise and the hare: Choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85: 1301–1315.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* 96: 336–348.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stewart C, Via LE. 1993. A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *Biotechniques* 14: 748–750.

Surcel A, Zhou X, Quan L, Ma H. 2008. Long-term maintenance of stable copy number in the eukaryotic SMC family: Origin of a vertebrate meiotic SMC1 and fate of recent segmental duplicates. *Journal of Systematics and Evolution* 46: 405–423.

Warwick SI, Mummenhoff K, Sauder CA, Koch MA, Al-Shehbaz IA. 2010. Closing the gaps: Phylogenetic relationships in the Brassicaceae based on DNA sequence data of nuclear ribosomal ITS region. *Plant Systematics and Evolution* 285: 209–232.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, Pokorny L, Shaw AJ, DeGironimo L, Stevenson DW, Surek B, Villarreal JC, Roure B, Philippe H, dePamphilis CW, Chen T, Deyholos MK, Baucom RS, Kutchan TM, Augustin MM, Wang J, Zhang Y, Tian Z, Yan Z, Wu X, Sun X, Wong GK-S, Leebens-Mack J. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences USA* 111: E4859–E4868.

Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI, Lim YP, Kang JJ, Hong JH, Kim CB, Bhak J, Bancroft I, Park BS. 2006. Sequence-level analysis of the diploidization process in the triplicated *FLOWERING LOCUS C* region of *Brassica rapa*. *Plant Cell* 18: 1339–1347.

Yang Y, Moore MJ, Brockington SF, Soltis DE, Wong GK-S, Carpenter EJ, Zhang Y, Chen L, Yan Z, Xie Y, Sage RF, Covshoff S, Hibberd JM, Nelson MN, Smith SA. 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Molecular Biology and Evolution* 32: 2001–2014.

Yang YW, Lai KN, Tai PY, Ma DP, Li WH. 1999. Molecular phylogenetic studies of *Brassica*, *Rorippa*, *Arabidopsis* and allied genera based on the internal transcribed spacer region of 18S-25S rDNA. *Molecular Phylogenetics and Evolution* 13: 455–462.

Yang YW, Tai PY, Chen Y, Li WH. 2002. A study of the phylogeny of *Brassica rapa*, *B. nigra*, *Raphanus sativus*, and their related genera using noncoding regions of chloroplast DNA. *Molecular Phylogenetics and Evolution* 23: 268–275.

Yuan YW, Liu C, Marx HE, Olmstead RG. 2009. The pentatricopeptide repeat (PPR) gene family, a tremendous resource for plant phylogenetic studies. *New Phytologist* 182: 272–283.

Zeng L, Zhang Q, Sun R, Kong H, Zhang N, Ma H. 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. *Nature Communications* 5: 4956.

Zhang N, Zeng L, Shan H, Ma H. 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytologist* 195: 923–937.

Zhu Q, Ge S. 2005. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* 167: 249–265.

Zimmer EA, Wen J. 2013. Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution* 66: 539–550.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12204/suppinfo:

Fig. S1. Maximum likelihood tree of *SMC2* paralogs from representative Brassicaceae species, with those from *Cleome serrulata* Pursh and *Populus trichocarpa* Torr. & A. Gray ex Hook. as outgroups. Values above branches are maximum likelihood bootstrap support values. Numbers after species names represent different paralogs.

Fig. S2. Fifty percent maximum likelihood majority-rule consensus tree of Brassicaceae based on single-gene phylogeny of *MLH1*. Values above branches are maximum likelihood bootstrap proportions (BP). Only branches with >50 BP are displayed.

Fig. S3. Fifty percent maximum likelihood majority-rule consensus tree of Brassicaceae based on single-gene phylogeny of *SMC2*. Values above branches are maximum likelihood bootstrap proportions (BP). Only branches with >50 BP are displayed.

Fig. S4. Fifty percent maximum likelihood (ML) majority-rule consensus tree of Brassicaceae based on single-gene phylogeny of *MCM5*. Values above branches are ML bootstrap proportions (BP). Only branches with >50 BP are displayed.

Table S1. GenBank accession numbers of *MLH1*, *MCM5*, and *SMC2* gene sequences from extracted DNA samples and transcriptomes.

Table S2. Positions and lengths of highly variable intron regions in *Arabidopsis thaliana*. The corresponding sites in the alignment were deleted in order to reduce phylogenetic noise.

Table S3. Result of Modeltest for *MLH1*, *MCM5*, and *SMC2* showing the top three substitution models with highest probability.

JSE Journal of Systematics and Evolution

doi: 10.1111/jse.12196

Research Article

Phylogeny and diversification of Chinese Araliaceae based on nuclear and plastid DNA sequence data

Rong Li[1] and Jun Wen[2]*

[1]Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

[2]Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012, USA

*Author for correspondence. E-mail: wenj@si.edu. Tel.: 202-633-4881. Fax: 202-786-2563.

Received 10 November 2015; Accepted 29 January 2016; Article first published online 14 March 2016

Abstract Chinese Araliaceae consist of 20 genera and ca. 175 species. To assess the evolutionary relationships of Araliaceae and their biogeographic diversification in China, the phylogeny of Chinese Araliaceae was constructed by sampling 96 accessions representing 20 genera and 50 species of Chinese Araliaceae and 45 closely related taxa using sequences of the nuclear ribosomal internal transcribed spacer (ITS) region and six plastid regions (the *ndhF* gene, the *trnL-trnF* region, the *rps16* intron, the *atpB-rbcL* intergenic spacer, the *rpl16* intron, and the *psbA-trnH* intergenic spacer). Phylogenetic analyses of the combined plastid and ITS data supported the results of the previously studies that the Chinese members of Araliaceae were scattered within the Asian Palmate group and the *Aralia-Panax* group with *Osmoxylon* at the base of core Araliaceae. The generic status of *Pentapanax* and *Tupidanthus* is not supported. Our analysis clearly places them in *Aralia* and Asian *Schefflera*, respectively. In a broader phylogenetic framework of Araliaceae, based on the fossil-calibrated Bayesian dating, Chinese Araliaceae was inferred to have originated in Asia and underwent a rapid radiation in its evolutionary history. Its diversification is hypothesized to have been driven largely by the orogenies in Asia during the Cenozoic. In China, the distribution pattern of the phylogenetic diversity of Araliaceae corresponds with its taxonomic diversity across the entire region.

Key words: *Aralia-Panax* group, Asian Palmate group, Chinese Araliaceae, diversification, phylogeny.

Araliaceae (the ginseng family) consist of approximately 45 genera and 1500 species with a wide distribution in tropical and subtropical Asia, the Pacific and the Indian Ocean basin, and the neotropics, with a few well-known genera from the north and south temperate zones (Wen et al., 2001; Plunkett et al., 2004). Members of Araliaceae are characterized by mostly woody habit and relatively conserved floral morphology, i.e., mostly 5-merous flowers with inferior ovaries, inflorescences commonly a compound umbel (rarely a raceme or head), and fruit a drupe with 2-5 (rarely >5) seeds, yet highly variable leaf morphology (simple, palmately compound, to variously pinnately compound) (Philipson, 1970; Wen et al., 2001; Yi et al., 2004). Phylogenetic studies of the Araliaceae based on analyses of sequence data from nuclear ribosomal DNA and chloroplast DNA have circumscribed four major monophyletic groups (the Asian Palmate group, the *Polyscias-Pseudopanax* group, the *Aralia-Panax* group, and the greater *Raukaua* group) with a few other genera placed in a basal polytomy (Wen et al., 2001; Plunkett et al., 2004; Mitchell et al., 2012).

In total, 20 genera and ca. 175 species of Araliaceae have been recognized in China (Table 1). Of these taxa, two genera and nearly 50% of the species are endemic (Shang & Lowry, 2007). Since Li (1942), taxonomic treatments for Chinese Araliaceae have been published, such as Hoo & Tseng (1965, 1978) and Shang & Lowry (2007). Tseng & Hoo (1982) divided Araliaceae into five tribes (Plerandreae Benth., Tetraplasandreae Hoo & Tseng, Mackinlayeae Benth., Aralieae Benth., and Panaceae Benth.) based on petal aestivation and leaf morphology. In their classification system, the Chinese members of Araliaceae were placed in four tribes (Plerandreae, Tetraplasandreae, Aralieae, and Panaceae) (Table 1). Recent molecular phylogenetic studies (Wen et al., 2001; Mitchell & Wen, 2004; Plunkett et al., 2004) revealed that the Chinese members of Araliaceae are scattered within two major lineages of the family (the Asian Palmate group and the *Aralia-Panax* group), with *Osmoxylon* Miq. at the base of core Araliaceae (Table 1). Even so, the phylogenetic relationships among the genera within each major group have been poorly resolved in the early studies, although most genera are well circumscribed by molecular data. With the Chinese Araliaceae accounting for nearly half of the total number of genera in the family, a phylogenetic analysis of Araliaceae from China using additional markers and expanding the sampling scheme is indispensable to facilitate a better understanding of the evolutionary diversification and classification of this family.

China harbors a broad geographic area (about 9.6 million km^2 and a span for more than 50° in latitudes), with enormous variations in geographic and topographical features, from mostly plateaus and mountains in the west to lower lands in

Table 1 Distribution and classification according to Tseng & Hoo (1982) and molecular phylogenetic studies (Wen et al., 2001; Plunkett et al., 2004) for genera of Chinese Araliaceae

Genus	Number of species (World/China)	Distribution	Tseng & Hoo	Molecular phylogenetic studies
Aralia L. (1753)	69/45	Asia, North America to South America	Tribe Aralieae	The *Aralia-Panax* group
Brassaiopsis Decne. & Planch. (1854)	ca. 30/24	The Himalayan region to C China, Indochina to W Malesia	Tribe Plerandreae	The Asian Palmate group
Chengiopanax Shang & J. Y. Huang (1993)	2/1	C China and Japan	Tribe Plerandreae	The Asian Palmate group
Dendropanax Decne. & Planch. (1854)	ca. 70/14	The Himalaya, E and SE Asia to W Malesia, C and S America	Tribe Plerandreae	The Asian Palmate group
Eleutherococcus Maxim. (1859)	38/18	E Asia and the Himalaya	Tribe Plerandreae	The Asian Palmate group
Fatsia Decne. & Planch. (1854)	3/1	E Asia (Japan, Taiwan of China, Ogasawara-shoto)	Tribe Plerandreae	The Asian Palmate group
Gamblea C. B. Clarke (1879)	4/2	E Asia, the Himalayan region to SE Asia	Tribe Plerandreae	The Asian Palmate group
Hedera L. (1753)	15/2	Temperate Eurasia	Tribe Plerandreae	The Asian Palmate group
Heteropanax Seem. (1866)	8/6	The Himalaya to S and SW China, Indochina to SE Asia	Tribe Tetraplasandreae	The Asian Palmate group
Kalopanax Miq. (1863)	1/1	E Asia	Tribe Plerandreae	The Asian Palmate group
Macropanax Miq. (1856)	17/7	The Himalaya to E and SE Asia, W Malesia	Tribe Plerandreae	The Asian Palmate group
Merrilliopanax H. L. Li (1942)	3/3	The Himalayan region to W China	Tribe Plerandreae	The Asian Palmate group
Metapanax J. Wen & Frodin (2001)	2/2	C and W China, N Vietnam	Tribe Plerandreae	The Asian Palmate group
Oplopanax Miq. (1863)	3/1	E Asia, NW North America	Tribe Plerandreae	The Asian Palmate group
Osmoxylon Miq. (1863)	60/1	Taiwan of China to Borneo, the Phillippines, to New Guinea and nearby Pacific islands	Tribe Plerandreae	The base of core Araliaceae
Panax L. (1753)	18/6	E Asia and E North America	Tribe Panaceae	The *Aralia-Panax* group
Schefflera J. R. Forst. & G. Forst. (1775)†	ca. 450/36	Asia, only the Asian group is referred to here	Tribe Plerandreae	The Asian Palmate group
Sinopanax H. L. Li (1949)‡	1/1	Taiwan of China	Tribe Plerandreae	The Asian Palmate group
Tetrapanax K. Koch (1859)‡	1/1	Taiwan of China (native), mainland China (introduced)	Tribe Plerandreae	The Asian Palmate group
Trevesia Vis. (1840)	10/1	S China, and NE India through SE Asia to W Malesia	Tribe Plerandreae	The Asian Palmate group

†*Schefflera* is the largest genus of Araliaceae, with 600–900 species. Recent studies have shown that *Schefflera* is polyphyletic (Plunkett et al., 2005; Li & Wen, 2014). The 450 or so species in Table 1 refer to only a group of closely related Asian species of *Schefflera*; ‡The genus is endemic to China.

the east (Wu & Wang, 1983). The diverse habitats and rich biodiversity in China have long intrigued biologists (Axelrod et al., 1998; Qian & Ricklefs, 1999; Wang et al., 2012). The past climate fluctuations and correlated change in the distribution of land and sea, plate tectonic activities, and the Quaternary glacial periods are seen as major contributors to the present distribution of plants and animals in this region (Shi et al., 1998; An et al., 2001). However, only a few phylogenetic studies with a biogeographic context in this region have focused on plants (e.g., Fan et al., 2009; Yu et al., 2011; Zhou et al., 2013). To better understand the biogeographic diversification of plants in this region, we need to evaluate the phylogenetic relationships and estimate divergence times in many lineages. The Chinese Araliaceae provide a good opportunity to examine the biogeographic pattern in this region because the family is distributed throughout China except Xinjiang Uyghur Autonomous Region.

In this study, we use a taxon sampling scheme throughout the range of the family, expanding the Chinese sampling used in previous studies up to 50 current recognized species in *Flora of China*. The main goals of the present study are to: (1) test the phylogenetic relationships of Araliaceae in China, with a particular emphasis on the generic status of *Pentapanax* Seem., which was merged into *Aralia* L. by Wen (2002), but recognized by Shang & Lowry (2007), and *Tupidanthus* J. D. Hooker & Thomson, which was included in Asian *Schefflera* (Frodin & Govaerts, 2003; Li & Wen, 2014), but recognized by Shang & Lowry (2007); and (2) elucidate the biogeographic diversification history of the Chinese Araliaceae. We have herein sequenced the nuclear internal transcribed spacer (ITS) regions of ribosomal DNA and six coding or non-coding plastid regions (the *ndhF* gene, the *trnL-trnF* intergenic spacer, the *rps16* intron, the *atpB-rbcL* intergenic spacer, the *rpl16* intron, and the *psbA-trnH* intergenic spacer), because these sequences have been shown to be useful for inferring relationships at the generic and specific levels of Araliaceae (e.g., Wen et al., 2008; Mitchell et al., 2012; Li & Wen, 2013, 2014; Valcárcel et al., 2014). We also used the "relaxed clock" analyses and fossil calibrations (Drummond et al., 2006) to obtain age estimates of the main clades of Chinese Araliaceae.

Material and Methods

Taxon sampling

Ninety-six plant accessions used in this study were sequenced for the nuclear ribosomal ITS regions, and the plastid *ndhF* gene, the *trnL-trnF* intergenic spacer, the *rps16* intron, the *atpB-rbcL* intergenic spacer, the *rpl16* intron, and the *psbA-trnH* intergenic spacer (Table 2). The sampling included 50 accessions representing 20 genera and 50 species of Chinese Araliaceae, which covers the morphological and geographic diversity of Chinese Araliaceae. We included 45 various taxa in the following genera within the core Araliaceae (Wen et al., 2001; Plunkett et al., 2004): *Arthrophyllum* Blume, *Astrotricha* DC., *Cussonia* Thunb., *Gastonia* Comm. & Lam., *Harmsiopanax* Warb., *Meryta* J. R. Forst. & G. Forst., *Oreopanax* Decne. & Planch., *Plerandra* A. Gray, *Polyscias* J. R. Forst. & G. Forst., *Pseudopanax* K. Koch, *Raukaua* Seem., and *Tetraplasandra* A. Gray. Because a close relationship of Pittosporaceae and Araliaceae has been shown (e.g., in Plunkett et al., 1996), *Pittosporum illicioides* Makino of Pittosporaceae was included as an outgroup taxon. The wide range of multiple taxa was selected to further test the evolutionary relationships of Chinese Araliaceae and to infer the biogeographic diversification with a broader phylogenetic framework.

DNA extraction, amplification and sequencing

Total DNA was extracted from about 15 mg silica-gel dried leaf material using the DNeasy plant mini kits (Qiagen, Mississauga, Canada) following the manufacturer's protocol or the modified CTAB extraction method (Doyle & Doyle, 1987).

The ITS region was amplified and sequenced using primers ITS4 and ITS5 (White et al., 1990). When amplification of the ITS region was unsuccessful, two other primers Nnc18S10 and C26A were used (Wen & Zimmer, 1996). The gene *ndhF* was amplified and sequenced in three segments with the following primer pairs: 1F and 972R, 803F and 1603R, 1318F and 2110R or 1995R (Olmstead & Sweere, 1994; Wen et al., 2003). The *trnL-trnF* region was amplified and sequenced using primers c and f (Taberlet et al., 1991). The *rps16* intron was amplified and sequenced using primers F and R2 (Oxelman et al., 1997). The *atpB-rbcL* region was amplified and sequenced following Manen et al. (1994). The *rpl16* intron was amplified and sequenced using primers *rpl16*-F and R (Asmussen, 1999). The *psbA-trnH* region was amplified and sequenced using primers *psbA* and *trnH* (Sang et al., 1997). Polymerase chain reaction (PCR) amplifications were carried out in a 25 μL volume containing 1.5 mmol/L $MgCl_2$, 0.2 mmol/L of each dNTP, 0.4 mmol/L of each primer, 1 U of *Taq* polymerase (Bioline, Taunton, MA, USA), and approximately 10–50 ng of DNA template under the following conditions: 3 min at 95 °C, followed by 37 cycles of 20 s at 94 °C, 30 s at 50 °C, and 40 s at 72 °C, and then a final 5 min extension at 72 °C.

The PCR products were purified using the polyethylene glycol precipitation procedure following the protocol of Sambrook et al. (1989). Cycle sequencing was carried out using BigDye 3.1 reagents using the following profile: 35 cycles of 97 °C for 15 s, 50 °C for 5 s, and 60 °C for 4 min. The products of cycle-sequencing reactions were cleaned using the Sephadex columns (Amersham Pharmacia Biotech, Piscataway, NJ, USA). The sequences were generated on an ABI prism 3730XL capillary sequencer (Applied Biosystems, Foster City, CA, USA).

Sequence alignment and phylogenetic analyses

The program Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA) was used to evaluate chromatograms for base confirmation and to edit contiguous sequences. Sequences were initially aligned with ClustalX version 1.83 (Thompson et al., 1997), followed by manual adjustments on Se-Al v2.0a11 (Rambaut, 2007).

Following the previous studies using the same DNA markers (Mitchell et al., 2012; Li & Wen, 2013, 2014), we combined the plastid (*ndhF*, *trnL-trnF*, *rps16*, *atpB-rbcL*, *rpl16*, and *psbA-trnH*) and the ITS data sets for phylogenetic analyses. Given that incongruence between nuclear and plastid data sets was detected (Plunkett et al., 2004; Valcárcel et al., 2014), our topology inferred from the combined data was checked with the BEAST species tree of Araliaceae used in a recent study (Valcárcel et al., 2014). We found the placements of Chinese

Table 2 Voucher information and GenBank accession numbers for Chinese Araliaceae and related taxa used in this study. Nomenclatural standard for the published species follows Frodin & Govaerts (2003). All collections are deposited at the US National Herbarium of the Smithsonian Institution (Washington DC)

Taxa	Locality	Voucher	*ndhF*	*trnL-F*	*rps16*	*atpB-rbcL*	*rpl16*	*psbA-trnH*	ITS
Aralia apioides Hand.-Mazz.	Yunnan, China	Wen et al. 1752	KU246162	KU246218	KU246204	KU246134	KU246190	KU246176	KU246148
Aralia bipinnata Blanco	Luzon, Philippines	Wen 8276	KU246158	KU246214	KU246200	KU246130	KU246186	KU246172	KU246144
Aralia castanopsisicola (Hayata) J. Wen	Taiwan, China	Wen 9428	KU246159	KU246215	KU246201	KU246131	KU246187	KU246173	KU246145
Aralia continentalis Kitag.	Jilin, China	Wen 5545	KU246154	KU246210	KU246196	KU246126	KU246182	KU246168	KU246140
Aralia echinocaulis Hand.-Mazz.	China	Wen 12035	KU246163	KU246219	KU246205	KU246135	KU246191	KU246177	KU246149
Aralia excelsa (Griseb.) J. Wen	Costa Rica	Wen 6779	KU246161	KU246217	KU246203	KU246133	KU246189	KU246175	KU246147
Aralia fargesii Franch.	China	Wen 12090	KU246164	KU246220	KU246206	KU246136	KU246192	KU246178	KU246150
Aralia glabra Matsum.	Japan	Soejima 1085	KU246155	KU246211	KU246197	KU246127	KU246183	KU246169	KU246141
Aralia henryi Harms[§]	Marlipo, Yunnan, China	Wen 10640	KC952426	KC952074	KC952162	KC952602	KC952250	KC952514	KC952338
Aralia leschenaultii (DC.) J. Wen	Gongshan, Yunnan, China	R. Li 499	KU246167	KU246223	KU246209	KU246139	KU246195	KU246181	KU246153
Aralia spinifolia Merr.	Taiwan, China	Wen 9449	KU246160	KU246216	KU246202	KU246132	KU246188	KU246174	KU246146
Aralia stellata (King) J. Wen[†]	Doi Luang Valley, Doichiang, Thailand	Maxwell 03-505	GU054745	GU055125	GU055030	GU054555	GU054935	GU054840	GU054650
Aralia vietnamensis Ha[†]	Lvchun, Yunnan, China	Shui 81844	GU054761	GU055141	GU055046	GU054571	GU054951	GU054856	GU054666
Arthrophyllum diversifolium Blume[‡]	Cameron, Pahang, Malaysia	Wen 8372	JX106024	JX106183	JX106144	JX106064	JX106224	JX106104	JX106265
Astrotricha latifolia Benth.[‡]	Queensland, Australia	P. I. Forster 7547	JX106026	JX106185	JX106146	JX106066	JX106226	JX106106	JX106267
Brassaiopsis hispida Seem.[‡]	Gongshan, Yunnan, China	Wen 5031	JX106027	JX106186	JX106147	JX106067	JX106227	JX106107	JX106268
Chengiopanax fargesii (Franch.) C. B. Shang & J. Y. Huang[†]	Xinning, Hunan, China	Wen 9316	GU054746	GU055126	GU055031	GU054556	GU054936	GU054841	GU054651
Cussonia paniculata Eckl. & Zeyh.[‡]	Eastern Cape, South Africa	Wen 10072	JX106030	JX106189	JX106149	JX106070	JX106230	JX106110	JX106271
Cussonia thyrsiflora Thunb.[‡]	Eastern Cape, South Africa	Wen 10057	JX106032	JX106191	JX106151	JX106072	JX106232	JX106112	JX106273
Dendropanax arboreus (L.) Decne. & Planch.[†]	Canton, San Jose, Costa Rica	Wen 7045	GU054787	GU055167	GU055072	GU054597	GU054977	GU054882	GU054692
Dendropanax burmanicus Merr.[†]	Gongshan, Yunnan, China	H. Li 33274	GU054781	GU055161	GU055066	GU054591	GU054971	GU054876	GU054686
Dendropanax caloneurus (Harms) Merr.[†]	Sa Pa, Lao Cai, Vietnam	Wen 6063	GU054712	GU055092	GU054997	GU054522	GU054902	GU054807	GU054617
Dendropanax caucanus Harms[†]	Canton, Puntarenas, Costa Rica	Wen 7018	GU054726	GU055106	GU055011	GU054536	GU054916	GU054821	GU054631
Dendropanax chevalieri (R. Vig.) Merr.[†]	Sa Pa, Lao Cai, Vietnam	Wen 6079	GU054783	GU055163	GU055068	GU054593	GU054973	GU054878	GU054688
Dendropanax dentiger (Harms) Merr.[†]	Xinning, Hunan, China	Wen 9306	GU054749	GU055129	GU055034	GU054559	GU054939	GU054844	GU054654
Dendropanax globosus M. J. Cannon & Cannon[†]	Monteverde, Puntarenas, Costa Rica	Wen 6848	GU054714	GU055094	GU054999	GU054524	GU054904	GU054809	GU054619
Dendropanax hainanensis (Merr. & Chun) Chun[†]	Mangshan, Hunan, China	Y. F. Deng 16240	GU054750	GU055130	GU055035	GU054560	GU054940	GU054845	GU054655
Dendropanax lancifolius Ridl.[†]	Cameron, Pahang, Malaysia	Wen 8362	GU054774	GU055154	GU055059	GU054584	GU054964	GU054869	GU054679
Dendropanax maingayi King[†]	Cameron, Pahang, Malaysia	Wen 8364	GU054728	GU055108	GU055013	GU054538	GU054918	GU054823	GU054633

Continued

Table 2 Continued

Taxa	Locality	Voucher	*ndhF*	*trnL-F*	*rps16*	*atpB-rbcL*	*rpl16*	*psbA-trnH*	ITS
Dendropanax oligodontus Merr. & Chun†	Lingshui, Hainan, China	Wen 6603	GU054784	GU055164	GU055069	GU054594	GU054974	GU054879	GU054689
Dendropanax poilanei Bui†	Lac Duong, Lam Dong, Vietnam	Wen 11049	GU054772	GU055152	GU055057	GU054582	GU054962	GU054867	GU054677
Dendropanax proteus (Champ. ex Benth.) Benth.†	Heishiding, Guangdong, China	Wen 5780	GU054716	GU055096	GU055001	GU054526	GU054906	GU054811	GU054621
Eleutherococcus sieboldianus (Makino) Koidz.†	Narita-shi, Chiba-ken, Japan	Wen 8538	GU054706	GU055086	GU054991	GU054516	GU054896	GU054801	GU054611
Eleutherococcus simonii (Simon-Louis ex Mouill) Hesse§	Kunming, Yunnan, China	Wen 10657	KC952427	KC952075	KC952163	KC952603	KC952251	KC952515	KC952339
Fatsia oligocarpella Koidz.†	Japan	H. Kato 30041	GU054755	GU055135	GU055040	GU054565	GU054945	GU054850	GU054660
Gamblea innovans (Siebold & Zucc.) C. B. Shang‡	Japan	Soejima 1094	JX106033	JX106192	JX106152	JX106073	JX106233	JX106113	JX106274
Gastonia custispongia Lam.†	Belgium Botanical Garden (cult.), Belgium	Wen s.n.	GU054756	GU055136	GU055041	GU054566	GU054946	GU054851	GU054661
Harmsiopanax ingens Philipson‡	Jayawijaja, Papua, Indonesia	Wen 10749	JX106034	JX106193	JX106153	JX106074	JX106234	JX106114	JX106275
Hedera sinensis (Tobler) Hand.-Mazz.†	Sa Pa, Lao Cai, Vietnam	Wen 5980	GU054732	GU055112	GU055017	GU054542	GU054922	GU054827	GU054637
Heteropanax fragrans (Roxb. ex DC.) Seem.‡	Chiang Dao, Chiang Mai, Thailand	Wen 7492	JX106035	JX106194	JX106154	JX106075	JX106235	JX106115	JX106276
Kalopanax septemlobus (Thunb.) Koidz.†	Xinning, Hunan, China	Wen 9341	GU054740	GU055120	GU055025	GU054550	GU054930	GU054835	GU054645
Macropanax dispermus (Blume) Kuntze†	Bandung, West Java, Indonesia	Wen 10137	GU054702	GU055082	GU054987	GU054512	GU054892	GU054797	GU054607
Macropanax maingayi (C. B. Clarke) Philipson†	Langat, Malaysia	Wen 8355	GU054741	GU055121	GU055026	GU054551	GU054931	GU054836	GU054646
Macropanax rosthornii (Harms) C. Y. Wu & G. Hoo†	Dujiangyan, Sichuan, China	Wen 9264	GU054708	GU055088	GU054993	GU054518	GU054898	GU054803	GU054613
Macropanax serratifolius K. M. Feng & Y. R. Li	Jingping, Yunnan, China	Wen 10530	KU246156	KU246212	KU246198	KU246128	KU246184	KU246170	KU246142
Macropanax undulatum Seem.	Hekou, Yunnan, China	Wen 10569	KU246157	KU246213	KU246199	KU246129	KU246185	KU246171	KU246143
Merrilliopanax chinensis H. L. Li§	Gongshan, Yunnan, China	Wen 5065	KC952457	KC952105	KC952193	KC952633	KC952281	KC952545	KC952369
Merrilliopanax listeri (King) H. L. Li‡	Gongshan, Yunnan, China	Wen 5038	JX106036	JX106195	JX106155	JX106076	JX106236	JX106116	JX106277
Meryta denhamii Seem.‡	Belgium Botanical Garden (cult.), Belgium	Wen s.n.	JX106037	JX106196	JX106156	JX106077	JX106237	JX106117	JX106278
Metapanax davidii (Franch.) J. Wen & Frodin†	Dujiangyan, Sichuan, China	Wen 9266	GU054720	GU055100	GU055005	GU054530	GU054910	GU054815	GU054625
Metapanax delavayi (Franch.) J. Wen & Frodin†	Lufeng, Yunnan, China	Wen 9146	GU054707	GU055087	GU054992	GU054517	GU054897	GU054802	GU054612

Continued

Table 2 Continued

Taxa	Locality	Voucher	*ndhF*	*trnL-F*	*rps16*	*atpB-rbcL*	*rpl16*	*psbA-trnH*	ITS
Oplopanax elatus (Nakai) Nakai†	Wusong, Jilin, China	Wen 5418	GU054757	GU055137	GU055042	GU054567	GU054947	GU054852	GU054662
Oreopanax polycephalus Harms†	Oxapampa, Dpto. Pasco, Peru	Wen 8595	GU054733	GU055113	GU055018	GU054543	GU054923	GU054828	GU054638
Oreopanax xalapense (Kunth) Decne. & Planch.†	Barva, Heredia, Costa Rica	Wen 6934	GU054734	GU055114	GU055019	GU054544	GU054924	GU054829	GU054639
Osmoxylon novoguineense (Scheff.) Becc.‡	West Papua, Indonesia	Wen 10706	JX106039	JX106199	JX106158	JX106079	JX106240	JX106119	JX106281
Osmoxylon pectinatum (Merr.) Philipson‡	Lutao, Taiwan, China	Wen 9411	JX106040	JX106200	JX106159	JX106080	JX106241	JX106120	JX106282
Panax bipinnatifidus Seem.	Gongshan, Yunnan, China	R. Li 714	KU246166	KU246222	KU246208	KU246138	KU246194	KU246180	KU246152
Panax trifolius L.†	Baltimore, Maryland, US	Wen 10099	GU054796	GU055176	GU055081	GU054606	GU054986	GU054891	GU054701
Panax wangianus S. C. Sun	China	Wen 12167	KU246165	KU246221	KU246207	KU246137	KU246193	KU246179	KU246151
Pittosporum illicioides Makino‡	Anji, Zhejiang, China	Wen 8486	JX106041	JX106201	JX106160	JX106081	JX106242	JX106121	JX106283
Plerandra insolita A. C. Sm.‡	Oahu (cult.), Hawaii, US	Wen 7076	JX106042	JX106202	JX106161	JX106082	JX106243	JX106122	JX106284
Polyscias australiana (F. Muell.) Philipson‡	Keerom, Papua, Indonesia	Wen 10707	JX106043	JX106203	JX106162	JX106083	JX106244	JX106123	JX106285
Polyscias nodosa (Blume) Seem.‡	Konawe, SE Sulawesi, Indonesia	Wen 10303	JX106044	JX106204	JX106163	JX106084	JX106245	JX106124	JX106286
Polyscias schultzei Harms‡	Jayawijaja, Papua, Indonesia	Wen 10730	JX106045	JX106205	JX106164	JX106085	JX106246	JX106125	JX106287
Pseudopanax arboreus (L. f.) K. Koch‡	Belgium Botanical Garden (cult.), Belgium	Wen s.n.	JX106047	JX106207	JX106166	JX106087	JX106248	JX106127	JX106289
Pseudopanax laetevirens (Gay) Franch.§	Valdivia, Chile	Wen 7728	KC952443	KC952091	KC952179	KC952619	KC952267	KC952531	KC952355
Pseudopanax valdiviensis (Gay) Harms§	Valdivia, Chile	Wen 7719	KC952444	KC952092	KC952180	KC952620	KC952268	KC952532	KC952356
Raukaua anomalus (Hook.) A. D. Mitch.‡	New Zealand (cult.)	A. Mitchell s.n.	JX106050	JX106210	JX106169	JX106090	JX106251	JX106130	JX106292
Schefflera actinophylla (Endl.) Harms§	Oahu (cult.), Hawaii, US	Wen 7056	KC952460	KC952108	KC952196	KC952636	KC952284	KC952548	KC952372
Schefflera angulata (Pav.) Harms†	Oxapampa, Peru	Wen 8589	GU054735	GU055115	GU055020	GU054545	GU054925	GU054830	GU054640
Schefflera arboricola (Hayata) Merr.§	Nantou, Taiwan, China	Wen 9448	KC952452	KC952100	KC952188	KC952628	KC952276	KC952540	KC952364
Schefflera aromatica (Blume) Harms§	Cibodas, West Java, Indonesia	Wen 10668	KC952437	KC952085	KC952173	KC952613	KC952261	KC952525	KC952349
Schefflera bodinieri (H. Lév.) Rehder§	Jinping, Yunnan, China	Shui 81850	KC952428	KC952076	KC952164	KC952604	KC952252	KC952516	KC952340
Schefflera delavayi (Franch.) Harms§	Xinning, Hunan, China	Wen 9347	KC952414	KC952062	KC952150	KC952590	KC952238	KC952502	KC952326

Continued

Table 2 Continued

Taxa	Locality	Voucher	*ndhF*	*trnL-F*	*rps16*	*atpB-rbcL*	*rpl16*	*psbA-trnH*	ITS
Schefflera digitata J. R. Forst. & G. Forst.‡	Lincoln, Canterbury, New Zealand	S. Oliver s.n.	JX106057	JX106217	JX106176	JX106097	JX106258	JX106137	JX106299
Schefflera heptaphylla (L.) Frodin§	Taoyuan, Taiwan, China	Wen 9386	KC952447	KC952095	KC952183	KC952623	KC952271	KC952535	KC952359
Schefflera hypoleucoides Harms§	Pingbian, Yunnan, China	Wen 8446	KC952419	KC952067	KC952155	KC952595	KC952243	KC952507	KC952331
Schefflera insularum (Seem.) Harms§	Laguna, Luzon, Philippines	Wen 8258	KC952448	KC952096	KC952184	KC952624	KC952272	KC952536	KC952360
Schefflera leucantha R. Vig.§	Chiang Mai (cult.), Thailand	Wen 7369	KC952455	KC952103	KC952191	KC952631	KC952279	KC952543	KC952367
Schefflera longipedicellata (Lecomte) Bernardi‡	Antsiranana, Madagascar	Wen 9564	JX106059	JX106219	JX106178	JX106099	JX106260	JX106139	JX106301
Schefflera myriantha (Baker) Drake‡	Antsiranana, Madagascar	Wen 9570	JX106060	JX106220	JX106179	JX106100	JX106261	JX106140	JX106302
Schefflera oxyphylla (Miq.) R. Vig.§	Kuala Langat, Selangor, Malaysia	Wen 8351	KC952478	KC952126	KC952214	KC952654	KC952302	KC952566	KC952390
Schefflera papuana Ridl.§	Bogor (cult.), West Java, Indonesia	Wen 10151	KC952411	KC952059	KC952147	KC952587	KC952235	KC952499	KC952323
Schefflera pentandra (Pav.) Harms†	Huampal, Peru	Wen 8619	GU054722	GU055102	GU055007	GU054532	GU054912	GU054817	GU054627
Schefflera petelotii Merr.§	Cuc Phuong, Ninh Binh, Vietnam	Wen 10918	KC952468	KC952116	KC952204	KC952644	KC952292	KC952556	KC952380
Schefflera pueckleri (K. Koch) Frodin§	Lac Duong, Lam Dong, Vietnam	Wen 11061	KC952469	KC952117	KC952205	KC952645	KC952293	KC952557	KC952381
Schefflera rugosa (Blume) Harms§	Bogor (cult.), West Java, Indonesia	Wen 10158	KC952412	KC952060	KC952148	KC952588	KC952236	KC952500	KC952324
Schefflera scandens (Blume) R. Vig.§	Cibodas, West Java, Indonesia	Wen 10114	KC952407	KC952055	KC952143	KC952583	KC952231	KC952495	KC952319
Schefflera sepikiana Harms§	Jayawijaja, Papua, Indonesia	Wen 10724	KC952436	KC952084	KC952172	KC952612	KC952260	KC952524	KC952348
Schefflera shweliensis W. W. Smith‡	Tzayu, Xizang, China	Wen 9199	JX106062	JX106222	JX106181	JX106102	JX106263	JX106142	JX106304
Schefflera wardii Marquand & Airy Shaw§	Linzhi, Xizang, China	Wen 9224	KC952420	KC952068	KC952156	KC952596	KC952244	KC952508	KC952332
Sinopanax formosanus (Hayata) H. L. Li†	Kaohsiung, Taiwan, China	Wen 9390	GU054723	GU055103	GU055008	GU054533	GU054913	GU054818	GU054628
Tetrapanax papyrifer (Hook.) K. Koch†	Bogor (cult.), West Java, Indonesia	Wen 10135	GU054758	GU055138	GU055043	GU054568	GU054948	GU054853	GU054663
Tetraplasandra hawaiensis A. Gray‡	Oahu (cult.), Hawaii, US	Wen 7075	JX106063	JX106223	JX106182	JX106103	JX106264	JX106143	JX106305
Trevesia palmata (Roxb. ex Lindl.) Vis.§	Hekou, Yunnan, China	Wen 8460	KC952421	KC952069	KC952157	KC952597	KC952245	KC952509	KC952333
Trevesia sundaica Miq.¶	Bogor, West Java, Indonesia	Wen 10162	KF591509	KF591527	KF591536	KF591545	KF591500	KF591518	KF591489

†Sequences of taxa from Li & Wen (2013); ‡Sequences of taxa from Mitchell et al. (2012); §Sequences of taxa from Li & Wen (2014); ¶Sequences of taxa from Valcárcel et al. (2014); and the remaining sequences were generated for the present study.

Araliaceae in the species tree are congruent with the one revealed by our topology based on the combined plastid and ITS data sets. For this reason, we carried out a combined analysis of all plastid and ITS data sets. In doing so, we noted that the resulting tree obtained from the combined data was better resolved and generally better supported.

Phylogenetic trees were constructed using maximum parsimony (MP) and Bayesian methods (Fig. 1). The MP analyses were conducted using PAUP* version 4.0b10 (Swofford, 2002). All characters were weighted equally and gaps were treated as missing data. The most parsimonious trees were obtained with heuristic searches of 1000 replicates with random stepwise sequence addition, tree bisection-reconnection (TBR) branch swapping, collapse of zero-length branches, multiple tree option in effect, saving 100 trees from each random sequence addition. Parsimony bootstrap values (PB) for the clades (Felsenstein, 1985) revealed in the maximally parsimonious trees (MPTs) were calculated with 500 bootstrap replicates. In each replicate, we carried out 100 random sequences addition replicates with tree bisection-reconnection (TBR) swapping algorithm and keeping no more than 10 trees per replicate. Tree statistics including consistency index (CI) and the retention index (RI) were calculated using PAUP*.

Modeltest 3.7 (Posada & Buckley, 2004) was used to determine the optimal model of molecular evolution and gamma rate heterogeneity using the Akaike Information Criterion (AIC). Bayesian inference was implemented with MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) using a mixed model Bayesian analysis strategy. We assigned model parameters for each gene partitions identified by AIC in Modeltest (Table 3). The Markov chain Monte Carlo (MCMC) algorithm was run for 4 000 000 generations with one cold and three heated chains, starting from random trees and sampling one out of every 100 generations. Runs were repeated twice. The resulting log-likelihood and number of generations were plotted to determine the point after which the log-likelihoods had stabilized. After discarding the trees saved prior to this point as burn-in, the remaining trees were imported into PAUP* and a 50% majority-rule consensus tree was producted to obtain posterior probabilities (PP) of the clades. Internodes with posterior probabilities $\geq$0.95 in the consensus trees were considered statistically significant.

Estimation of divergence times

To estimate divergence times within Chinese Araliaceae, the combined plastid and ITS data set was used with gaps treated as missing data. The Bayesian dating method with a relaxed molecular clock was implemented with the program BEAST 1.5.3 (Drummond & Rambaut, 2007) using the strategy of different nucleotide substitution models for each gene region suggested by Modeltest (Table 3). The Yule process for the tree prior model was employed using uncorrelated rates drawn from a lognormal distribution (Drummond et al., 2006). A normal distribution was specified for the priors. Posterior distributions of parameters were approximated using two independent MCMC analyses of 40 000 000 generations with 10% burn-in. Results were checked using the program Tracer 1.5 (Rambaut & Drummond, 2007) to ensure that plots of the two analyses were converging on the same area and the value of the effective sample size for each statistic was above 100.

We constrained the ages of two nodes in the phylogeny of Chinese Araliaceae and its close relatives (Fig. 2). First, the stem lineage of *Metapanax* was constrained to be 44 mya old (node A in Fig. 2) based on the fruit fossil of *Paleopanax oregonensis* Manchester from the Nut Beds flora of the Clarno Formation (north-central Oregon in the middle Eocene). This fossil was comparable to the Asian "*Pseudopanax*" (= *Metapanax*) (Manchester, 1994). Secondly, the crown age of Araliaceae was constrained to be 84 mya old (node B in Fig. 2) based on the estimates by Mitchell et al. (2012).

Results

DNA sequence data and phylogenetic relationships

We excluded the poly A, poly T or poly A/T regions from the data sets (*trnL-F*, 5 bp between 716 and 720; *rps16*, 21 bp between 744 and 764; *atpB-rbcL*, 7 bp between 611 and 617 and 13 bp between 626 and 638; *rpl16*, 9 bp between 764 and 772 and 4 bp between 846 and 849; and *psbA-trnH*, 6 bp between 61 and 66, 13 bp between 341 and 353, and 3 bp between 560 and 562). The statistics of the plastid and ITS data sets are shown in Table 3. Treating gaps as missing data, the maximum parsimony analysis based on the combined plastid and ITS data produced 30 814 MPTs of 3457 steps, with a CI of 0.61, a CI excluding uninformative characters of 0.45, a RI of 0.70, and a RC of 0.42. The 50% majority-rule consensus tree resulting from the Bayesian analysis was largely congruent with the trees of the parsimony analysis except that the genera *Oreopanax*, *Sinopanax*, and *Fatsia* formed a clade with the posterior probability (PP) value of 0.98; *Raukaua anomalus*, *Schefflera digitata*, *Pseudopanax laetevirens*, and *P. valdiviensis* formed a monophyletic group (PP = 1.0). The Bayesian tree with parsimony bootstrap (PB) and posterior probability (PP) support is shown in Fig. 1.

The combined plastid and ITS data strongly supported the previously identified four groups (Wen et al., 2001; Plunkett et al., 2004; Mitchell et al., 2012): the Asian Palmate group (PB = 96%, PP = 1.0), the *Polyscias-Pseudopanax* group (PB = 98%, PP = 1.0), the *Aralia-Panax* group (PB = 100%, PP = 1.0), and the greater *Raukaua* group (PB < 50%, PP = 1.0) (Fig. 1). The Chinese members (19 genera) of Araliaceae were scattered within the Asian Palmate group and the *Aralia-Panax* group except that the placement of *Osmoxylon* remains unclear (Fig. 1). The Asian Palmate group includes *Dendropanax*, *Macropanax*, *Metapanax*, *Kalopanax*, *Trevesia*, *Brassaiopsis*, *Eleutherococcus*, *Merrilliopanax*, *Sinopanax*, *Hedera*, *Chengiopanax*, *Fatsia*, *Gamblea*, Asian *Schefflera*, *Heteropanax*, *Tetrapanax*, and *Oplopanax*. The *Aralia-Panax* group comprises *Aralia* and *Panax*.

Molecular dating

The chronogram and results of divergence time estimation based on the combined platid and ITS data set from the Bayesian approach are shown in Fig. 2. The crown of the Asian Palmate group was dated to be at 60.16 mya (95% HPD: 50.24–70.33 mya; node 1 in Fig. 2). The crown *Aralia–Panax* group was estimated at 57.49 mya (95% HPD: 44.18–69.18 mya; node 2 in Fig. 2).

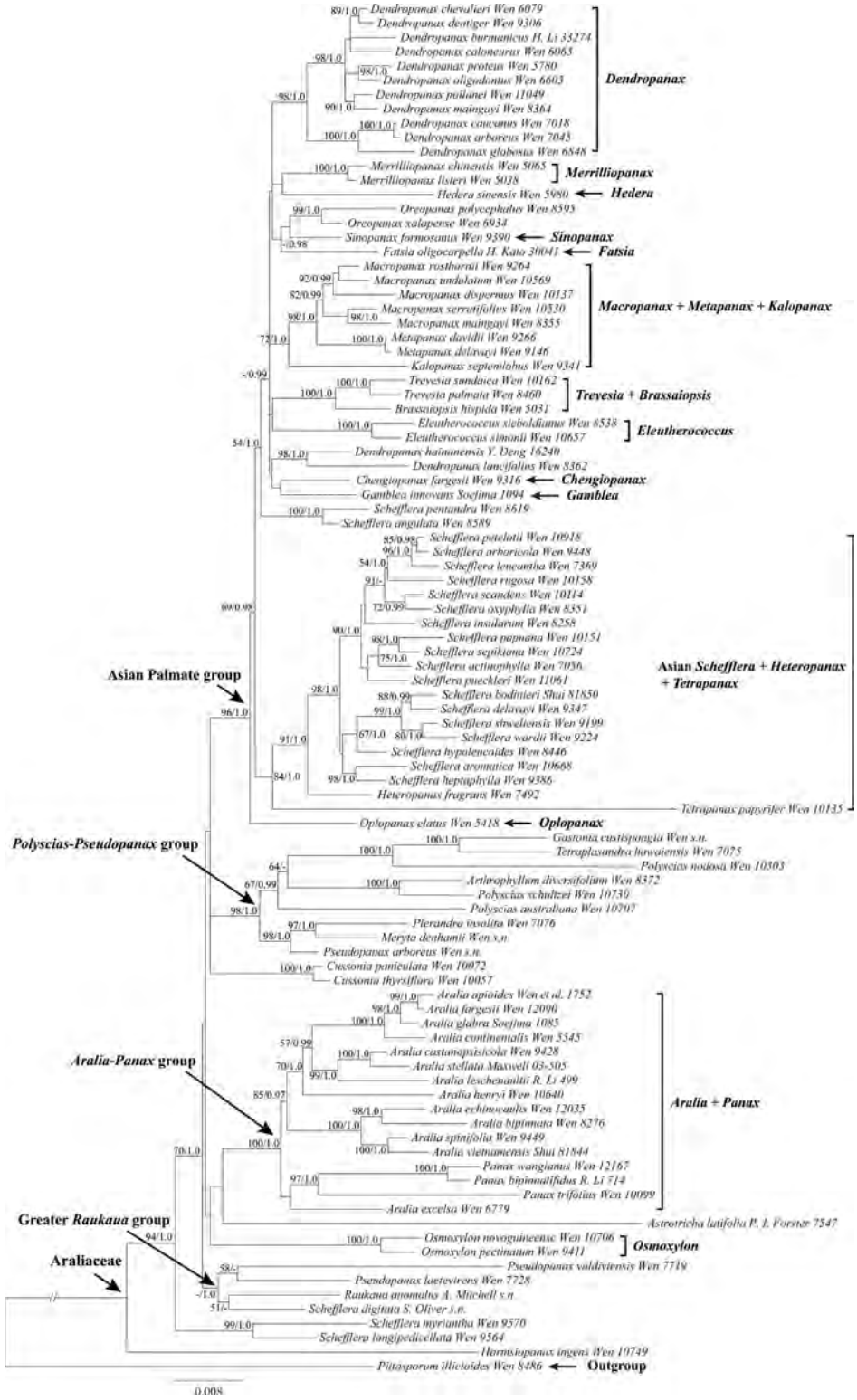


Fig. 1. The Bayesian tree of Chinese Araliaceae derived from analysis of the combined plastid and ITS data. Parsimony bootstrap values (PB) for maximum parsimony analysis in 500 replicates $>$ 50% are shown in the left and Bayesian posterior probabilities (PP) $\geq$ 0.95 are indicated in the right. Dash shows that the PB value lower than 50% or the PP value lower than 0.95. The branch to Araliaceae is truncated to allow for better display of the topology of the tree.

Table 3 Characters of the plastid and the nuclear ITS data sets

	Aligned length (bp)	Number of variable sites (%)	Number of informative sites (%)	Model selected by AIC
ndhF	1918	370 (19.3%)	163 (8.5%)	TVM+I+G
trnL-F	1013	213 (21.0%)	85 (8.4%)	TVM+G
rps16	937	210 (22.4%)	76 (8.1%)	GTR+G
atpB-rbcL	841	143 (17.0%)	61 (7.3%)	TVM+G
rpl16	1184	240 (20.3%)	105 (8.9%)	TVM+I+G
psbA-trnH	567	181 (31.9%)	97 (17.1%)	K81uf+I+G
ITS	652	334 (51.2%)	228 (35.0%)	GTR+I+G
Combined plastid and ITS data matrix	7112	1691 (23.8%)	815 (11.5%)	–

Discussion

Phylogenetic relationships in Chinese Araliaceae

Comparison of the findings of our molecular data to the traditional classification of Chinese Araliaceae based on morphology reveals little agreement (Table 1). The closely related genera *Aralia* and *Panax* were placed in two distinct tribes (Aralieae and Panaceae) in Tseng & Hoo's (1982) system. Also, the members from tribe Tetraplasandreae (e.g., *Heteropanax*) and tribe Plerandreae (e.g., *Dendropanax*, *Oplopanax*) in Tseng & Hoo's (1982) system form a group together with Neotropical representatives (e.g., *Oreopanax*, *Schefflera angulata*) in the molecular tree (Fig. 1). Those showed that the morphological characters (petal aestivation and leaf morphology) employed to delimit infrafamilial taxa within Chinese Araliaceae are highly homoplastic and have little utility at the tribal level.

Generic evaluation of Chinese Araliaceae

The *Aralia–Panax* group

The close relationship between *Aralia* and *Panax* was strongly supported (PB = 100%, PP = 1.0) in the present study (Fig. 1). Overall, *Aralia* and *Panax* share many characters, including imbricate floral aestivation, uniform endosperm, similar pollen morphology and ultrastructure, and similar floral vasculature (Eyde & Tseng, 1971; Wen, 1993; Wen & Nowicke, 1999). The current definition of *Aralia* recognized the following morphological synapomorphies: pinnately compound leaf architecture, presence of stipules, 5-12-locular ovaries, smooth seed surface, and flattened seeds (Wen, 1993, 2011). In contrast, its closest relative, the genus *Panax* possesses palmately compound leaves, whorled leaf arrangement, 2-4-locular ovaries, rough seed surfaces, and non-flattened seeds (Wen & Zimmer, 1996; Wen, 2001a, 2001b; Zuo et al., 2011, 2015). Historically, the phylogenetic position of *Panax* has long been controversial, Decaisne & Planchon (1854) and Clarke (1879) placed *Panax* within *Aralia*. Furthermore, Harms (1896) and Hoo (1961) regarded *Panax* as derived from an herbaceous member of *Aralia*.

Aralia stellata, *A. castanopsisicola*, and *A. leschenaultii*, included in this study, were previously classified as species of *Pentapanax* (Seemann, 1864), which was established primarily based on undivided styles, once pinnately compound leaves with 3-5 leaflets, and racemose to umbellate inflorescence units. Based on the minor morphological character variation (leaf structure within the pinnate architecture and style division) between *Pentapanax* and *Aralia*, Wen (1993) treated *Pentapanax* as a section of *Aralia* (also see Wen, 2002, 2004, 2011). Our phylogenetic analysis supported the merge of *Pentapanax* into *Aralia* (Wen, 1993), because the taxa of *Pentapanax* were nested within *Aralia* (Fig. 1).

Asian *Schefflera* + *Heteropanax* + *Tetrapanax* clade

Chinese members of *Schefflera* belong to the Asian *Schefflera*. The monophyly of Asian *Schefflera* + *Heteropanax* + *Tetrapanax* is supported (PB = 84%, PP = 1.0) by our analysis. Our phylogenetic study confirmed the earlier finding (Li & Wen, 2014) that the sister of Asian *Schefflera* is *Heteropanax*. The Asian *Schefflera–Heteropanax* subclade is then sister to the Asian monotypic *Tetrapanax* (Fig. 1). Asian *Schefflera* is similar to *Heteropanax* and *Tetrapanax* in habit, lack of prickles, lack of pedicle articulations, inflorescence architecture, and valvate aestivation. Their close relationships were also recognized by Harms (1894), who assigned those three taxa to his broadly defined tribe Schefflereae based on the shared character of valvate petals. However, Asian *Schefflera* can be easily distinguished from *Heteropanax* by its palmately compound or rarely simple to double digitately compound leaves, stipules united within the base of petiole and extending into a ligular appendage, ovaries with 5 or more locules, and styles united into a column or absent, rarely base united and free apically. On the contrary, *Heteropanax* possesses 2–5-pinnately compound leaves, inconspicuous stipules, 2-locular ovaries, and styles free or united to middle. *Tetrapanax* has simple and palmately lobed leaves, cone-shaped stipules, often 4-merous flowers, 2-locular ovaries, and free styles.

Historically, *Schefflera pueckleri* was treated as a member of the distinct genus *Tupidanthus* Hook. f. & Thomson (1856), which was established based on its very high numbers of stamens and locules in the ovary. However, the evolution of polymery in Araliaceae has been shown to have occurred independently multiple times (Wen et al., 2001; Plunkett et al., 2004). In the present study, *Tupidanthus* is clearly nested within the Asian *Schefflera* clade, supporting its transfer to *Schefflera* by Frodin (Stone, 1978), which also shares the character of the absence of styles like its close relatives.

Macropanax + *Metapanax* + *Kalopanax* clade

The close relationship between *Macropanax* and *Metapanax* recognized by the present and previous studies (Wen et al., 2001; Plunkett et al., 2004) is supported by the shared presence of evergreen leaves, mostly dentate leaflet margins, articulated pedicels, and bicarpellate ovaries. However, the two genera can be distinguished by their different endosperm

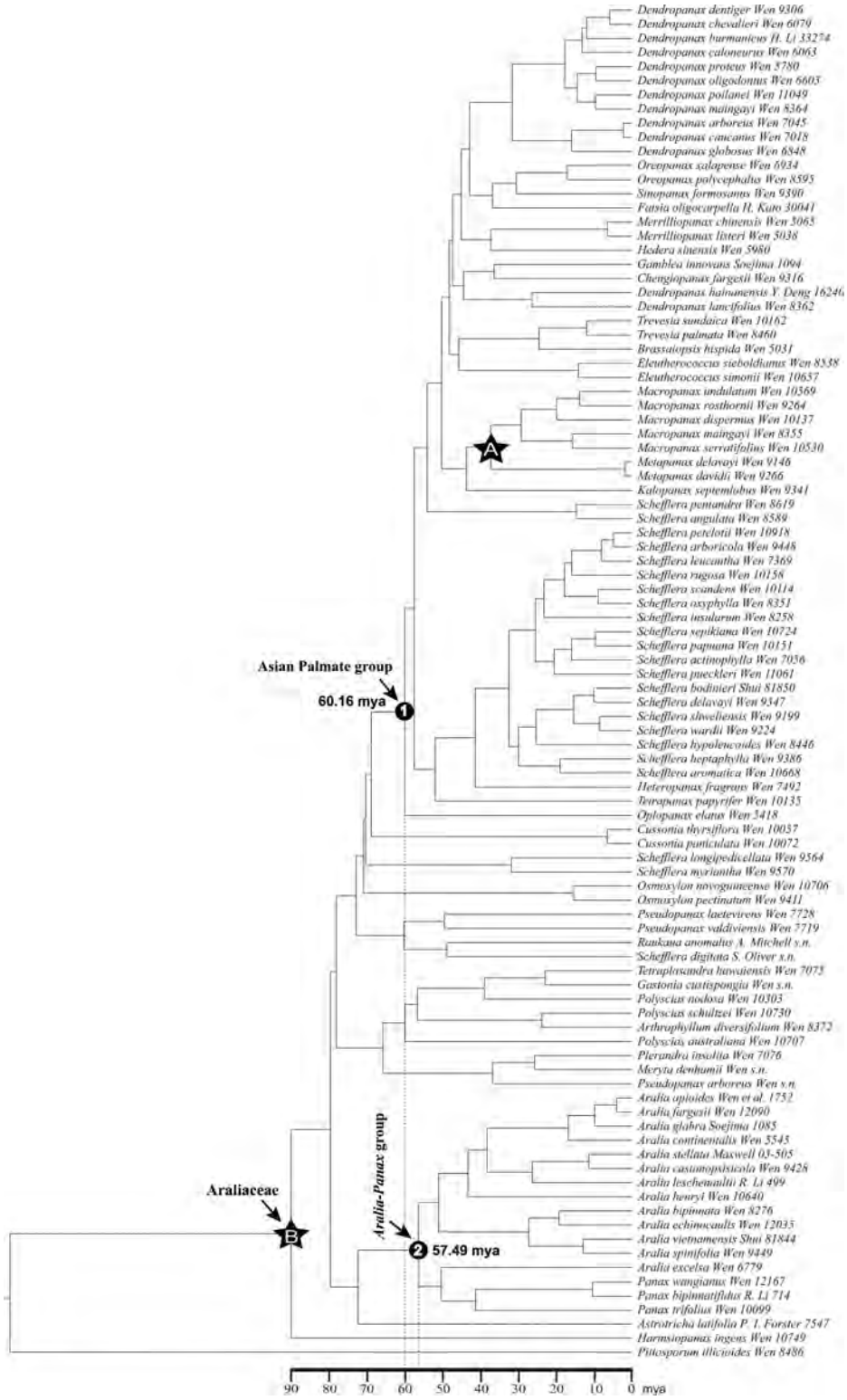


Fig. 2. Chronogram of Chinese Araliaceae inferred from combined plastid and internal transcribed spacer data using BEAST. Nodes labeled A or B indicate fossil calibration points. Nodes labeled 1–2 are indicated with age estimates and are discussed in the text.

types (ruminate in *Macropanax*, smooth in *Metapanax*), the division of styles (undivided in *Macropanax*, divided in *Metapanax*), and fruit shape (subglobose or ovoid in *Macropanax*, somewhat flattened in *Metapanax*) (Shang, 1985; Wen et al., 2001; Wen & Frodin, 2001). Our phylogeny presented here suggests that the sister of the *Macropanax–Metapanax* subclade is *Kalopanax* (PB = 72%, PP = 1.0) (Fig. 1). These three genera are similar in both inflorescence structure (terminal

panicle with the umbels as the basic inflorescence units) and floral morphyology (hermaphrodite flowers), but they may also be clearly differentiated: *Kalopanax* is deciduous and generally have prickles on the stem, whereas *Macropanax* and *Metapanax* are evergreen and unarmed. However, the ITS phylogeny suggests *Eleutherococcus* as the sister of *Macropanax-Metapanax* instead of *Kalopanax* possibly due to an early radiation with inter-lineage hybridizations for the Asian Palmate group (Valcárcel et al., 2014).

Brassaiopsis and *Trevesia*

The close relationship between *Brassaiopsis* and *Trevesia* (PB = 100%, PP = 1.0) (Fig. 1) was found in the present phylogeny and previous molecular analyses (Wen et al., 2001; Plunkett et al., 2004). Previous workers (e.g., Harms, 1894; Li, 1942; Hoo & Tseng, 1978) regarded the two genera as distantly related, placing emphasis on differences in the number of locules in the ovary. However, Jebb (1998) discussed the morphological similarities between *Brassaiopsis* and *Trevesia*, both of which contain species with more or less prickly stems, non-articulated pedicels, semi-inferior ovaries, undivided styles, and projected floral disks. Both genera can be distinguished by their different numbers of floral parts (2 or rarely 3–5 in *Brassaiopsis*, 6–12 in *Trevesia*) and endosperm types (ruminate in *Brassaiopsis*, smooth in *Trevesia*).

Hedera and its putative relationship to *Dendropanax* or *Merrilliopanax*

The sister-groups of *Hedera* remain unclear because the origin of the *Hedera* lineage may fit in a temperate niche conservatism scenario where the combination of the radiation with lineage admixtures prevents us from discovering its relatives (Valcárcel et al., 2014). Several workers (e.g., Hutchinson, 1967; Tseng & Hoo, 1982; Shang & Callen, 1988) regarded *Dendropanax* as closely related to *Hedera*, and this close relationship was supported by the ITS phylogeny of Araliaceae (Wen et al., 2001). Morphology does also support the association between *Hedera* and *Dendropanax* since they both have a single style, 3–5 carpels, and entire to 3-5-lobed leaves (Li & Wen, 2013; Valcárcel et al., 2014). The putative sister-group relationship between *Hedera* and *Merrilliopanax* was supported by previous phylogenetic study (Mitchell & Wen, 2004). Morphologically, both genera share simple and lobed leaves, while major differences regarding carpels (5 in *Hedera*, 2 in *Merrilliopanax*) and style (united into a short column in *Hedera*, free or united at base in *Merrilliopanax*).

The sister relationship between the Asian *Sinopanax* and the Neotropical *Oreopanax*

A sister-group relationship between the monotypic *Sinopanax* and the much larger Neotropical genus *Oreopanax* is suggested by the previous phylogenetic analyses (Wen et al., 2001; Plunkett et al., 2004). However, the sister relationship is shown, but not strongly supported in our analysis (Fig. 1). *Fatsia* sometimes appears as a sister to *Sinopanax* in the plastid analyses (see Valcárcel et al., 2014). *Sinopanax formosana* is endemic to Taiwan and was originally described by Hayata as *Oreopanax formosana*, with which it shares a number of characters, including palmately-lobed simple leaves, large terminal panicles of small capitate inflorescences, and ruminate endosperm. However, Li (1949) argued that the Taiwanese species differs in having a 2-carpellate (vs. 5-carpellate) ovary, a hermaphroditic (vs. polygamo-dioecious or polygamo-monoecious) sexual system, and rather short (vs. long) styles. The Bayesian analysis in the present study further indicates that *Sinopanax* and *Oreopanax* form a clade with the Asian *Fatsia* (Fig. 1). These three genera can be distinguished by their basic inflorescence units (capitulum in *Sinopanax* and *Oreopanax*, umbel in *Fatsia*).

The placement of *Oplopanax*

The genus *Oplopanax* with three species shows an intercontinental disjunct distribution between eastern Asia and western North America (Shang & Lowry, 2007). *Oplopanax* is sister to a large clade that includes the remaining genera of the Asian Palmate group (PB = 96%, PP = 1.0) (Fig. 1), which is consistent with the earlier findings (Mitchell et al., 2012; the plastid topology in Valcárcel et al., 2014). *Oplopanax* is characterized by a combination of characters including prickly shrubs with palmate lobed leaves, terminal inflorescences, 2-locular ovaries, 2-free or united below and persistent styles, red fruits at maturity (Frodin & Govaerts, 2003).

Eleutherococcus, *Chengiopanax*, and *Gamblea*

Chengiopanax and *Gamblea* have been separated from *Eleutherococcus* (Shang & Huang, 1993; Shang et al., 2000). *Eleutherococcus* has prickles on their stems, whereas *Chengiopanax* and *Gamblea* have no prickles. *Chengiopanax* has undivided styles and large inflorescences, but *Gamblea* has divided styles and small inflorescences. Each of these genera is supported as distinct based on our data, however, the relationships among them need to be explored further with a broader sampling of *Eleutherococcus*.

Osmoxylon

The genus has a wide distribution in the Malesian region, western Melanesia to Vanuatu and is especially well developed in the Philippines and Solomon islands (Frodin & Govaerts, 2003). Morphologically, it is characterized by the ligulate stipules and the marked petiole base with several spiral or transversal crests. The inflorescence is a terminal compound umbel with primary rays terminating into three branches. The genus is a phylogenetically isolated member of Araliaceae (Fig. 1). Only one species occurs in Lan Yu of Taiwan (Shang & Lowry, 2007).

Origin and biogeographic diversification of Araliaceae in China

Phylogenetic analyses of Araliaceae suggested that the family originated in the Australasian region of the paleotropics and then migrated into different tropical and subtropical regions (Wen et al., 2001; Plunkett et al., 2004). The origin of the Asian Palmate group seems to be Asia based on the ancestral area reconstruction performed by Mitchell et al. (2012). Considering the Chinese members of Araliaceae account for 85% and 100% of the total number of genera in the Asian Palmate Group and the *Aralia–Panax* group, respectively, we propose that Chinese Araliaceae most likely originated in Asia. The present phylogenetic analyses using multiple markers have not generated a well resolved phylogeny of Chinese Araliaceae. One possible explanation is that Chinese Araliaceae underwent a rapid radiation in its evolutionary history. Our

divergence time estimates place the crown of the Asian Palmate group and the *Aralia–Panax* group in the Paleocene (Fig. 2), consistent with the results in Mitchell et al. (2012) and Valcárcel et al. (2014). The orogenies in China (e.g., the formation of the Nanling Mountains during the Cretaceous, the uplift of the Yunnan-Guizhou Plateau and the mountain ranges in southeastern China in the early Cenozoic, and the major uplifts of the Tibetan Plateau in the late Tertiary) had led to changes in habitats and climates, accompanied by the changes of land and sea (Hsü, 1983; An et al., 2001), which may have facilitated the radiation of Chinese Araliaceae throughout this region. The similar scenario has been proposed for the Asian Palmate group where an early radiation with interlineage hybridizations and genome doubling has been detected and linked to the cooling occurred during the Upper Cretaceous (Valcárcel et al., 2014). Given that the short internal branches retrieved at the base of the Araliaceae tree and the divergence times inferred for the base of the tree, we suggest that the diversification pattern is not a particular situation for the Asian Palmate group, but possibly to the whole family. Future studies are needed to use phylogenomic and analytical biogeographic approaches (Wen et al., 2013, 2015; Zimmer & Wen, 2015) to unravel the history of evolutionary radiations in Araliaceae.

The present center of diversity of the Chinese Araliaceae is the mountains of southwestern China, which is equal to the Sino–Himalayan region, the southwestern China plateau region, and the Mid–Mekong region proposed by Li (1944). There are 134 species of Araliaceae (belonging to 16 genera) in this region, 54 of which are endemic to China. Congruence with taxonomic diversity, all five main lineages of Chinese Araliaceae (the *Aralia–Panax* group, the Asian *Schefflera* + *Heteropanax* + *Tetrapanax* clade, the *Macropanax* + *Metapanax* + *Kalopanax* clade, the *Brassaiopsis–Trevesia* clade, and the *Dendropanax* clade) occur in this region, in which, the *Aralia–Panax* group has the richest species diversity (37 taxa). The high species richness and endemism may have been generated by the rising of the Himalaya resulted from the collision of the Indian and Asian plates in the early Tertiary (An et al., 2001; Spicer et al., 2003). The mountain building processes accompanying the uplift of the Himalaya created regional topographic complexities in southwestern China (Shi et al., 1998). The topographic diversity and large river systems in the region created a wide range of habitats. The diverse habitats combined with the climatic changes of the late Tertiary and the Quaternary may have facilitated the diversification of Chinese Araliaceae in this region. The similar diversification pattern was reported in the species-rich genus *Rhodiola* (Crassulaceae) by Zhang et al. (2014), who suggested that rapid radiation was promoted by the uplifts of the Himalaya in the Tertiary (see Wen et al., 2014 for additional examples).

Another major area of diversification is the monsoon realm in South China, where 74 species of Araliaceae (belonging to 13 genera) occur, 33 are endemic to China. There are four main evolutionary lineages of Chinese Araliaceae represented in this region. However, the *Macropanax* + *Metapanax* + *Kalopanax* clade does not appear here. This region is equal to the southern China maritime region and the Gulf of Tonkin region as proposed by Li (1944). Physically this region is characterized by hills and low mountains extending more or less over the whole area because of the formation of the Nanling Mountains during the Cretaceous (Hsü, 1983). This region is also well known for its tropical and subtropical monsoon climate with adequate moisture in summer and warm dry in winter (Zheng, 2013). These favourable conditions maintain greater variety of habitats in South China and thus probably accelerate the speciation, diversification, and preservation for the species of Chinese Araliaceae in this region.

The remaining areas in China include 52 species of araliaceous plants (belonging to 14 genera), 29 of which are endemic to China. Only three main lineages of Chinese Araliaceae occur in these regions, and the *Brassaiopsis–Trevesia* clade and the *Dendropanax* clade are entirely absent. The flora of these regions, as a whole, is temperate in nature (Li, 1944). The mountainous topography (e.g., Qinling Mountains, Taihang Mountains, and Changbai Mountains) and humid monsoon climate (except northwestern China and Qinghai–Tibet Plateau) have resulted in a wide variety of habitats within these regions (Zheng, 2013), which may have provided excellent opportunities for the diversification of the temperate members of the Chinese Araliaceae there (e.g., the species-rich temperate genus *Eleutherococcus*). The temperate regions of China (primarily eastern China, Central China and northeastern China) also show strong biogeographic affinity with Japan and North America (Wen, 1999, 2001c; Donoghue & Smith, 2004; Xiang et al., 2015) and these biogeographic affinities are clearly shown in *Aralia*, *Panax* and *Oplopanax* (Li, 1944, 1952; Wen, 1998; Wen et al., 2010).

Acknowledgements

The study was supported by a grant from the John D. and Catherine T. MacArthur Foundation, the Laboratory of Analytical Biology at the National Museum of Natural History of the Smithsonian Institution, and the Talent Project of Yunnan Province (Grant No. 2015HB092). Laboratory assistance was provided by Jeff Hunt, Lei Xie, Xinwei Xu, and Yunjuan Zuo. We thank S. Oliver, Anthony Mitchell, H. Kato, Yunfei Deng, and Yumin Shui for collecting samples and field assistance, and Professor Zhiduan Chen for invitation to contribute to this special issue.

References

An ZS, Kutzbach JE, Prell WL, Porter SC. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times. *Nature* 411: 62–66.

Asmussen CB. 1999. Toward a chloroplast DNA phylogeny of the tribe Geonomeae (Palmeae). *Memoirs of the New York Botanical Garden* 83: 121–129.

Axelrod DI, Al-Shehbaz I, Raven PH. 1998. History of the modern flora of China. In: Zhang AL, Wu SG eds. *Floristic characteristics and diversity of East Asian plants*. Beijing: Higher Education Press; Berlin: Springer-Verlag. 43–55.

Clarke CB. 1879. Araliaceae. In: Hooker JD ed. *Flora of British India*, vol. 2. London: L. Reeve & Co. 134–179.

Decaisne J, Planchon JE. 1854. Esquisse d'une monographie des Araliacées. *Revue Horticole*, Séries 4, 3: 104–109.

Donoghue MJ, Smith SA. 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 1633–1644.

Doyle JJ, Doyle JL. 1987. A rapid isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: 699–710.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.

Eyde RH, Tseng CC. 1971. What is the primitive flora structure of Araliaceae? *Journal of the Arnold Arboretum* 52: 205–239.

Fan J, Qin HN, Li DZ, Jin XH. 2009. Molecular phylogeny and biogeography of *Holcoglossum* (Orchidaceae: Aeridinae) based on nuclear ITS, and chloroplast *trnL-F* and *matk*. *Taxon* 58: 849–861.

Felsenstein J. 1985. Confidence intervals on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.

Frodin DG, Govaerts R. 2003. *World checklist and bibliography of Araliaceae*. London: The Royal Botanic Gardens, Kew.

Harms H. 1894. Araliaceae. In: Engler A, Prantl K eds. *Die natürlichen Planzenfamilien*. Leipzig: Engelmann. 3(8): 1–62.

Harms H. 1896. Zur Kenntnis der Gattungen *Aralia* und *Panax*. *Botanische Jahrbücher fur Systematik* 23: 1–23.

Hoo G. 1961. The systematics, relationship and distribution of the Araliaceae of China. *Bulletin of Amoi University (Natural Sciences)* 8: 1–11.

Hoo G, Tseng CJ. 1965. Contributions to the Araliaceae of China. *Acta Phytotaxonomica Sinica* 10 (Suppl. 1): 129–176.

Hoo G, Tseng CJ. 1978. Araliaceae. In: *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press. 54: 1–188.

Hook JD, Thomson T. 1856. *Tupidanthus calyptratus. Botanical Magazine* 82: tab. 4908.

Hsü J. 1983. Late Cretaceous and Cenozoic vegetation in China, emphasizing their connections with North America. *Annals of the Missouri Botanical Garden* 70: 490–508.

Hutchinson J. 1967. *The genera of flowering plants*, vol. 2. London: Oxford University Press.

Jebb MHP. 1998. A revision of the genus *Trevesia* (Araliaceae). *Glasra* 3: 85–113.

Li HL. 1942. The Araliaceae of China. *Sargentia* 2: 1–134.

Li HL. 1944. The phytogeographical division of China, with special reference to Araliaceae. *Proceedings of the Academy of Natural Sciences of Philadelphia* 96: 249–277.

Li HL. 1949. A new genus of the Araliaceae. *Journal of the Arnold Arboretum* 30: 231–232.

Li HL. 1952. Floristic relationships between eastern Asia and eastern North-America. *Transactions of the American Philosophical Society* 42: 371–429.

Li R, Wen J. 2013. Phylogeny and biogeography of *Dendropanax* (Araliaceae), a genus disjunct between tropical/subtropical Asia and the Neotropics. *Systematic Botany* 38: 536–551.

Li R, Wen J. 2014. Phylogeny and biogeography of Asian *Schefflera* (Araliaceae) based on nuclear and plastid DNA sequence data. *Journal of Systematics and Evolution* 52: 431–449.

Manchester SR. 1994. Fruits and seeds of the middle Eocene Nut Beds Flora, Clarno Formation, Oregon. *Palaeontographica Americana* 58: 38–39.

Manen JF, Natali A, Ehrendorfer F. 1994. Phylogeny of Rubiaceae-Rubieae inferred from the sequence of a cpDNA intergene region. *Plant Systematics and Evolution* 190: 195–211.

Mitchell A, Li R, Brown JW, Schöenberger I, Wen J. 2012. Ancient divergence and biogeography of *Raukaua* (Araliaceae) and close relatives in the southern hemisphere. *Australian Systematic Botany* 25: 432–446.

Mitchell A, Wen J. 2004. Phylogenetic utility and evidence for multiple copies of Granule-Bound Starch Synthase I (GBSSI) in Araliaceae. *Taxon* 53: 29–41.

Olmstead RG, Sweere JS. 1994. Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* 43: 467–481.

Oxelman B, Liden M, Berglund D. 1997. Chloroplast *rps16* intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution* 206: 393–410.

Philipson WR. 1970. Constant and variable features of the Araliaceae. *Botanical Journal of the Linnean Society* 63 (Suppl. 1): 87–100.

Plunkett GM, Lowry II PP, Frodin DG, Wen J. 2005. Phylogeny and geography of *Schefflera*: Pervasive polyphyly in the largest genus of Araliaceae. *Annals of Missouri Botanical Garden* 92: 202–224.

Plunkett GM, Soltis DE, Soltis PS. 1996. Higher level relationships of Apiales (Apiaceae and Araliaceae) based on phylogenetic analysis of *rbcL* sequences. *American Journal of Botany* 83: 499–515.

Plunkett GM, Wen J, Lowry II PP. 2004. Infrafamilial classifications and characters in Araliaceae: Insights from the phylogenetic analysis of nuclear (ITS) and plastid (*trnL-trnF*) sequences data. *Plant Systematics and Evolution* 245: 1–39.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53: 793–808.

Qian H, Ricklefs RE. 1999. A comparison of the taxonomic richness of vascular plants in China and the United States. *American Naturalist* 154: 160–181.

Rambaut A. 2007. Se-Al version 2.0a11 [online]. Available from http://tree.bio.ed.ac.uk/software/seal/ [accessed 02 February 2015].

Rambaut A, Drummond AJ. 2007. Tracer v1.5 [online]. Availabel from http://beast.bio.ed.ac.uk/Tracer [accessed 02 February 2015].

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. 2nd ed. New York: Cold Spring Harbor Laboratory Press.

Sang T, Crawford DJ, Stuseey TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* 84: 1120–1136.

Seemann B. 1864. Revision of the natural order Hederaceae. II. On the genera with a single style. *The Journal of Botany* 2: 289–309.

Shang CB. 1985. The study of genus *Macropanax* Miq. (Araliaceae). *Journal of Nanjing Institute of Forestry* 1: 12–29.

Shang CB, Callen D. 1988. Pollen morphology of the family Araliaceae in China. *Bulletin of Botanical Research* 8: 13–35.

Shang CB, Huang JY. 1993. *Chengiopanax* – a new genus of Araliaceae. *Bulletin of Botanical Research* 13: 44–49.

Shang CB, Lowry II PP, Frodin DG. 2000. A taxonomic revision and re-definition of the genus *Gamblea* (Araliaceae). *Adansonia* 22: 45–55.

Shang CB, Lowry II PP. 2007. Araliaceae. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 13: 435–491.

Shi YF, Li JJ, Li BY. 1998. *Uplift and environmental changes of Qinghai-Tibetan Plateau in the Late Cenozoic*. Guangzhou: Guangdong Science and Technology Press.

Spicer RA, Harris NBW, Widdowson M, Herman AB, Guo S, Valdes PJ, Wolfe JA, Kelley SP. 2003. Constant elevation of southern Tibet over the past 15 million years. *Nature* 421: 622–624.

Stone BC. 1978. *Schefflera*. In: Ng FSP ed. *Tree flora of Malaya*. London: Longman. 3: 25–32.

Swofford DL. 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sunderland: Sinauer Associates.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The Clustalx windows interface. Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.

Tseng CJ, Hoo G. 1982. A new classification scheme for the family Araliacee. *Acta Phytotaxonomica Sinica* 20: 125–130.

Valcárcel V, Fiz-Palacios O, Wen J. 2014. The origin of the early differentiation of Ivies (*Hedera* L.) and the radiation of the Asian Palmate group (Araliaceae). *Molecular Phylogenetics and Evolution* 70: 492–503.

Wang Z, Fang J, Tang Z, Lin X. 2012. Relative role of contemporary environment versus history in shaping diversity patterns of China's woody plants. *Ecography* 35: 1124–1133.

Wen J. 1993. Generic delimitation of *Aralia* (Araliaceae). *Brittonia* 45: 47–55.

Wen J. 1998. Evolution of the eastern Asian and eastern North American disjunct pattern: Insights from phylogenetic studies. *Korean Journal of Plant Taxonomy* 28: 63–81.

Wen J. 1999. Evolution of eastern Asian and eastern North American disjunct pattern in flowering plants. *Annual Review of Ecology and Systematics* 30: 421–455.

Wen J. 2001a. Evolution of the *Aralia–Panax* complex (Araliaceae) as inferred from nuclear ribosomal ITS sequences. *Edinburgh Journal of Botany* 58: 183–200.

Wen J. 2001b. Species diversity, nomenclature, phylogeny, biogeography, and classification of the ginseng genus (*Panax* L., Araliaceae). In: Punja ZK ed. *Utilization of biotechnological, genetic and cultural approaches for North American and Asian ginseng improvement*. Vancouver: Simon Fraser University Press. 67–88.

. Wen J. 2001c. Evolution of eastern Asian-eastern North American biogeographic disjunctions: A few additional issues. *International Journal of Plant Sciences* 162: S117–S122.

Wen J. 2002. Revision of *Aralia* sect. *Pentapanax* (Seem.) J. Wen (Araliaceae). *Cathaya* 13–14: 1–116.

Wen J. 2004. Systematics and biogeography of *Aralia* L. sect. *Dimorphanthus* (Miq.) Miq. (Araliaceae). *Cathaya* 15–16: 1–187.

Wen J. 2011. Systematics and biogeography of *Aralia* L. (Araliaceae): Revision of *Aralia* sects. *Aralia*, *Humiles*, *Nanae*, and *Sciadodendron*. *Contributions from the United States National Herbarium* 57: 5–35.

Wen J, Frodin DG. 2001. *Metapanax*, a new genus of Araliaceae from China and Vietnam. *Brittonia* 53: 116–121.

Wen J, Ickert-Bond SM, Nie ZL, Li R. 2010. Timing and modes of evolution of eastern Asian–North American biogeographic disjunctions in seed plants. In: Long M, Gu H, Zhou Z eds. *Darwin's heritage today: Proceedings of the Darwin 200 Beijing International Conference*. Beijing: Higher Education Press. 252–269.

Wen J, Lee C, Lowry PP, Hiep NT. 2003. Inclusion of the Vietnamese endemic genus *Grushvitzkya* in *Brassaiopsis* (Araliaceae): Evidence from nuclear ribosomal ITS and chloroplast *ndhF* sequences. *Botanical Journal of the Linnean Society* 142: 455–463.

Wen J, Liu JQ, Ge S, Xiang QY, Zimmer EA. 2015. Phylogenomic approaches to deciphering the tree of life. *Journal of Systematics and Evolution* 53: 369–370.

Wen J, Nowicke JW. 1999. Pollen ultrastructure of *Panax* (the ginseng genus, Araliaceae), an eastern Asian and eastern North American disjunct genus. *American Journal of Botany* 86: 1624–1636.

Wen J, Plunkett GM, Mitchell AD, Wagstaff SJ. 2001. The evolution of Araliaceae: A phylogenetic analysis based on ITS sequences of nuclear ribosomal DNA. *Systematic Botany* 26: 144–167.

Wen J, Ree RH, Ickert-Bond SM, Nie ZL, Funk V. 2013. Biogeography: Where do we go from here? *Taxon* 62: 912–927.

Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H. 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Frontiers in Genetics* 5: 4.

Wen J, Zhu YP, Lee CH, Widjaja E, Saw LG. 2008. Evolutionary relationships of Araliaceae in the Malesian region: A preliminary analysis. *Acta Botanica Yunnanica* 30: 391–399.

Wen J, Zimmer EA. 1996. Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): Inferences from ITS sequences of nuclear ribosomal DNA. *Molecular Phylogenetics and Evolution* 6: 167–177.

White TJ, Bruns TD, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ eds. *PCR protocols, a guide to methods and applications*. San Diego: Academic Press. 315–322.

Wu ZY, Wang HS. 1983. *Physical geography of China: Phytogeography*. Beijing: Science Press.

Xiang JY, Wen J, Peng H. 2015. Evolution of the eastern Asian–North American biogeographic disjunctions in ferns and lycophytes. *Journal of Systematics and Evolution* 53: 2–32.

Yi TS, Lowry PP, Plunkett GM, Wen J. 2004. Chromosomal evolution in Araliaceae and close relatives. *Taxon* 53: 987–1005.

Yu Y, Downie SR, He X, Deng X, Yan L. 2011. Phylogeny and biogeography of Chinese *Heracleum* (Apiaceae tribe Tordylieae) with comments on their fruit morphology. *Plant Systematics and Evolution* 296: 179–203.

Zhang JQ, Meng SY, Allen GA, Wen J, Rao GY. 2014. Rapid radiation and dispersal out of the Qinghai-Tibetan Plateau of an alpine plant lineage *Rhodiola* (Crassulaceae). *Molecular Phylogenetics and Evolution* 77: 147–158.

Zheng D. 2013. Physical geography, climate. In: Hong DY, Blackmore S eds. *Plants of China, a companion to the Flora of China*. Beijing: Science Press. 35–70.

Zhou Z, Hong D, Niu Y, Li G, Nie ZL, Wen J, Sun H. 2013. Phylogenetic and biogeographic analyses of the Sino-Himalayan endemic genus *Cyananthus* (Campanulaceae) and implications for the evolution of its sexual system. *Molecular Phylogenetics and Evolution* 68: 482–497.

Zimmer EA, Wen J. 2015. Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. *Journal of Systematics and Evolution* 53: 371–379.

Zuo YJ, Chen ZJ, Kondo K, Funamoto T, Wen J, Zhou SL. 2011. DNA barcoding of *Panax* species. *Planta Medica* 72: 182–187.

Zuo YJ, Wen J, Ma JS, Zhou SL. 2015. Evolutionary radiation of the *Panax bipinnatifidus* species complex (Araliaceae) in the Sino-Himalayan region of eastern Asia as inferred from AFLP analysis. *Journal of Systematics and Evolution* 15: 210–220.

Anywhere Article.

Any format, any device, any time.

Today, more than ever, we need access to information that is immediate, clear and communicable. As a member of the community we serve, you will know how important it is to have access to that data, whenever you need it, and wherever you are.

What is Anywhere Article?

Anywhere Article is focused on making our online journal content on Wiley Online Library more readable and portable, whilst also allowing rich information to be brought to the surface. It achieves these goals in the following ways:

1. Readability

Clean design. Superfluous information and unnecessary distractions have been removed so that readers can focus on the article. Figures can be viewed in context or separately, and easily navigated, browsed or downloaded.

2. Functionality

The new layout and sidebar tray allow readers access to important information, ie; references, figures, publication history at any point in the reading experience, without losing their place on the main page.

3. Mobility

Whatever device you use - desktop, tablet, or mobile - the article will be presented to take best advantage of that device, always readable, always easy to use, wherever you are.

When Can I Start Using Anywhere Article?

Enhanced Article (HTML) You can view an article in the new 'Anywhere Article' format wherever you see this link. You'll be able to view it easily on the device of your choice, at your convenience.

Visit **www.wileyonlinelibrary.com** today and look out for the new links underneath each journal article, try it, and see the difference for yourself.

WILEY

Aims and Scope: *Journal of Systematics and Evolution* (JSE) is a bi-monthly international English language journal dedicated to the understanding of biodiversity and evolution. Papers in any of the following and related fields will be considered: phylogenetics and phylogenomics, biogeography (including phylogeography), biosystematics, description of well documented new taxa, taxonomic revisions and monographs (incorporating data from morphology, molecules, anatomy and ecology), molecular evolution, evolutionary development, evolutionary ecology, population biology, conservation biology, paleobiology, and related methods and theory.

Publisher

Journal of Systematics and Evolution is published by
John Wiley & Sons Australia, Ltd
42 McDougall Street, Milton
QLD 4064 Australia
(PO Box 1226, Milton, QLD 4064 Australia)
Tel: +61 (0)7 3859 9755

Journal Customer Services

For ordering information, claims and any enquiry concerning your journal subscription please go to www.wileycustomerhelp.com/ask or contact your nearest offi ce.

Americas: Email: cs-journals@wiley.com;
Tel: 1 781 388 8598 or 1 800 835 6770
(toll free in the USA & Canada).

Europe, Middle East and Africa:
Email: cs-journals@wiley.com;
Tel: 44 (0) 1865 778315.

Asia Pacific: Email: cs-journals@wiley.com
Tel: 65 6511 8000.

Japan: For Japanese speaking support, Email: cs-japan@wiley.com; Tel: +65 6511 8010 or Tel (tollfree): 005 316 50 480.

Visit our Online Customer Help available in 7 languages at www.wileycustomerhelp.com/ask

Information for Subscribers

Journal of Systematics and Evolution is published in 6 issues per year. FTE-based pricing, based on the overall demand and on the size and type of the subscribing institution (academic, government, hospital or corporate), applies for this Journal. Institutional subscription prices for 2016 are:

FTE small print & online: US$ 929 (US and Rest of World), € 755 (Europe), £ 652 (UK). FTE medium print & online: US$ 1324 (US and Rest of World), € 1076 (Europe), £ 929 (UK). FTE large print & online: US$ 1892 (US and Rest of World), € 1539 (Europe), £ 1325 (UK). Subscribers in China may purchase print subscriptions from the Editorial Office, Tel: +86 10 6283 6572, Email: jse@ibcas.ac.cn. Prices are exclusive of tax. Australian GST, Canadian GST/HST and European VAT will be applied at the appropriate rates. For more information on current tax rates, please go to www.wileyonlinelibrary.com/tax-vat. The price includes online access to the current and all online back files to January 1st 2012, where available. For other pricing options, including access information and terms and conditions, please visit www.wileyonlinelibrary.com/access.

Delivery Terms and Legal Title

Where the subscription price includes print issues and delivery is to the recipient's address, delivery terms are Delivered at Place (DAP); the recipient is responsible for paying any import duty or taxes. Title to all issues transfers FOB our shipping point, freight prepaid. We will endeavour to fulfil claims for missing or damaged copies within six months of publication, within our reasonable discretion and subject to availability.

Offprints

Within China: The Editorial Office. Email: jse@ibcas.ac.cn;
Tel: 86 10 6283 6572, 86 10 6283 6132;
Fax: 86 10 6283 6132.
Outside of China: C.O.S. Printers Pte Ltd., 9 Kian Teck Crescent, Singapore 628875. Email: offprint@cosprinters.com;
Fax: 65 6265 9074.

Back issues

Single issues from current and prior year volumes are available at the current single issue price from jse@ibcas.ac.cn. Tel: 86 10 6283 6572, 86 10 6283 6132; Fax: 86 10 6283 6132

Copyright and Photocopying

Abstracting and Indexing Services: The Journal is indexed by CC, SCIE.

Disclaimer: The Publisher, Institute of Botany, Chinese Academy of Sciences, Botanical Society of China, and Editors cannot be held responsible for errors or any consequences arising from the use of information contained in this journal; the views and opinions expressed do not necessarily reflect those of the Publisher, Institute, Society and Editors, neither does the publication of advertisements constitute any endorsement by the Publisher, Institute, Society and Editors of the products advertised.

For submission instructions, subscription and all other information visit: http://wileyonlinelibrary.com/journal/jse

JSE accepts articles for Open Access publication. Please visit http://olabout.wiley.com/WileyCDA/Section/id-406241.html for further information about OnlineOpen.

View this journal online at wileyonlinelibrary.com/journal/jse

Wiley's Corporate Citizenship initiative seeks to address the environmental, social, economic, and ethical challenges faced in our business and which are important to our diverse stakeholder groups. Since launching the initiative, we have focused on sharing our content with those in need, enhancing community philanthropy, reducing our carbon impact, creating global guidelines and best practices for paper use, establishing a vendor code of ethics, and engaging our colleagues and other stakeholders in our efforts.

Follow our progress at www.wiley.com/go/citizenship

Edited by:	Journal of Systematics and Evolution Editorial Office of Institute of Botany, Chinese Academy of Sciences 20 Nanxincun, Xiangshan, Beijing 100093, China Tel: 86 10 6283 6572, 86 10 6283 6132 Fax: 86 10 6283 6132 E-mail: jse@ibcas.ac.cn
Chief editors:	Song Ge and Jun Wen
Monitored by:	Chinese Academy of Sciences
Sponsored by:	Institute of Botany, Chinese Academy of Sciences Botanical Society of China
Published by:	John Wiley & Sons Australia, Ltd
Supported by:	Shenzhen Fairy Lake Botanical Garden, Shenzhen, China
Printed by:	Jingmei Color Printing Co. Ltd, Beijing, China

Postal No. 2-226 Code No. BM71

JSE
Journal of Systematics and Evolution
Volume 54 Number 4 July 2016

Special Issue: The Tree of Life: China Project

Cover illustration: Current framework of rosid phylogeny is displayed as a big tree with sturdy trunk and wide-spreading branches. A series of snapshots of flowers or fruits are displayed on the top of each clade, reflecting their morphological characters and economic importance. Different colors represent 17 orders of *Rosidae*. The super-ordinal relationships within *Rosidae* are outlined, sketching out a basic structure of the tree trunk (please see Sun et al., pp. 363–391 in this issue). This cover image shapes like a fan in Chinese painting, attempting to reflect that this special issue is a major regional effort toward the whole tree of life project.

ISSN 1674-4918
CN 11-5779/Q

WILEY

ISSN 1674-4918
07>
9 771674 491166

WILEY