RESEARCH ARTICLE

American Journal of Botany

INVITED SPECIAL ARTICLE

For the Special Issue: Using and Navigating the Plant Tree of Life

Character evolution and missing (morphological) data across *Asteridae*

Gregory W. Stull[1,5], Melanie Schori[2], Douglas E. Soltis[3,4], and Pamela S. Soltis[4]

Manuscript received 11 July 2017; revision accepted 8 December 2017.

[1] Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

[2] United States Department of Agriculture, National Germplasm Resources Laboratory, Beltsville, MD 20705-2305, USA

[3] Department of Biology, University of Florida, Gainesville, FL 32611-8525, USA

[4] Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA

[5] Author for correspondence (e-mail: gwstull@gmail.com)

Citation: Stull, G. W., M. Schori, D. E. Soltis, and P. S. Soltis. 2018. Character evolution and missing (morphological) data across *Asteridae*. *American Journal of Botany* 105(3): 470–479.

doi:10.1002/ajb2.1050

PREMISE OF THE STUDY: Our current understanding of flowering plant phylogeny provides an excellent framework for exploring various aspects of character evolution through comparative analyses. However, attempts to synthesize this phylogenetic framework with extensive morphological data sets have been surprisingly rare. Here, we explore character evolution in *Asteridae* (asterids), a major angiosperm clade, using an extensive morphological data set and a well-resolved phylogeny.

METHODS: We scored 15 phenotypic characters (spanning chemistry, vegetative anatomy, and floral, fruit, and seed features) across 248 species for ancestral state reconstruction using a phylogenetic framework based on 73 plastid genes and the same 248 species.

KEY RESULTS: Iridoid production, unitegmic ovules, and cellular endosperm were all reconstructed as synapomorphic for *Asteridae*. Sympetaly, long associated with asterids, shows complex patterns of evolution, suggesting it arose several times independently within the clade. Stamens equal in number to the petals is likely a synapomorphy for *Gentianidae*, a major asterid subclade. Members of *Lamianae*, a major gentianid subclade, are potentially diagnosed by adnate stamens, unilacunar nodes, and simple perforation plates.

CONCLUSIONS: The analyses presented here provide a greatly improved understanding of character evolution across *Asteridae*, highlighting multiple characters potentially synapomorphic for major clades. However, several important parts of the asterid tree are poorly known for several key phenotypic features (e.g., degree of petal fusion, integument number, nucellus type, endosperm type, iridoid production). Further morphological, anatomical, developmental, and chemical investigations of these poorly known asterids are critical for a more detailed understanding of early asterid evolution.

KEY WORDS angiosperm synapomorphies; *Asteridae*; *Campanulidae*; character evolution; *Gentianidae*; iridoids; *Lamiidae*; morphology.

Our understanding of angiosperm phylogeny arguably has increased more in the past three decades than in the preceding three centuries (Soltis et al., 2005; Judd et al., 2016). This advancement is, in part, a result of large-scale collaborative efforts using Sanger sequencing to construct taxon-rich (but generally gene-poor) phylogenies spanning angiosperms or major subclades (e.g., Olmstead et al., 1992, 1993, 2000; Chase et al., 1993; Soltis et al., 1998, 1999, 2000, 2011; Savolainen et al., 2000a, 2000b; Albach et al., 2001b; Bremer et al., 2002; Wurdack and Davis, 2009; Refulio-Rodriguez and Olmstead, 2014). However, while these studies made great progress, many deep-level relationships proved difficult to resolve using only a handful of mostly chloroplast genes (e.g., Bremer et al., 2002). More recently, studies have assembled genome-scale data sets, primarily generated with next-generation sequencing (NGS)

technologies, aimed at resolving recalcitrant nodes across the angiosperm tree (e.g., Jansen et al., 2007; Moore et al., 2006, 2007, 2010; Wang et al., 2009; Xi et al., 2012, 2014; Soltis et al., 2013a; Wickett et al., 2014; Zeng et al., 2014; Stull et al., 2015). Although some relationships remain uncertain (e.g., the positions of Dilleniaceae, Caryophyllales, Santalales, and Berberidopsidales, and relationships within Lamiales), our current understanding of angiosperm phylogeny has facilitated vast improvements in classification (APG, 1998, 2003, 2009, 2016; Cantino et al., 2007; Soltis et al., 2011; Stull et al., 2015) and constitutes an invaluable tool for investigating various aspects of flowering plant evolution (Soltis et al., 1999).

However, given the continued focus on resolving the framework of angiosperm phylogeny (e.g., Soltis et al., 2011; Zeng et al., 2014; Liu et al., 2017), the pace of sequence generation has greatly surpassed efforts to accumulate and analyze morphological data for understanding broader patterns of angiosperm evolution. Broad-scale morphological data sets are essential for determining synapomorphies, assessing patterns of character evolution across major clades, and incorporating information from the fossil record. Unfortunately, too few studies have synthesized morphological data with available phylogenetic frameworks to elucidate broader patterns of flowering plant evolution (but see, e.g., Albach et al., 2001a; Ronse De Craene et al., 2003; Doyle, 2005, 2007; Ronse De Craene, 2008; Endress and Doyle, 2009, 2015; Endress, 2010, 2011a, 2011b; Ronse De Craene and Brockington, 2013; Soltis et al., 2013b; Zanne et al., 2014). Although a wealth of morphological data is available in the older literature, even basic morphological information is missing for many taxa (Stevens, 2001 onward). Consequently, we still have much to learn about flowering plant evolution in light of our improved knowledge of phylogenetic relationships (see other articles in this Special Issue).

Here, we synthesize our current understanding of phylogeny and available morphological data to explore patterns of character evolution in asterids, with particular emphases on the large subclade *Gentianidae* (Cantino et al., 2007; also known as core asterids or euasterids: e.g., APG, 1998). The clade *Gentianidae* (names with phylogenetic definitions following the PhyloCode [e.g., Cantino et al. 2007] are presented in italics throughout) represents a major angiosperm radiation including ~80,000 species or ~30% of flowering plant species richness (assuming ~250,000 species, as per Judd et al., 2016; this proportion is expected to hold for larger estimates of angiosperm species, e.g., Govaerts, 2001, 2003). This group, at least in part, has been recognized by botanists for at least 200 years (e.g., de Jussieu, 1789), although phylogenetic studies have expanded its circumscription. Gentianids—which encompass most angiosperms with fused corollas (sympetaly)—were recognized by Takhtajan (1980), Cronquist (1981), and other leading authors of the 20th century as Asteridae, but this name is currently used in a broader sense (*Asteridae* sensu Cantino et al., 2007), encompassing Ericales and Cornales as well as *Gentianidae*.

Members of *Gentianidae*, as currently recognized, fall into two major clades, *Lamiidae* and *Campanulidae*, each with ~40,000 species. Several previous studies have investigated character evolution in the asterids (Albach, 2001a; Bremer et al., 2001), but the identification of clear synapomorphies for the asterids as a whole (as well as major subgroups, e.g., *Gentianidae*, *Lamiidae*, and *Campanulidae*) has proven difficult in light of poor phylogenetic resolution, a limited sampling of morphological and other non-DNA characters, and potentially complicated patterns of character evolution, including frequent parallelisms (Endress, 1996; Judd and Olmstead, 2004). For example, Albach et al. (2001a), focusing on several embryological and biochemical characters, documented in *Asteridae* the prevalence of unitegmic and tenuinucellate ovules, as well as iridoid production, but each character showed complicated patterns of gain/loss. Other major features associated with asterids—e.g., sympetaly and cellular endosperm formation—are also not ubiquitous, especially among Cornales, Ericales, and "early-diverging" lamiids and campanulids (Stevens, 2001 onward).

This study explores patterns of character evolution in *Asteridae* (sensu Cantino et al., 2007) using an improved phylogenetic framework (e.g., Tank and Donoghue, 2010; Refulio-Rodriguez and Olmstead, 2014; Stull et al., 2015) with expanded sampling of key, "early-diverging" lamiids and campanulids (Stull et al., 2015) and a broad set of phenotypic characters (spanning chemistry, vegetative anatomy, and floral, fruit, and seed features). In addition to identifying the ancestral morphological features for major asterid clades, we highlight areas of the asterid tree with critical missing phenotypic data and hope to spur efforts to assemble phenotypic data across asterids for more in-depth future comparative studies.

MATERIALS AND METHODS

Taxonomic and molecular sampling

We sampled 248 species across core eudicots (*Gunneridae*), 227 of which are asterids. Our sampling was designed to represent all major asterid lineages while maintaining species-level compatibility with available morphological data sets for asterids (e.g., Albach et al., 2001a; Bremer et al., 2001). Following the APG IV classification (APG, 2016), all asterid orders were represented, as were all 69 core asterid (gentianid) families; five (of six) families of Cornales were included as were 16 (of 22) families of Ericales. Our sampling of the gentianid order Boraginales includes 10 species, representing seven of the 11 families recognized in more recent treatments (Leubert et al., 2016). We also included an extensive sampling of "basal lamiid" genera, which have been under-sampled in most previous large-scale phylogenetic studies (e.g., Soltis et al., 2011; Refulio-Rodriguez and Olmstead, 2014). In particular, we included *Oncotheca* (Oncothecaceae), *Metteniusa*, 10 genera of Icacinaceae s.l. now placed in Metteniusaceae (Stull et al., 2015), 21 of the 23 genera remaining in Icacinaecae s.s. (Stull et al., 2015), and all three genera of Garryales (*Aucuba*, *Garrya*, and *Eucommia*). We also included at least one representative from each of the five families of Aquifoliales, which is positioned sister to the rest of the campanulid clade.

Given the prevalence of chloroplast DNA sequence data from previous phylogenetic studies (e.g., Moore et al., 2010; Soltis et al. 2011; Stull et al., 2015), we sampled 73 chloroplast genes for phylogenetic analyses to provide a solid framework for subsequent character reconstructions. Although the resulting data set includes ~61% missing data, taxa with relatively complete gene sampling are well distributed across the asterids, owing to previous studies employing chloroplast genomes to resolve major angiosperm and asterid relationships (e.g., Moore et al., 2010; Stull et al., 2015). Thus, the sequence-rich species should provide a well-resolved scaffold to place the remaining species sampled for fewer loci. This matrix is available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.7783j. GenBank numbers for the included sequences are presented in Appendix S1 (see Supplemental Data with this article).

Phylogenetic analyses

We conducted phylogenetic analyses using both maximum likelihood (ML) and Bayesian approaches. The ML analyses were conducted in RAxML v 8.2.10 (Stamatakis, 2014), including a rapid bootstrap analysis as well as a search for the best-scoring ML tree using the GTR+GAMMA model, with model parameters partitioned by gene region. The Bayesian analyses were implemented in MrBayes v. 3.2.1 (Huelsenbeck and Ronquist, 2001; Ronquist et al., 2012), including 15 million generations with four chains sampling the posterior every 1000 generations. The Bayesian analysis also used the GTR+GAMMA model partitioned by gene. Convergence of the MrBayes analysis was determined by visually inspecting the outputs of the program using Tracer v. 1.5 (Rambaut and Drummond, 2009). We conducted the Bayesian analyses primarily to obtain the posterior distribution, which provides a convenient source of trees for integrating phylogenetic uncertainty in downstream analyses—in this case, ancestral state reconstruction. Phylogenetic trees are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.7783j.

Phenotypic character sampling

We selected 15 morphological characters (Table 1)—many of which have been previously emphasized in broad-scale discussions and classifications of asterids in particular and angiosperms in general (e.g., Cronquist, 1981, 1988; Stevens, 2001 onwards; Takhtajan, 2009)—for ancestral state reconstruction across asterids. Some of these characters (ovule integument number, nucellus type, endosperm formation type, iridoid compound production) have been analyzed previously (e.g., Albach et al., 2001a), whereas others (e.g., fruit type, seed number, and habit) have been highlighted as potentially important in early asterid evolution, given recent phylogenetic analyses that clarified the basal branching order of *Lamiidae* (Stull et al., 2015). Table 1 outlines the characters and corresponding states explored in this study; Appendix S2 discusses these characters further and provides rationale for the states employed in our reconstructions. The complete matrix of phenotypic characters is available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.7783j.

Character data were obtained from the following sources: Mauritzon (1936), Bailey and Howard (1941a,b), Howard (1940, 1942a–d, 1992), Sleumer (1942, 1969, 1971), Fagerlind (1945), Dickison (1986), *Flora of North America* (Flora of North America Editorial Committee, 1993 onward), Takhtajan (1997, 2009), Jensen (2000), Albach et al. (2001a), Bremer et al. (2001), Kårehed (2001), Knapp (2002), eFloras (2008; Flora of North America and Flora of China), Peng and Howard (2008), Lens et al. (2008), González and Rudall (2010), Endress and Rapini (2014), Dickison and Bittrich (2016), Potgieter and Duno (2016), Potgieter et al. (2016), and Schori (2016). We scored character states for each species, rather than for each genus (e.g., Bremer et al., 2001). However, in many cases, character states were not reported for individual species but instead for genera or families. If, for a given character, the state was reported as invariant across the broader group (e.g., genus or family) containing the species, and if the circumscription of the group as presented reflects our current understanding of phylogeny, we scored the species accordingly. However, if variation was noted across the group, we scored the character as missing data for the species. For the characters sympetaly and synsepaly (petal and sepal fusion, respectively), states were scored based on the mature condition of the flower, except in cases where available developmental data provided evidence for an alternative state (e.g., petals in Araliaceae are fused early in development and free at maturity; Erbar et al., 2004; Leins and Erbar, 2004). In such cases, the developmental state was used as this should correspond to the original/ancestral state of the character for the taxon in question (Stevens, 2001 onward).

Character reconstructions

Ancestral state reconstructions were conducted individually for each character using maximum likelihood in Mesquite v. 3.2 (Maddison and Maddison, 2017). The reconstructions were performed using the Mk1 model and the best-scoring tree from the RAxML analysis, but without branch-length information (as some of the internal branch lengths were extremely small—effectively zero—causing errors in Mesquite; the branch lengths were thus transformed to equivalent lengths to permit ancestral state reconstruction). Under

TABLE 1. Characters and character states scored and analyzed for this study (see Appendix S2 for definitions).

Characters	States						
Habit	Woody (0)	Herbaceous (1)	Suffrutescent (2)				
Perforation plates	Simple (0)	Scalariform (1)					
Nodal anatomy	Unilacunar (0)	Trilacunar (1)	Multilacunar (2)				
Iridoids	Absent (0)	Present (1)					
Synsepaly	Free (0)	Fused (1)					
Sympetaly	Free (0)	Fused (1)					
Stamen number: petal number	More stamens than petals (0)	Equal in number (1)	Stamens fewer (2)	Both numerous (3)			
Stamen adnation	Free (0)	Adnate to corolla (1)					
Ovary position	Superior (0)	Inferior (1)	Half-inferior (2)				
Carpel number	Many: more than ten (0)	One (1)	Two (2)	Three (3)	Four (4)	Five (5)	Six to ten (10)
Fruit type	Drupe (0)	Berry (1)	Capsule (2)	Schizocarp (3)	Nut/achene (4)	Follicle (5)	Samara (6)
Seed number (per fruit)	Many: more than ten (0)	One (1)	Two (2)	Three (3)	Four (4)	Five (5)	Six to ten (10)
Ovule integument	Bitegmic (0)	Partially bitegmic (1)	Unitegmic (2)	Ategmic (3)			
Ovule nucellus	Tenuinucellate (0)	Crassinucellate (1)	Weakly crassinucellate (2)	Pseudocrassinucellate (3)			
Endosperm*	Nuclear (0)	Cellular (1)	Absent (2)				

the Mk1 model, transitions between all combinations of states are permitted. We also reconstructed ancestral states of key nodes (i.e., *Asteridae*, Ericales + *Gentianidae*, *Gentianidae*, *Lamiidae*, *Lamianae*, and *Campanulidae*) using ML and Bayesian Inference in BayesTraits v. 3.0 (Pagel and Meade, 2006; http://www.evolution.rdg.ac.uk/BayesTraitsV3/BayesTraitsV3.html). These analyses included branch-length information and incorporated phylogenetic uncertainty by using a random set of 200 post-burnin trees from the MrBayes analysis described above. The BayesTraits analyses included multi-state Markov models with either ML or Markov chain Monte Carlo (MCMC) analyses implemented, in both cases using the default parameters of the program; for ML, this included 10 ML attempts per tree (with 20,000 ML maximum evaluations); for MCMC, this included 1,010,000 iterations (with a sample period of 1000 and a burnin of 10,000).

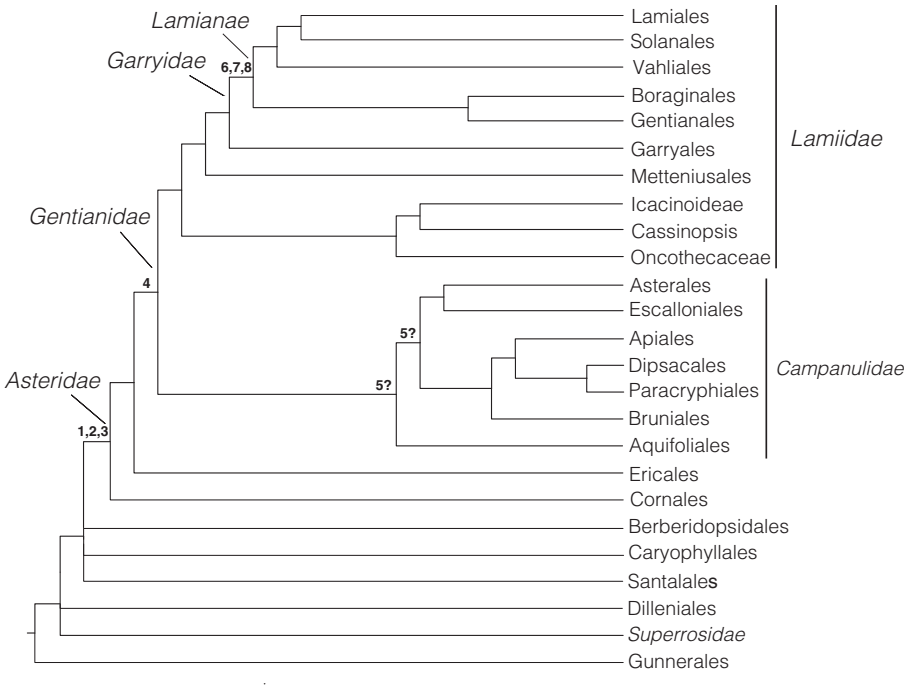

FIGURE 1. Summary tree for *Asteridae* showing potential synapomorphies of major clades recovered from the ancestral state reconstructions. The original reconstructions for each character are presented in Appendix S5–S19. The numbers denote the location of apomorphies and correspond to the following character states: 1. iridoid production, 2. unitegmic ovules, 3. cellular endosperm, 4. stamens equal in number to the petals, 5. inferior ovaries, 6. stamens adnate to the corolla, 7. unilacunar nodes, and 8. simple perforation plates.

RESULTS

The phylogenetic trees recovered from both ML (Appendix S3) and Bayesian (Appendix S4) approaches are largely congruent with each other and with previous studies of asterid phylogeny (e.g., Soltis et al., 2011; Refulio-Rodriguez and Olmstead, 2014; Stull et al., 2015). Areas of the tree with poor support include the positions of Dilleniaceae, Berberidopsidales, Santalales, and Caryophyllales. The relationships recovered here among major clades of asterids are consistent with previous studies: Cornales and Ericales are successively sister with 100% bootstrap support (BSS) to the *Gentianidae* clade, which comprises two well-supported subclades, *Lamiidae* and *Campanulidae* (each with 100% BSS). Within *Lamiidae*, Icacinaceae s.s. (Stull et al., 2015) and *Oncotheca* (Oncothecaceae) form a clade (BSS 100%) sister to the rest of the lamiids (= *Metteniusidae*; Stull et al., 2015). Within *Metteniusidae,* Metteniusaceae (Metteniusales) and Garryales are successively sister (with 96% and 99% BSS for the respective nodes) to the core lamiids, also known as *Lamianae*, which received maximal BSS as monophyletic.

Relationships among core lamiid clades recognized as orders in APG IV (2016) are less well supported: Boraginales + Gentianales (79% BSS); Solanales + Lamiales (63% BSS); and Vahliales + (Solanales + Lamiales) (91% BSS). Within *Campanulidae*, Aquifoliales were placed with maximal BSS as sister to the rest of the clade, but relationships within the latter clade (i.e., campanulids excluding Aquifoliales) are less well supported. Escalloniales and Asterales form a clade (95% BSS) sister to a clade of Bruniales, Apiales, Paracryphiales, and Dipsacales (83% BSS), among which relationships are poorly supported.

Figure 1 summarizes synapomorphies of major clades as recovered from the ancestral state reconstructions. The individual Mesquite ML reconstructions for each character are presented in Appendix S5–S19. The ML and Bayesian results from BayesTraits (incorporating branch-length information and phylogenetic uncertainty) are presented in Table 2. The results for particular characters are presented below. Although the Mesquite and BayesTraits analyses were largely congruent, some notable differences were recovered; these are highlighted below.

Habit and vegetative anatomy

A woody habit is reconstructed as ancestral for *Asteridae*—Mesquite proportional likelihood (MPL), BayesTraits proportional likelihood (BPL), BayesTraits posterior probability (BPP): 0.96/0.97/0.93—as well as for the deeper nodes leading to the root of the tree (in the Mesquite analyses). Within *Asteridae*, and especially within *Gentianidae*, there are complicated patterns of transition between woody and herbaceous habits. Both major gentianid clades, lamiids and campanulids, are reconstructed as ancestrally woody (lamiids: 0.99/0.99/0.99; campanulids: 0.99/0.92/0.86). The clade *Lamianae*, which comprises the bulk of lamiid diversity, is predominantly herbaceous, but the ancestral state of this clade is ambiguous (i.e., no state recovered at 0.50 or greater). It seems likely that there have been numerous transitions between woody and herbaceous habits (including probably numerous reversals). Within campanulids, the two basal-most nodes are reconstructed as woody, suggesting a similar pattern of numerous woody–herbaceous transitions within the clade.

Trilacunar nodes were reconstructed as ancestral for *Asteridae* (0.88/0.99/0.65). The Mesquite ML analysis reconstructed unilacunar nodes as synapomorphic for Ericales + *Gentianidae* (MPL: 0.83), while the BayesTraits analyses reconstructed trilacunar nodes as

TABLE 2. Ancestral state reconstructions for major clades inferred from maximum likelihood and Bayesian inference (in BayesTraits v 3.0; Pagel and Meade, 2006; http://www.evolution.rdg.ac.uk/BayesTraitsV3/BayesTraitsV3.html). The parentheses following each reconstructed state contain ML proportional likelihoods and Bayesian posterior probabilities for each state. Several characters (i.e., carpel number, fruit type, and seed number) could not be reconstructed using BayesTraits because they possessed too many free parameters to estimate.

Character	*Asteridae*	Ericales + *Gentianidae*	*Gentianidae*	*Lamiidae*	*Lamianae*	*Campanulidae*
Habit	Woody (0.97/0.93)	Woody (0.93/0.87)	Woody (0.98/0.95)	Woody (0.99/0.99)	Ambiguous	Woody (0.92/0.86)
Perforation plates	Scalariform (0.93/0.99)	Scalariform (0.99/0.90)	Scalariform (0.99/0.95)	Scalariform (0.99/0.95)	Simple (0.99/0.82)	Scalariform (0.99/0.90)
Nodal anatomy	Trilacunar (0.99/0.65)	Trilacunar (0.93/0.50)	Trilacunar (0.99/0.74)	Trilacunar (0.99/0.80)	Unilacunar (0.68/0.89)	Trilacunar (0.99/0.60)
Iridoids	Present (0.99/0.99)	Present (0.99/0.98)	Present (0.99/0.99)	Present (1.0/0.99)	Present (0.99/0.96)	Present (0.99/0.99)
Synsepaly	Fused (0.99/0.73)	Fused (0.99/0.57)	Fused (0.99/0.88)	Fused (0.99/0.80)	Fused (0.69/0.59)	Fused (0.97/0.83)
Sympetaly	Free (1.0/0.99)	Free (1.0/0.99)	Free (1.0/0.99)	Free (1.0/0.72)	Fused (1.0/0.70)	Free (1.0/0.99)
Stamen vs. petal number	More stamens (0.79/0.71)	More stamens (0.61/0.45)	Equal in number (0.99/0.91)	Equal in number (0.99/0.99)	Equal in number (0.99/0.80)	Equal in number (0.99/0.84)
Stamen Adnation	Free (0.99/0.99)	Free (0.99/0.99)	Free (0.99/0.99)	Free (0.99/0.98)	Fused (0.56/0.87)	Free (0.99/0.99)
Ovary position	Superior (0.65/0.28)	Superior (0.97/0.71)	Superior (0.85/0.50)	Superior (0.93/0.87)	Superior (0.89/0.49)	Inferior (0.79/0.37)
Ovule integument	Unitegmic (0.71/0.57)	Bitegmic (0.81/0.77)	Unitegmic (0.99/0.99)	Unitegmic (0.99/0.99)	Unitegmic (0.99/0.99)	Unitegmic (0.99/0.99)
Ovule nucellus	Crassinucellate (0.99/0.95)	Crassinucellate (0.99/0.89)	Crassinucellate (0.99/0.96)	Crassinucellate (0.99/0.99)	Tenuinucellate (0.99/0.99)	Crassinucellate (0.95/0.72)
Endosperm	Cellular (0.99/0.94)	Cellular (0.99/0.97)	Cellular (0.99/0.97)	Cellular (0.93/0.87)	Cellular (0.98/0.95)	Cellular (0.98/0.95)

ancestral for all major clades except the *Lamianae* clade, which has unilacunar nodes (BPL: 0.68; BPP: 0.89). Scalariform perforation plates appear to be ancestral for *Asteridae* (0.82/0.93/0.99), but this state may have arisen along an earlier branch (i.e., somewhere between the nodes defining *Superasteridae* and *Asteridae*). However, within asterids, reversals to simple perforation plates might represent synapomorphies for a number of clades: e.g., *Icacinoideae* (Stull et al. 2015), *Lamianae*, Campanulaceae, and within Apiales.

Chemistry

Although the presence of iridoids is almost entirely confined to *Asteridae,* with rare exceptions (e.g., *Liquidambar* and *Daphniphyllum*: Kaplan and Gottlieib, 1982), many members of this clade do not produce iridoids. The presence/absence of iridoids is poorly documented across basal lamiids and basal campanulids, and this lack of chemical data might explain ambiguity in the Mesquite reconstruction of several nodes (e.g., *Asteridae*: absent, 0.57; present, 0.43). The BayesTraits reconstructions, however, clearly recovered iridoid production as ancestral for asterids (BPL: 0.99; BPP: 0.99). Within gentianids, however, iridoids appear to have been lost on numerous occasions; in some cases, loss of iridoids seems to characterize major clades (e.g., Boraginales and Solanales).

Floral morphology

Free petals are reconstructed as the ancestral state of *Asteridae* (0.90/1.0/0.99), Ericales + *Gentianidae* (0.83/1.0/0.99), *Gentianidae* (0.77/1.0/0.99), and *Campanulidae* (0.92/1.0/0.99). The Mesquite analysis reconstructed fused petals as the ancestral state of lamiids (0.91), in contrast with the BayesTraits analyses, which recovered free petals as ancestral for lamiids (BPL: 1.0; BPP: 0.72). All analyses recovered fused petals as ancestral for *Lamianae* (0.97/1.0/0.70). Within the campanulids, most of the deeper nodes are reconstructed with free petals, suggesting that sympetaly arose on numerous occasions within the clade.

The fusion of staminal filaments to the petals (stamen adnation) shows a similar pattern to sympetaly. The ancestral state of *Asteridae* (0.99/0.99/0.99) and its major subclades, e.g., Cornales (MPL: 100), Ericales (MPL: 100), lamiids (0.93/0.99/0.98), and campanulids (0.99/0.99/0.99), is reconstructed as free. Within lamiids, the point at which stamen adnation arose is ambiguous, but potentially it evolved in the common ancestor of *Lamianae* (0.52/0.56/0.87). As with sympetaly, stamen adnation is reconstructed as arising multiple times within campanulids. The relative number of stamens and petals was ambiguous across the different analyses, but BayesTraits recovered stamen number equal to petal number as synapomorphic for *Gentianidae* (BPL: 0.99; BPP: 0.91). Having fewer stamens than petals is likely a synapomorphy for Lamiales or a subclade within the order.

Superior ovaries (i.e., hypogynous flowers) are reconstructed as ancestral for *Asteridae* (0.99/0.65/0.28) and predominate across the clade, but multiple subclades show inferior ovaries (i.e., epigynous flowers) as a possible synapomorphy, e.g., Cornales (MPL: 0.47) and Rubiaceae (MPL: 0.97). The BayesTraits analyses recovered inferior ovaries as ancestral for *Campanulidae*, albeit with some ambiguity (BPL: 0.79, BPP: 0.37), while Mesquite recovered inferior ovaries as ancestral for a subclade, i.e., all campanulids excluding Aquifoliales (MPL: 0.94).

Fruit type

The ancestral fruit type of *Asteridae* is ambiguous, reconstructed as either drupes (MPL: 0.57) or capsules (MPL: 0.43); this character could not be reconstructed using BayesTraits because the number of states required the estimation of too many free parameters. Drupaceous fruits, however, appear ancestral for *Gentianidae* (MPL: 0.85) and its two major subclades, lamiids (MPL: 0.96) and campanulids (MPL: 0.86). Within *Lamianae*, fruit type is highly diverse, but capsules are reconstructed as ancestral for this clade (MPL: 0.93). Capsules are also reconstructed as ancestral for Asterales (MPL: 0.90), a major clade of campanulids.

Ovule features

Unitegmic ovules were reconstructed as synapomorphic for *Asteridae* (0.95/0.71/0.57). However, unitegmic ovules are not ubiquitous across *Asteridae*: some members of Ericales possess bitegmic ovules, and bitegmic or partially bitegmic ovules have been documented in Icacinaceae and Metteniusaceae (although in general this character is poorly documented across basal lamiids). The Mesquite and BayesTraits analyses showed major discrepancies in the reconstructions of nucellus condition. Mesquite recovered tenuinucellate ovules as possibly synapomorphic for *Asteridae* (PL: 0.63), whereas the BayesTraits analyses recovered crassinucellate ovules as ancestral for *Asteridae* (0.99/0.95) and all major subclades except *Lamianae* (tenuinucellate ovules: 0.99/0.99). This character does indeed show considerable variation within the asterids, with Cornales showing both crassinucellate and tenuinucellate ovules, and some members of Icacinaceae showing weakly crassinucellate ovules; this character is also very poorly documented across the basal lamiids.

Cellular endosperm is reconstructed as synapomorphic for *Asteridae* (0.96/0.99/0.94), but multiple clades within *Asteridae* appear to show reversals to nuclear endosperm (e.g., Boraginales, Gentianales, some Aquifoliales, and Apiales). Endosperm type, however, like the other ovule features noted above, is poorly documented in basal lamiids and basal campanulids (i.e., Aquifoliales).

DISCUSSION

Our ancestral state reconstructions, employing an expanded phylogenetic framework, provide an improved understanding of character evolution across *Asteridae*, both confirming and challenging long-held ideas about asterid morphological evolution. Below we outline likely synapomorphies of major clades and also provide a more detailed discussion of the ancestral morphological features of *Gentianidae*, one of the largest angiosperm clades (~80,000 spp.). We then highlight areas of the phylogeny with considerable missing data, potentially causing ambiguity in our understanding of the evolutionary history of certain phenotypic characters. We hope this section will guide future studies on asterid character evolution. Finally, we discuss some inherent difficulties of conducting ancestral state reconstruction (e.g., incomplete sampling, ambiguity in delimiting and/or coding character states); these factors are important to consider when interpreting results given the fundamental influence that they can have on ancestral state reconstruction.

Synapomorphies of major clades

A number of the characters we examined show considerable variability across the phylogeny, potentially due to parallel evolution and/or reversals, leading to ambiguity in the reconstruction of ancestral states and synapomorphies. Nevertheless, we documented multiple unambiguous synapomorphies for major clades. Ovules with a single integument (unitegmy) and cellular endosperm appear to be synapomorphies of *Asteridae*. Patterns of nucellus evolution vary across analyses; therefore, we could not determine whether tenuinucellate ovules are a synapomorphy of asterids. Iridoid production is also likely a synapomorphy of asterids, but this character shows complex patterns of evolution within asterids, possibly owing to multiple instances of loss.

Sympetaly, which has long been associated with asterids, does not appear to be a synapomorphy of either *Asteridae* or the subclade *Gentianidae*. The ancestral state of *Lamiidae* differed across analysis type; Mesquite reconstructed fused petals as ancestral, whereas BayesTraits recovered free petals as ancestral. These differences might stem from the absence of branch-length information (and/or phylogenetic uncertainty) in the Mesquite analyses.

Nevertheless, this character appears to show a much more complicated pattern of evolution across asterids than traditionally thought, potentially with multiple independent transitions from free to fused petals. Two developmental patterns, generally referred to as "early" and "late" sympetaly (Erbar and Leins, 1996; Erbar, 1991), have been documented for sympetalous corollas in asterids. Although early and late sympetaly generally correspond to flowers in *Campanulidae* and *Lamiidae*, respectively, there are notable exceptions (e.g., Rubiaceae, a major clade of *Lamiidae*, has early sympetaly). Our results, combined with the developmental complexity of sympetalous corollas across asterids, support the possibility that sympetalous corollas evolved on multiple occasions in *Asteridae*. More detailed morphological and developmental investigations, within a phylogenetic context, will be necessary to better understand the evolution of sympetalous corollas in asterids.

Stamens equal in number to the petals might represent a synapomorphy for *Gentianidae*, but this character showed some differences across the different reconstructions. Stamen adnation appears to have also arisen independently on multiple occasions, but it nevertheless serves as a synapomorphy for particular clades. Stamen adnation is probably a synapomorphy of *Lamianae*, for example, and potentially for several campanulid subclades (although those patterns are more complex). Capsular fruits might also represent a synapomorphy of *Lamianae*, but numerous transitions to other fruit types occur within this clade.

Simple perforation plates and unilacunar nodes may represent synapomorphies for *Lamianae*. However, these states also predominate in Icacinaceae (in particular, they are ubiquitous across *Icacinoideae*, which includes all genera except *Cassinopsis*), which is positioned outside *Lamianae* in the broader *Lamiidae*. This suggests that these features either evolved independently in both *Icacinoideae* and *Lamianae* or instead arose on a deeper branch (i.e., the common ancestor of *Lamiidae*).

Ancestral morphology of Gentianidae

Although synapomorphies for *Gentianidae* remain unclear or few (e.g., stamens and petals equal in number), the character reconstructions presented here have greatly clarified the morphological features ancestral to this clade. Ancestral gentianids were most likely woody with scalariform perforation plates and trilacunar nodes; iridoids were present; the flowers potentially had free (or only slightly fused) petals, free stamens, and a superior ovary; fruits were drupes; ovules were unitegmic; and endosperm production was cellular. Most of these features are also supported as ancestral for the two major gentianid subclades, lamiids and campanulids. This view contrasts with previous research suggesting, for example, that capsular fruits were ancestral for campanulids (Beaulieu and Donoghue, 2013).

Our analyses provide a solid foundation for exploring character transitions (or key innovations: Miller, 1949; Galis, 2001) within *Gentianidae* that might be associated with increased diversification. Although the gentianids represent one of the largest angiosperm

clades, with approximately one-third of angiosperm species richness—i.e., one third of 250,000 to 300,000+ species, depending on estimates (Govaerts, 2001, 2003; Christenhuszn and Byng, 2016; Judd et al., 2016)—much of that diversity can be attributed to particular species-rich subclades, e.g., Lamiales (or subclades within), Rubiaceae, Asteraceae, Solanaceae, Apocynaceae, and Boraginales. Future studies could explore, for example, whether transitions to sympetaly, capsular or achene fruits, and/or an herbaceous habit within independent gentianid subclades are associated with increased diversification (e.g., Beaulieu and Donoghue, 2013).

Although the ancestral habit of *Asteridae* and *Gentianidae* was reconstructed as woody, *Gentianidae* appears to show complicated patterns of transition between woody and herbaceous habits. Several notably species-rich gentianid clades have high proportions (~50% or greater) of herbaceous taxa (e.g., Lamiales, Solanales, Asterales, Apiales; FitzJohn et al., 2014), suggesting that herbaceousness might be linked to increased diversification rates. This is not necessarily surprising as the herbaceous habit has been linked to increased rates of molecular evolution (Smith and Donoghue, 2008) and the ability to inhabit diverse climates (Zanne et al., 2014), but additional research will be necessary to better understand the relationship of habit to diversification rate (Gianoli, 2004).

Missing (morphological) data and future work

Several characters long considered important in asterid evolution and classification (e.g., sympetaly, integument number, nucellus type, endosperm type, iridoid production) are poorly documented in several key parts of the asterid tree. Genera formerly included in Icacinaceae s.l. (Howard, 1940; Sleumer, 1942)—and now placed variously in the lamiids and campanulids—are very poorly documented for these particular characters. Also, some of the genera of Icacinaceae s.l. that have been studied possess strange combinations of features not found in other gentianids (e.g., *Emmotum*, which is now in Metteniusaceae, has bitegmic and crassinucellate ovules; Endress and Rapini, 2014). More detailed morphological, anatomical, and developmental investigations of poorly studied lamiids (e.g., Oncothecaceae, Metteniusaceae, Icacinaceae s.s.) and campanulids (e.g., members of Aquifoliales) will therefore be critical to amass the data necessary to better understand the evolution of these characters in asterids.

Difficulties of ancestral state reconstruction

Ancestral state reconstruction is a challenging enterprise, and reconstructions are ultimately influenced by a number of factors that deserve careful attention from the researcher. Phylogenies are seldom reconstructed with complete resolution and confidence, and therefore the incorporation of phylogenetic uncertainty is critical in ancestral state reconstruction; use of a single tree may convey more confidence than warranted (Pagel et al., 2004). To accommodate phylogenetic uncertainty in our analyses (e.g., among core lamiid order, [Stull et al., 2015] and among certain campanulid orders [Soltis et al., 2011]), we used the program BayesTraits (Pagel and Meade, 2006) to conduct reconstructions across a distribution of trees. In these analyses, the ancestral states of several characters (e.g., habit, nodal anatomy, and stamen adnation) were ambiguous for the core lamiids (i.e., *Lamianae*), but this might be a consequence of complicated evolutionary patterns or missing data, in addition to phylogenetic uncertainty.

Sampling is another critical consideration for ancestral state reconstruction. It is often impractical to sample all species within a clade of interest (especially when the clade is large, as in the present study; see also the paper in this issue by Folk et al. [2018]), and there are several approaches to accommodate incomplete sampling. One is to sample and score individual species scattered across the clade of interest (e.g., Soltis et al., 2013b); as long as the sampling of species is sufficient to capture the morphological variation present in the clade, this should be an effective approach for reconstructing ancestral states of the clade as a whole (Salisbury and Kim, 2001). Another approach is to choose exemplar species of subclades and to assign states to the exemplars based on previous studies or knowledge of variation across the entire subclade (Donoghue and Ackerly, 1996); when the subclade is polymorphic for a given character, the exemplar is either scored as polymorphic or assigned the presumed ancestral state for the clade (Nixon and Davis, 1991). The latter approach (using exemplars) is beneficial in that it potentially incorporates more complete information about the morphological variation within a clade, but problematic in that it requires more assumptions and/or prior knowledge that might not be readily available (e.g., the ancestral state of each subclade included in the analysis). To avoid such complications, we opted to score individual species sampled broadly across the clade of interest (asterids), which is the approach typically employed in contemporary studies of character evolution (e.g., Beaulieu and Donoghue, 2013; Soltis et al., 2013b; Zanne et al., 2014; Wu et al., 2015; Sauquet et al., 2017).

Missing phenotypic data (for poorly studied taxa) can lead to ambiguity in ancestral state reconstruction, especially when data are missing from crucial parts of the tree (e.g., across a basal grade; Donoghue and Acklerly, 1996; Sauquet et al., 2017). A perhaps more vexing issue lies in the delimitation and coding of morphological character states—which can be problematic for both morphology-based phylogenetics and ancestral state reconstruction (Stevens, 1991; Wiens, 2001). The delimitation of characters into character states is rarely straightforward; often taxa will exhibit expressions of a character that are difficult to bin into one qualitative state or another (Stevens, 1991; Wiens, 2001; Scotland et al., 2003). Fruits are a classic example of a character that is challenging to break into meaningful categories, given the prevalence of fruits intermediate between frequently used types, as well as ambiguous homology between superficially or functionally similar fruit morphologies (e.g., Judd, 1985; Stevens, 2001 onward; Beaulieu and Donoghue, 2013; Judd et al., 2016). It is therefore important to outline clearly the characters and states used for reconstructions—as we have attempted to do (Appendix S2)—so that future research can reproduce, build on, or critique previous studies of character evolution.

ACKNOWLEDGEMENTS

G.W.S. thanks Peter Stevens for helpful discussion regarding the characters and states used for this study. The authors thank the associate editor, Richard Olmstead, and an anonymous reviewer, who provided helpful comments for improving the manuscript. This work was supported by NSF grant DEB-1310805 (Doctoral Dissertation Improvement Grant to S. R. Manchester, P.S.S., and G.W.S.) and NSF grant DBI-1612032 (Postdoctoral Fellowship to G.W.S.).

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

LITERATURE CITED

Albach, D. C., P. S. Soltis, and D. E. Soltis. 2001a. Patterns of embryological and biochemical evolution in the asterids. *Systematic Botany* 26: 242–262.

Albach, D. C., P. S. Soltis, D. E. Soltis, and R. G. Olmstead. 2001b. Phylogenetic analysis of asterids based on sequences of four genes. *Annals of the Missouri Botanical Garden* 88: 162–212.

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

APG, II [Angiosperm Phylogeny Group II]. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

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 Linnaean Society* 161: 105–121.

APG, IV [Angiosperm Phylogeny Group 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.

Bailey, I. W., and R. A. Howard. 1941a. The comparative morphology of the Icacinaceae I. Anatomy of the node and internode. *Journal of the Arnold Arboretum* 22: 125–132.

Bailey, I. W., and R. A. Howard. 1941b. The comparative morphology of the Icacinaceae II. Vessels. *Journal of the Arnold Arboretum* 22: 171–187.

Beaulieu, J. M., and M. J. Donoghue. 2013. Fruit evolution and diversification in campanulid angiosperms. *Evolution* 67: 3132–3144.

Bremer, K., A. Backlund, B. Sennblad, U. Swenson, K. Andreasen, M. Hjertson, J. Lundberg, M. Backlund, and B. Bremer. 2001. A phylogenetic analysis of 100+ genera and 50+ families of euasterids based on morphological and molecular data with notes on possible higher level morphological synapomorphies. *Plant Systematics and Evolution* 229: 137–169.

Bremer, B., K. Bremer, N. Heidari, P. Erixon, R. G. Olmstead, A. A. Anderberg, M. Källersjö, and E. Barkhordarian. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molecular Phylogenetics and Evolution* 24: 274–301.

Cantino, P. D., J. A. Doyle, S. W. Graham, W. S. Judd, R. G. Olmstead, D. E. Soltis, P. S. Soltis, and M. J. Donoghue. 2007. Towards a phylogenetic nomenclature of *Tracheophyta*. *Taxon* 56: 822–846.

Chase, M. W, D. E. Soltis, R. G. Olmstead, D. Morgan, D. H. Les, B. D. Mishler, M. R. Duvall, et al. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. *Annals of the Missouri Botanical Garden* 80: 528–548+550–580.

Christenhuszn, M. J. M., and J. W. Byng. 2016. The number of known plant species in the world and its annual increase. *Phytotaxa* 261: 201–217.

Cronquist, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, NY, NY, USA.

Cronquist, A. 1988. The evolution and classification of flowering plants, 2nd ed. New York Botanical Garden, NY, NY, USA.

de Jussieu, A.-L. 1789. Genera plantarum. Herissant & Barrois, Paris.

Dickison, W. C. 1986. Further observations on the oral anatomy and pollen morphology of *Oncotheca* (Oncothecaceae). *Brittonia* 38: 249–259.

Dickison, W. C., and V. Bittrich. 2016. Metteniusaceae. *In* J. W. Kadereit and V. Bittrich [eds.], The families and genera of vascular plants, vol. 14, Flowering plants. Eudicots. Springer International Publishing, Cham, Switzerland.

Donoghue, M. J., and D. D. Ackerly. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philosophical Transactions of the Royal Society* 351: 1241–1249. (Reprinted in: Silvertown, J., M. Franco, and J. L. Harper. 1997. Plant Life Histories: Ecology, Phylogeny and Evolution. Cambridge Univ. Press).

Doyle, J. A. 2005. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. *Grana* 44: 227–251.

Doyle, J. A. 2007. Systematic value and evolution of leaf architecture across the angiosperms in light of molecular phylogenetic analyses. *Cfs Courier Forschungsinstitut Senckenberg* 21–37.

eFloras. 2008. http://www.efloras.org [accessed April 2017]. Missouri Botanical Garden, St. Louis, MO, USA; Harvard University Herbaria, Cambridge, MA, USA.

Endress, P. K. 1996. Homoplasy in angiosperm flowers. *In* M. J. Sanderson and L. Hufford [eds.], Homoplasy: The recurrence of similarity in evolution. Academic Press, San Diego, CA, USA.

Endress, P. K. 2010. Flower structure and trends of evolution in eudicots and their major subclades. *Annals of the Missouri Botanical Garden* 97: 541–583.

Endress, P. K. 2011a. Angiosperm ovules: diversity, development, and evolution. *Annals of Botany* 107: 1465–1489.

Endress, P. K. 2011b. Evolutionary diversification of the flowers in angiosperms. *American Journal of Botany* 98: 370–396.

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

Endress, P. K., and J. A. Doyle. 2015. Ancestral traits and specializations in the flowers of the basal grade of living angiosperms. *Taxon* 64: 1093–1116.

Endress, P. K., and A. Rapini. 2014. Floral structure of *Emmotum* (Icacinaceae sensu stricto or Emmotaceae), a phylogenetically isolated genus of lamiids with a unique pseudotrimerous gynoecium, bitegmic ovules and monosporangiate thecae. *Annals of Botany* 114: 945–959.

Erbar, C. 1991. Sympetalae—a systematic character? *Botanische Jahrbücher für Systematik* 112: 417–451.

Erbar, C., and P. Leins. 1996. Distribution of the character states "early" and "late sympetaly" within the "Sympetalae Tetracyclicae" and presumably related groups. *Botanica Acta* 109: 427–440.

Erbar, C., P. Leins, B.-E. van Wyk, and P. M. Tilney. 2004. Sympetaly in Apiales (Apiaceae, Araliaceae, Pittosporaceae). *South African Journal of Botany* 70: 458–467.

Fagerlind, F. 1945. Bau des Gynöceums, der Samenanlage und des Embryosackes bei einigen Repräsentanten der Familie Icacinaceae. *Svensk Botanisk Tidskrift* 39: 346–364.

FitzJohn, R. G., M. W. Pennell, A. E. Zanne, P. F. Stevens, D. C. Tank, and W. K. Cornwell. 2014. How much of the world is woody? *Journal of Ecology* 102: 1266–1272.

Flora of North America Editorial Committee [eds.]. 1993 onward. Flora of North America North of Mexico. 20+ vols. Oxford University Press, New York, USA; Oxford, UK.

Folk, R. A., P. S. Soltis, D. E. Soltis, and R. Guralnick. 2018. New prospects in the detection and comparative analysis of hybridization in the tree of life. *American Journal of Botany*. https://doi.org/10.1002/ajb2.1018.

Galis, F. 2001. Kew innovations and radiations. *In* G. P. Wagner [ed.], The character concept in evolutionary biology. Academic Press, San Diego, CA, USA.

Gianoli, E. 2004. Evolution of a climbing habit promotes diversification in flowering plants. *Proceedings of the Royal Society of London, B, Biological Science* 271: 2011–2015.

González, F. A., and P. J. Rudall. 2010. Flower and fruit characters in the early-divergent lamiid family Metteniusaceae, with particular reference to the evolution of pseudomonomery. *American Journal of Botany* 97: 191–206.

Govaerts, R. 2001. How many species of seed plants are there? *Taxon* 50: 1085–1090.

Govaerts, R. 2003. How many species of seed plants are there? – a response. *Taxon* 52: 583–584.

Howard, R. A. 1940. Studies of the Icacinaceae. I. Preliminary taxonomic notes. *Journal of the Arnold Arboretum* 21: 461–489.

Howard, R. A. 1942a. Studies of the Icacinaceae. II. *Humirianthera, Leretia, Mappia,* and *Nothapodytes*, valid genera of the Icacineae. *Journal of the Arnold Arboretum* 23: 55–78.

Howard, R. A. 1942b. Studies of the Icacinaceae. III. A revision of *Emmotum*. *Journal of the Arnold Arboretum* 23: 479–494.

Howard, R. A. 1942c. Studies of the Icacinaceae. IV. Considerations of the New World genera. *Contributions from the Gray Herbarium of Harvard University* 142: 3–60.

Howard, R. A. 1942d. Studies of the Icacinaceae. V. A revision of the genus *Citronella* D. Don. *Contributions from the Gray Herbarium of Harvard University* 142: 60–89.

Howard, R. A. 1992. A revision of *Casimirella*, including *Humirianthera* (Icacinaceae). *Brittonia* 44: 166–172.

Huelsenbeck, J. P., and F. R. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. *Biometrics* 17: 754–755.

Jansen, R. K., Z. Cai, L. A. Raubeson, H. Daniell, C. W. dePamphilis, J. Leebens-Mack, K. F. Müller, et al. 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.

Jensen, S. R. 2000. Chemical relationships of *Polypremum procumbens*, *Tetrachondra hamiltonii*, and *Peltanthera floribunda*. *Biochemical Systematics and Ecology* 28: 45–51.

Judd, W. S. 1985. A revised traditional/descriptive classification of fruits for use in floristics and teaching. *Phytologia* 58: 232–242.

Judd, W. S., and R. G. Olmstead. 2004. A survey of tricolpate (eudicot) phylogenetic relationships. *American Journal of Botany* 91: 1627–1644.

Judd, W. S., C. S. Campbell, E. A. Kellogg, P. F. Stevens, and M. J. Donoghue. 2016. Plant systematics: a phylogenetic approach, 4th ed. Sinauer, Sunderland, MA, USA.

Kaplan, M. A. C., and O. R. Gottlieib. 1982. Iridoids as systematic markers in dicotyledons. *Biochemical Systematics and Ecology* 10: 329–347.

Kårehed, J. 2001. Multiple origins of the tropical forest tree family Icacinaceae. *American Journal of Botany* 88: 2259–2274.

Knapp, S. 2002. Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in Solanaceae. *Journal of Experimental Botany* 53: 2001–2022.

Leins, P., and C. Erbar. 2004. Floral organ sequences in Apiales (Apiaceae, Araliaceae, Pittosporaceae). *South African Journal of Botany* 70: 468–474.

Lens, F., J. Kårehed, P. Baas, S. Jansen, D. Rabaey, S. Huysmans, T. Hamann, and E. Smets. 2008. The wood anatomy of the polyphyletic Icacinaceae s.l., and their relationships within asterids. *Taxon* 57: 525–552.

Leubert, F., L. Cecchi, M. W. Frohlich, M. Gottschling, C. M. Guilliams, K. E. Hasenstab-Lehman, and H. H. Hilger. 2016. Familial classification of the Boraginales. *Taxon* 65: 502–522.

Liu, M., J. Zhao, J. Wang, Z. Liu, and G. Liu. 2017. Phylogenetic analysis of 25 plant species representing 19 angiosperm families and one gymnosperm family based on 390 orthologous genes. *Plant Systematics and Evolution* 303: 413–417.

Maddison, W. P., and D. R. Maddison. 2017. Mesquite: a modular system for evolutionary analysis, version 3.2. Computer program and documentation distributed by the author, website http://mesquiteproject.org [accessed 20 April 2017].

Mauritzon, J. 1936. Embryologische Angaben über Stackhousiaceae, Hippocrateaceae, und Icacinaceae. *Svensk Botanisk Tidskrift* 30: 541–550.

Miller, A. 1949. Some ecologic and morphologic considerations in the evolution of higher taxonomic categories. *In* E. Mayr and E. Schüz [eds.], Ornithologie als Biologische Wissenschaft, 84–88. Universitätsverlag, Heidelberg, Germany.

Moore, M. J., C. D. Bell, P. S. Soltis, and D. E. Soltis. 2007. Using plastid genomic-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences, USA* 104: 19363–19368.

Moore, M. J., A. Dhingra, P. S. Soltis, R. Shaw, W. G. Farmerie, K. M. Folta, and D. E. Soltis. 2006. Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biology* 6: 17–30.

Moore, M. J., P. S. Soltis, C. D. Bell, J. G. Burleigh, and D. E. Soltis. 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.

Nixon, K. C., and J. I. Davis. 1991. Polymorphic taxa, missing values and cladistic analyses. *Cladistics* 7: 233–241.

Olmstead, R. G., B. Bremer, K. M. Scott, and J. D. Palmer. 1993. A parsimony analysis of the Asteridae sensu lato based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 80: 700–722.

Olmstead, R. G., K.-J. Kim, R. K. Jansen, and S. J. Wagstaff. 2000. The phylogeny of the *Asteridae* sensu lato based on chloroplast *ndhF* gene sequences. *Molecular Phylogenetics and Evolution* 16: 96–112.

Olmstead, R. G., H. J. Michaels, K. M. Scott, and J. D. Palmer. 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.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53: 673–684.

Pagel, M., and A. Meade. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *American Naturalist* 167: 808–825.

Peng, H., and R. A. Howard. 2008. Icacinaceae. *In* Z. Y. Wu, P. H. Raven, and D. Y. Hong [eds.], Flora of China, vol. 11, 505–514. Science Press, Beijing, China; Missouri Botanical Garden Press, St. Louis, MO, USA.

Potgieter, M. J., and R. Duno. 2016. Icacinaceae. *In* J. W. Kadereit and V. Bittrich [eds.], The families and genera of vascular plants, vol. 14, Flowering plants. Eudicots. Springer International Publishing, Cham, Switzerland.

Potgieter, M. J., M. Schori, and T. M. A. Utteridge. 2016. Stemonuraceae. *In* J. W. Kadereit and V. Bittrich [eds.], The families and genera of vascular plants, vol. 14, Flowering plants. Eudicots. Springer International Publishing, Cham, Switzerland.

Rambaut, A., and A. J. Drummond. 2009. Tracer, version 1.5 for Macintosh. Computer program and documentation distributed by the author, website http://beast.bio.ed.ac.uk/Tracer.

Refulio-Rodriguez, N. F., and R. G. Olmstead. 2014. Phylogeny of Lamiidae. *American Journal of Botany* 101: 287–299.

Ronquist, F., M. Teslenko, P. van der Mark, D. Ayres, A. Darling, S. Höhna, B. Larget, et al. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Ronse De Craene, L. P. 2008. Homology and evolution of petals in the core eudicots. *Systematic Botany* 33: 301–325.

Ronse De Craene, L. P., and S. Brockington. 2013. Origin and evolution of petals in the angiosperms. *Plant Ecology and Evolution* 146: 5–25.

Ronse De Craene, L. P., D. E. Soltis, and P. S. Soltis. 2003. Evolution of floral structures in the basal angiosperms. *International Journal of Plant Sciences* 164(Supplement 5): S329–S363.

Salisbury, B. A., and J. Kim. 2001. Ancestral state estimation and taxon sampling density. *Systematic Biology* 50: 557–564.

Sauquet, H., M. von Balthazar, S. Magallón, J. A. Doyle, P. K. Endress, E. J. Bailes, E. Barroso, et al. 2017. The ancestral flower of angiosperms and its early diversification. *Nature Communications* 8: 16047. https://doi.org/10.1038/ncomms16047.

Savolainen, V., M. W. Chase, S. B. Hoot, C. M. Morton, D. E. Soltis, C. Bayer, M. F. Fay, A. Y. de Bruijn, S. Sullivan, and Y.-L. Qiu. 2000b. Phylogenetics of flowering plants based upon a combined analysis of plastid *atpB* and *rbcL* gene sequences. *Systematic Biology* 49: 306–362.

Savolainen, V., M. F. Fay, D. C. Albach, A. Backlund, M. van der Bank, K. M. Cameron, S. A. Johnson, et al. 2000a. Phylogeny of eudicots: a nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bulletin* 55: 257–309.

Schori, M. 2016. Cardiopteridaceae. *In* J. W. Kadereit and V. Bittrich [eds.], The families and genera of vascular plants, vol. 14, Flowering plants. Eudicots. Springer International Publishing, Cham, Switzerland.

Scotland, R. W., R. G. Olmstead, and J. R. Bennett. 2003. Phylogeny reconstruction: the role of morphology. *Systematic Biology* 52: 539–548.

Sleumer, H. 1942. Icacinaceae. *In* A. Engler [ed.], Die natürlichen P anzenfamilien, 2nd ed., vol. 20b, 322–396. Wilhelm Engelmann, Leipzig, Germany.

Sleumer, H. 1969. Materials towards the knowledge of the Icacinaceae of Asia, Malesia, and adjacent areas. *Blumea* 17: 181–264.

Sleumer, H. 1971. Icacinaceae. *In* C. G. G. J. van Steenis [ed.], Flora Malesiana, series I, vol. 7, 1–87. Noordho, Leyden, Netherlands.

Smith, S. A., and M. J. Donoghue. 2008. Rates of molecular evolution are linked to life history in flowering plants. *Science* 322: 86–89.

Soltis, D. E., M. A. Gitzendanner, G. W. Stull, M. Chester, A. Chanderbali, S. Chamala, I. Jordon-Thaden, P. S. Soltis, P. S. Schnable, and W. B. Barbazuk. 2013a. The potential of genomics in plant systematics. *Taxon* 62: 886–898.

Soltis, D. E., M. E. Mort, M. Latvis, E. V. Mavrodiev, B. C. O'Meara, P. S. Soltis, J. G. Burleigh, and R. Rubio de Casas. 2013b. Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany* 100: 916–929.

Soltis, D. E., S. A. Smith, N. Cellinese, K. J. Wurdack, D. C. Tank, S. F. Brockington, N. F. Refulio-Rodriguez, et al. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* 98: 704–730.

Soltis, D. E., P. S. Soltis, M. W. Chase, M. E. Mort, D. C. Albach, M. Zanis, V. Savolainen, et al. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL,* and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

Soltis, D. E., P. S. Soltis, P. K. Endress, and M. W. Chase. 2005. Phylogeny and evolution of angiosperms. Sinauer, Sunderland, MA, USA.

Soltis, D. E., P. S. Soltis, D. L. Nickrent, L. A. Johnson, W. J. Hahn, S. B. Hoot, J. A. Sweere, et al. 1998. Phylogenetic relationships among angiosperms inferred from 18S rDNA sequences. *Annals of the Missouri Botanical Garden* 84: 1–49.

Soltis, P. S., D. E. Soltis, and M. W. Chase. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.

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

Stevens, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: a review. *Systematic Botany* 16: 553–583.

Stevens, P. F. 2001 onward. Angiosperm Phylogeny Website, version 12. Website http://www.mobot.org/MOBOT/research/APweb/ [accessed 20 April 2017].

Stull, G. W., R. Duno de Stefano, D. E. Soltis, and P. S. Soltis. 2015. Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *American Journal of Botany* 102: 1794–1813.

Takhtajan, A. L. 1980. Outline of the classification of flowering plants (Magnoliophyta). *Botanical Review* 46: 225–359.

Takhtajan, A. L. 1997. Diversity and classification of flowering plants. Columbia University Press, NY, NY, USA.

Takhtajan, A. L. 2009. Flowering plants, 2nd ed. Springer, Dordrecht, Netherlands.

Tank, D. C., and M. J. Donoghue. 2010. Phylogeny and phylogenetic nomenclature of the Campanulidae based on an expanded sample of the genes and taxa. *Systematic Botany* 35: 425–441.

Wang, H., M. J. Moore, P. S. Soltis, C. D. Bell, S. F. Brockington, R. Alexandre, C. C. Davis, et al. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences, USA* 106: 3853–3858.

Wickett, N. J., S. Mirarab, N. Nguyen, T. Warnow, E. Carpenter, N. Matasci, S. Ayyampalayam, et al. 2014. A phylotranscriptomics analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences, USA* 111: E4859–E4868.

Wiens, J. J. 2001. Character analysis in morphological phylogenetics: problems and solutions. *Systematic Biology* 50: 689–699.

Wu, Z.-Y., R. I. Milne, C.-J. Chen, J. Liu, H. Wang, and D.-Z. Li. 2015. Ancestral state reconstruction reveals rampant homoplasy of diagnostic morphological characters in Urticaceae, conflicting with current classification. *PLoS ONE* 10: e0141821.

Wurdack, K. J., and C. C. Davis. 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.

Xi, Z., L. Liu, J. S. Rest, and C. C. Davis. 2014. Coalescent versus concatenation methods and the placement of Amborella as sister to water lilies. *Systematic Biology* 63: 919–932.

Xi, Z., B. R. Ruhfel, H. Schaefer, A. M. Amorim, M. Sugumaran, K. J. Wurdack, P. K. Endress, et al. 2012. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proceedings of the National Academy of Sciences, USA* 109: 17519–17524.

Zanne, A. E., D. C. Tank, W. K. Cornwell, J. M. Eastman, S. A. Smith, R. G. FitzJohn, D. J. McGlinn, et al. 2014. Three keys to the radiation of angiosperms into freezing environments. *Nature* 506: 89–92.

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