RESEARCH ARTICLE

American Journal of Botany

Botanical Society of America
leading scientists and educators since 1893

INVITED SPECIAL ARTICLE

For the Special Issue: Exploring Angiosperms353: a Universal Toolkit for Flowering Plant Phylogenomics

Reconstructing Dipsacales phylogeny using Angiosperms353: issues and insights

Aaron K. Lee[1,2,*], Ian S. Gilman[3,*], Mansa Srivastav[3], Ariel D. Lerner[1], Michael J. Donoghue[3], and Wendy L. Clement[1,4]

Manuscript received 25 October 2020; revision accepted 12 May 2021.

[1] Department of Biology, The College of New Jersey, Ewing, NJ 08628, USA

[2] Department of Plant and Microbial Biology, University of Minnesota - Twin Cities, Saint Paul, MN 55108, USA

[3] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA

[4] Author for correspondence (e-mail: clementw@tcnj.edu)

*Indicates equal authorship.

Citation: Lee, A. K., I. S. Gilman, M. Srivastav, A. D. Lerner, M. J. Donoghue, and W. L. Clement. 2021. Reconstructing Dipsacales phylogeny using Angiosperms353: issues and insights. *American Journal of Botany* 108(7): 1122–1142.

doi:10.1002/ajb2.1695

PREMISE: Phylogenetic relationships within major angiosperm clades are increasingly well resolved, but largely informed by plastid data. Areas of poor resolution persist within the Dipsacales, including placement of *Heptacodium* and *Zabelia*, and relationships within the Caprifolieae and Linnaeeae, hindering our interpretation of morphological evolution. Here, we sampled a significant number of nuclear loci using a Hyb-Seq approach and used these data to infer the Dipsacales phylogeny and estimate divergence times.

METHODS: Sampling all major clades within the Dipsacales, we applied the Angiosperms353 probe set to 96 species. Data were filtered based on locus completeness and taxon recovery per locus, and trees were inferred using RAxML and ASTRAL. Plastid loci were assembled from off-target reads, and 10 fossils were used to calibrate dated trees.

RESULTS: Varying numbers of targeted loci and off-target plastomes were recovered from most taxa. Nuclear and plastid data confidently place *Heptacodium* with Caprifolieae, implying homoplasy in calyx morphology, ovary development, and fruit type. Placement of *Zabelia*, and relationships within the Caprifolieae and Linnaeeae, remain uncertain. Dipsacales diversification began earlier than suggested by previous angiosperm-wide dating analyses, but many major splitting events date to the Eocene.

CONCLUSIONS: The Angiosperms353 probe set facilitated the assembly of a large, single-copy nuclear dataset for the Dipsacales. Nevertheless, many relationships remain unresolved, and resolution was poor for woody clades with low rates of molecular evolution. We favor expanding the Angiosperms353 probe set to include more variable loci and loci of special interest, such as developmental genes, within particular clades.

KEY WORDS Adoxaceae; Caprifoliaceae; divergence times; fossils; *Heptacodium*; Hyb-Seq; Morinoideae; phylogenomics; target enrichment; *Zabelia*.

Despite enormous progress in reconstructing the phylogeny of major angiosperm clades, key relationships within many groups have yet to be confidently resolved, and the timing of diversification remains contentious. Additionally, the evidence for plant phylogenetics has largely been based on a readily amplified set of plastid gene regions, and more recently whole plastid genomes, leaving open the possibility of undiscovered conflicts between plastid and nuclear genomes (Stull et al., 2020).

The Dipsacales provide a case in point. Much effort has been devoted over the past 40 years to inferring the phylogeny of this clade of ca. 1100 species of campanulid angiosperms, many of which are horticulturally important and ecologically significant in mesic forests around the northern hemisphere (Fig. 1). Morphological data sets were analyzed at first (e.g., Donoghue, 1983; Wagenitz and Liang, 1984; Judd et al., 1994; Backlund and Donoghue, 1996), but were soon complimented by studies of cpDNA (in some cases using

nuclear ribosomal ITS and mitochondrial sequences) that also supported two major clades within the Dipsacales, the Adoxaceae and the Caprifoliaceae s.l., including the traditional Morinaceae, Valerianaceae, and Dipsacaceae (e.g., Donoghue et al., 1992; Pyck et al., 1999; Bell et al., 2001; Zhang et al., 2002a; Donoghue et al., 2001, 2003; Winkworth et al., 2008). Over the past few years, several major studies of the Dipsacales phylogeny based on whole plastid genomes have been published (Fan et al., 2018; Wang et al., 2020; Xiang et al., 2020). These largely support earlier conclusions, but with far more data. Meanwhile, studies focused on major clades within the Dipsacales, such as the Valerianaceae (Bell and Donoghue, 2005b; Bell et al., 2012), the Morinaceae (Bell and Donoghue, 2003), the Dipsacaceae (Avino et al., 2009; Carlson et al., 2009), *Lonicera* (Theis et al., 2008; Smith, 2009; Smith and Donoghue, 2008), and *Viburnum* (Clement et al., 2014; Spriggs et al., 2015; Eaton et al., 2017; Landis et al., 2020), have further clarified relationships.

Taken together, these studies support the following narrative. The first Dipsacales were woody plants with simple opposite leaves, and their flowers were sympetalous with five stamens alternating with five corolla lobes. Adoxaceae then probably evolved smaller, rotate corollas, and fleshy fruits with a reduced number of seeds enclosed in endocarps (Fig. 1A, E). Caprifoliaceae, in contrast, evolved long tubular, bilaterally symmetrical corollas (with several reversals to radial symmetry), a long style with a capitate stigma, and nectaries of small hairs at the base of the corolla tube (Donoghue et al., 2003; Howarth and Donoghue, 2005; Howarth et al., 2011; Fig. 1B–D, J, M). Within the Caprifoliaceae, the large Linnina clade, including the Linnaeeae, and the traditional Morinaceae, Valerianaceae, and Dipsacaceae, is marked by a reduction from five to four stamens (Donoghue et al., 2003; Fig. 1J, M). In this clade, there was also a reduction to one functional carpel (by the abortion of two of the three carpels) and the production of an achene fruit (Fig. 1H). Within this clade, supernumerary inflorescence bracts were variously modified in relation to dispersal, and an epicalyx closely enveloping the ovary probably evolved independently several times (Roels and Smets, 1996; Donoghue et al., 2003).

Despite this progress, the relationships of several key lineages remain unresolved. The monotypic *Heptacodium* is a prime example. There is uncertainty as to whether *Heptacodium* is more closely related to the Caprifolieae (as supported by most cpDNA analyses; Fan et al., 2018; Pyck and Smets, 2000; Wang et al., 2020; Xiang et al., 2020) or to the large Linnina clade that includes the major herbaceous lineages (as supported by several conspicuous morphological characters; Zhang et al., 2002b; Jacobs et al., 2011). This lack of clarity has impeded our interpretation of a number of key morphological features. For example, we are unable to determine whether achene fruits, which have been implicated as triggers of diversification (Beaulieu and Donoghue, 2013), evolved once or more than once within the Dipsacales. Likewise, how many times did enlarged wing-like calyx lobes evolve?

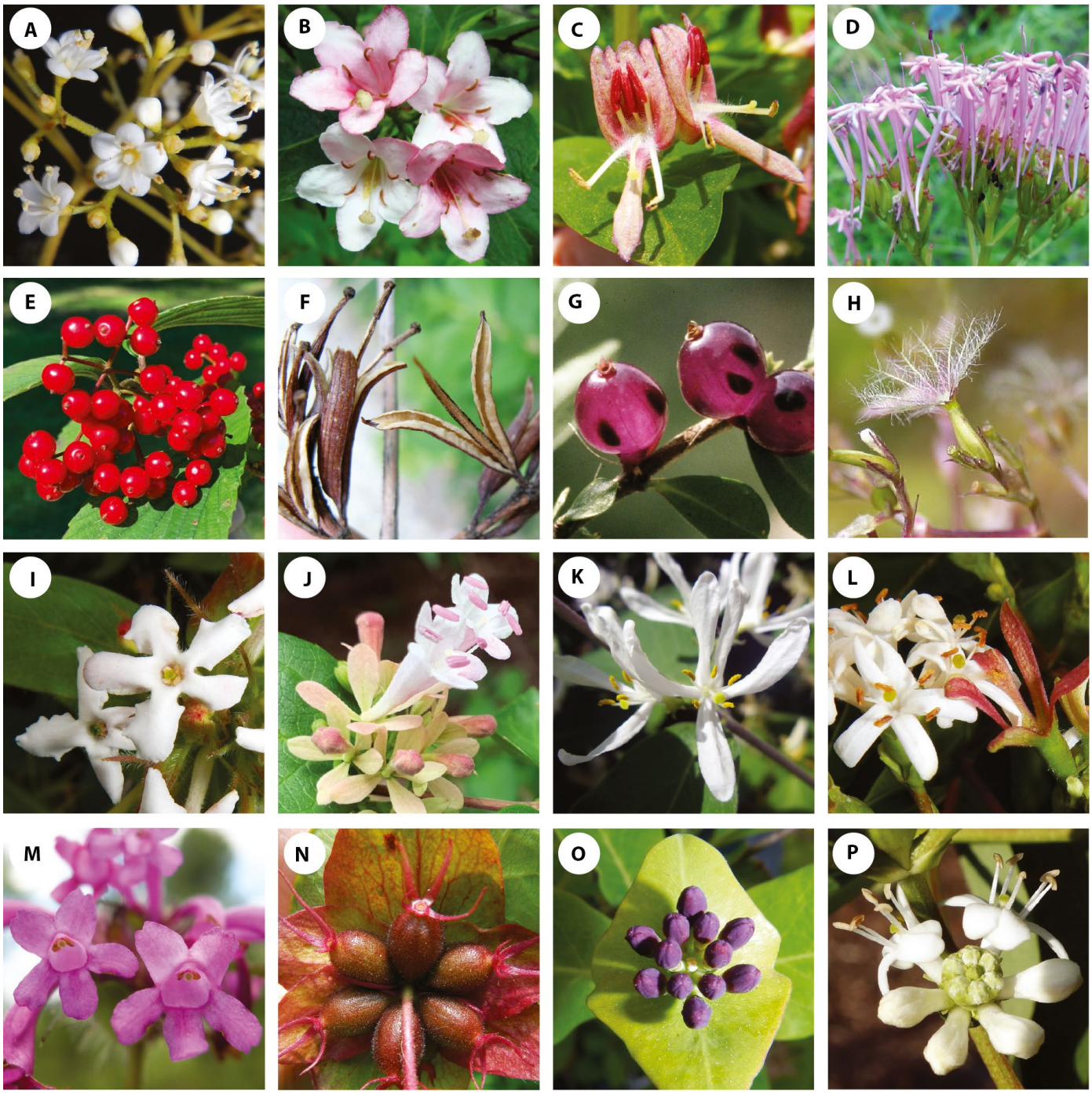
Zabelia provides a second example. It is unclear whether it belongs with the Linnaeeae, where it has traditionally been placed based on morphological characters (Jacobs et al., 2011; Wang et al., 2015), or with the Morinoideae, as supported by most analyses of plastid data (Jacobs et al., 2010; Fan et al., 2018; Wang et al., 2020; Xiang et al., 2020). This uncertainty prevents a full understanding of the evolution of enlarged calyx lobes, corolla form, inclusion of stamens within the corolla tube, the epicalyx, and the herbaceous habit, and how these traits might relate to diversification rate (Smith and Donoghue, 2008; Beaulieu et al., 2013).

Previous studies have also failed to convincingly resolve relationships within the Caprifolieae among *Lonicera*, *Leycesteria*, *Symphoricarpos*, and *Triosteum*, and, consequently, we have been unable to determine the direction of evolution of inflorescence, flower, and fruit traits in this clade. Finally, recent plastid-based analyses have been shifting our understanding of relationships within the Linnaeeae (Fan et al., 2018; Wang et al., 2020; Xiang et al., 2020), especially regarding the position of the newly recognized *Vesalea* and *Diabelia* (Landrein, 2010). The placement of taxa within Linnaeeae will have consequences for our interpretation of the evolution of inflorescence architecture, nectaries, and enlarged bracts and calyx lobes.

Finally, the timing of the Dipsacales radiation remains unclear, with multiple studies yielding quite different age estimates (cf. Bell and Donoghue, 2005; Fan et al., 2018; Wang et al., 2020). This uncertainty stems in part from calibrations using only a few fossils, some of which have not been critically evaluated or correctly placed in the tree. And, again, prior estimates have been based almost entirely on chloroplast sequences.

The development of the Angiosperms353 probe set (Johnson et al., 2019) provides a new opportunity to sequence and assemble

FIGURE 1. Dipsacales flower and fruit diversity. (A) *Viburnum triphyllum* (Adoxaceae) with small, rotate, radially symmetrical flowers, each with a short style and trilobed stigma. (B) *Weigela florida* (Diervillioideae, Caprifoliaceae) with large, tubular, bilaterally symmetrical flowers, each with an elongate style and capitate stigma. (C) *Lonicera alpigena* (Caprifolioideae, Caprifoliaceae) with bilaterally symmetrical flowers, here with four corolla lobes pointing up and one down (4:1). (D) *Centranthus angustifolius* (Valerianoideae, Caprifoliaceae); each flower with one stamen and a ventral nectar spur. (E) *Viburnum setigerum* (Adoxaceae) showing fleshy drupe fruits, each with one endocarp/seed within. (F) *Weigela florida* (Diervillioideae, Caprifoliaceae) showing septicidally dehiscent capsules with two carpels and a persistent septum/style. (G) *Lonicera ligustrina* (Caprifolioideae, Caprifoliaceae) with berry fruits, each with multiple seeds. (H) *Valeriana officinalis* (Valerianoideae, Caprifoliaceae) with an achene fruit with one seed and a pappus-like calyx (photograph cropped and courtesy of M. Bendel, CC license, https://creativecommons.org/licenses/by-sa/3.0/). (I) *Zabelia triflora* (Caprifoliaceae) showing a salverform corolla with stamens included. (J) *Abelia chinensis* (Linnaeeae, Caprifoliaceae) with expanded pink calyx lobes, funnelform corolla, and four exerted stamens. (K) *Lonicera morrowii* (Caprifolioideae, Caprifoliaceae) with five stamens and a corolla with two dorsal lobes, two lateral lobes, and one ventral lobe. (L) *Heptacodium miconioides* (Caprifolioideae, Caprifoliaceae) showing flowers with five stamens and a corolla with two dorsal lobes, two lateral lobes, and one ventral lobe, as well as an achene fruit topped by expanded red calyx lobes. (M) *Acanthocalyx delavayi* (Morinoideae, Caprifoliaceae) with salverform corollas and stamens included (photograph by X. Bo, Hengduan Biodiversity Project). (N) *Leycesteria formosa* (Caprifolioideae, Caprifoliaceae) showing a portion of a polytelic inflorescence with six inferior ovaries and subtending bracts, each set with two axillary groups of three ovaries. (O) *Lonicera etrusca* (Caprifolioideae, Caprifoliaceae) with a polytelic inflorescence showing two sets of six flower buds, each set with two axillary groups of three flowers. (P) *Heptacodium miconioides* (Caprifolioideae, Caprifoliaceae) with an inflorescence structure similar to *Leycesteria* (N) and *Lonicera* (O) (photograph of Arnold Arboretum of Harvard University accession 181-96*C courtesy of William (Ned) Friedman). Note that the central bud is not a flower, but an unexpanded set of flowers. All photographs by M. J. Donoghue, unless otherwise noted.



a large number of putatively single-copy nuclear genes. These new data hold the promise of finally resolving the remaining phylogenetic problems noted above, and they should allow us to identify concordant or discordant patterns of evolution in nuclear versus plastid DNA. Here we present analyses of the Dipsacales phylogeny based on our use of the Angiosperms353 probe set. Our primary objective was to improve our understanding of the evolution of these plants, but, as early adopters, we also wanted to more generally explore the usefulness of this Hyb-Seq approach and, thereby, orient further methodological developments along these lines. We begin by describing our efforts to assemble an Angiosperms353 data set for the Dipsacales, and then present a set of phylogenetic analyses of the taxa and genes that yielded data of sufficient quality. These analyses highlight several regions in which individual gene trees show patterns of discordance. Because we were also able to obtain whole plastome sequences for many accessions from the off-target reads, we analyzed these data for comparison both with previous cpDNA findings and with the trees from our Angiosperms353 data set. Using both our nuclear and cpDNA trees, we also present new analyses of the absolute timing of the Dipsacales diversification, using an expanded set of carefully vetted fossils that now constrain the ages of most of the major clades. Finally, we discuss the implications

of our new phylogenetic and dating analyses for the evolution of a set of key morphological traits.

MATERIALS AND METHODS

Taxon sampling and nomenclature

We selected 94 Dipsacales species, representing 43 of the 44 commonly recognized genera (all but *Pseudoscabiosa*), and two outgroup taxa, *Ilex crenata* (Aquifoliaceae) and *Paracryphia alticola* (Paracryphiaceae), for a total of 96 samples (Appendix 1). Though we acknowledge recent taxonomic revision of *Valeriana* (specifically, *Valeriana lobata* = *V. crispa* and *Plectritis congesta* = *V. congesta*; Kutschker, 2011), we chose to apply the names associated with previously published Dipsacales literature for continuity.

The names used here for major clades within the Dipsacales mainly follow the nomenclatural scheme and phylogenetic definitions of Donoghue et al. (2001), but with several modifications reflecting advances in phylogeny and the adoption of several names proposed since that time. Here, we recognized two sister clades within the Dipsacales, Adoxaceae and Caprifoliaceae. Note, that for continuity with prior literature, we here reject Viburnaceae as a substitute for Adoxaceae. Within Caprifoliaceae, we refer to five mutually exclusive clades with names ending in "-oideae": Caprifolioideae, Diervillioideae, Dipsacoideae, Morinoideae, and Valerianoideae. These names follow the usage of the Angiosperm Phylogeny website (Stevens, 2001), which reduces confusion associated with having taxon names ending in "-aceae" nested within one another (as in Donoghue et al., 2001). Further, we refer to well-supported subclades within the "-oideae" clades that have featured prominently in prior publications. Within Caprifolioideae, Caprifolieae includes *Leycesteria*, *Lonicera*, *Symphoricarpos*, and *Triosteum* but excludes *Heptacodium*. Within Dipsacoideae, Dipsaceae corresponds to the traditional Dipsacaceae but excludes *Triplostegia*. We also refer to two other major clades within Caprifoliaceae. Linnaeeae includes *Abelia*, *Diabelia*, *Dipelta*, *Kolkwitzia*, *Linnaea*, and *Vesalea*. And Linnina is the large clade that includes Linnaeeae, Morinoideae, Dipsacoideae, Valerianoideae, and *Zabelia*. In no cases do we associate these names with traditional taxonomic ranks (family, subfamily, etc.). Forthcoming work will formally define and register these clade names following conventions outlined in the PhyloCode (Cantino and de Queiroz, 2020).

Our sampling strategy was designed to span the phylogenetic diversity of the Dipsacales while more deeply sampling two larger clades, *Viburnum* and *Lonicera*, in an effort to evaluate the efficacy of the Angiosperms353 probe set at multiple phylogenetic scales. We also targeted our sampling to address recalcitrant relationships within the Dipsacales including the placements of *Heptacodium* and *Zabelia*, and evolutionary relationships within the Caprifolieae and the Linnaeeae.

DNA extraction and library preparation

DNA was extracted using a modified CTAB protocol (Doyle and Doyle, 1987) mostly from silica-dried leaf material. Sample concentration was assessed using a Qubit 2.0 fluorometer (Life Technologies, Carlsbad, CA, USA), and, when necessary, samples were re-concentrated to meet the minimum recommended target mass of 100 ng of genomic DNA in a 52.5-μL reaction volume (myBaits Hybridization Capture for Targeted NGS version 4.01, Daicel Arbor Biosciences, Ann Arbor, MI, USA). One sample (*Bassecoia siamensis*) had less than 100 ng of DNA. To create more uniform fragment lengths appropriate for sequencing, we sheared samples using a Covaris M220 Focused-ultrasonicator (Covaris, Woburn, MA, USA) following the manufacturer's protocol for 550-bp insert size. Fragment lengths were assessed using a 2100 Bioanalyzer Instrument (Agilent Technologies, Santa Clara, CA, USA). We size-selected fragments via a SPRI-bead cleanup with Ampure XP magnetic beads (Beckman Coulter Life Sciences, Indianapolis, IN, USA). Finally, end repair and A-tailing was completed using KAPA End Repair & A-Tailing Buffer and Enzyme (Roche, Basel, Switzerland) following the manufacturer's instructions.

Libraries were constructed using a KAPA Hyper Prep Kit with PCR Library Amplification/Illumina series KR0961 version 5.16 (Roche, Basel, Switzerland) following the manufacturer's instructions, but with 1/3 reaction volumes at all steps. After adapter ligation, samples were amplified for 12 cycles with a target of 100 ng DNA for hybridization; however, four taxa (*Bassecoia siamensis*, *Pterocephalus strictus*, *Symphoricarpos sinensis*, and *Valeriana lobata*) required an additional four cycles (16 cycles total). After amplification, we combined samples into eight equimolar hybridization reaction pools. To minimize possible enrichment biases due to shared genomic variation (e.g., GC content, locus dropout), we pooled our taxa based on presumed phylogenetic relationships. We then proceeded with hybridization capture using the myBaits® Target Capture Kit "Angiosperm 353 version 1" (product #308108, Daicel Arbor Biosciences, Ann Arbor, Michigan, USA) following the manufacturer's protocol (Hybridization Capture for Targeted NGS version 4.01), but with 1/4 baits volume (1.375 uL instead of 5.5 uL) during hybridization. Hybridization reactions were held at 65°C for 17.5 h. The resulting enriched products were amplified using KAPA HiFi 2X HotStart ReadyMix for 12 cycles. Final PCR product quality was assessed using a 2100 Bioanalyzer Instrument and sequenced on an Illumina MiSeq with version 3 (600-cycle) chemistry (Illumina, San Diego, CA, USA) in the DNA Discovery Center at the Field Museum of Natural History (Chicago, IL, USA).

Data cleaning and targeted loci assembly, filtering, and alignment

Raw sequencing reads were demultiplexed in BaseSpace (Illumina) and Illumina adapters were removed (raw reads available on NCBI SRA BioProject PRJNA723358). The raw reads were quality trimmed using Trimmomatic version 0.38 (Bolger et al., 2014) with a quality cutoff of 15 in a 4-bp sliding window. Leading and trailing bases below a quality score of three were also trimmed, and sequences shorter than 50 bp were discarded.

To assemble the 353 loci targeted, we used a two-step assembly approach with each step using the hybrid read mapping and de novo assembly implemented in HybPiper version 1.3.1 (Johnson et al., 2016). In the first step, we used the original Angiosperms353 probe set as the target sequence file (Johnson et al., 2019). In the second step, we expanded the target sequence file using the longest sequence assembled per locus from a sample within each major clade. These samples were *Acanthocalyx nepalensis*, *Diervilla sessilifolia*, *Heptacodium miconioides*, *Ilex crenata*, *Linnaea borealis*, *Lonicera*

caerulea, *Patrinia triloba*, *Triplostegia glandulifera*, *Viburnum vernicosum*, and *Zabelia dielsii*. Using lineage-specific target sequences has been shown to improve locus recovery (Murphy et al., 2020; McLay et al., 2021). Each HybPiper run used default parameters with BWA version 0.7.17 (Li and Durbin, 2009) as the mapper and SPAdes version 3.13.0 (Bankevich et al., 2012) as the de novo assembler.

Our filtering strategy sought to remove loci present in just a few taxa as well as remove taxa with poor locus recovery, but without greatly reducing the locus and sample pool overall. Alignments and gene trees were built using MAFFT version 7.149b (Katoh et al., 2002; Katoh and Standley, 2013) and RAxML version 8.2.12 (Stamatakis, 2014), respectively. We inferred and visually inspected trees from all assembled sequences for each exonic locus flagged as paralogous by "paralog_investigator.py" from HybPiper. Loci with paralogs emerging from shallow, tip-specific duplications were pruned down to one arbitrary representative per sample while maintaining relationships to their closest orthologs. Loci with paralogs arising from deeper duplication events were removed altogether from the locus pool. Separating alignments into orthologous subsets of these loci would have produced subsets with uneven taxon sampling due to variable paralog recovery. Off-target flanking regions were then assembled for this paralog-filtered set of loci using "intronerate.py" from HybPiper. We generated two sets of loci from these assemblies: exon-only and exon+flanking (also referred to as supercontig) regions (data sets available on GitHub: https://github.com/aaklee/DipsacalesHybSeq).

Further locus filtering consisted primarily of tree-based methods to refine alignments and remove spurious sequences, modeled after the homolog tree building step of the Yang and Smith (2014) orthology assessment pipeline. This approach allowed us to retain more sequence data for each locus, especially for samples missing data in off-target flanking regions. We chose this approach over site-based trimming approaches, such as GUIDANCE (Landan and Graur, 2008; Sela et al., 2015) or trimAL (Capella-Gutiérrez et al., 2009) because species tree inference can be adversely affected by site-based trimming (Tan et al., 2015). In preliminary testing, we found that trimming with more relaxed parameters removed too many potentially informative sites depending on site conservation and the amount and distribution of missing data. In contrast, more conservative site-based trimming mainly reduced computational analysis times, which is analogous to simply trimming sites based on missing data (Salichos and Rokas, 2013; Tan et al., 2015; Steenwyck et al., 2020).

In parallel analyses, exon-only and exon+flanking loci were preliminarily aligned with PASTA version 1.8.5 (Mirarab et al., 2015) using default parameters and MAFFT, Opal version 2.1.3 (Wheeler and Kececioglu, 2007), and FastTree version 2.1.7 (Price et al., 2010). PASTA uses an initial guide tree and a divide-and-conquer strategy to refine the alignment between groups of closely related sequences. Sites with >95% missing data were removed from alignments using Phyutility version 2.2.6 (Smith and Dunn, 2008), and preliminary gene trees for each locus were built using RAxML with 100 bootstrap replicates under a GTR+GAMMA model. Outlier long branches that changed the longest distance between tips within a gene tree by more than 10% were identified and removed using TreeShrink version 1.3.8b (Mai and Mirarab, 2018). Finally, sites with >70% missing data were removed from the alignments using Phyutility. Custom scripts for this pipeline are available as part of the aforementioned repository on GitHub.

Species tree inference and concordance analyses from targeted loci

Species trees were inferred using concatenated maximum likelihood (ML) and coalescent methods for both the exon-only and exon+flanking region matrices. In our ML analyses, alignments were concatenated using AMAS version 0.98 (Borowiec, 2016) and partitioned by locus. Trees were constructed in RAxML with 100 rapid bootstrap replicates under a GTR+GAMMA model applied to all partitions. Concatenated alignment matrix statistics were summarized using AMAS. RAxML was also used to estimate final gene trees for each locus, again with 100 rapid bootstrap replicates under a GTR+GAMMA model for input into ASTRAL-III version 5.7.3 (Zhang et al., 2018). Poorly supported nodes within gene trees (ML bootstrap support < 33) were collapsed into polytomies using DendroPy version 4.4.0 (Sukumaran and Holder, 2010) to reduce the influence of uncertain relationships in gene trees on species tree estimation. Node support for ASTRAL trees was estimated as local posterior probability (LPP).

We assessed concordance between gene trees and ASTRAL species trees using PhyParts (Smith et al., 2015), which calculates the number of bipartitions in rooted gene trees that are in concordance or conflict with the rooted species tree as well as ASTRAL quartet support. All trees were rooted by *Ilex crenata* and *Paracryphia alticola*, when possible, but gene trees that did not contain either species were rooted based on previous Dipsacales topologies (e.g., Donoghue et al., 2003). We also used the ASTRAL-III polytomy test (Sayyari and Mirarab, 2018) to assess concordance, which uses quartets from gene tree topologies to test the null hypothesis that all three nearest-neighbor interchanges at a given node are equally frequent.

Finally, we used SNAQ within PhyloNetworks version 0.11.0 (Solís-Lemus and Ané, 2016) in Julia version 1.5.3 to assess the likelihoods of strictly bifurcating versus network, or hybrid, topologies (Appendix S1). We first reduced our exon+flanking gene trees to those that contained the 13 most data-rich taxa representing each major clade and pruned all but these taxa from the gene trees. We chose to use exon+flanking rather than exon-only gene trees because these trees were generally more resolved and contained greater numbers of parsimony informative sites (see Results). Quartet frequencies from the pruned gene trees and the ASTRAL exon+flanking region species tree were used in an initial SNAQ analysis to estimate the likelihood of the strictly bifurcating topology. The resulting network with zero hybridization events was then used in subsequent network estimations that allowed for one or two hybridization events.

Plastid tree inference

Plastid genomes were assembled from off-target reads from the Hyb-Seq data. First, plastid reads were filtered from the raw data by performing a reference-based assembly in Bowtie 2 (Langmead and Salzberg, 2012; Langmead et al., 2018) against plastid genomes from closely related taxa (Appendix S2). Then, using the filtered plastid reads, we conducted a reference-based assembly in Bowtie 2 using a closely related reference sequence (Appendix S2), and generated a consensus sequence of the assembly in Geneious R9 (Biomatters Ltd, Auckland, New Zealand). The percentage of the plastid genome assembled as compared to the reference is reported in Appendix S2.

To extract genes and intergenic regions, first, MAFFT was used to align plastid assemblies to an annotated reference (Appendix S2). Annotations were transferred to the assembled genomes and custom

scripts (available on GitHub) were used to identify coordinates for exon-intergenic spacer and plastid subunit boundaries to ultimately extract genes and intergenic spacer regions separately. Intergenic spacer regions spanning subunit boundaries were not extracted. Alignments of genic and intergenic regions were concatenated using AMAS.

Plastid alignments were generated in a number of ways and involved either whole genome or gene region alignments. First, whole plastid genomes were aligned using MAFFT. Second, we took a clade-based approach that identified and aligned species assigned to one of eight Dipsacales clades or the outgroup using MAFFT (Appendix S2). The merge function in MAFFT was then used to align all nine alignments to each other preserving features of the original clade alignments. Finally, to overcome any potential issues with gene rearrangement among plastid genomes, we analyzed concatenated matrices of exons and intergenic spacer regions, separately and in combination, in partitioned analyses. All plastid trees were reconstructed using RAxML under a GTR+GAMMA model of sequence evolution and 500 bootstrap replicates.

Divergence time estimation

Divergence times were separately estimated using both nuclear and plastid data sets. The plastid tree was pruned to have the same taxon set as the nuclear tree. However, unlike the nuclear tree, *Viburnum coriaceum*, *V. vernicosum*, and *V. cinnamomifolium* could not be included in the plastid tree because sufficient plastid scaffolds could not be assembled from the off-target reads. They were replaced by three other recovered and closely related species (*V. sargentii*, *V. dentatum*, and *V. amplificatum*).

Since the complete data sets were too large for dating analyses, we selected genes using an approach that sorts loci based on a PCA of seven gene properties (Mongiardino Koch and Thompson, 2021; Mongiardino Koch, 2021 [Preprint]). The top set of loci have increased phylogenetic signal and low values of known sources of systematic bias. We used exon+flanking regions for the nuclear analyses and genic regions for the plastid analyses. We rooted gene trees using Phyx (Brown et al., 2017) and concatenated alignments using AMAS. Gene sorting was done on only those loci that had at least one of the two outgroups for rooting. The nuclear topology was fully constrained, including clades with low support, using the ASTRAL topology from the exon+flanking region sequence matrix, and the plastid topology from the clade-based alignment approach (alignments available on GitHub). Dating analyses were carried out in BEAST version 2.6.3 (Bouckaert et al., 2019) under a birth death tree prior, a GTR+GAMMA model of sequence evolution, and an uncorrelated relaxed log-normal clock (Drummond et al., 2006).

Previous Dipsacales dating analyses have used only two fossils, one of which was a *Viburnum* leaf fossil from the Cretaceous that was used by Wang et al. (2020). However, *Viburnum* leaf fossils from the Cretaceous and Paleocene have been reassigned to other clades in a series of studies (Manchester, 2002; Landis et al., 2020, and references therein). Hence, we conducted a thorough review of the Dipsacales fossil literature to identify additional credible fossils. This literature review led to the identification of 10 additional fossils suitable for calibration (Appendix S3), all of which we apply to the Dipsacales for the first time (Ozaki, 1980; Manchester et al., 2009, 2015; Liang et al. 2013; Pavlyutkin, 2015; Svetlana et al., 2019). We were able to place one to three fossils in all of the major clades except Morinoideae and Dipsacoideae. We also placed a secondary calibration on the Dipsacales stem that ranged from 49 to 121 Ma. The maximum bound was based on a previous estimate of the maximum possible age of the Dipsacales stem as inferred using BEAST (Ramírez-Barahona et al., 2020), and the minimum bound corresponded to the age of the oldest known Dipsacales fossil.

A log-normal prior probability distribution was assigned to all fossils and a uniform distribution was assigned to the secondary calibration. Each fossil calibration was roughly allowed to take a maximum age of ~110 Ma based on a previous maximum age estimate for the Dipsacales crown (Ramírez-Barahona et al., 2020). Minimum ages and maximum probability densities of fossil priors were set according to their respective ages, which were selected on the basis of available geochronological data (Appendix S3; Li and Xiao, 1980; Evanoff et al., 2001; Blanchard et al., 2016; Yabe et al., 2019). Details of fossil assessment, calibration dates, and placement justification are provided in Appendix S3.

Nuclear and plastid analyses were each run for 400 million generations, with parameters sampled every 10,000 generations. Log and tree files were combined from two runs for both nuclear and plastid analyses using LogCombiner version 2.6.3, with the first 10% of samples discarded as burnin; the resampling state frequency was set to 10,000 (Drummond and Rambaut, 2007). The effective sample size (>200), mixing, and convergence were determined using Tracer version 1.7.1 (Rambaut et al., 2018). We used TreeAnnotator version 1.10.4 to generate maximum clade credibility (MCC) trees (Drummond and Rambaut, 2007). Results were visualized in FigTree version 1.4.4 (Rambaut, 2012) and figures were produced using the ggtree package (Yu et al., 2017) in R (R Core Team, 2020).

RESULTS

Locus recovery

Expanding the target sequence file using lineage-specific sequences yielded slightly longer loci and more loci per sample, but these differences were not significant ($p > 0.05$; Appendix S4). Globally, loci had a median recovery length of 50.1% of their expected length (interquartile range [IQR]: 37.9–64.0%), and the median maximum recovery length was 106.0% (IQR: 91.6–116.1%). These metrics suggest that when coding sequences could be assembled, loci were generally assembled to half their expected length, with many surpassing the expected length and extending into intronic and extragenic flanking regions. Recovered loci generally spanned dozens of taxa (median: 47 taxa/locus, IQR: 31–57%), which translated into low rates of missing data across sequence matrices. Furthermore, all but two loci (target numbers 6430 and 6705) were recovered in at least one taxon, which confirmed that successful enrichment of nearly all Angiosperms353 loci is possible within the Dipsacales and relatives.

A total of 256 loci were flagged as paralogous by HybPiper, with a notable increase when using our expanded target recovery file (Appendix S4). The median number of putative paralogs within each sample was three, and most paralog warnings were restricted to Morinoideae (see Discussion) where paralogs were recovered for dozens of targeted loci (e.g., 134 in *Cryptothladia kokonorica* as the maximum flags for any sample in the data set; Appendix S4I). After pruning tip-specific duplicates, 224 of these loci were retained and the remaining 32 were excluded from phylogenetic analyses (Fig. 2). After removal of loci with deep paralogs and tree-based filtering, the locus pool was reduced from 351 total loci to 313 exonic and 308 exon+flanking regions (Appendix S5).

Although we recovered upward of hundreds of loci for dozens of taxa, many taxa had too little data to warrant inclusion in phylogenetic analyses. It was apparent that variation in locus recovery was linked to hybridization pools. Group differences in sequencing reads, enrichment, locus recovery, etc., among hybridization pools were assessed using SciPy version 1.0 (Virtanen et al., 2020), scikit-posthocs

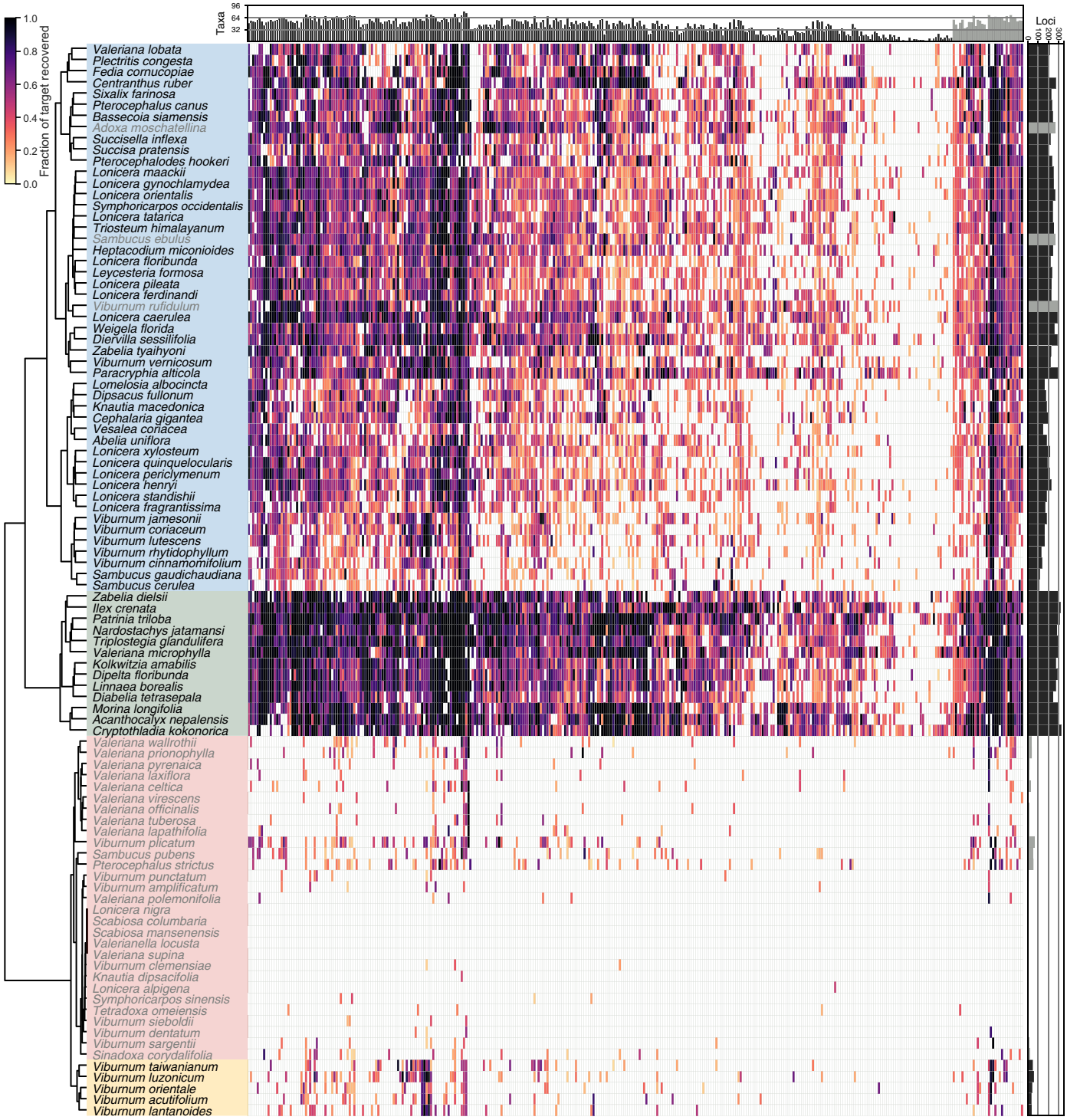


FIGURE 2. Heat map showing the fraction of the target exon length recovered for each locus (columns) for each taxon (rows). A cluster dendrogram to the left of the heatmap shows clustering of rows based on fraction target length recovered into four clusters: "high", green; "moderate", blue; "acceptable", yellow; and "low", red. The histogram above the heatmap shows the number of taxa that recovered each locus. The histogram to the right of the heatmap shows the number of loci recovered for each taxon. Loci and taxa in black were retained in the final analyses; those in light gray were dropped (see Methods). The order of loci has been rearranged relative to their order in the Angiosperms353 target file for visualization.

version 0.6.4 (Terpilowski, 2019), and Pingouin version 0.3.4 (Vallat, 2018) in Python version 3.5.7. When examining differences between library pools, we found that the numbers of recovered loci in samples from different pools were significantly different (Kruskal–Wallis test, $p < 0.0001$, df = 7, $H = 38.369$). These differences were primarily due to two pools (pools 2 and 3) with very poor recovery (Appendix S6C, F; Dunn's test $p < 0.001$), which largely corresponded to the taxa in the lowest recovery cluster (Fig. 2). While there were no significant differences between the number of reads that each pool received (KW test $p = 0.145$, df = 7, $H = 10.858$), pools 2 and 3 were significantly less enriched (KW $p < 0.0001$, df = 7, $H = 58.876$) than all others (Appendix S6B, E; Dunn's test $p < 0.01$ for all comparisons except pool 2 vs pool 4, $p < 0.05$). To test whether these differences could have been caused by variation in sample quality, we further tested for differences among pools in sample mass and concentration, but found the only significant difference to be that pool 7 samples generally had higher sample concentration than pool 2, and higher sample mass than most other pools (Appendix S7). Furthermore, regression analyses found no significant relationships between sample quality (mass and concentration) and the number of reads per sample (Appendix S8A, B), sample enrichment (Appendix S8C, D), or the number of recovered loci (Appendix S8G, H). However, there were significant positive relationships between enrichment and both sample mass and concentration, albeit with low explanatory power (adjusted $R^2_{\text{mass}} = 0.085$, $R^2_{\text{concentration}} = 0.069$, $p_{\text{slope}} < 0.01$; Appendix S8E, F). Together, we believe that these results imply that locus recovery using the Angiosperms353 targets was not strongly influenced by sample quality, despite a slight decrease in enrichment in lower-quality samples. The batch effects linked to these samples were likely due to benchtop human errors that occurred when setting up hybridization reactions.

Dipsacales species phylogeny

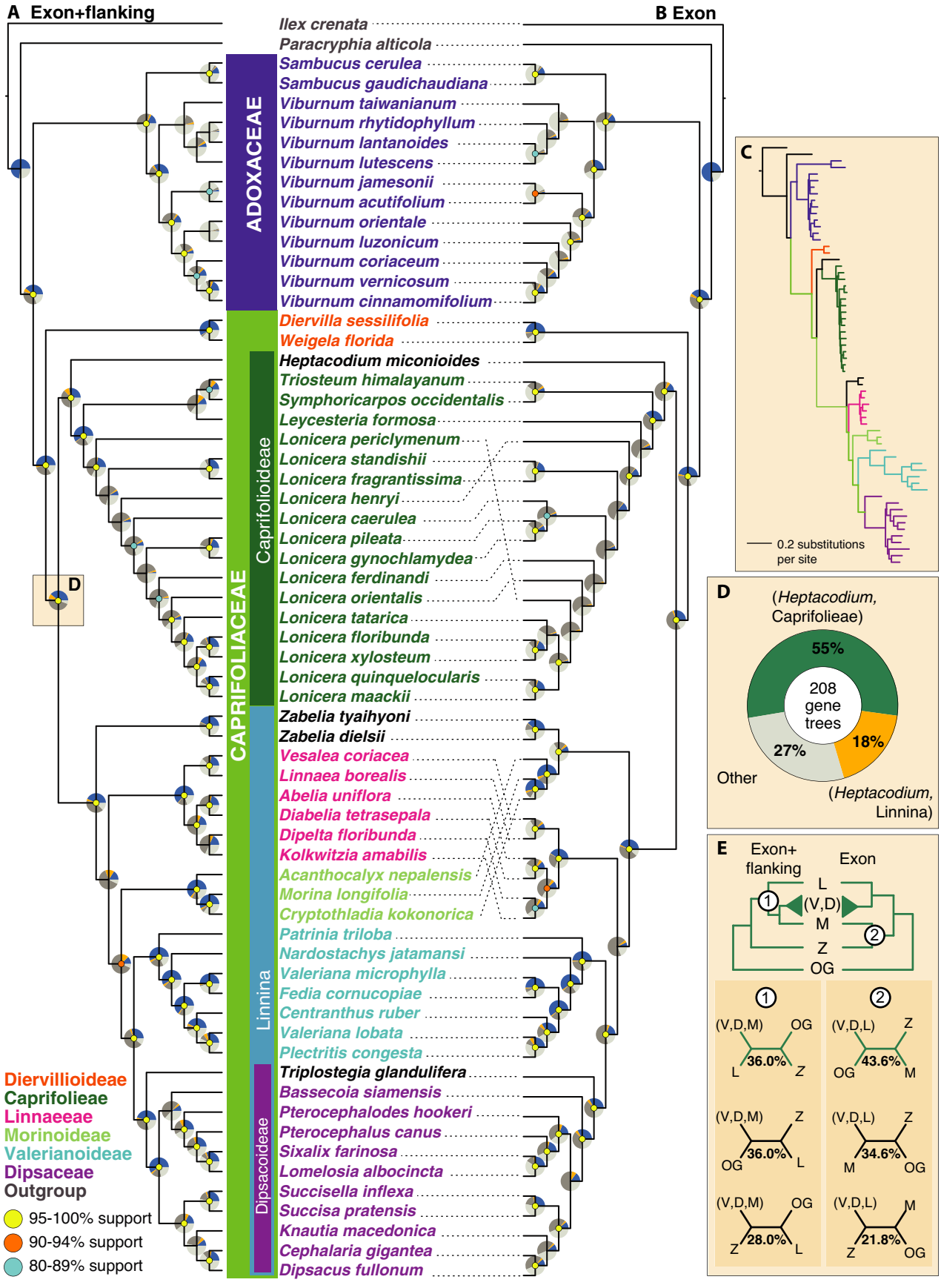
Initial inspection using the HybPiper accessory script hybpiper_stats.py revealed that locus recovery varied greatly across samples (Fig. 2). To set more objective criteria for taxon inclusion, we used Ward clustering of samples based on mean exon length recovered using the 353 target loci as features in SciPy version 1.0 (Virtanen et al., 2020). Ward clustering is a hierarchical, or agglomerative, clustering method that minimizes the within-cluster variance as clusters are iteratively merged. Ward clustering of samples and recovered coding sequences revealed three distinct clusters: (1) 13 taxa with many loci recovered and long median recovery lengths, (2) 49 taxa with slightly fewer loci and moderate recovery lengths, and (3) 34 taxa with few to no loci recovered and short recovery lengths (Fig. 2; Appendix S9). The number of loci with reads mapping was high across all taxa (median: 334, IQR: 310–345), but slightly bimodal enrichment resulted in highly bimodal distributions of recovered loci (Appendix S4B–E). The median number of recovered loci per taxon was 196 (IQR: 33–256; Appendix S4E), ranging from zero loci recovered in four taxa (*Lonicera nigra*, *Scabiosa mansenensis*, *Valeriana supina*, and *Valerianella locusta*) and a maximum of 326 loci recovered in *Cryptothladia kokonorica* (Appendix S4E). Recovery was biased downward by our lowest recovery cluster, and taxa within the high and moderate recovery clusters recovered medians of 295 and 225 loci, respectively (Appendix S9).

We retained samples in recovery clusters 1 and 2, but further investigated the deepest split within cluster 3 (Fig. 2). This cluster contained a subgroup of five samples with a few dozen loci (median = 27, IQR = 20–40) but low recovery lengths (median = 36.4%, IQR = 24.7–49.0%; Fig. 2). Although this subgroup was nested within our lowest recovery cluster, it contained our only tractable representatives of multiple members of *Viburnum* that were of particular interest for assessing the utility of Angiosperms353 kit for resolving closely related species. We therefore proceeded with a sequence matrix including 67 samples consisting of all members of clusters 1 and 2, and the five members of 3 (yellow subgroup; Fig. 2).

Preliminary phylogenetic analyses identified three taxa with moderate to high locus recovery (*Adoxa moschatellina*, *Sambucus ebulus*, and *Viburnum rufidulum*) that were consistently placed in aberrant positions (e.g., *A. moschatellina* nested within Dipsacoideae with high bootstrap and LPP support) in both nuclear and plastid analyses. Furthermore, branch lengths and alignments did not indicate that identical or nearly identical sequences were assembled for multiple samples. Nor did we assemble multiple copies of each locus that yielded consistent alternative placements, as would be expected if contamination occurred during library construction. We believe that these samples may have been contaminated or mislabeled prior to inclusion in the present study, and we removed these samples from further downstream analyses after confirming that their removal did not affect the relationships at nearby nodes. Thus, the final nuclear taxon sets included 64 taxa (Fig. 2).

Topologies inferred from the targeted nuclear loci were largely congruent regardless of method (ASTRAL or RAxML) or locus set (exon-only or exon+flanking). Exon+flanking region alignments included nearly triple the number of bases and double the proportion of parsimony informative sites compared to exon-only alignments (Appendix S5). Species trees generated generally recovered high LPP (>0.95, Fig. 3A) and high bootstrap support values (>80%, Appendix S10). All trees supported the monophyly of the Dipsacales and of most of the major clades within the Dipsacales, including the split between Adoxaceae and Caprifoliaceae (Fig.

FIGURE 3. Species tree of Dipsacales reconstructed using ASTRAL-III for (A) the exon+flanking region alignment with 308 loci and 376,948 bp (B) and the exon-only alignment with 313 loci and 142,825 bp representing 64 taxa. Clade membership is noted by taxon name color or bars running the length of the tree. Local posterior probabilities are represented by yellow (95–100%), orange (90–94%), and blue (80–89%) circles placed at nodes, and nodes without circles have clade support less than 80%. On each ASTRAL tree, results of the concordance analysis are summarized with pie charts placed on each node. Specifically, blue represents the proportion of gene trees that support the shown species tree, yellow represents the proportion of trees that support the most common conflicting bipartition, dark gray supports other conflicting bipartitions, and light gray represents the proportion of gene trees uninformative at that node. The species tree generated from the exon+flanking region is largely congruent with the exon-only species tree except for notable differences within Caprifolieae and the placement of *Zabelia*. (C) RAxML tree of the exon+flanking region data showing branch lengths. Branch colors represent clades as indicated on the ASTRAL trees. (D) Proportion of the 208 exon+flanking region gene trees that support two hypotheses based on morphology for the placement of *Heptacodium*: a clade containing *Heptacodium* and Caprifolieae, or *Heptacodium* and Linnina. "Other" refers to the ambiguous placement of *Heptacodium* in the gene tree. (E) Proportion of quartets supporting alternative Linnina relationships as highlight by nodes 1 and 2 for both the exon+flanking and exon-only topologies. *Abbreviations:* D, Dipsacoideae; L, Linnaeeae; M, Morinoideae; OG, outgroup; V, Valerianoideae; *Z*, *Zabelia*.


A Exon+flanking
B Exon
Ilex crenata
Paracryphia alticola
Sambucus cerulea
Sambucus gaudichaudiana
Viburnum taiwanianum
Viburnum rhytidophyllum
Viburnum lantanoides
Viburnum lutescens
Viburnum jamesonii
Viburnum acutifolium
Viburnum orientale
Viburnum luzonicum
Viburnum coriaceum
Viburnum vernicosum
Viburnum cinnamomifolium
Diervilla sessilifolia
Weigela florida
Heptacodium miconioides
Triosteum himalayanum
Symphoricarpos occidentalis
Leycesteria formosa
Lonicera periclymenum
Lonicera standishii
Lonicera fragrantissima
Lonicera henryi
Lonicera caerulea
Lonicera pileata
Lonicera gynochlamydea
Lonicera ferdinandi
Lonicera orientalis
Lonicera tatarica
Lonicera floribunda
Lonicera xylosteum
Lonicera quinquelocularis
Lonicera maackii
Zabelia tyaihyoni
Zabelia dielsii
Vesalea coriacea
Linnaea borealis
Abelia uniflora
Diabelia tetrasepala
Dipelta floribunda
Kolkwitzia amabilis
Acanthocalyx nepalensis
Morina longifolia
Cryptothladia kokonorica
Patrinia triloba
Nardostachys jatamansi
Valeriana microphylla
Fedia cornucopiae
Centranthus ruber
Valeriana lobata
Plectritis congesta
Triplostegia glandulifera
Bassecoia siamensis
Pterocephalodes hookeri
Pterocephalus canus
Sixalix farinosa
Lomelosia albocincta
Succisella inflexa
Succisa pratensis
Knautia macedonica
Cephalaria gigantea
Dipsacus fullonum
ADOXACEAE
CAPRIFOLIACEAE
Caprifolioideae
Linnina
Dipsacoideae
D
Diervillioideae
Caprifolieae
Linnaeeae
Morinoideae
Valerianoideae
Dipsaceae
Outgroup
95-100% support
90-94% support
80-89% support
C
0.2 substitutions per site
D
(Heptacodium, Caprifolieae)
55%
208 gene trees
18%
27%
Other
(Heptacodium, Linnina)
E
Exon+ flanking
Exon
L
(V,D)
M
Z
OG
1
2
(V,D,M) OG
36.0%
L Z
(V,D,L) Z
43.6%
OG M
(V,D,M) Z
36.0%
OG L
(V,D,L) Z
34.6%
M OG
(V,D,M) OG
28.0%
Z L
(V,D,L) M
21.8%
Z OG

3). Within Caprifoliaceae, Diervillioideae was sister to the rest of the clade; *Heptacodium* was recovered as sister to Caprifolieae, and this clade, the Caprifolioideae, was sister to the Linnina clade (Fig. 3). Within Linnina, we found that Linnaeeae, Morinoideae, Valerianoideae, and Dipsacoideae were each monophyletic and that Valerianoideae + Dipsacoideae consistently formed a clade. *Patrinia* and *Nardostachys* formed a grade relative to the core Valerianoideae, and *Triplostegia* was strongly supported as sister to the Dipsaceae within Dipsacoideae (Fig. 3).

Topological differences within the major clades were uncovered between the ASTRAL exon+flanking (Fig. 3A) and exon-only (Fig. 3B) approaches to tree reconstruction. There was a minor conflict in Dipsacoideae involving the placement of *Bassecoia*. Additionally, we recovered topological differences within the Caprifolieae and differences regarding the relationship of *Zabelia* to either the rest of Linnina or to Morinoideae (see concordance analysis below). Finally, there were many poorly supported clades and some conflicts regarding species relationships within *Lonicera* and *Viburnum,* and branch lengths within both of these clades were notably short (Fig. 3C; Appendix S10).

To recover the plastid tree, we explored a number of different alignment schemes and here present the results of the clade-based alignment (Fig. 4; Appendix S11). We included 81 taxa in these analyses and were able to retain taxa such as *Sinadoxa* among others that were lost altogether in the filtering processes associated with the nuclear data. Many consistent relationships were recovered regardless of alignment strategy such as *Heptacodium* sister to Caprifolieae, *Zabelia* sister to Morinoideae, and *Triplostegia* sister to Dipsaceae. Few notable topological differences were recovered among trees resulting from the various alignment strategies. In particular, relationships within Caprifolieae differed among the various alignment schemes with notably short branches separating *Leycesteria*, *Symphoricarpos*, and *Triosteum* (Appendix S11.3).

The tree topologies reconstructed from nuclear and plastid data were similar in many respects. Importantly, all data and analyses agreed on the placement of *Heptacodium* as sister to Caprifolieae. However, the plastid and nuclear topologies differed on relationships within Caprifolieae. In the nuclear tree, *Symphoricarpos* and *Triosteum* were consistently linked, while the position of *Leycesteria* varied (Fig. 3, Appendix S10). This topology was not recovered by plastid data. In contrast, the plastid analyses either supported *Leycesteria* + *Triosteum* or *Leycesteria* + *Symphoricarpos*. (Fig. 4; Appendix S11). Within *Lonicera*, there were multiple consistently resolved species pairs but poor resolution of relationships among those pairs in both plastid and nuclear analyses. In our nuclear analyses, *Zabelia* was recovered either as sister to the rest of the entire Linnina clade (Fig. 3A; Appendix S10) or as sister to Morinoideae (Fig. 3B). However, in our plastid analyses (Fig. 4), *Zabelia* was linked with Morinoideae. Within Linnaeeae, the plastid data recovered a phylogenetic grade of *Abelia* and *Kolkwitzia* relative to *Dipelta* + *Diabelia*, while the nuclear data supported *Abelia* as sister to *Diabelia*. Finally, many differences were recovered within the Dipsacoideae. For example, the exon+flanking data supported *Bassecoia* and *Pterocephalodes*, forming a grade with respect to *Pterocephalus*, *Sixalix* and *Lomelosia*, while in plastid analyses, *Bassecoia* was strongly supported as sister to *Pterocephalodes*.

Within the better sampled *Viburnum* and *Lonicera* clades, a number of relationships were poorly supported and differed depending on inference approach and data type. In *Viburnum*, our nuclear analyses recovered the basal split for these species found in RAD-seq analyses (Eaton et al., 2017; Landis et al., 2020). However, we note that the placement of *V. cinnamomifolium* of the *Tinus* clade is at odds with all prior studies, which place it instead as sister to *V. vernicosum* in Fig. 3 (Clement et al. 2014; Landis et al., 2020). In our plastid analyses, we were able to include *V. clemensiae*, and it emerged as sister to the rest of *Viburnum*, as in most prior plastid studies (e.g., Clement et al., 2014). We were also able to include *Sinadoxa*, which we recovered as sister to *Sambucus*, consistent with previous studies (Eriksson and Donoghue, 1997; Donoghue et al., 2001). Within *Lonicera*, there were multiple consistently resolved species pairs but poor resolution of relationships among those pairs. *Lonicera periclymenum* was recovered as sister to the rest of the *Lonicera* (reflecting the two major subclades of *Lonicera*) in our ASTRAL exon+flanking (Fig. 3A) and plastid analyses (Fig. 4), consistent with previous studies (Theis et al., 2008; Smith, 2009), but this was not the case in our remaining nuclear trees.

Concordance analysis

The results of our concordance analyses were generally congruent with bootstrap support, LPP values, and quartet scores and provided additional details on the genomic conflicts underlying poorly supported nodes (Fig. 3). Bootstrap support values tended to be high where concordance was also high and also at nodes where concordance was low but no alternative topology was frequent among gene trees (e.g., *Sambucus* and Adoxaceae crown nodes; Appendix S10). Nodes in the exon-only and exon+flanking species trees were informed by similar numbers of gene trees ($t = 0.453$, df = 122, $p = 0.651$) with similar distributions of concordance ($t = 1.506$, df = 122, $p = 0.135$) and conflict ($t = -0.954$, df = 122, $p = 0.342$). Polytomy tests reiterated these findings and often did not reject the null hypothesis that all possible bifurcations were equally supported by gene tree quartets at low supported node.

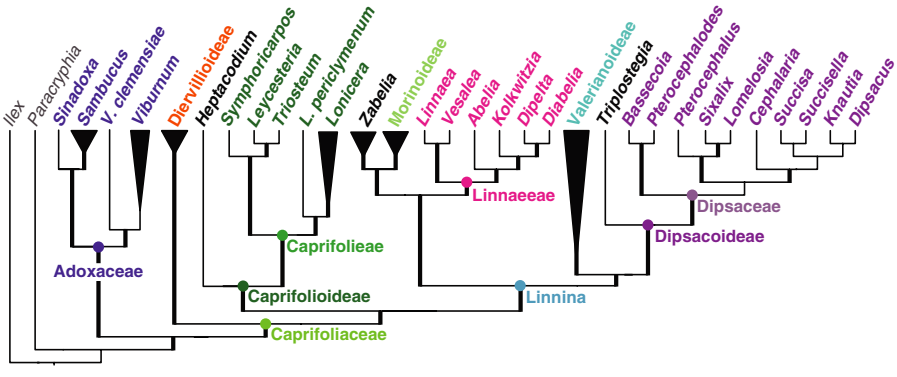


FIGURE 4. Cladogram summarizing the evolutionary relationships of 81 Dipsacales taxa as reconstructed from whole plastid genomes aligned using a clade-based alignment strategy. Clades supported by bootstrap support greater than 90% are indicated by thickened branches; all other clades had less that 70% bootstrap support. Major clades are either represented as tips or by circles labeled on nodes within the tree.

Across the exon+flanking region tree, nodes were informed by a median of 186 gene trees (IQR = 138–224), but relationships within the Adoxaceae were informed by significantly less (median = 86, IQR=43–127; $t = 6.91$, df = 60, $p < 0.0001$), and as few as 14. The relatively low number of gene trees bearing on relationships within the Adoxaceae likely explain why nodes were poorly to moderately supported in our ML species tree (Appendix S10). Given that previous reduced representation phylogenomic studies using restriction-site-associated DNA sequencing (RAD-seq) data in the Adoxaceae have provided much greater resolution (e.g., Landis et al., 2020), support for relationships within the Adoxaceae may greatly increase with target recovery in this area of the tree. The lack of gene trees and concordance within the Adoxaceae sharply contrasted with the Valerianoideae and Morinoideae, where locus recovery was high and gene trees were overwhelmingly concordant.

Concordance was not high in all clades with high locus recovery. In the Caprifolieae, nodes were generally well informed, but concordance was almost uniformly low. With the exception of a few pairs of sister species within *Lonicera*, conflict was the dominant signal throughout. We found high support and concordance for a monophyletic Caprifolieae and for a monophyletic *Lonicera*, but lower concordance for relationships among *Leycesteria*, *Lonicera*, *Symphoricarpos*, and *Triosteum*, and nearly ubiquitous conflict within *Lonicera*. Only 87 of 224 exon+flanking region gene trees (39%) supported a monophyletic Caprifolioideae (including *Heptacodium*). However, concordance analysis does not assess the magnitude of disparity between a given gene tree and the species tree topology, and it was possible that alternative topologies did not significantly violate the placement of *Heptacodium* as sister to the Caprifolieae. To better assess possible placements of *Heptacodium*, we tabulated the number of gene trees that supported the two hypothesized placements of *Heptacodium*: (1) sister to the Caprifolieae, or (2) a member of Linnina. Of the 208 exon+flanking region gene trees that contained *Heptacodium* and at least one member each of the Caprifolieae and Linnina, a clade containing *Heptacodium* and Caprifolieae was three times as frequent (114 gene trees) as a clade containing *Heptacodium* and Linnina (38 gene trees) (Fig. 3D). These ratios were very similar to quartet support for this node in the ASTRAL exon+flanking region tree: 59% of quartets support *Heptacodium* + Caprifolieae and only 22% support *Heptacodium* + Linnina (Appendix S12). No introgression events involving *Heptacodium* were proposed by SNAQ, and, more generally, allowing for network topologies (vis-a-vis strictly bifurcating) produced minimal likelihood gains (Appendix S1). We therefore did not find evidence supporting a hybrid origin of *Heptacodium*, or any other lineages, in our analyses. Quartet support was relatively equal for alternative relationships among Caprifolieae genera, mirroring the similar numbers of gene trees supporting the species tree and best alternative topologies (Appendix S12).

Concordance analysis also highlighted the conflicting signals between major Linnina clades: Dipsacoideae, Linnaeeae, Morinoideae, Valerianoideae, and *Zabelia*. There was high concordance among gene trees supporting the reciprocal monophyly of each of these clades, as well as the sister relationship of Valerianoideae + Dipsacoideae, but moderate to strong conflict along the backbone of Linnina (Fig. 3; Appendix S12). Parsing discordant gene trees showed only moderate support for one alternative topology, which placed Morinoideae sister to *Zabelia*, in the exon-only topology (Fig. 3). The exon+flanking species tree concordance analysis showed 49 gene trees supporting the species tree topology and 46 gene trees supporting the best alternative, which placed Morinoideae sister to *Zabelia* (Fig. 3A; Appendix S12). Thus, there was greater signal for the *Zabelia* + Morinoideae relationship in the exon+flanking region gene trees than in the exon-only data (33 gene trees; Fig. 3A, B).

Analysis of recovered quartets from the exon+flanking region gene trees showed nearly equal support for *Zabelia* as (1) sister to all other Linnina taxa (as in the species tree; 36.0%), (2) as sister to Linnaeeae (36.0%), and (3) as sister to Valerianoideae + Dipsacoideae + Morinoideae (28.0%; Fig. 3E node 1; Appendix S12). The exon-only quartet analysis was slightly skewed toward the exon-only species tree topology (43.6%), which supported the sister relationship *Zabelia* + Morinoideae, but the main alternative relationship of *Zabelia* as sister to the rest of Linnina (34.6%), as in the exon+flanking region species tree, received substantial quartet support (Fig. 3E node 2; Appendix S12). Furthermore, if we assumed that Morinoideae and *Zabelia* were sister, there was still ambiguity about the clade sister to core Linnina; the species tree relationship Morinoideae + *Zabelia* (35.0%) was only marginally more frequent than Valerianoideae + Dipsacoideae (33.7%) and Linnaeeae (31.1%) (Appendix S12).

Divergence time estimation

To select loci for divergence time estimation with nuclear and plastid data, gene sorting was conducted based on six gene properties (Appendix S13.2, S13.3). For the nuclear analysis, we initially had 308 exon+flanking regions. Of these, 296 had at least one outgroup taxon for rooting. Hence, we carried out gene sorting on this smaller subset of gene trees with an outgroup. Since some top-sorted loci had poor taxon representation, we selected the top 10 loci that had at least 55 taxa. Similarly for the plastid analysis, we carried out gene sorting on the 96 (from a total of 101) genic regions that had at least one outgroup taxon. From among those, we selected the top 10 loci that were present across at least 61 taxa. The final nuclear matrix had 21,505 bp with 6.9% missing data, and the plastid matrix had 21,453 bp with 0.47% missing data, for 64 taxa each; missing data was calculated as the number of taxa missing per sub-setted locus.

Using the Angiosperms353 kit, we present the first dated Dipsacales phylogeny based on nuclear loci (Fig. 5). It is also the most densely sampled Dipsacales chronogram to date, with 64 species representing 36 genera, calibrated with 10 fossils representing all major clades except Morinoideae and Dipsacoideae. Our estimate for the Dipsacales crown was centered on ca. 107 Ma, and its two major clades, Adoxaceae and Caprifoliaceae, were inferred to have originated in the Late Cretaceous. Subsequent splitting of these lineages began in the Late Cretaceous-Palaeocene but most of the major lineages arose from the Eocene onward, and most diversification (here simply the number of splitting events in our tree during given time periods; Fig. 5) occurred within the last 30 Myr.

Our dated plastid tree (Appendix S13.1) was roughly concordant with the nuclear tree, but a few clades were found to have notably different ages (Fig. 5). Some difference in clade ages was expected given the somewhat different tree topologies and, hence, differences in the placement of the calibrations. However, clades that were recovered in different positions between the two trees were not found to be remarkably different ages (*Zabelia* and Morinoideae). Instead, three clades that ended up in similar positions were found to have notably different ages. The most remarkable difference was in the age of the most recent common ancestor of *Linnaea* and *Vesalea* – a difference of 15 Myr (ca. 33 Ma in the nuclear tree and ca. 48

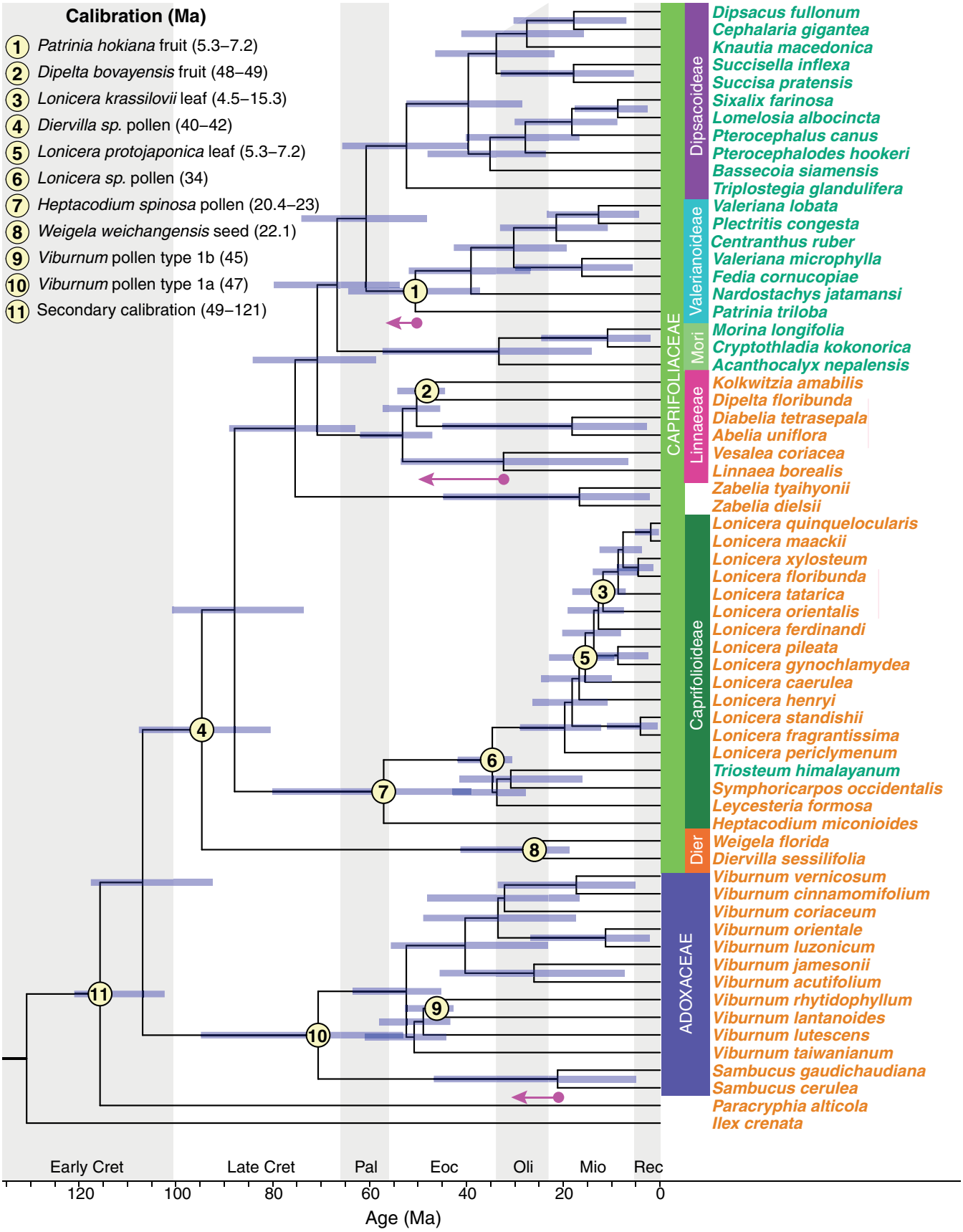

Calibration (Ma)
1 Patrinia hokiana fruit (5.3–7.2)
2 Dipelta bovayensis fruit (48–49)
3 Lonicera krassilovii leaf (4.5–15.3)
4 Diervilla sp. pollen (40–42)
5 Lonicera protojaponica leaf (5.3–7.2)
6 Lonicera sp. pollen (34)
7 Heptacodium spinosa pollen (20.4–23)
8 Weigela weichangensis seed (22.1)
9 Viburnum pollen type 1b (45)
10 Viburnum pollen type 1a (47)
11 Secondary calibration (49–121)
CAPRIFOLIACEAE
Dipsacoideae
Valerianoideae
Mori
Linnaeeae
Caprifolioideae
Dier
ADOXACEAE
Dipsacus fullonum
Cephalaria gigantea
Knautia macedonica
Succisella inflexa
Succisa pratensis
Sixalix farinosa
Lomelosia albocincta
Pterocephalus canus
Pterocephalodes hookeri
Bassecoia siamensis
Triplostegia glandulifera
Valeriana lobata
Plectritis congesta
Centranthus ruber
Valeriana microphylla
Fedia cornucopiae
Nardostachys jatamansi
Patrinia triloba
Morina longifolia
Cryptothladia kokonorica
Acanthocalyx nepalensis
Kolkwitzia amabilis
Dipelta floribunda
Diabelia tetrasepala
Abelia uniflora
Vesalea coriacea
Linnaea borealis
Zabelia tyaihyonii
Zabelia dielsii
Lonicera quinquelocularis
Lonicera maackii
Lonicera xylosteum
Lonicera floribunda
Lonicera tatarica
Lonicera orientalis
Lonicera ferdinandi
Lonicera pileata
Lonicera gynochlamydea
Lonicera caerulea
Lonicera henryi
Lonicera standishii
Lonicera fragrantissima
Lonicera periclymenum
Triosteum himalayanum
Symphoricarpos occidentalis
Leycesteria formosa
Heptacodium miconioides
Weigela florida
Diervilla sessilifolia
Viburnum vernicosum
Viburnum cinnamomifolium
Viburnum coriaceum
Viburnum orientale
Viburnum luzonicum
Viburnum jamesonii
Viburnum acutifolium
Viburnum rhytidophyllum
Viburnum lantanoides
Viburnum lutescens
Viburnum taiwanianum
Sambucus gaudichaudiana
Sambucus cerulea
Paracryphia alticola
Ilex crenata
Early Cret
Late Cret
Pal
Eoc
Oli
Mio
Rec
120
100
80
60
40
20
0
Age (Ma)

FIGURE 5. Time-calibrated nuclear tree of Dipsacales showing divergence time estimates. Blue node bars represent 95% highest posterior density age estimates. Taxon labels for herbaceous taxa are colored in green while those for woody taxa are colored in orange. Light yellow circles represent calibrations; numbers within the circles indicate calibration points described in the legend (details provided in Appendix S3). Geological epochs are represented as follows: Rec, Recent (Holocene+Pleistocene+Pliocene); Mio, Miocene; Oli, Oligocene; Eoc, Eocene; Pal, Paleocene; Late Cret, Late Cretaceous; Early Cret, Early Cretaceous. Clade name abbreviations are as follows: Dier, Diervillioideae; Mori, Morinoideae. Pink arrows have been placed above nodes that were found to have an age difference of 8 Myr or more compared to the plastid tree; the circles are placed at the ages found in the nuclear tree and the arrowheads extend to the age recovered in the analysis of plastid data.

Ma in the plastid tree). The *Sambucus* crown was found to be notably younger in the nuclear tree (ca. 20 Ma versus ca. 31 Ma; Fig. 5; Appendix S13.1). Similarly, the Valerianoideae crown was found to have originated ca. 50 Ma according to the nuclear tree, an age 8 Myr younger than that in the plastid tree.

Our taxon set contained woody as well as nonwoody lineages, and we found that age estimates differed accordingly. Specifically, we found that some woody clades were inferred to be only slightly older than the calibrations applied to them, while the age of the only calibrated herbaceous clade was much older than its calibration. For example, the most recent common ancestor (MRCA) of *Viburnum rhytidophyllum* and *V. lantanoides* was found to be ca. 46 Myr old, not much older than the 45 Ma calibration applied. Similarly, the MRCA of *Dipelta floribunda* and *Kolkwitzia amabilis* in the nuclear tree, and that of *Diabelia tetrasepala* and *Dipelta floribunda* in the plastid tree, was estimated to have evolved ca. 45–47 Ma, very close to the fossil constraint assigned to them (48–49 Ma). In contrast, the herbaceous Valerianoideae was found to be ca. 50–58 Myr old, in spite of its fossil calibration of 5–7 Ma.

DISCUSSION

Angiosperms353 efficacy and utility

Our results show the utility of the Angiosperms353 probe set but also highlight nuances of its efficacy that we believe researchers should consider when selecting a reduced representation sequencing strategy. We recovered loci on par with those reported by Johnson et al. (2019) for our moderate- and high-performing clusters of taxa. Despite enrichment problems in two of our hybridization pools, we were indeed able to map reads to nearly all targets across all samples. Our statistical inquiries and the success of congenerics in other library pools suggest that re-prepping and resequencing our failed pools would likely rescue these samples and that deeper sequencing would boost locus recovery more generally. Furthermore, the success of our "recycling" strategy, which uses representative loci from a preliminary iteration of HybPiper to increase recovery during a second round of assembly, demonstrates how bioinformatic elaborations can increase the utility of existing sequence data. However, it is unclear whether increasing efforts would resolve all outstanding questions, specifically those relationships among closely related species with low rates of molecular evolution.

Within the Dipsacales, the Angiosperms353 targets confidently resolved the backbone of the tree (e.g., Adoxaceae and Caprifoliaceae as sister clades, monophyly of Linnina). These trees also confirmed a number of important relationships within the major clades. For example, within the Dipsacoideae, we find that *Triplostegia* is sister to the Dipsaceae (Bell et al., 2001; Avino et al., 2009; Pyck and Smets, 2004; Carlson et al., 2009) and that within Valerianoideae, *Patrinia* and *Nardostachys* are successive sisters to core Valerianoideae (Bell et al., 2001, 2012; Donoghue et al., 2001; Bell and Donoghue, 2005). Within the Morinoideae, we confirm relationships among the three traditionally recognized genera (i.e., *Acanthocalyx* as sister to *Morina* + *Cryptothladia*; Bell and Donoghue, 2003).

We note that the members of Morinoideae sampled here were responsible for most of the paralog warnings we received. These findings are congruent with the increase in base chromosome observed in *Morina longifolia* (x = 17, Temsch and Geilhuber, 2010) relative to Valerianoideae (x = (7–) 8 (9–13); Hidalgo et al., 2010) and Dipsacoideae (x = 7–10; Temsch and Geilhuber, 2010), while the Dipsacales more generally are based on x = 8 or x = 9 (Sax and Kribs, 1930; Benko-Iseppon and Morawetz, 2000). The base number of 17 suggests the possibility of ancient hybridization between parental species with x = 8 and x = 9. However, although many paralogs were recovered in *Acanthocalyx*, *Cryptothladia*, and *Morina*, these paralogs most often formed lineage specific clades that appear to conflict with the hypothesis of a shared Morinoideae duplication event. Duplication events are common within Dipsacales; for example, genome size has been dynamic in *Viburnum*, with multiple duplication and some evidence of subsequent downsizing (Moeglein et al., 2020). It is also noteworthy that, while plastid and some nuclear data suggest that *Zabelia* is sister to Morinoideae, *Zabelia* does not share this increase in base chromosome number, and instead has x = 8–9 (Kim et al., 2001), as is found throughout Linnaeeae (Sax and Kribs, 1930; Zhang et al., 2002). However, chromosome numbers are variable within *Zabelia*, and all taxa surveyed thus far are polyploid, with $2n = 4x = 32$ (36) up to $2n = 12x = 108$ in *Z. biflora* (Kim et al., 2001). Polyploidy in *Zabelia* suggest the possibility of hybridization in the ancestry of this lineage as well. We recovered roughly four times more paralog warnings for *Z. dielsii* (86) than *Z. tyaihoynii* (22), but without genome size estimates for the former, we cannot say whether the increase in paralog warnings may be related to shifts in ploidy. Polyploidy has been reported in other taxa for which we received numerous paralog warnings (*Centranthus ruber* and *Fedia cornucopiae* are both $2n = 4x = 32$; Hidalgo et al., 2010), and these samples notably were in our best-performing cluster (Fig. 2). The effects of duplication and genome downsizing on locus recovery should be prioritized in future studies, as they appear to have impacted our recovery in real but complex ways.

With respect to the several outstanding phylogenetic issues within the Dipsacales, our results are mixed. As the first phylogenetic work on the Dipsacales to sample a significant number of nuclear loci, our analyses fully agree with previous findings, based largely on plastid data, in that *Heptacodium* belongs to the Caprifolioideae as sister to the Caprifolieae (Pyck and Smets, 2000; Winkworth et al., 2008; Fan et al., 2018; Wang et al., 2020; Xiang et al., 2020). We confidently dismiss the hypothesis that *Heptacodium* originated as an ancient hybrid between a member of the Caprifolieae and a member of the Linnaeeae (Zhang et al., 2002b; Jacobs et al., 2011), based on concordance of a majority of gene trees and quartets and a lack of signal in hybridization analyses. Instead, the shared morphological

similarities between *Heptacodium* and Linnaeeae are probably best explained as straightforward homoplasy (see below).

Despite this important success, our analyses have failed to convincingly resolve relationships among the major clades of Linnina and among the genera of the Caprifolieae and the Linnaeeae (Appendix S12). Within Linnina, *Zabelia* has been placed with the Morinoideae in our plastid trees (Fig. 4; Appendix S11), consistent with several previous plastid analyses (Fan et al., 2018; Wang et al., 2020; Xiang et al., 2020). However, our nuclear trees are split between this former hypothesis and one resolving *Zabelia*, Linnaeeae, and Morinoideae as successive sisters to Dipsacoideae + Valerianoideae (Fig. 3; Appendix S10). Further complicating these hypotheses, we note that among gene trees that support the *Zabelia* + Morinoideae sister relationship, the most frequent nuclear topology is incongruent with the plastid data and instead places Linnaeeae as sister to Dipsacoideae and Valerianoideae (Appendix S12). These conflicts within nuclear data and across cellular compartments are mostly informed by slim topological majorities in our ASTRAL analyses, particularly when using exon+flanking matrices (Appendix S12). While we did not exhaustively test hypotheses for the sources of conflict, we believe our findings reflect a more complex history than can be explained by incomplete lineage sorting alone. Although we did not find an appreciable signal for introgression or hybridization within Linnina, it is possible that deeply conserved, single copy loci may not be ideal for detecting such events. Deeper exploration of recent or ancestral whole genome duplications may also shed light on diversification among these clades. At this point, we consider the issue of the placements of *Zabelia*, Linnaeeae, and Morinoideae unsettled.

In the case of the Caprifolieae, the nuclear data are inconclusive, with different gene trees at odds with one another and with the results from our plastid analysis. The connection in our nuclear trees between *Symphoricarpos* and *Triosteum* perhaps better aligns with morphological characters (see below). A similar level of ambiguity exists within the Linnaeeae, where our nuclear analyses are incongruent with our plastid tree. Nuclear data support *Linnaea* and *Vesalea* sister to a clade containing sister pairs of *Abelia* + *Diabelia* and *Dipelta* + *Kolkwitzia*. The connection of *Abelia* to *Diabelia*, in particular, is strongly supported by nuclear gene trees (Fig 3; Appendix S12). Our plastid topology, which agrees with that of Wang et al. (2020), also recovers *Vesalea* and *Linnaea* as a clade sister to the remaining lineages, but finds *Diabelia* sister to *Dipelta*. Finally, Xiang et al. (2020) offered a third topology in which *Diabelia* is linked with *Kolkwitzia*. Unlike *Zabelia*, where support for two alternative placements lends credence to introgression, lack of resolution in the Linnaeeae appears more likely due to incomplete lineage sorting. We found that despite our collecting genome-scale data, nuclear and plastid DNA show relatively little variation among the members of Linnaeeae, as evidenced by short branch lengths.

We included multiple species of *Viburnum* and of *Lonicera* in our analyses to test the ability of the Angiosperms353 kit to resolve relationships among more closely related species. Given the high levels of gene tree ambiguity and discordance within *Viburnum*, we were happy to see support, albeit weak, in our ASTRAL analyses for a set of relationships that match the two major *Viburnum* lineages based on RAD-seq data (Eaton et al., 2017; Landis et al., 2020), though in some cases at odds with plastid-based trees (Clement et al., 2014; Spriggs et al., 2015). We note that *V. clemensiae* emerged as sister to the remainder of *Viburnum* in our expanded plastid analyses (Fig. 4), as in some, but not all, previous analyses (cf. Spriggs et al., 2015; and Lens et al., 2016). Unfortunately, too few genes were recovered for this key species to include it in our nuclear analyses. Overall, we conclude that our analyses of the nuclear data failed to provide confident resolution or new insights into relationships within *Viburnum*. The same can be said for *Lonicera*; nodes within *Lonicera* were informed by a median of 196 exon+flanking region gene trees (minimum of 102), yet these gene trees were overwhelmingly ambiguous or discordant with respect to species relationships. In both ASTRAL and RAxML analyses we found support for several species pairs that are also well supported in plastid analyses and morphological classification systems: for example, *L. standishii* + *L. fragrantissima* and *L. maackii* + *L. quinquelocularis*. In addition, both ASTRAL exon+flanking region and plastid species trees placed *L. periclymenum* as sister to the rest of *Lonicera*, as has been found in previous plastid analyses (Theis et al., 2008; Smith, 2009).

Viburnum and *Lonicera* may be worst-cases for the Angiosperms353 approach. These are woody clades with relatively high species diversity but low rates of molecular evolution. Other approaches, such as RAD-seq, have been successful in *Viburnum* (Eaton et al., 2017; Landis et al., 2020), and in other woody plant groups, including *Quercus* (Hipp et al., 2014) and *Salix* (Wagner et al., 2020). The major advantage of these probe-less methods is the enormous number of SNPs that they typically recover. However, methods that provide more data must be weighed against the limited combinability of these types of data across distant clades. Transcriptomes have also become popular tools for building broad phylogenies (e.g., Wang et al., 2019; One Thousand Plant Transcriptomes Initiative, 2019); while combinable, orthology inference from transcriptomes is less straightforward than using highly conserved low- or single-copy targets like the Angiosperms353 set. Finally, probe sets merging universal and lineage-specific target loci increase combinability across data sets (Jantzen et al., 2020; Christe et al., 2021; Eserman et al., 2021; Ogutcen et al., 2021). In direct comparisons between these sets, lineage-specific loci may not provide additional resolving power in the case of rapid radiations, as in Cyperaceae (Larridon et al., 2020), but can still be effective in clades of comparable size and age to *Viburnum* and *Lonicera* (e.g., *Buddleja*, Chau et al., 2018; *Memecylon* and *Tibouchina*, Jantzen et al., 2020). Suffice it to say that there is no one approach to confidently reconstruct flowering plant phylogeny. However, the Angiosperms353 probe set fills a significant gap, perhaps particularly in scaffolding areas of the overall flowering plant tree with little existing data.

Implications for diversification

Previous efforts to date the Dipsacales have relied on a small number of plastid loci (Bell and Donoghue, 2005a) or have used whole plastid genomes with less representative taxon sampling (Fan et al., 2018; Wang et al., 2020). Now, with our dated nuclear tree we are able to compare dates based on nuclear and plastid trees. The most straight forward of these comparisons is with our own plastid data set as it includes almost the same tip species and uses all of the same fossil calibrations.

Using six chloroplast markers, one fossil calibration, one secondary calibration, and 30 species belonging to 30 genera, Bell and Donoghue (2005a) found that the Dipsacales crown originated ca. 102 Ma. Using whole plastomes of 47 species belonging to 20 genera, two fossil calibrations, and one secondary calibration, Wang

et al. (2020) found the Dipsacales crown to be somewhat older (ca. 112 Ma). We note, however, that it is difficult to compare the results of Wang et al. (2020) to those of Bell and Donoghue (2005a), or to our analyses because Wang et al. relied on a discredited *Viburnum* fossil leaf from the Cretaceous.

Regarding the age of the Dipsacales, at least, the results of Bell and Donoghue (2005a) and of Wang et al. (2020) broadly coincide with our own estimates. However, all of these results differ significantly from those of Fan et al. (2018), who used plastomes of 16 species (lacking Dipsacoideae, Morinoideae, and Valerianoideae), and constraints of 36 Ma for *Dipelta* and 80 Ma for the Dipsacales node. Not surprisingly, they recovered much younger ages for the Dipsacales crown (ca. 81 Ma) and subsequent splitting events (e.g., ca. 47 Ma for Adoxaceae; 62 Ma for Caprifoliaceae). We also note that our estimates of ca. 104–107 Myr for the Dipsacales crown age are considerably older than those based on angiosperm-wide dating analyses (ca. 71 Myr in Magallón et al., 2015; 90 Myr in Ramírez-Barahona et al., 2020; and less than 79 Myr in Li et al., 2019). In many cases the dates from such broad studies are used as constraints in studies of particular clades of interest. While this does provide a reasonable starting point for such studies, and we have, in fact, relied here on a maximum age for Dipsacales obtained from Ramírez-Barahona et al. (2020), our analyses suggest that such constraints be used cautiously as the broader studies may often include fewer species from the focal clade and fewer internal fossil constraints. Based on our findings (see also Landis et al., 2020), we predict that studies of particular clades that include more species and multiple fossils will tend to push divergence dates back in time.

Aside from better taxon sampling, our study also used an expanded calibration set, consisting of 10 fossils (as opposed to one or two in previous studies) and one secondary calibration. Past analyses have failed to utilize these fossils, including the oldest known and accepted Dipsacales fossil (*Dipelta* fruit from the Eocene). Our choice of a birth–death process model reflects our interpretation that lineage extinctions have played an important role in shaping the modern diversity of Dipsacales. We note that some of the fossils used here (e.g., *Diplodipelta*, see Manchester and Donoghue, 1995) imply that entire clades may have been lost. However, in future studies of this problem, a host of alternative models should be explored, including the use of a simpler Yule process as a tree prior.

We find that the different age estimates for at least some clades are influenced by whether they include woody or herbaceous plants. Organisms with different life histories have been shown to exhibit heterogeneous rates of molecular evolution due in part to differences in generation times (e.g., Gaut et al., 1996; Laroche et al., 1997; Kay et al., 2006; De La Torre et al., 2017). Smith and Donoghue (2008) found that the DNA sequences of woody taxa evolved ~2.5 times slower than herbaceous lineages across five major angiosperm clades, including the Dipsacales. Due to this life-history dependent heterogeneity, relaxed clock methods are likely to estimate younger than true ages for slowly evolving woody groups, and older ages for more rapidly evolving herbaceous groups (Smith and Donoghue, 2008; Beaulieu et al., 2013). Our findings using only the uncorrelated lognormal (UCLN) model suggest the need for more detailed comparisons using, for example, random local clock models (Drummond and Suchard, 2010) or an approach that more directly incorporates prior information on evolutionary rates in woody versus herbaceous plants.

As noted above, we found some of our age estimates to be consistent with this hypothesis, but it is hard to gauge the magnitude of this effect because the true ages of individual clades are unknown and few individual clades have been both well sampled and calibrated with multiple fossils. An exception is the woody *Viburnum*, where dates have been estimated in multiple studies. Ages much younger than the oldest accepted *Viburnum* fossils were inferred by Fan et al. (2018; ca. 17 Myr) and by Wang et al. (2020; ca. 22 Myr). Both of these studies included just two *Viburnum* species and no internal calibrations. Using one fossil pollen grain, Moore and Donoghue (2007, 2009) found that the crown age of *Viburnum* was ca. 28 Myr, again much younger than the oldest *Viburnum* fossils that we now accept. Dating with many more tips and the same one fossil, Spriggs et al. (2015) found it to be roughly twice as old (ca. 55 Ma). Lens et al. (2016) found the *Viburnum* crown to be 56 Myr old, but their analysis used doubtful Eocene fossil leaves (Manchester et al., 2002). Dating with five pollen fossils, nearly complete species sampling, and a fossilized birth death process, Landis et al. (2020) obtained an age for crown *Viburnum* of ca. 70 Myr. Our estimated age for the *Viburnum* crown (ca. 48–52 Myr) is much younger, which we attribute to our small sample of *Viburnum* species/clades, the use of just two fossils within the group, and an overall analysis that included both woody and herbaceous clades. The lesson that we draw from this example is that estimated dates depend very strongly on the number and distribution of fossils within the group in question and on the number of species included and their representation of major subclades (Hug and Roger, 2007; Soares and Schrago, 2015). In *Viburnum*, it appears that increased sampling has played a major role in progressively pushing back the age of the group and, in so doing, ameliorating the tendency for the slow rate of molecular evolution to underestimate clade ages. Inference of a much older age in the Landis et al. (2020) study, with near complete sampling and carefully chosen fossil constraints supports this point.

Although our study improves upon prior Dipsacales analyses by increasing taxon sampling and calibrating with more carefully vetted fossils, there is still much work to be done, including, as noted above, the exploration of a wider variety of models. At this stage, we suspect that the ages estimated for the major woody clades are probably still too young, while those inferred for the herbaceous clades may be too old. Of course, these potential distortions will translate into potentially erroneous estimates of diversification rates. In the case of the Dipsacales, we suspect that the rates of diversification have actually been much higher in the herbaceous clades, especially in the Valerianoideae and the Dipsacoideae, with over 300 species each (cf. Smith and Donoghue, 2008; Beaulieu and Donoghue, 2013). Our taxon sampling also leaves room for improvement as a few major clades are missing from our nuclear analyses (e.g., *Adoxa* and its relatives).

Implications for morphological evolution and classification

Our analyses shed new light on the relationships of a number of lineages that have resisted resolution. Even in cases where we remain uncertain, the nuclear and plastid phylogenies reconstructed here give us a chance to reflect on the implications of alternative phylogenetic hypotheses for morphological evolution. Consideration of different possible evolutionary relationships within the Dipsacales puts a finer point on exactly what is at stake in terms of homology and patterns of homoplasy and highlights the need for further study not only of phylogenetic relationships, but also of the genetics and development of the morphological traits themselves. In the long

run, we hope that such considerations will promote the development of better models of morphological evolution.

The strongly supported placement of *Heptacodium* with the Caprifolieae in all of our analyses has especially interesting implications for morphological evolution. Despite substantial support from both nuclear and plastid data linking *Heptacodium* with Caprifolieae, *Heptacodium* shares several striking traits with the Linnina clade, including a reduction to a single fertile carpel and ovule and the production of achene fruits. Furthermore, *Heptacodium* shares greatly enlarged calyx lobes at the apex of the fruits with *Abelia*, *Diabelia*, and *Vesalea* of the Linnaeeae, as well as with *Zabelia* (Fig. 1J, L). Our studies imply that achenes and enlarged calyx lobes evolved independently in *Heptacodium* and in Linnina, thus refuting the homology of these traits despite their evident similarity. A roughly equally parsimonious alternative interpretation would be that achenes evolved in the common ancestor of the large clade that includes all of the Caprifoliaceae except Diervillioideae, and that these were then lost several times independently. In our estimation, the evolution of achenes early in the evolutionary history of Caprifoliaceae followed by multiple transitions to other fruit types is less likely as this hypothesis implies the reversal of carpel abortion and the evolution of multiseeded fleshy fruits in the Caprifolieae. We also note that these are transitions that have been shown to be rare in the Campanulidae at large (Beaulieu and Donoghue, 2013). While homoplasy in these traits is inevitable when *Heptacodium* is linked with the Caprifolieae, this placement of *Heptacodium* is consistent with the derived inflorescence structure of the Caprifolioideae. Specifically, these plants produce polytelic thyrse-like inflorescences in which (at least ancestrally) a set of six flowers is borne at each node—a cyme of three flowers in the axil of each of the opposite bracts (Landrein et al., 2012; Fig. 1N–P). Under our present interpretation, this inflorescence form is homologous in *Heptacodium* and Caprifolieae, whereas it would need to have evolved independently if *Heptacodium* were linked with Linnina. Finally, the form of the flower in *Heptacodium* (with two spreading dorsal, two, lateral, and one ventral corolla lobe) resembles a common form in the Caprifolieae (Fig. 1K, L).

We note that this discordance among morphological characters prompted the hypothesis that *Heptacodium* was the product of an ancient hybridization event between some member of the Caprifolieae and some member of the Linnaeeae (Zhang et al., 2002b; Jacobs et al., 2011), but we find little support for this idea. Our nuclear and chloroplast data sets do not conflict with respect to the placement of *Heptacodium*, as might have been expected, nor was any introgression detected in SNAQ analyses. Likewise, we do not observe strong conflict among the genes in our nuclear data set with respect to these nodes: most of the relevant gene trees (114) support the connection between *Heptacodium* and Caprifolieae, while many fewer (38) link *Heptacodium* with any members of the Linnina clade.

Comparing our nuclear results to recently published plastid analyses, and to the plastid results reported here, we are struck by the disagreement in the relationships within Linnina, especially by conflicting placements of *Zabelia*. In Wang et al. (2020), Xiang et al. (2020), our plastid analyses (Fig. 4), and our ASTRAL exon-only analysis (Fig. 3B), *Zabelia* is placed with Morinoideae. In contrast, our ASTRAL exon+flanking and RAxML analyses place *Zabelia* as sister to the entire Linnina clade (Fig. 3A; Appendix S10). On the surface, this would appear to be a strong conflict between nuclear and plastid data; inclusion of flanking regions seems to have bolstered the signal in the summary coalescent ASTRAL analysis for relationships already present in concatenated RAxML matrices, further strengthening support for cytonuclear conflicts at these nodes. However, we also note that considerable gene tree conflicts underlie our nuclear results. Therefore, we remain uncertain about these relationships within Linnina. The placement of *Zabelia* and Linnaeeae in a successive grade is perhaps most congruent with morphological data (in fact, *Zabelia* has usually been treated as a segregate of *Abelia* based primarily on the similarly expanded calyx lobes in some species), but either result entails conflict with morphological characters. Specifically, we note that *Zabelia* and Morinoideae share a similar salverform corolla (Fig. 1D,I,M) as compared to the funnelform corolla in *Abelia* and other Linnaeeae (Fig. 1J). *Zabelia* and Morinoideae also share derived pollen characters, including a smooth (psilate) exine and a continuous ring-shaped endoaperture in the equatorial plane (endocingulum) (Verlaque, 1983; Jacobs et al., 2011). These traits accord better with the plastid-based phylogeny. On the other hand, the placement of *Zabelia* with Morinoideae would imply the independent evolution of the herbaceous habit in the Morinoideae and again in the Valerianoideae + Dipsacoideae clade.

Within the Caprifolieae, our nuclear data support a connection between *Symphoricarpos* and *Triosteum*, which was not recovered among our plastid trees (Appendix S11). Here, the morphological data appear to favor the *Symphoricarpos-Triosteum* connection as reconstructed by the nuclear data. In both of them there are four carpels (two of which abort in *Triosteum*) and drupe fruits with relatively dry, mealy mesocarp tissue. Their pollen grains are also similar in lacking echinae on the tectum (Donoghue, 1985; Jacobs et al., 2011; Xu et al., 2011).

Within the Linnaeeae, there are noteworthy similarities and differences between our nuclear analyses and those based on plastid data, including our own. One point of agreement is the strongly supported link between the circumboreal *Linnaea* and the Mexican *Vesalea* (Wang et al., 2015). These share raceme-like inflorescences and similar nectary structure (Landrein et al., 2012; Landrein and Prenner, 2016). However, a clear disagreement concerns the placement of the recently segregated *Diabelia* (Landrein, 2010; Landrein et al., 2012). In our nuclear analyses, *Diabelia* is linked directly with *Abelia*, which squares well with previous classification systems in which *Diabelia* was treated as a subgroup within *Abelia*. However, in all recent plastid analyses, including our own, *Diabelia* is separated from *Abelia*, and placed instead with *Dipelta* (or possibly *Kolkwitzia*). This difference has consequences for our understanding of the evolution of expanded calyx lobes in particular. Our results with respect to *Heptacodium* and *Zabelia* clearly indicate that expanded calyx lobes evolved independently in these two lineages and then again within the Linnaeeae. Our nuclear results are consistent with the evolution of an expanded calyx in *Vesalea* and independently in the *Abelia-Diabelia* clade. The plastid analyses, by separating *Abelia* and *Diabelia*, would require at least one additional origin or multiple losses.

CONCLUSIONS

Here, we present the first view of Dipsacales evolution based on over 300 low-copy nuclear loci. The Angiosperms353 probe set provided a ready-to-use genomic tool to efficiently generate sequence

data for hundreds of orthologous loci, and we successfully recovered the majority of targeted loci across the Dipsacales. While the phylogeny inferred from these data is largely well-supported, it does not answer several of the long-standing phylogenetic questions in the Dipsacales, and resolution was notably poor in woody clades in which multiple species were included. Additionally, despite our use of many more fossils, we suspect that our exploratory divergence time estimates may still be yielding ages that are too young for woody clades and too old for herbaceous ones. This tree, along with our targeted assessment of gene tree concordance at key nodes, enables, for the first time, a critical comparison of plastid and nuclear trees and their implications for the homology or homoplasy of a host of morphological characters. We confidently support *Heptacodium* as sister to the Caprifolieae, though relationships within this clade otherwise remain unclear. Given significant conflict among our nuclear gene trees and with the plastid data, it appears that *Zabelia* and Morinoideae had more complicated evolutionary histories than previously imagined. In the future, it might be especially fruitful to extend the Angiosperms353 probe set to include additional loci to assist with the resolution of recalcitrant groups, such as those with long generation times, as well as genes involved in the development of floral form and the fusion of parts to investigate molecular homology between characters in particular clades.

ACKNOWLEDGMENTS

The authors thank the many people who collected and processed the specimens that made this work possible including Chuck Bell, Dave Boufford, Sara Carlson, Nancy Pyck, Stephen Smith, and Patrick Sweeney. We also thank Michael Dossman and Kathryn Richardson and the Arnold Arboretum for permission to work in the living collections, and Ned Friedman and Xu Bo for the use of their photographs. Our gratitude to Santiago Cardenas for assistance with plastid assemblies and Michael Landis for advice regarding divergence time analysis. Norm Wickett, Matt Johnson, and Lindsey Bechen greatly facilitated this work by providing us with early access to the Angiosperms353 probe set and providing space to carry out the initial stages of benchwork in his laboratory at the Chicago Botanic Garden. We thank Lisa Pokorny and two anonymous reviewers for providing insightful comments on the manuscript. The authors acknowledge using the ELSA high performance computing cluster at The College of New Jersey for the research reported in this paper. This cluster is funded by the National Science Foundation under grant number OAC-1828163. This work was funded by National Science Foundation grants to W.L.C. (DEB-1929670) and M.J.D. (DEB-1929533).

AUTHOR CONTRIBUTIONS

M.J.D., W.L.C., and I.S.G. designed the study. I.S.G. performed the hybridization protocol and sequencing, analyzed concordance, and performed statistical analyses. A.K.L. assembled the sequence data, completed nuclear tree inference, and supported plastid tree reconstruction. A.D.L. and W.L.C. assembled plastid data and reconstructed plastid trees. M.S. carried out the dating analyses, and M.J.D. provided information on morphological characters. All authors contributed to writing and editing the manuscript.

DATA AVAILABILITY STATEMENT

All sequence data collected for this work are available on the NCBI Sequence Read Archive (SRA) under BioProject PRJNA723358. All assemblies, alignments, trees, and custom scripts associated with the methods of this paper are available on GitHub at https://github.com/aaklee/DipsacalesHybSeq.

SUPPORTING INFORMATION

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

APPENDIX S1. SNAQ hybridization analysis.

APPENDIX S2. Plastid assembly and alignment.

APPENDIX S3. Fossil and calibration information.

APPENDIX S4. Locus-based recovery statistics.

APPENDIX S5. Nuclear alignment statistics.

APPENDIX S6. Pool quality group differences.

APPENDIX S7. Pool recovery group differences.

APPENDIX S8. Sample quality and recovery regressions.

APPENDIX S9. Cluster-based recovery statistics.

APPENDIX S10. Maximum likelihood analysis of nuclear alignments.

APPENDIX S11. Alternative plastid alignment strategies and trees.

APPENDIX S12. Quartet scores for focal nodes.

APPENDIX S13. Time calibrated plastid tree.

LITERATURE CITED

Avino, M., G. Tortoriello, and P. Caputo. 2009. A phylogenetic analysis of Dipsacaceae based on four DNA regions. *Plant Systematics and Evolution* 279: 69–86.

Backlund, A., and M. J. Donoghue. 1996. Morphology and phylogeny of the order Dipsacales. *In* Phylogeny of the Dipsacales, A. Backlund, Ph.D. dissertation. Department of Systematic Botany, Uppsala University, Uppsala, Sweden.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.

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

Beaulieu, J. M., B. C. O'Meara, and M. J. Donoghue. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62: 725–737.

Bell, C. D., and M. J. Donoghue. 2003. Phylogeny and biogeography of Morinaceae (Dipsacales) based on nuclear and chloroplast DNA sequences. *Organisms, Diversity, and Evolution* 3: 227–237.

Bell, C. D., and M. J. Donoghue. 2005a. Dating the Dipsacales: comparing models, genes, and evolutionary implications. *American Journal of Botany* 92: 284–296.

Bell, C. D., and M. J. Donoghue. 2005b. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms, Diversity, and Evolution* 5: 147–159.

Bell, C. D., E. J. Edwards, S. T. Kim, and M. J. Donoghue. 2001. Dipsacales phylogeny based on chloroplast DNA sequences. *Harvard Papers in Botany* 6: 481–499.

Bell, C. D., A. Kutschker, and M. T. K. Arroyo. 2012. Phylogeny and diversification of Valerianaceae (Dipsacales) in the southern Andes. *Molecular Phylogenetics and Evolution* 63: 724–737.

Benko-Iseppon, A. M., and W. Morawetz. 2000. Viburnales: cytological features and a new circumscription. *Taxon* 49: 5–16.

Blanchard, B., H. Wang, and D. L. Dilchard. 2016. Fruits, seeds and flowers from the Bovay and Bolden clay pits (early Eocene Tallahatta Formation, Claiborne Group), northern Mississippi, USA. *Palaeontologia Electronica* 19.3.52A: 1–59 palaeo-electronica.org/content/2016/1596-fossil-plants-from-mississippi.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Borowiec, M. L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4: e1660.

Brown, J. W., J. F. Walker, and S. A. Smith. 2017. Phyx: phylogenetic tools for unix. *Bioinformatics* 33: 1886–1888.

Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* 15: e1006650.

Cantino, P. D., and K. de Queiroz. 2020. International code of phylogenetic nomenclature (PhyloCode). version 6, updated June 2020. http://phylonames.org/code/ [accessed October 2020].

Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.

Carlson, S. E., V. Mayer, and M. J. Donoghue. 2009. Phylogenetic relationships, taxonomy, and morphological evolution in Dipsacaceae (Dipsacales) inferred by DNA sequence data. *Taxon* 58: 1075–1091.

Chau, J. H., W. A. Rahfeldt, and R. G. Olmstead. 2018. Comparison of taxon-specific versus general locus sets for targeted sequence capture in plant phylogenomics. *Applications in Plant Sciences* 6: e1032.

Christe, C., C. G. Boluda, D. Koubínová, L. Gautier, and Y. Naciri. 2021. New genetic markers for Sapotaceae phylogenomics: More than 600 nuclear genes applicable from family to population levels. *Molecular Phylogenetics and Evolution* 160: 107123.

Clement, W. L., M. Arakaki, P. W. Sweeney, E. J. Edwards, and M. J. Donoghue. 2014. A chloroplast tree for *Viburnum* (Adoxaceae) and its implications for phylogenetic classification and character evolution. *American Journal of Botany* 101: 1029–1049.

De La Torre, A. R., Z. Li, Y. Van de Peer, and P. K. Ingvarsson. 2017. Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution* 34: 1363–1377.

Donoghue, M. J. 1983. The phylogenetic relationships of *Viburnum*. *In* N. I. Platnick and V. A. Funk [eds.], Advances in cladistics, vol. 2, 143–166. Columbia University Press, NY, NY, USA.

Donoghue, M. J. 1985. Pollen diversity and exine evolution in *Viburnum* and the Caprifoliaceae *sensu lato*. *Journal of the Arnold Arboretum* 66: 421–469.

Donoghue, M. J., C. D. Bell, and R. C. Winkworth. 2003. The evolution of reproductive characters in Dipsacales. *International Journal of Plant Sciences* 164: S453–S464.

Donoghue, M. J., T. Eriksson, P. A. Reeves, and R. G. Olmstead. 2001. Phylogeny and phylogenetic taxonomy of Dipsacales, with special reference to *Sinadoxa* and *Tetradoxa* (Adoxaceae). *Harvard Papers in Botany* 6: 459–479.

Donoghue, M. J., R. G. Olmstead, J. F. Smith, and J. D. Palmer. 1992. Phylogenetic relationships of Dipsacales based on *rbc*L sequences. *Annals of the Missouri Botanical Garden* 79: 333–345.

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

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: e88.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.

Drummond, A. J., and M. A. Suchard. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology* 8: 114.

Eaton, D. A. R., E. L. Spriggs, B. Park, and M. J. Donoghue. 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology* 66: 399–412.

Eriksson, T., and M. J. Donoghue. 1997. Phylogenetic relationships of *Sambucus* and *Adoxa* (Adoxoideae, Adoxaceae) based on nuclear ribosomal ITS sequences and preliminary morphological data. *Systematic Botany* 22: 555–573.

Eserman, L. A., S. K. Thomas, E. E. D. Coffey, and J. H. Leebens-Mack. 2021. Target sequence capture in orchids: developing a kit to sequence hundreds of single copy loci. *Applications in Plant Sciences* 9: e11416.

Evanoff, E., W. C. McIntosh, and P. Murphey. 2001. Stratigraphic summary and 40Ar/39Ar geochronology of the Florissant Formation, Colorado. *In* E. Evanoff, K. M. Gregory-Wodzicki, and K. R. Johnson [eds.], Fossil flora and stratigraphy of the Florissant Formation, Colorado, 1–16. Proceedings of the Denver Museum of Nature and Science, ser. 4, no. 1, Denver, CO, USA.

Fan, W.-B., Y. Wu, J. Yang, K. Shahzad, and Z.-H. Li. 2018. Comparative chloroplast genomics of Dipsacales species: insights into sequence variation, adaptive evolution, and phylogenetic relationships. *Frontiers in Plant Science* 9: 689.

Gaut, B. S., B. R. Morton, C. C. McCaig, and M. T. Clegg. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. *Proceedings of the National Academy of Sciences, USA* 93: 10274–10279.

Hidalgo, O., J. Mathez, S. Garcia, T. Garnatje, J. Pellicer, and J. Vallès. 2010. Genome size study in the Valerianaceae: first results and new hypotheses. *Journal of Botany* 2010: 1–19.

Hipp, A., D. Eaton, J. Cavender-Bares, E. Fitzek, R. Nipper, and P. Manos. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS One* 9: e93975.

Howarth, D. G., and M. J. Donoghue. 2005. Duplications in CYC-like genes from Dipsacales correlate with floral form. *International Journal of Plant Sciences* 166: 357–370.

Howarth, D. G., T. Martins, E. Chimney, and M. J. Donoghue. 2011. Diversification of *CYCLOIDEA* expression in the evolution of bilateral flower symmetry in Caprifoliaceae and *Lonicera* (Dipsacales). *Annals of Botany* 107: 1521–1532.

Hug, L. A., and A. J. Roger. 2007. the impact of fossils and taxon sampling on ancient molecular dating analyses. *Molecular Biology and Evolution* 24: 1889–1897.

Jacobs, B., K. Geuten, N. Pyck, S. Huysmans, S. Jansen, and E. Smets. 2011. Unraveling the phylogeny of *Heptacodium* and *Zabelia* (Caprifoliaceae): An interdisciplinary approach. *Systematic Botany* 36: 231–252.

Jacobs, B., N. Pyck, and E. Smets. 2010. Phylogeny of the Linnaea clade: Are *Abelia* and *Zabelia* closely related? *Molecular Phylogenetics and Evolution* 57: 741–752.

Jantzen, J. R., P. Amarasinghe, R. A. Folk, M. Reginato, F. A. Michelangeli, D. E. Soltis, N. Cellinese, and P. S. Soltis. 2020. A two-tier bioinformatic pipeline to develop probes for target capture of nuclear loci with applications in Melastomataceae. *Applications in Plant Sciences* 8: e11345.

Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1600016.

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigué, R. S. Cowan, W. L. Eiserhardt, N. Epitawalage, et al. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* 68: 594–606.

Judd, W. S., R. W. Sanders, and M. J. Donoghue. 1994. Angiosperm family pairs: preliminary phylogenetic analyses. *Harvard Papers in Botany* 1: 1–51.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30: 3059–3066.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Kay, K. M., J. B. Whittall, and S. A. Hodges. 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. *BMC Evolutionary Biology* 6: 36.

Kim, T. J., B. Y. Sun, and Y. B. Suh. 2001. Palynology and cytotaxonomy of the genus *Abelia* s.l. *Caprifoliaceae. Korean Journal of Plant Taxonomy* 31: 91–106.

Kutschker, A. 2011. Revisión del género *Valeriana* (Valerianaceae) en Sudamérica austral. *Gayana Botanical Journal* 68: 244–296.

Landan, G., and D. Graur. 2008. Local reliability measures from sets of co-optimal multiple sequence alignments. *Pacific Symposium on Biocomputing* 13: 15–24.

Landis, M. J., D. A. R. Eaton, W. L. Clement, B. Park, E. L. Spriggs, P. W. Sweeney, E. J. Edwards, and M. J. Donoghue. 2020. Joint phylogenetic estimation of geographic movements and biome shifts during the global diversification of *Viburnum*. *Systematic Biology* 70: 67–85.

Landrein, S. 2010. *Diabelia*, a new genus of tribe Linnaeeae subtribe Linnaeinae (Caprifoliaceae). *Phytotaxa* 3: 34–38.

Landrein, S., and G. Prenner. 2016. Structure, ultrastructure and evolution of floral nectaries in the twinflower tribe Linnaeeae and related taxa (Caprifoliaceae). *Botanical Journal of the Linnean Society* 181: 37–69.

Landrein, S., G. Prenner, M. W. Chase, and J. J. Clarkson. 2012. *Abelia* and relatives: phylogenetics of Linnaeeae (Dipsacales–Caprifoliaceae *s.l.*) and a new interpretation of their inflorescence morphology. *Botanical Journal of the Linnean Society* 169: 692–713.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9: 357–359.

Langmead, B., C. Wilks, V. Antonescu, and R. Charles. 2018. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 35: 421–432.

Laroche, J., P. Li, L. Maggia, and J. Bousquet. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. *Proceedings of the National Academy of Sciences, USA* 94: 5722–5727.

Larridon, I., T. Villaverde, A. R. Zuntini, L. Pokorny, G. E. Brewer, N. Epitawalage, I. Fairlie, et al. 2020. Tackling rapid radiations with targeted sequencing. *Frontiers in Plant Science* 10: 1655.

Lens, F., R. A. Vos, G. Charrier, T. van der Niet, V. Merckx, P. Baas, J. Aguirre Gutierrez, et al. 2016. Scalariform-to-simple transition in vessel perforation plates triggered by differences in climate during the evolution of Adoxaceae. *Annals of Botany* 118: 1043–1056.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.

Li, H.-T., T.-S. Yi, L.-M. Gao, P.-F. Ma, T. Zhang, J.-B. Yang, M. A. Gitzendanner, et al. 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* 5: 461–470.

Li, J., and Z. G. Xiao. 1980. Introduction of Qipanshan section (K-50-16). *In* 1/200000 geological map of People's Republic of China. Geological Map Printing Plant of China, Beijing, China [in Chinese].

Liang, X.-Q., Y. Li, Z. Kvacek, V. Wilde, and C.-S. Li. 2013. Seeds of *Weigela* (Caprifoliaceae) from the Early Miocene of Weichang, China and the biogeographical history of the genus. *Taxon* 62: 1009–1018.

Magallón, S., S. Gómez-Acevedo, L. L. Sánchez-Reyes, and T. Hernández-Hernández. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* 207: 437–453.

Mai, U., and S. Mirarab. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* 19: 272.

Manchester, S. R. 2002. Leaves and fruits of *Davidia* (Cornales) from the Paleocene of North America. *Systematic Botany* 27: 368–382.

Manchester, S. R., M. A. Akhmetiev, and T. M. Kodrul. 2002. Leaves and fruits of *Celtis aspera* (Newberrt) comb. nov. (Celtidaceae) from the Paleocene of North America and Eastern Asia. *International Journal of Plant Sciences* 163: 725–736.

Manchester, S. R., Z.-D. Chen, A.-M. Lu, and K. Uemura. 2009. Eastern Asian endemic seed plant genera and their paleogeographic history throughout the Northern Hemisphere. *Journal of Systematics and Evolution* 47: 1–42.

Manchester, S. R., and M. J. Donoghue. 1995. Winged fruits of Linnaeeae (Caprifoliaceae) in the Tertiary of Western North America: *Diplodipelta* gen. nov. *International Journal of Plant Sciences* 156: 709–722.

Manchester, S. R., F. Grímsson, and R. Zetter. 2015. Assessing the fossil record of asterids in the context of our current phylogenetic framework. *Annals of the Missouri Botanical Garden* 100: 329–363.

McLay, T. G. B., J. L. Birch, B. F. Gunn, W. Ning, J. A. Tate, L. Nauheimer, E. M. Joyce, et al. 2021. New targets acquired: improving locus recovery from the Angiosperms353 probe set. *Applications in Plant Sciences* 9: e11420.

Mirarab, S., N. Nguyen, S. Guo, L.-S. Wang, J. Kim, and T. Warnow. 2015. PASTA: Ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. *Journal of Computational Biology* 22: 377–386.

Moeglein, M. K., D. S. Chatelet, M. J. Donoghue, and E. J. Edwards. 2020. Evolutionary dynamics of genome size in a radiation of woody plants. *American Journal of Botany* 107: 1527–1541.

Mongiardino Koch, N. M. 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. *bioRxiv* https://doi.org/10.1101/2021.02.13.431075 [Preprint].

Mongiardino Koch, N. M., and J. R. Thompson. 2021. A total-evidence dated phylogeny of Echinoidea combining phylogenomic and paleontological data. *Systematic Biology* 70: 421–439.

Moore, B. R., and M. J. Donoghue. 2007. Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *American Naturalist* 170(Supplement 20): S28–55.

Moore, B. R., and M. J. Donoghue. 2009. A Bayesian approach for evaluating the impact of historical events on rates of diversification. *Proceedings of the National Academy of Sciences, USA* 106: 4307–4312.

Murphy, B., F. Forest, T. Barraclough, J. Rosindell, S. Bellot, R. Cowan, M. Golos, et al. 2020. A phylogenomic analysis of Nepenthes (Nepenthaceae). *Molecular Phylogenetics and Evolution* 144: 106668.

Ogutcen, E., C. Christe, K. Nishii, N. Salamin, M. Möller, and M. Perret. 2021. Phylogenomics of Gesneriaceae using targeted capture of nuclear genes. *Molecular Phylogenetics and Evolution* 157: 107068.

One Thousand Plant Transcriptome Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574: 679–685.

Ozaki, K. 1980. Late Miocene Tatsumitoge Flora of Tottori Prefecture, Southwest Honshu, Japan (III). *Science Reports of the Yokohama National University, Section 2 Biology and Geology* 27: 19–45.

Pavlyutkin, B. I. 2015. A new species of *Lonicera* (Caprifoliaceae) from the Miocene of Primorye region (the Russian Far East). *Botanica Pacifica* 4: 157–160.

Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5: e9490.

Pyck, N., P. Roels, and E. Smets. 1999. Tribal relationships in Caprifoliaceae: evidence from a cladistic analysis using *ndh*F sequences. *Systematics and Geography of Plants* 69: 145–159.

Pyck, N., and E. Smets. 2000. A search for the phylogenetic position of the sevenson flower (*Heptacodium*, Dipsacales): combining molecular and morphological evidence. *Plant Systematics and Evolution* 225: 185–199.

Pyck, N., and E. Smets. 2004. On the systematic position of *Triplostegia* (Dipsacales): a combined molecular and morphological approach. *Belgian Journal of Botany* 137: 125–139.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: https://www.R-project.org/ [accessed March 2021].

Rambaut, A. 2012. FigTree v1. 4. Website: http://tree.bio.ed.ac.uk/software/figtree [accessed March 2021].

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Ramírez-Barahona, S., H. Sauquet, and S. Magallón. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology and Evolution* 4: 1232–1238.

Roels, P., and E. Smets. 1996. A floral ontogenetic study in the Dipsacales. *International Journal of Plant Sciences* 157: 203–218.

Salichos, L., and A. Rokas. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature* 497: 327–331.

Sax, K., and D. A. Kribs. 1930. Chromosomes and phylogeny in Caprifoliaceae. *Journal of the Arnold Arboretum* 11: 147–153.

Sayyari, E., and S. Mirarab. 2018. Testing for polytomies in phylogenetic species trees using quartet frequencies. *Genes* 9: 132.

Sela, I., H. Ashkenazy, K. Katoh, and T. Pupko. 2015. GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Research* 43: W7–W14.

Smith, S. A. 2009. Taking into account phylogenetic and divergence-time uncertainty in a parametric biogeographical analysis of the Northern Hemisphere plant clade Caprifolieae. *Journal of Biogeography* 36: 2324–2337.

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

Smith, S. A., and C. W. Dunn. 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* 24: 715–716.

Smith, S. A., M. J. Moore, J. W. Brown, and Y. Yang. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evolutionary Biology* 15: 1–15.

Solís-Lemus, C., and C. Ané. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics* 12: e1005896.

Soares, A. E. R., and C. G. Schrago. 2015. The influence of taxon sampling on Bayesian divergence time inference under scenarios of rate heterogeneity among lineages. *Journal of Theoretical Biology* 364: 31–39.

Spriggs, E. L., W. L. Clement, P. W. Sweeney, S. Madriñán, E. J. Edwards, and M. J. Donoghue. 2015. Temperate radiations and dying embers of a tropical past: the diversification of *Viburnum*. *New Phytologist* 207: 340–354.

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

Steenwyk, J. L., T. J. Buida III, Y. Li, X.-X. Shen, and A. Rokas. 2020. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biology* 18: e3001007.

Stevens, P. F. 2001. onward. Angiosperm phylogeny website, version 15, April 2019 [more or less continuously updated]. Website: http://www.mobot.org/MOBOT/research/APweb/ [accessed October 2020].

Stull, G., P. Soltis, D. Soltis, M. Gitzendanner, and S. Smith. 2020. Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany* 107: 790–805.

Sukumaran, J., and M. T. Holder. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26: 1569–1571.

Svetlana, P., U. Torsten, A. Anna, T. Valentina, T. Polina, and X. Yaowu. 2019. Early Miocene flora of central Kazakhstan (Turgai Plateau) and its paleoenvironmental implications. *Plant Diversity* 41: 183–197.

Tan, G., M. Muffato, C. Ledergerber, J. Herrero, N. Goldman, M. Gil, and C. Dessimoz. 2015. Current methods for automated filtering of multiple sequence alignments frequently worsen single-gene phylogenetic inference. *Systematic Biology* 64: 778–791.

Temsch, E. M., and J. Greilhuber. 2010. Genome size in Dipsacaceae and *Morina longifolia* (Morinaceae). *Plant Systematics and Evolution* 289: 45–56.

Terpilowski, M. 2019. scikit-posthocs: pairwise multiple comparison tests in Python. *Journal of Open Source Software* 4: 1169.

Theis, N., M. J. Donoghue, and J. Li. 2008. Phylogenetics of the Caprifolieae and *Lonicera* (Dipsacales) based on nuclear and chloroplast DNA sequences. *Systematic Botany* 33: 776–783.

Thiers, B. 2017. [continuously updated] Index Herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. Website: http://sweetgum.nybg.org/science/ih/ [accessed October 2020].

Vallat, R. 2018. Pingouin: statistics in Python. *Journal of Open Source Software* 3: 1026.

Verlaque, R. 1983. Contribution a l'etude du genre *Morina* L. *Pollen et Spores* 25: 143–162.

Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, et al. 2020. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods* 17: 261–272.

Wagenitz, G., and B. Liang. 1984. Die Nektarien der Dipsacales und ihre systematische Bedeutung. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie* 104: 91–113.

Wagner, N., L. He, and E. Hörandl. 2020. Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Frontiers in Plant Science* 11: 1077.

Wang, H.-F., S. Landrein, W.-P. Dong, Z.-L. Nie, K. Kondo, T. Funamoto, J. Wen, and S.-L. Zhou. 2015. Molecular phylogeny and biogeographic diversification of Linnaeoideae (Caprifoliaceae s. l.) disjunctly distributed in Eurasia, North America and Mexico. *PLoS One* 10: e0116485.

Wang, H.-X., H. Liu, J. J. Moore, S. Landrein, B. Liu, Z.-X. Zhu, and H.-F. Wang. 2020. Plastid phylogenomic insights into the evolution of the Caprifoliaceae s.l. (Dipsacales). *Molecular Phylogenetics and Evolution* 142: 106641.

Wang, N., Y. Yang, M. J. Moore, S. F. Brockington, J. F. Walker, J. W. Brown, B. Liang, et al. 2019. Evolution of Portulacineae marked by gene tree conflict and gene family expansion associated with adaptation to harsh environments. *Molecular Biology and Evolution* 36: 112–126.

Wheeler, T. J., and J. D. Kececioglu. 2007. Multiple alignment by aligning alignments. *Bioinformatics* 23: i559–568.

Winkworth, R. C., C. D. Bell, and M. J. Donoghue. 2008. Mitochondrial sequence data and Dipsacales phylogeny: mixed models, partitioned Bayesian analyses, and model selection. *Molecular Phylogenetics and Evolution* 46: 830–843.

Xiang, C. L., H. J. Dong, S. Landrein, F. Zhao, W. B. Yu, D. E. Soltis, P. S. Soltis, et al. 2020. Revisiting the phylogeny of Dipsacales: new insights from phylogenomic analyses of complete plastomic sequences. *Journal of Systematics and Evolution* 58: 103–117.

Xu, L., L. Lu, D.-Z. Li, and H. Wang. 2011. Evolution of pollen in the Dipsacales. *Plant Diversity and Resources* 33: 249–259.

Yabe, A., E. Jeong, K. Kim, and K. Uemura. 2019. Oligocene-Neogene fossil history of Asian endemic conifer genera in Japan and Korea. *Journal of Systematics and Evolution* 57: 114–128.

Yang, Y., and S. A. Smith. 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. *Molecular Biology and Evolution* 31: 3081–3092.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T.-Y. Lam. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* 8: 28–36.

Zhang, W. H., Z. D. Chen, J. H. Li, H. B. Chen, and Y. C. Tang. 2002a. Phylogeny of the Dipsacales s.l. based on chloroplast *trn*L-F and *ndh*F sequences. *Molecular Phylogenetics and Evolution* 26: 176–189.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

Zhang, Z. Y., Z. K. Zhou, and Z. J. Gu. 2002b. Karyomorphology of *Heptacodium* (Caprifoliaceae s. str.) and its phylogenetic implications. *Taxon* 51: 499–505.

APPENDIX 1. Specimen voucher information.

List of voucher or source material for Dipsacales species sequenced in this study. For each accession, the collector, collector number, and herbarium or living collection information is provided. *Index Herbarium* (Thiers, 2017) abbreviations are used throughout, except the Weberling personal herbarium (to be incorporated in M) is abbreviated as "Web". Also, Arnold Arboretum, Meise Botanic Garden, Belgium, and the Marsh Botanical Gardens at Yale University are abbreviated AA, MBG, and Marsh, respectively. Species are arranged alphabetically by clade (Fig. 3).

Adoxaceae: ***Adoxa moschatellina*** L.: *D. Boufford et al. 28906*, A; ***Sambucus cerulea*** Raf.: no voucher; ***S. ebulus*** L.: *Sundin s.n.*, GH; ***S. gaudichaudiana*** DC.: *Crisp & Telford 8540*, CGB; ***S. pubens*** Michx.: *D. Tank 103*, no voucher; ***Sinadoxa corydalifolia*** C.Y.Wu, Z.L.Wu & R.F.Huang: *D. Boufford et al. 26555*, GH; ***Tetradoxa omeiensis*** (H.Hara) C.Y.Wu: *Donoghue et al. 4000*, A; ***Viburnum acutifolium*** Benth.: *M.J. Donoghue 96*, YU; ***V. amplificatum*** J.Kern: *156003*, SAN; ***V. cinnamomifolium*** Rehder: *P.W. Sweeney et. al 2105*, YU; ***V. clemensiae*** J.Kern: *P.W. Sweeney et. al 2135*, YU; ***V. coriaceum*** Bl.: *P.W. Sweeney et. al 2088*, YU; ***V. dentatum*** L.: *5070-1A, 193779*, A; ***V. jamesonii*** (Oersted) Killip & A.C.Smith: *P.W. Sweeney et al. 1636*, YU; ***V. lantanoides*** Michx.: *M.J. Donoghue s.n.*, YU; ***V. lutescens*** Bl.: *P.W. Sweeney et al. 2104*, YU; ***V. luzonicum*** Rolfe: *Shen et al. 673*, A; ***V. orientale*** Pall.: *Merello et al. 2291*, MO; ***V. plicatum*** Thunberg: *M.J. Donoghue & K.-F. Chung KFC1940*, YU; ***V. punctatum*** Buch.-Ham. ex D.Don: *M.J. Donoghue 1:2017*, BHPL; ***V. rhytidophyllum*** Hemsl. ex Forb. & Hemsl.: *1386-82B*, living collections, AA; ***V. rufidulum*** Raf.: *21418A, 174348*, A; ***V. sargentii*** L.: *M.J. Donoghue and R.C. Winkworth 17*, YU; ***V. sieboldii*** Miq.: *616-6B, 174465*, A; ***V. taiwanianum*** Hayata: *W.-H. Hu et al. 2186*, MO; ***V. vernicosum*** Gibbs: *P.W. Sweeney 2123*, YU.

Caprifolioideae: ***Heptacodium miconioides*** Rehder: *1549-80A*, living collections, AA; ***Leycesteria formosa*** Wall.: *D. Boufford 29815*, A; ***Lonicera alpigena*** L.: *1310-84A*, living collections, AA; ***L. caerulea*** L.: *249-96B, 00152525*, A; ***L. ferdinandi*** Franch.: *563-87A, 00192872*, A; ***L. floribunda*** Boiss. & Buhse: *240-96B, 00152632*, A; ***L. fragrantissima*** Lindl. & J. Paxton: living collections, Marsh.; ***L. gynochlamydea*** Hemsl.: *15032A, 00152666*, A; ***L. henryi*** Hemsl.: *1696-80MASS, 00152695*, A; ***L. maackii*** (Rupr.) Maxim.: *7190B, 00152848*, A; ***L. nigra*** L.: *118-2000A*, living collections, AA; ***L. orientalis*** Lam.: *1240-84A, 00161070*, A; ***L. periclymenum*** L.: *276-2007MASS*, living collections, AA; ***L. pileata*** Oliv.: *18-92A*, living collections, AA; ***L. quinquelocularis*** Hardw.: *436-72A, 00161127*, A; ***L. standishii*** Jacques: *6669-1B, 00161205*, A; ***L. tatarica*** L.: *299-78A*, living collections, AA; ***L. xylosteum*** L.: *856-76A*, living collections, AA; ***Symphoricarpos occidentalis*** Hook.: *103-78A*, living collections, AA; ***S. sinensis*** Rehder: *Sweeney et al. 2605*, YU; ***Triosteum himalayanum*** Wall.: *Sweeney et al 2347*, YU.

Dievillioideae: ***Diervilla sessilifolia*** Buckley: *201-2004A, 00227422*, A; ***Weigela florida*** (Bunge)A. DC.: living collections, Marsh.

Dipsacoideae: ***Bassecoia siamensis*** (Craib) B.L. Burtt: *Smithiand 620049*, E; ***Cephalaria gigantea*** (Ledeb.) Bobrov: *19743238*, living collections, MBG; ***Dipsacus fullonum*** L.: *19801959*, living collections, MBG; ***Knautia dipsacifolia*** Kreutzer: *19762161*, living collections, MBG; ***K. macedonica*** Griseb.: *19931628-68*, living collections, MBG; ***Lomelosia albocincta*** (Greuter) Gretuer & Burdet: *Strid & Papanicolaou 15109*, G; ***Pterocephalodes hookeri*** (C.B.Clarke) V.Mayer & Ehrend.: *D. Boufford 33905*, A; ***Pterocephalus strictus*** Boiss. & Hohen.: *Archibald 8316*, E; ***P. canus*** Coult. ex DC.: *Assadi et al. 1776*, E; ***Scabiosa columbaria*** L.: *1000-1354*, living collections, Royal Botanic Garden Edinburgh.; ***S. mansenensis*** Nakai: *Hyun s.n.*, WU; ***Sixalix farinosa*** (Coss.) Greuter & Burdet: no voucher; ***Succisa pratensis*** Moench: *19752365*, living collections, MBG; ***Succisella inflexa*** (Kluk) Beck: *19761319*, living collections, MBG; ***Triplostegia glandulifera*** Wall. ex DC.: *D. Boufford 27738*, A.

Linnaeeae: ***Abelia uniflora*** R.Br. ex Wall.: N. Pyck DNA Collections, no voucher; ***Diabelia tetrasepala*** (Koidz.) Landrein: N. Pyck DNA Collections, no voucher; ***Dipelta floribunda*** Maxim.: *1650-80E, 00158743*, A; ***Kolkwitzia amabilis*** Graebn.: *20447-B*, living collections, AA; ***Linnaea borealis*** L.: *198-2015A*, living collections, AA; ***Vesalea coriacea*** (Hemsl.) T.Kim & B.Sun ex Landrein: N. Pyck DNA Collections, no voucher.

Morinoideae: ***Acanthocalyx nepalensis*** subsp. ***delevaya*** (Franch.) D.Y.Hong: *D. Boufford et al. 28374*, A; ***Cryptothladia kokonorica*** (K.S.Hao) M.J.Cannon: *D. Boufford et al. 26762*, A; ***Morina longifolia*** Wall. ex DC.: living collections, Marsh.

Valerianoideae: ***Centranthus ruber*** (L.) DC.: living collections, Marsh; ***Fedia cornucopiae*** (L.) Gaertn.: *Weberling s.n.*, Web, M; ***Nardostachys jatamansi*** (D.Don) DC.: *D. Boufford 29371*, A; ***Patrinia triloba*** Miq.: cultivated, Royal Botanic Garden Edinburgh; ***Plectritis congesta*** (Lindl.) DC.: *Shenk 308*, YU; ***Valeriana celtica*** L.: *Rau s.n.*, no voucher; ***V. lapathifolia*** Vahl.: *Weberling 10736*, Web., M; ***V. laxiflora*** DC.: *Weberling 10712*, Web., M; ***V. lobata*** Hoech amend. Bors.: *Weberling 10938*, Web., M; ***V. microphylla*** L.: *Liede & Meve 3481*, UBT; ***V. officinalis*** L.: living collections, Marsh; ***V. polemonifolia*** Phil.: *Weberling 10666*, Web., M; ***V. prionophylla*** Standl.: *Kurz and Koppilz s.n.*, Web., M; ***V. pyrenaica*** L.: *Weberling 89155*, Web., M; ***V. supina*** Ard.: *Nyffeler 1076*, YU; ***V. tuberosa*** L.: *Glöcker s.n.*, Web., M; ***V. virescens*** Clos.: *Weberling 10828*, Web., M; ***V. wallrothii*** Kreyer: *19801546*, living collections, MBG; ***Valerianella locusta*** (L.) Laterr.: *Patterson 2001*, SFSU.

Zabelia: ***Zabelia dielsii*** (Graebn.) Makino: N. Pyck DNA Collections, no voucher; ***Zabelia tyaihyoni*** (Nakai) Hisauti & H.Hara: N. Pyck DNA Collections, no voucher.

Outgroups: ***Ilex crenata*** Thunb.: ***Chase 119***, NCU; ***Paracryphia alticola*** (Schltr.) Steenis: ***Pintaud 561***, K.