A phylogenomic perspective on gene tree conflict and character evolution in Caprifoliaceae using target enrichment data, with Zabelioideae recognized as a new subfamily

Hong-Xin Wang[a,#], Diego F. Morales-Briones[b,#], Michael J. Moore[c], Jun Wen[d], Hua-Feng Wang[a*]

[a] *Key Laboratory of Tropical Biological Resources of Ministry of Education, College of Tropical Crops, Hainan University, Haikou 570228, China*

[b] *Department of Plant and Microbial Biology, College of Biological Sciences, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, Saint Paul, MN 55108, USA*

[c] *Department of Biology, Oberlin College, Oberlin, OH44074, USA*

[d] *Department of Botany, National Museum of Natural History, MRC-166, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA*

*Authors for correspondence:

Hua-Feng Wang, E-mail: hfwang@hainanu.edu.cn

[#] Authors contributed equally to this study.

Abstract

The use of diverse datasets in phylogenetic studies aiming for understanding evolutionary histories of species can yield conflicting inference. Phylogenetic conflicts observed in animal and plant systems have often been explained by hybridization, incomplete lineage sorting (ILS), or horizontal gene transfer. Here, we employed target enrichment data, species tree and species network approaches to infer the backbone phylogeny of the family Caprifoliaceae, while distinguishing among sources of incongruence. We used 713 nuclear loci and 46 complete plastome sequence data from 43 samples representing 38 species from all major clades to reconstruct the phylogeny of the family using concatenation and coalescence approaches. We found significant nuclear gene tree conflict as well as cytonuclear discordance. Additionally, coalescent simulations and phylogenetic species network analyses suggested putative ancient hybridization among subfamilies of Caprifoliaceae, which seems to be the main source of phylogenetic discordance. Ancestral state reconstruction of six morphological characters revealed some homoplasy for each character examined. By dating the branching events, we inferred the origin of Caprifoliaceae at approximately 66.65 Ma in the late Cretaceous. By integrating evidence from molecular phylogeny, divergence times, and morphology, we herein recognize Zabelioideae as a new subfamily in Caprifoliaceae. This work shows the necessity of using a combination of multiple approaches to identify the sources of gene tree discordance. Our study also highlights the importance of using data from both nuclear and chloroplast genomes to reconstruct deep and shallow phylogenies of plants.

Keywords: Caprifoliaceae; Hybridization; Introgression, Phylogenetic networks, *Zabelia*; Zabelioideae.

1 Introduction

Gene tree discordance is a ubiquitous feature of phylogenomic data sets (Galtier and Daubin, 2008; Degnan and Rosenberg, 2009; Szöllősi et al., 2015; Sun et al., 2015; Lin et al., 2019). Many studies have shown that incomplete lineage sorting (ILS), hybridization, and other processes such as horizontal gene transfer, gene duplication, or recombination, may be contributing to discordance among gene trees (Degnan and Rosenberg, 2009; Linder and Naciri, 2015). Among these potential sources of discordance, hybridization has been especially important in plant systematics research (e.g., Morales-Briones et al., 2018; Lee-Yaw et al., 2019; Stull et al., 2020; Morales-Briones et al., 2021). Hybridization may be expected to be prevalent in rapidly radiating groups, which is increasingly recognized as a major force in evolutionary biology, in many cases leading to new species and lineages (Mallet, 2007; Abbott et al., 2010; Yakimowski and Rieseberg, 2014; Konowalik et al., 2015). ILS is one of the prime sources of gene tree discordance, which has attracted increasing attention in the past decades as phylogenetic reconstruction methods allowed its modeling (Edwards 2009; Liu et al., 2015). Despite that, distinguishing ILS from hybridization is still challenging (Linder and Naciri, 2015). More recently, methods to estimate phylogenetic networks that account simultaneously for ILS and hybridization have been developed (Solís-Lemus and Ané, 2016; Wen et al., 2018). At the same time, empirical studies using phylogenetic networks to identify the sources gene tree discordance are increasing (e.g., Morales-Briones et al., 2018, 2021; Widhelm et al., 2019; Feng et al., 2020).

Caprifoliaceae is a medium-sized family with about 960 plant species belonging to 41 extant genera that are mainly distributed in eastern Asia and eastern North America (Donoghue et al., 2001; Bell, 2004; Wang et al., 2020; Xiang et al., 2020). The family has long been the focus of phylogenetic studies of character evolution, especially regarding its tremendous diversity in reproductive structures (Backlund, 1996; Donoghue et al., 2003). Caprifoliaceae has five corolla lobes and five stamens as ancestral states, which are retained in Diervilleae C.A.Mey., *Heptacodium* Rehd., and

Caprifolieae (though in some *Symphoricarpos* Duhamel and *Lonicera* L. there are four corolla lobes and four stamens). However, for other genera, the number of stamens is reduced to four or even one. Caprifoliaceae shows even greater variation in fruit types (e.g., achene in *Abelia* R. Br., berry in *Lonicera*, and capsule in *Weigela* Thunb.; Manchester & Donoghue, 1995; Donoghue et al., 2003). Some genera possess highly specialized morphological characters (e.g., the spiny leaves of *Acanthocalyx* (DC.) Tiegh., *Morina* L. and *Dipsacus* L.) that have likely played key roles in lineage-specific adaptive radiation (Blackmore and Cannon, 1983; Caputo and Cozzolino, 1994; Donoghue et al., 2003) (Fig. 1).

Circumscriptions of Caprifoliaceae have been controversial. Backlund & Pyck (1998) suggested that Caprifoliaceae should be defined narrowly to include only five genera, *Heptacodium* Rehder, *Leycesteria* Wall., *Lonicera*, *Symphoricarpos*, and *Triosteum* L. This narrowly circumscribed concept of the family has been also accepted by some authors (e.g., APG, 1998; Yang & Landrein, 2011; Xiang et al., 2020). By contrast, some researchers proposed to integrate Morinaceae, Dipsacaceae, Valerianaceae, and Caprifoliaceae s. str. into the Caprifoliaceae s.l. (e.g., Judd et al.,1994; Donoghue et al., 2001; Stevens, 2001 onwards; Wang et al., 2015; Wang et al., 2020). To maximize stability and ease identification based on recent phylogenetic studies (e.g., Li et al., 2019; Wang et al., 2020; Xiang et al., 2020), we prefer the Caprifoliaceae s.l. concept that includes seven major clades: Linnaeoideae, *Zabelia*, Morinoideae, Valerianoideae, Dipsacoideae, Caprifolioideae and Diervilloideae (Stevens, 2001 onwards; Wang et al., 2015; APG, 2016; Wang et al., 2020). Phylogenetic relationships within Caprifoliaceae have been studied extensively during the past two decades using plastid and nuclear DNA data (Fig. 2), but the placement of *Zabelia* (Rehder) Makino has never been resolved confidently using either morphological characters (Backlund, 1996; Donoghue et al., 2003) or molecular data (Donoghue et al., 1992; Jacobs et al., 2010; Smith et al., 2010; Landrein et al., 2012; Stevens, 2019; Wang et al., 2020; Xiang et al., 2020). Based on nuclear (ITS) and chloroplast DNA (cpDNA) data (*trnK*, *matK*, *atpB-rbcL*, *trnL-F*) of 51 taxa, Jacobs et al. (2010) found moderate support (bootstrap support [BS] = 62%) for the placement of

Zabelia (formerly part of *Abelia*) in a clade with Morinoideae, Dipsacoideae, and Valerianoideae. Based on the same data set, Jacobs et al. (2010) raised *Abelia* sect. *Zabelia* to the genus level as *Zabelia*, and more recent studies have confirmed the distinctiveness of *Zabelia* (Landrein et al., 2012; Wang et al., 2015), often finding it sister to Morinoideae, although with low (BS ≤ 50%) to moderate (50% < BS ≤ 70%) support (Donoghue et al., 1992; Jacobs et al., 2010; Tank and Donoghue, 2010; Wang et al., 2015). Based on cpDNA data (*rbcL*, *trnL-K*, *matK* and *ndhF*) of 14 taxa, Landrein et al. (2012) suggested that *Zabelia* and *Diabelia* Landrein (Linnaeoideae) had similar "primitive" inflorescences of reduced simple thyrses. Landrein et al. (2012) conducted phylogenetic analyses of the Caprifoliaceae based on the structural characters of reproductive organs. In these analyses, *Zabelia* was sister to the clade of Morinoideae, and Valerianoideae + Dipsacoideae. Recently, Xiang et al. (2020) carried out analyses of complete plastomes of 32 species in this clade, demonstrating that *Heptacodium* and *Triplostegia* Wall. ex DC. are sister to Caprifoliaceae *s.s.* and Dipsacaceae, respectively, and have thus been included as members of those groups. Furthermore, *Zabelia* was found to be sister to Morinaceae in all analyses (Xiang et al., 2020). Likewise, using complete plastomes from 56 accessions representing 47 species of Caprifoliaceae, Wang et al. (2020) recovered the clade composed of Linnaeoideae, and Morinoideae + *Zabelia* as sister to Dipsacoideae + Valerianoideae) with maximum support (BS = 100%).

In this study, we assembled and analyzed a custom target enrichment dataset of Caprifoliaceae to: (1) evaluate sources of gene tree discordance, in order to clarify the backbone phylogeny of Caprifoliaceae with special attention to positions of recalcitrant taxa (i.e., *Zabelia* and Morinoideae); and (2) determine the evolutionary patterns of key morphological characters of Caprifoliaceae.

2 Materials and methods

2.1 Taxon sampling

We sampled 43 individuals from 38 species of Caprifoliaceae, including representatives of all seven major clades (including *Zabelia*) of Caprifoliaceae sensu

Stevens (2001 onwards) and Wang et al. (2020). Additionally, three species of Adoxaceae were included as outgroups. Most samples (38) were collected in the field, where leaf tissue was preserved in silica gel. The remaining samples were obtained from the United States National Herbarium (US) at the Smithsonian Institution (Table S1). Vouchers of newly collected samples were deposited in the herbarium of the Institute of Tropical Agriculture and Forestry (HUTB), Hainan University, Haikou, China. Complete voucher information is listed in Supporting Information Table S1.

2.2 DNA extraction, target enrichment, and sequencing

We extracted total genomic DNA from silica gel-dried tissue or herbarium tissue using the CTAB method of Doyle and Doyle (1987). We checked the quantity of each extraction with a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and sonicated 400 ng of DNA using a Covaris S2 (Covaris, Woburn, MA) to produce fragments ~150-350 bp in length for library preparations. To ensure that genomic DNA was sheared at approximately the selected fragment size, we evaluated all samples on a 1.2% (w/v) agarose gel.

We identified putative single copy nuclear (SCN) genes with MarkerMiner v.1.2 (Chamala et al., 2015) with default settings, using the transcriptomes of *Dipsacus asper*, *Lonicera japonica*, *Sambucus canadensis*, *Valeriana officinalis*, and *Viburnum odoratissimum* from 1KP (Matasci et al., 2014), and the genome of *Arabidopsis thaliana* (L.) Heynh. (Gan et al., 2011) as a reference. SCN genes identified with MarkerMiner were further filtered using GoldFinder (Vargas et al., 2019), requiring loci with at least 400 bp and a coverage of at least three species. This resulted in 428 SCN genes for phylogenetic analyses. A custom set of 80 bp biotinylated RNA baits (MYbaits) based on exon sequences were manufactured by Arbor Biosciences (Ann Arbor, MI, USA), with a 2× tiling density. The bait sequences are available as a supplemental file (Appendix 1).

Library preparation was done with the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, MA, USA) following the manufacturer's protocol. Library concentrations were quantified using a Qubit 2.0, with a dsDNA HS Assay Kit

(Thermo Fisher Scientific). Fragment size distribution was determined with a High Sensitivity D1000 ScreenTape run on the Agilent 2200 TapeStation system (Agilent Technologies, Inc., Santa Clara, California, United States). Solution-based hybridization and enrichment with MYbaits followed Weitemier et al. (2014). Libraries pools were and sequenced by Novogene Corporation (Sacramento, California, U.S.A.) on one lane using the Illumina HiSeq 4000 sequencing platform (Illumina Inc, San Diego, California, U.S.A.) producing 150 bp paired-end reads.

Given the low recovery of plastome reads from target enrichment libraries, we used a genome skimming approach to ensure recovery of full plastomes. Following Wang et al. (2020) with minor modifications, we built separate libraries for total genomic DNA. These libraries were sequenced using the BGISEQ-500 platform at BGI Shenzhen (China) with 100 bp paired-end reads.

2.3 Read processing and assembly

Sequencing adapters and low-quality bases were removed with Trimmomatic v0.36 (ILLUMINACLIP: TruSeq_ADAPTER: 2:30:10 SLIDINGWINDOW: 4:5 LEADING: 5 TRAILING: 5 MINLEN: 25; Bolger et al., 2014). Assembly of nuclear loci was carried out with HybPiper v.1.3.1 (Johnson et al., 2016), on an exon basis to avoid chimeric sequences in multi-exon genes that may be produced by potential paralogy (Morales-Briones et al., 2018). Only exons with a reference length of ≥ 150 bp were assembled (1220 exons from 442 genes). Paralog detection was carried out for all exons with the 'paralog_investigator' option of HybPiper. All assembled loci (with and without paralogs detected) were processed following Morales-Briones et al. (2020) to obtain 'monophyletic outgroup' (MO) orthologs (Yang & Smith, 2014).

Plastome assembly followed Wang et al. (2020). Briefly, raw reads were filtered with SOAPfilter_v2.2 (BGI-Shenzhen, China) and ddapter sequences and low-quality reads were removed. Plastome assembly was carryout using MITObim v1.8 (Hahn et al. 2013) following Wang et al. (2020).

2.4 Phylogenetic analyses

We used concatenation and coalescent-based methods to reconstruct the phylogeny of Caprifoliaceae. We performed phylogenetic analyses on the nuclear and plastid data sets separately. Individual nuclear exons were aligned with MAFFT v7.407 (Katoh & Standley, 2013) and aligned columns with more than 90% missing data were removed using Phyutility (Smith & Dunn, 2008). A maximum likelihood (ML) tree was estimated from the concatenated matrix, partitioning by gene, using RAxML v8.2.12 (Stamatakis, 2014) and the GTRGAMMA model for each partition. Clade support was assessed with 200 rapid bootstrap replicates (BS). We also estimated a species tree with ASTRAL-III v5.7.1 (Zhang et al., 2018) from individual ML gene trees inferred using RAxML with a GTRGAMMA model. Local posterior probabilities (LPP; Sayyari & Mirarab, 2016) were used to assess clade support.

Gene tree discordance was evaluated using two approaches. First, we mapped the individual nuclear gene trees onto the species tree and calculated the internode certainty all (ICA; Salichos et al., 2014) and number of conflicting and concordant bipartitions on each node of the species trees using Phyparts (Smith et al., 2015). Then we used Quartet Sampling (QS; Pease et al., 2018) to distinguish strong conflict from weakly supported branches in the nuclear tree. We carried out QS with 1000 replicates.

The plastomes sequences were aligned with MAFFT. A ML tree was estimated with RAxML using the GTR + I + G model and 1000 bootstrap replicates for clade support. Additionally, we used QS with 1000 replicates to evaluate branch support.

2.5 Assessment of hybridization

To test whether ILS alone could explain cytonuclear discordance, we used coalescent simulations similar to Folk et al. (2017) and García et al. (2017). We simulated 10,000 gene trees under the coalescent with DENDROPY v.4.1.0 (Sukumaran & Holder, 2010) using the ASTRAL species trees as a guide tree with branch lengths scaled by four to account for organellar inheritance. We summarized the simulated gene trees on the cpDNA tree. Under a scenario of ILS alone, any relationships in the empirical chloroplast tree should be present in the simulated trees

and have a high frequency; under a hybridization scenario, relationships unique to the cpDNA tree should be at low (or zero) frequency (García et al., 2017).

2.6 Species network analysis

We inferred species networks using a maximum pseudo-likelihood approach (Yu et al., 2012). Due to computational restrictions and given our main focus on potential reticulation among major clades of Caprifoliaceae (i.e., along the backbone), except Caprifolioideae and Diervilloideae, which did not show major signal of conflict with respect to the rest of Caprifoliaceae [i.e., the remaining five major groups formed a clade with maximum support (see section 3.4)]. First, we reduced our 46-taxon data set to one outgroup and 10 ingroup taxa to include Dipsacoideae, Linnaeoeideae, Moroinoideae and *Zabelia* (11-taxa data set). To disentangle nested hybridization, we created a reduced, 9-taxa data set by removing Dipsacoideae (because *Dipsacus* and *Scabiosa* were found to be involved in several inferred hybridization events) and a 7-taxa data set that excluded these two taxa as well as *Morina* and *Zabelia* (which were found to be involved in reticulation events in both the 11-taxa and 9-taxa networks). Species network searches were carried out with PHYLONET v.3.6.1 (Than et al., 2008) with the command 'InferNetwork_MPL' and using individual ML gene trees. Network searches were performed using only nodes in the gene trees that had BS support of at least 50%, allowing for up to five hybridization events and optimizing the branch lengths and inheritance probabilities of the returned species networks under the full likelihood. To estimate the optimal number of hybridizations and test whether the species network fit our gene trees better than a strictly bifurcating species tree, we computed the likelihood scores of concatenated RAxML, ASTRAL and plastid DNA trees, given the individual gene trees, as implemented in Yu et al. (2012), using the command 'CalGTProb' in PHYLONET. Finally, we performed model selection using the Akaike information criterion (Akaike, 1973), the bias-corrected Akaike information criterion (AICc; Sugiura, 1978), and the Bayesian information criterion (Schwarz, 1978). The number of parameters equals the number of branch lengths being estimated,

plus the number of hybridization probabilities being estimated, and number of gene trees used to estimate the likelihood, to correct for finite sample size.

2.7 Divergence time estimation

Divergence times were inferred using BEAST v.2.4.0 (Bouckaert et al., 2014). There is potential ancient hybridization in Caprifoliaceae, and therefore we estimated diversification dates separately for the nuclear and chloroplast gene tree. The root age was set to 78.9 Ma (mean 78.9 Ma, normal prior distribution 76.3–82.2 Ma) following Li et al. (2019). We selected two fossils as calibration points. First, the fossil seeds of *Weigela* Thunb. from the Miocene and Pliocene in Poland (Lańcucka-Rodoniowa, 1967), and the Miocene in Denmark (Friis, 1985) were used to constrain its stem age (offset 23.0 Ma, lognormal prior distribution 23.0 – 28.4 Ma). Second, the fruit fossil *Diplodipelta* S.R.Manchester & M.J.Donoghue, from the late Eocene Florissant flora of Colorado (Manchester, 2000; Bell & Donoghue, 2005), was used as a constraint in three different positions. In each case, the *Diplodipelta* constraint was set as an offset of 36 Ma, with a lognormal prior distribution of 34.07–37.20 Ma. Wang et al. (2015) considered three placements of *Diplodipelta* because it is possible that *Diplodipelta* represents a common ancestor of *Diabelia* and *Dipelta*, and because the sepal of *Diplodipelta* is similar to *Diabelia*, while the fruit wing of *Diplodipelta* is similar to *Dipelta* (Manchester & Donoghue, 2005; Wang et al., 2015). Hence, following Wang et al. (2015), we tested three placements of the *Diplodipelta* constraint: we constrained the common ancestor of *Diabelia* and *Dipelta* (Analysis I), we constrained the common ancestor (crown group) of *Dipelta* (Analysis II), and we constrained the common ancestor (crown group) of *Diabelia* (Analysis III). For each of these constraint positions, we carried out divergence time estimations for the nuclear and chloroplast trees separately.

All dating analyses were performed with an uncorrelated lognormal relaxed clock (Drummond et al., 2012), GTR + G substitution model (Posada, 2008), estimated base frequencies, and a Yule process for the tree prior. The RAxML tree was used as the

starting tree, and two independent MCMC analyses of 300,000,000 generations with 10% burn-in and sampling every 3000 generations were conducted to evaluate the credibility of posterior distributions of parameters. BEAST log files were analyzed with Tracer v.1.7 (Drummond et al., 2012) for convergence with the first 10% of trees removed as burn-in. Parameter convergence was assessed using an effective sample size (ESS) of 200. Log files where combined with LogCombiner and a maximum clade credibility tree with median heights was generated with TreeAnnotator v.1.8.4 (Drummond et al., 2012).

2.8 Analysis of character evolution

Character states were coded from the literature, particularly from Backlund (1996), Donoghue et al. (2003), Jacobs et al. (2011) and Landrein (2017). The number of stamens was scored as follows: (0), 1; (1), 2; (2), 3; (3), 4; (4), 5. Two-character states were scored for style exertion: (0), not exceeding corolla; (1), exceeding corolla. Four fruit types were scored: (0), achene; (1), capsule, (2), berry; (3), drupe. The number of carpels was scored as: (0), 2; (1), 3; (2), 4. Number of seeds was scored as: (0), 1; (1), 2; (2), 4-5; (3), 6-20; (4), 20+; Two epicalyx types were scored: (0), no; (1), yes. All the morphological characters analyzed here are presented in Supplementary Fig. S1. Ancestral character state reconstruction was performed using ML as implemented in Mesquite v.3.51 (Maddison and Maddison, 2018) with the 'Trace character history' option based on the topology of the chloroplast trees. To explore differences caused by differing topologies, we also reconstructed ancestral character states onto the nuclear tree. The Markov k-state one-parameter model of evolution for discrete unordered characters (Lewis, 2001) was used.

2.9 Data accessibility

Raw Illumina data from sequence capture is available at the Sequence Read Archive (SRA) under accession SUB7674585 (see Table S1 for individual sample SRA accession numbers). DNA alignments, phylogenetic trees and results from all analyses and datasets can be found in the Dryad data repository.

3 Results

3.1 Assembly

The number of assembled exons per species (with > 75% of the target length) ranged from 130 (*Vesalea floribunda*) to 989 (*Diabelia sanguinea*) out of 1220 single-copy exon references, with an average of 725 exons (Table S2; Fig. S2). The number of exons with paralog warnings ranged from 1 in *Vesalea floribunda* to 619 in *Diabelia sanguinea* (Table S2). After paralog pruning and removal of exons with poor coverage across samples (at least 25 ingroup taxa), we kept 707 exons from 367 different genes. The resulting concatenated matrix had an aligned length of 343,609 bp with 96,479 parsimony-informative sites, a minimum locus size of 150 bp, and a maximum locus size of 3,503 bp, with an average of 486 bp. The plastome alignment resulted in a matrix of 208,607 bp with 32,960 parsimony-informative sites (Table 1).

3.2 Phylogenetic reconstruction

In analyses of both nuclear and plastid data, Diervilloideae and Caprifolioideae were successively sister to remaining Caprifoliaceae, which were resolved into five main groups: Diervilloideae, Caprifolioideae, Valerianoideae, *Zabelia* and Morinoideae. (Figs. 3–4). However, the relationships among the seven groups within Caprifoliaceae differed between nuclear and plastid analyses.

Nuclear dataset. The ASTRAL analysis (Fig. 3) recovered maximum support (LPP = 1) for relationships within Caprifoliaceae and its seven major clades, except for the placement of *Kolkwitzia amabilis* (LPP = 0.7). Diervilloideae was resolved as sister to the rest of Caprifoliaceae, followed by Caprifolioideae as successive sister. *Zabelia* and Morinoideae formed a clade that was placed as sister to the remaining major groups. Dipsacoideae was recovered as polyphyletic, where *Scabiosa* was sister of Valerianoideae, and together with *Dipsacus japonicus* formed a grade sister to Linnaeoideae. Within Linnaeoideae, the clade of *Vesalea* M.Martens & Galeotti + *Linnaea* Gronov. ex L. was recovered as sister to a clade of all other Linnaeoideae.

The topology of the nuclear concatenated RAxML tree (Fig. 4) was mostly similar to that of the ASTRAL trees regarding major clades and their relationships. Most major clades and relationship among them had maximum support (BS = 100). The two differences were that RAxML recovered a monophyletic Dipsacoideae (*Dipascus* + *Scabiosa*) sister to Valerianoideae and *Kolkwitzia amabilis* sister of *Dipelta.*

The conflict analyses (Fig. 3, Figs. S3-S7) confirmed the monophyly of Caprifoliaceae with 266 out of 289 informative gene trees being concordant (ICA = 0.73) and having full QS support (1/–/1; i.e., all sampled quartets supported that branch). Within Caprifoliaceae, major clades and the the relationships among them had low to strong support. Diervilloideae was supported by 294 gene trees (out of 325; ICA = 0.79) and full QS support. Caprifolioideae was supported by 161 gene trees (out of 216; ICA = 0.50) and strong QS support, with signal of an alternative topology (0.87/0.077/1). The relationship of Caprifolioideae to the remaining five major clades was supported by 306 gene trees (out of 398; ICA = 0.55) and strong QS support, with signal of an alternative topology (0.82/0.043/1). The remaining five major groups formed a clade supported by 417 gene trees (out of 598; ICA = 0.32) and full QS support. Morinoideae was supported by 409 gene trees (out of 448; ICA = 0.72) and full QS support, *Zabelia* was supported by 453 gene trees (out of 484; ICA = 0.81) and also full QS support. The clade composed of *Zabelia* + Morinoideae was supported by only 86 gene trees (out of 354) and moderate QS support, with signal of a possible alternative topology (0.25/0/0.99). In the ASTRAL topology (Fig 3; Figs S3–S4), *Scabiosa* (Dipsacoideae) had 233 supporting gene trees (out of 308; ICA = 0.37) and full QS support, Valerianoideae was supported by 314 gene trees (out of 358; ICA = 0.58) with full QS support. The clade composed of Valerianoideae + *Scabiosa* was supported by only 63 gene trees (out of 370; ICA = 0.11) and moderate QS support, with signal of a possible alternative topology (0.39/0/0.99). The sister relationship of *Dipsacus japonicus* and Linnaeoideae was supported by only 48 gene trees (out of 468; ICA = 0.08) and had weak QS support, with signal of a possible alternative topology (0.073/0/1). The sister relationship of the clade Valerianoideae + *Scabiosa* and the clade *Dipsacus japonicus* + Linnaeoideae was supported only by 86

gene trees (out of 461; ICA 0.09) with moderate QS support but no signal of an alternative topology (0.17/0.73/0.97). In turn, for the RAxML topology (Figs S5 –6), a monophyletic Dipsacoideae was supported by only 80 gene trees (out of 300; ICA = 0.13) but had strong QS support with signal of an alternative topology (0.74/0/1). Linnaeoideae was supported by only 106 gene trees (out of 507; ICA = 0.08) but had strong QS support with no signal of an alternative topology (0.83/0.95/1). Within Linnaeoideae, *Linnaea* was supported by 301 gene trees (out of 312; ICA = 0.88) and full QS support, *Vesalea* was supported by 101 gene trees (out of 223; ICA = 0.33) and strong QS support but there was signal of a possible alternative topology (0.92/0/1). The clade *Linnaea* +*Vesalea* was supported by only 128 gene trees (out of 343; ICA = 0.20) and had strong QS support with signal of a possible alternative topology (0.92/0/1). In the case of the remaining Linnaeoideae, *Dipelta* was supported by 276 gene trees (out of 326; ICA = 0.69) and full QS support, *Diabelia* was supported by 149 gene trees (out of 345; ICA = 0.20) and had strong QS support with signal of a possible alternative topology (0.62/0/1), and *Abelia* was supported by only 54 gene trees (out of 349; ICA= 0.20) but with strong QS support and signal of a possible alternative topology (0.79/0/1). The clade formed by *Abelia* + *Diabelia* was supported only by 45 gene trees (out of 394; ICA = 0.07) but with strong QS support and signal of a possible alternative topology (0.67/0.2/1). In the ASTRAL analysis, the sister relationship of *Kolkwitzia amabilis* and the clade of *Abelia* + *Diabelia* was supported only by 15 gene trees (out of 411; ICA = -0.06) with QS counter support and clear signal for an alternative topology (-0.22/0.19/0.94). In turn, for the RAxML topology, *Kolkwitzia amabilis* was placed as sister to *Dipelta* with the support of only 39 gene trees (out of 288; ICA = 0.08) and moderate QS support, with signal for a possible alternative topology (0.22/0.29/0.95). Finally, the clade composed of *Dipelta*, *Kolkwitzia*, *Abelia*, and *Diabelia* was supported by only 20 gene trees (out of 413; ICA = 0.06) and moderate QS support with signal for a possible alternative topology (0.37/0.059/0.95).

Plastid dataset. Phylogenetic analysis of the cpDNA dataset also recovered the same seven major clades in Caprifoliaceae, although relationships differed between

plastid and nuclear trees (Fig. 4). In the plastid tree all major clades and most relationships among them had full support (BS = 100; QS = 1/–/1; Fig. 4 and Fig. S7). The clade Valerianoideae + Dipsacoideae was recovered as sister to all remaining Caprifolicaeae. The clade composed of *Zabelia* + Morinoideae was recovered as sister to Linnaeoideae with strong BS support (75) and moderate QS support, with signal of an alternative topology (0.36/0.4/0/94). Within Linnaeoideae, *Diabelia* and *Dipelta* formed a clade with strong BS support (89) and moderate QS support, with signal of an alternative topology (0.36/0.24/0/93). *Kolkwitzia amabilis* was sister to *Diabelia* + *Dipelta* with full BS support and strong QS support, but with a signal of an alternative topology (-/94/0.33/0.99). *Abelia* was recovered with full support as the sister of the clade composed of *Diabelia*, *Dipelta* and *Kolkwitzia*. The main differences between the nuclear and cpDNA trees were the placement of *Zabelia* + Morinoideae as sister of Linnaeoideae, and the placement of *Kolkwitzia amabilis* as sister of the clade of *Diabelia* + *Dipelta*.

3.3 Coalescent simulations

Coalescent simulations under the organellar model did not produce gene trees that resembled the observed cpDNA tree. When the simulated gene trees were summarized on the observed chloroplast tree, most clade frequencies were near zero, as for instance *Kolkwitzia amabilis* and the clade Valerianoideae + Dipsacoideae, *Zabelia* + Morinoideae and the clade Linnaeoideae + Valerianoideae + Dipsacoideae (Fig. S8). This suggested that ILS alone cannot explain the high level of cytonuclear discordance observed in Caprifoliaceae.

3.4 Species network analysis

For all three data sets analyzed (11-taxa, 9-taxa, 7-taxa), any of the networks with one to five reticulations events was a better model than a strictly bifurcating tree (Table 2; Fig. S9). In the 11-taxa data set, the best network had four reticulations events, whereas three reticulation events were inferred for the best networks in the 9-taxa and 7-taxa trees (Fig. 5; Table 2). In the 11-taxa network (Fig. 5a), which included

Dipsacoideae, both species of Dipsacoideae were inferred to result from hybridization events involving members of Linnaeoideae. *Dipsacus* was inferred to have contributions from three lineages, including an inferred hybridization between the lineages leading to *Linnaea* and *Vesalea*. In the 9-taxa tree (excluding Dipsacoideae but including *Morina* and *Zabelia*; Fig. 5b), both *Morina* and *Zabelia* were inferred to have genetic contributions from Linnaeoideae, while the clade of *Abelia* + *Diabelia* was inferred to have received a contribution from the lineage leading to *Kolkwitzia*. In the 7-taxa tree (including only Linnaeoideae + *Viburnum*; Fig. 5c), the clade of *Abelia* + *Diabelia* was again inferred to have resulted from a reticulation event, as were *Dipelta* and *Linnea*, although in the case of *Linnea*, there was only a small contribution (1.2%) from the lineage leading to *Vesalea*.

3.5 Divergence time estimation

Divergence time estimates based on the nuclear data set suggested that the deepest divergences in Caprifoliaceae occurred in the early Paleocene, whereas most generic-level diversification occurred in the middle Eocene to middle Oligocene (Fig. 6). The divergence between Dipsacoideae and Valerianoideae was dated to 41.83 Ma (95% Highest Posterior Density (HPD) = 34.44–49.35 Ma) (Fig. 6). The diversification of Linnaeoideae was inferred to be at 48.33 Ma (95% HPD = 42.37–51.70 Ma) (Fig. 6). Within Linnaeoideae, both *Abelia* and *Kolkwitzia* originated almost contemporaneously in the mid-late Eocene. The onset of *Zabelia* and Morinoideae diversification occurred between 32.10 and 39.89 Ma (Fig. 6). A comparison of the time estimates for selected nodes under the six different analyses is available in Figs. S10 and S11. As would be expected, Analysis I (placing the *Diplodipelta* fossil constraint at the ancestor of *Diabelia* and *Dipelta*) resulted in younger ages for much of Linnaeaoideae, but otherwise did not significantly affect divergence dates. We found that our estimated ages were generally younger with nuclear vs. plastid data (Figs. S10 - S11). For instance, our analyses showed that Caprifoliaceae date to 66.65 Ma (95% HPD = 56.31- 69.44 Ma) in analysis I based on nuclear datasets (Fig. S 10), whereas in the two other analyses, the divergence time was estimated as 69.38 Ma

(95% HPD = 53.26- 79.45 Ma) and 69.87 Ma (95% HPD = 64.20- 76.34 Ma, Fig. S10). Based on chloroplast datasets, the divergence time of Caprifoliaceae was estimated at 76.43 Ma (95% HPD = 64.81- 82.10 Ma), 78.61 Ma (95% HPD = 71.33- 82.45 Ma) and 77.43 Ma (95% HPD = 58.33- 81.60 Ma) in the three analyses respectively (Fig. S10).

3.6 Character evolution

The likelihood inference of character evolution using the cpDNA tree detected some homoplasy in each of the six morphological characters examined (Figs. 8 - 10), with style exertion relative to corolla showing particularly high homoplasy (Fig. 7). Stamen number exhibited little homoplasy, with an inferred shift from five stamens ancestrally to four in the bulk of Caprifoliaceae, and within Valerianoideae, it further reduced to 3 and 1 (Fig. 7). In contrast, style exertion exhibited a high level of homoplasy in the early diversification of the family (Fig. 7). Even within Linnaeoideae, the state “not exceeding corolla” was inferred to have originated twice, in *Vesalea* and in the *Diabelia-Dipelta-Kolwitzia-Abelia* clade. The ancestral fruit type for Caprifoliaceae was uncertain due to the diversity in fruit types among the major early-diverging clades, but was most likely an achene (Fig. 8). Carpel number evolution was similarly complex (Fig. 8), with three carpels inferred to be ancestral in the family. For seed number, one seed was inferred as the ancestral character state for Caprifoliaceae, with independent shifts to two seeds in *Symphoricarpos* and *Dipelta* (Fig. 9). Independent origins of the epicalyx in *Dipsacus* and the clade of *Moina* and *Acanthocalyx* was also inferred (Fig. 9).

A summary of character state evolution using the nuclear gene tree and character states that are relevant for the taxonomy of the group is shown in Figs. S12, S13 and S14. We found that the patterns of character evolution from cpDNA tree and nuclear gene tree were similar.

4 Discussion

4.1 Phylogenetic incongruence and putative hybridization

Although both our nuclear and plastid phylogenies supported the same seven major clades of Caprifoliaceae, the relationships among these clades are incongruent between data sets (Figs. 5 and 6). For instance, in the nuclear ASTRAL tree, Linnaeoideae is recovered as sister to Dipsacoideae (except for *Dipsacus japonicus*) + Valerianoideae (Fig. 3), while in the plastid tree Linnaeoideae is sister to *Zabelia* + Morinoideae (Fig. 4). In contrast, in the nuclear RAxML concatenated tree (Fig. 3), Linnaeoideae is recovered as sister to Dipsacoideae +Valerianoideae. Some of these points of conflict pertain to areas of Caprifoliaceae phylogeny that have long been problematic—for example, the relationships between *Zabelia* and other subfamilies. The inclusion of extensive nuclear genome sampling for Caprifoliaceae in this study is important because plastome-only phylogenies may not fully capture evolutionary processes such as ILS or organellar capture via hybridization. Three main processes will lead to gene tree heterogeneity and cytonuclear discordance: gene duplication/extinction, horizontal gene transfer/hybridization, and ILS. Currently, there are many methods to detect gene discordance (e.g., Smith et al., 2015; Pease et al., 2018), but sources of such discordance remain hard to disentangle, especially when multiple processes co-occur (e.g., Morales-Briones et al. 2021).

Zabelia was long thought to be closely related to *Abelia* (Hara, 1983; Tang & Lu, 2005). However, based on molecular datasets, Tank and Donoghue (2010) and Jacobs et al. (2011) found that *Zabelia* was sister to Morinaceae (=Morinoideae in this study). Using six molecular loci and inflorescence morphology, Landrein et al. (2012) concluded that the position of *Zabelia* remained unclear. The molecular investigation of Xiang et al. (2020) found that the sister relationships between *Zabelia* + Morinaceae and Linnaceae + Valerianaceae + Dipsacaceae were not highly supported. Such phylogenetic incongruence provides the opportunity to test causal hypotheses of cytonuclear discordance, e.g., ILS or hybridization. Further, in our analyses (Fig. 4), widespread cytonuclear discordance exists across Caprifoliaceae, especially at genus levels, with a high level of conflict within genera. Regarding deep Caprifoliaceae relationships, the results from the nuclear analyses (Figs. 4 and 5) showed multiple instances (at least two) of well-supported conflict with the results from the plastome

(Fig. 4), and the plastid results were largely consistent with previous plastid and large-scale analyses of Caprifoliaceae (Wang et al., 2020).

It is worth mentioning that Dipsacoideae was not recovered as monophyletic only in the ASTRAL species tree (Fig. 3), in which *Dipsacus japonicus* had a sister relationship with Linnaeoideae. Still, nodes with the strong LPP support also had low ICA and QS support values, which suggests that ILS and/or unidentified hybrid lineages continue to obscure our understanding of relationships in Dipsacoideae.

Previous studies reported that hybridization has shaped the evolutionary history of Caprifoliaceae (e.g., *Heptacodium miconioides*) (Zhang et al., 2003; Landrein et al., 2002). The conflict analyses of the nuclear dataset revealed strong signals of gene tree discordance among the seven major clades of Caprifoliaceae. The coalescent simulations also suggested that the observed cytonuclear discordance cannot be explained by ILS alone, which along with the phylogenetic network analyses point to several potential reticulation events along the backbone of Caprifoliaceae (Fig. 5). *Morina* and *Zabelia* are frequently involved in inferred reticulations, and the two Dipsacoideae are involved in 3 of the four events in the 11-taxa tree. The clade of *Abelia* + *Diabelia* are also involved in inferred reticulation events in the 9-taxa and 7-taxa networks. If these inferred events correspond to actual past instances of gene flow (which can only be confirmed by more detailed genomic analyses), it would help to explain the high amount of phylogenetic conflict observed in our analyses. There is some potential morphological support for ancient hybridization. For example, the leaves of *Morina* have stiff spines, while the leaves of *Zabelia* have no spines. In part because of this, Wang et al. (2015) suggested that *Zabelia* may be of allopolyploid origin.

4.2 Temporal divergence of Caprifoliaceae

Our estimated ages using nuclear and chloroplast trees are generally younger than those of Wang et al. (2015) and Wang et al. (2020) based on two reliable fossils (Li et al., 2019). We found that the diversification and global spread of the subfamilies of Caprifoliaceae occurred during the late Cretaceous, Paleocene and Eocene (Figs. 6, S10,

S11), similar to the results of Beaulieu et al. (2013). The divergence times of Caprifoliaceae have been estimated to be around the Cretaceous–Paleogene (K-Pg) boundary (Figs. 6, S10). Our results are congruent with the phenomena reported in several other plant groups such as Amaranthaceae *s.l.* (Morales-Briones et al. 2021) and legumes (Koenen et al., 2020), and in lichenized fungi such as Lobariaceae (Ascomycota) (Widhelm et al., 2019). It is generally accepted that because of the mass extinctions that occurred around the K-Pg boundary, new habitats became available and diverse organisms experienced rapid diversifications (Schulte et al., 2010). As a result of later tectonic movements and climate fluctuations from the Paleocene to the Eocene, major Caprifoliaceae lineages subsequently underwent rapid diversifications.

The divergence times among the major lineages of the Caprifoliaceae were dated to the Oligocene and Eocene, and within-genus diversification was dated to the Miocene and Pliocene (Figs. 6, S10, S11). Our results may be explained by the hypothesis that members of the Caprifoliaceae are well adapted to relatively cool environments (Friis, 1985; Manchester and Donoghue, 1995; Manchester, 2000), and an increase in the earth's temperature in the late Paleocene and early Eocene may have forced them to move to higher elevations or latitudes. As plants moved to higher elevations, their distribution was likely to be fragmented, resulting in isolation between populations. We have some evidence to support this hypothesis: (1) This family is mainly distributed in north temperate zones, and some genera even reach areas near the Arctic Circle (such as *Linnaea*); (2) there are numerous species (such as *Valeriana officinalis*, *Lonicera rupicola*, and *L. spinos*a) with island-like distributional patterns at relatively high elevations. As global climates cooled beginning in the late Eocene and especially the Oligocene and Miocene, cold-adapted survivors of warmer climates may have flourished and shifted into new geographic areas, especially mountainous areas, but may have struggled in northern regions during the Pliocene and Pleistocene glacial cycles (Moore & Donoghue, 2007). These global climatic events (e.g., ancient orogeny and monsoon-driven events) that might have driven diversification in Caprifoliaceae have also been reported in other taxa (Lu et al., 2018; Ding et al., 2020). For example, some genera or taxa with tiny, narrow or needle-like leaves (e.g., *Linnaea*, *Lonicera*

590 *myrtillus*) may have benefited from the global cooling and drying of the Miocene and
591 Pliocene by expanding their ranges, while other lineages more adapted to the wetter,
592 warmer parts of the world (*Abelia, Diabelia,* and *Dipelta*) may not have contracted
593 during the same time period.
594
595 **4.3 Evolution of morphological characters**
596 Character state reconstruction was conducted using ML (Figs. 8-10) because of the
597 potential hemiplasy and xenoplasy produced by the discordance and hybridization
598 detected in the nuclear backbone (Avies & Robinson, 2008; Robinson et al., 2008;
599 Copetti et al., 2017; Wang et al., 2020). A consequence of this discordance may be
600 elevated levels of apparent homoplasy in the species tree (Copetti et al., 2017; Hahn &
601 Nakhleh, 2017).
602 Stamen number, fruit type, style exertion, number of carpels, number of seeds and
603 epicalyx presence have been traditionally used for generic recognition within
604 Caprifoliaceae (Backlund, 1996; Donoghue et al., 2003; Yang & Landrein, 2011;
605 Landrein et al., 2020). Discordance among morphological traits might plausibly arise
606 due to either variable convergent selection pressures or other phenomena such as
607 hemiplasy. The evidence indicates that the probability of hemiplasy is high for four
608 morphological characters in Caprifoliaceae: the branch lengths leading to lineages with
609 derived character states are uniformly short with high levels of gene tree discordance. It
610 is possible that gene flow has contributed to these patterns. For example, the ancestral
611 stamen number states (i.e., 2 and 4) found in *Morina longifolia* and *Acanthocalyx alba*
612 within the Morinoideae clade could be due to introgressed alleles, as we identified
613 putative introgression between those lineages (Fig. 5). Morphological and anatomical
614 studies showed that the earliest Caprifoliaceae had monosymmetric flowers (probably
615 weakly so at first) with larger calyx lobes, tubular corollas, elongate styles, and capitate
616 stigmas (Donoghue et al., 2003). Within Caprifoliaceae, the main change in stamen
617 number is a reduction from five to four stamens. Subsequently, there was a reduction to
618 two stamens within Morinoideae and to three, two, and one within Valerianoideae
619 (Figs. 8 and S11). These variations may be related to an underlying change in floral

symmetry (Donoghue et al., 2013), which may relate to carpel abortion or to differences in the arrangement of flowers at the level of the inflorescence.

Our results suggest that multiple independent evolutionary events of the carpel evolution in Caprifoliaceae have occurred (Figs. 9 and S12). In Caprifoliaceae, the abortion of two of the three carpels and the development of a single ovule within the remaining fertile carpel was evidently correlated with fruit type (Wilkinson, 1949). For some subfamilies of Caprifoliaceae, carpel abortion occurs at a relatively late stage of ovary development, so many species have two empty chambers at fruit maturity (e.g., Linnaeoideae, Morinoideae, and Valerianoideae). In fact, in some species, these empty compartments have been co-opted in various ways in connection with dispersal (e.g., inflated for water dispersal in some *Valeriana*).

Caprifoliaceae shows great variation in fruit types. Fleshy, bird-dispersed fruits are limited to the Caprifolioideae (Donoghue et al., 2003). *Lonicera* has berries, though generally with just a few seeds embedded in copious pulp. Based on our analyses, it is important to note that the ancestral carpel number for Caprifoliaceae is most likely 3. There is programmed carpel abortion and the number of seeds corresponds to the number of fertile carpels (Donoghue et al., 2003). For *Symphoricarpos*, two of the four carpels abort, and there are two stones. The mesocarp in the cases is rather dry and mealy in texture. In the Caprifoliaceae, achenes with a single seed are present in *Heptacodium* and in the large Linnaeoideae clade (though in *Dipelta*, and in *Linnaea* there are two seeds at maturity). From the standpoint of fruit evolution, the linkage of *Heptacodium* within Caprifolioideae implies either the independent evolution of achenes or a transition from achenes to fleshy fruits in the line leading to Caprifolioideae. Among the achene-producing Caprifoliaceae, there are various adaptations for wind dispersal. One of the most striking of these modifications is enlargement of the calyx lobes into wings as the fruits mature (e.g. in *Abelia*, *Dipelta* and *Diabelia*). Especially well known is the production of a feathery pappus-like structure in species such as *Valeriana officianalis* and *Centranthus ruber* in Valerianoideae. This modification facilitates passive external transport by animals. A similar case is also found in *Kolkwitzia*.

650 The reconstruction of character evolution thus shows that some characters that
651 were once considered important for taxonomy within the family have been inferred to
652 be the results of homoplasious evolution (Gould 2000; Pyck, 2001; Bell 2001, 2004;
653 Carlson et al., 2009; Zhai et al., 2019). In analysis of character evolution, homoplasy is
654 regarded as noise that, if not properly accommodated, jeopardizes phylogenetic
655 reconstructions using morphological characters. At the same time, hemiplasy is one of
656 the causes of homoplasy (Copetti et al., 2017). The phenomenon of hemiplasy is most
657 plausible when the internodal distances in a phylogenetic tree are short (relative to
658 effective population sizes) (Robinson et al., 2008). Furthermore, the extensive
659 hybridization detected in the backbone of Caprifoliaceae might further contribute to
660 hemiplasy and xenoplasy (Wang et al. 2020). This may explain why it has been
661 difficult to reconstruct the relationships and character evolution among the major
662 lineages and genera of the family. Eventually, more extensive sampling and
663 developmental studies will be needed to elucidate the mechanisms underlying the
664 morphological evolutionary patterns outlined here.
665
666 **4.4 Recognition of Zabelioideae as a new subfamily in Caprifoliaceae**
667 Despite the strong signals of gene tree discordance, our nuclear and plastid
668 phylogenies strongly support seven major clades in Caprifoliaceae: Linnaeoideae,
669 *Zabelia*, Morinoideae, Valerianoideae, Dipsacoideae, Caprifolioideae and
670 Diervilloideae, and show *Zabelia* as the sister to the morphologically highly distinct
671 Morinoideae (Figs. 4 - 5). Our analyses support reticulate evolution concerning the
672 origins of both the *Zabelia* lineage as well as the Morinoideae. Based on the
673 phylogenomic and morphological analyses, we herein propose to recognize *Zabelia* as
674 representing a new subfamily of Caprifoliaceae.
675 **Zabelioideae B. Liu & S. Liu ex H.F. Wang, H.X. Wang, D.F. Morales-B, M.J.**
676 **Moore & J. Wen, subfam. nov.**
677 **Type: *Zabelia* (Rehder) Makino.**
678 **Description:** Shrubs, deciduous; old branches often with six deep longitudinal
679 grooves. Leaves opposite, entire or dentate at margin; estipulate; petioles of opposite

680 leaf pairs dilated and connate at base, enclosing axillary buds. Inflorescence a
681 congested thyrse of cymes; cymes 1-3-flowered. Calyx 4- or 5-lobed, persistent,
682 spreading. Corolla 4- or 5-lobed, hypocrateriform, ± zygomorphic; corolla tube
683 cylindrical. Stamens 4, included, didynamous. Ovary 3-locular, 2 locules with 2 series
684 of sterile ovules and 1 locule with a single fertile ovule; stigmas green, capitate,
685 mucilaginous. Fruit an achene crowned with persistent and slightly enlarged sepals.
686 Basic chromosome number $x = 9$.

687 One genus and six species distributed in China, Japan, Korea, Afghanistan, NW
688 India, Kyrgyzstan, Nepal, and Russian Far East.

689 Zabelioideae is highly distinct morphologically from its sister Morinoideae. They
690 can be easily distinguished by their habit (with Zabelioideae as shrubs, and
691 Morinoideae as herbs), the six distinct, longitudinal grooves on twigs and branches of
692 Zabelioideae (the six grooves absent in Morinoideae), and the epicalyx (absent in
693 Zabelioideae and present in Morinoideae). Zabelioideae and Morinoideae share some
694 similarities in pollen micromorphology, as both have psilate pollen grains with an
695 endocingulum (Verlaque, 1983; Kim et al., 2001; Jacobs et al., 2011). The two
696 subfamilies diverged in the early-mid Eocene (Figs. 6, S7), and their long evolutionary
697 history associated with deep hybridization events, ILS and extinctions likely have made
698 it difficult to determine their phylogenetic placements.
699

700 **5 Conclusions**

701 Gene tree discordance has been commonly observed in phylogenetic studies.
702 Moreover, species tree estimation has been shown to be inconsistent in the presence of
703 gene flow (Solís-Lemus et al., 2016; Long & Kubatko, 2018), which suggests that both
704 ILS and gene flow simultaneously need to be considered in constructing phylogenetic
705 relationships. Here, our results show clear evidence of cytonuclear discordance and
706 extensive conflict between individual gene trees and species trees in Caprifoliaceae.
707 We also show that there has been widespread hybridization and/or introgression
708 among the major clades of Caprifoliaceae, which can explain most the gene tree
709 conflict and the long history of phylogenetic uncertainty in the family. Furthermore,

710 the temporal diversification of Caprifoliaceae provides a good case to support the
711 evolutionary radiation of a dominantly north temperate plant family in response to
712 climatic changes from the late Cretaceous to the present. Finally, based on evidence
713 from molecular phylogeny, divergence times, and morphological characters, we herein
714 recognize the *Zabelia* clade as representing a new subfamily, Zabelioideae, in
715 Caprifoliaceae. The phylogenetic framework presented here also sheds important
716 insights into character evolution in Caprifoliaceae.
717

718 **Acknowledgements**
719 The work was funded by National Scientific Foundation of China (31660055). We
720 thank Gabriel Johnson for his help with the target enrichment experiment, and the
721 United States National Herbarium for permission to sample some collections. We
722 acknowledge the staff in the Laboratories of Analytical Biology at the National
723 Museum of Natural History, the Smithsonian Institution for support and assistance.
724

725 **Author contributions**
726 H.F.W. and J.W. conceived the study. H.F.W. and D.F.M-B. performed the research
727 and analyzed the data. H.X.W., H.F.W., D.F.M-B., J.W. and M.J.M wrote and revised
728 the manuscript.
729

730 **ORCID**
731 **Hong-Xin Wang:** https://orcid.org/0000-0002-2283-7368
732 **Diego F. Morales-Briones:** https://orcid.org/0000-0003-1535-5739
733 **Michael J. Moore**: https://orcid.org/0000-0003-2222-8332
734 **Jun Wen:** https://orcid.org/0000-0001-6353-522X
735 **Hua-Feng Wang**: https://orcid.org/0000-0003-3331-2898
736
737

738 **References**

739 Abbott RJ, Hegarty MJ, Hiscock SJ, Brennan AC. 2010. Homoploid hybrid
740 speciation in action. *Taxon* 59: 1375–1386.
741 Akaike H. 1973. Information theory and an extension of the maximum likelihood
742 principle. In: Petrov BN, Csaki F, eds. *Second international symposium on*
743 *information theory*. Budapest, Hungary: Akademiai Kiado, 267–281.
744 APG, 1998. An ordinal classification for the families of flowering plants. *Annals of the*
745 *Missouri Botanical Garden* 85: 531–553.
746 APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the
747 orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean*
748 *Society* 181: 1–20.
749 Avise JC, Robinson TJ. 2008. Hemiplasy: A New Term in the Lexicon of
750 Phylogenetics. *Systematic biology* 57: 503-507.
751 Backlund, A. 1996. Phylogeny of the order Dipsacales. Ph. D. Dissertation.
752 Uppsala: Uppsala University.
753 Backlund A, Pyck N. 1998. Diervillaceae and Linnaeaceae, two new families of
754 caprifolioids. *Taxon* 47: 657–661.
755 Beaulieu JM, Tank DC, Donoghue MJ. 2013**.** A Southern Hemisphere origin for
756 campanulid angiosperms, with traces of the break-up of Gondwana. *BMC*
757 *Evolutionary Biology*. 13: 80.
758 Bell CD. 2004. Preliminary phylogeny of Valerianaceae (Dipsacales) inferred from
759 nuclear and chloroplast DNA sequence data. *Molecular Phylogenetics and*
760 *Evolution* 31: 340–350.
761 Bell, CD, Donoghue MJ. 2005. Dating the Dipsacales: comparing models, genes, and
762 evolutionary implications. *American Journal of Botany* 92: 284–296.
763 Bell CD, Edwards EJ, Kim ST, Donoghue MJ. 2001. Dipsacales phylogeny based on
764 chloroplast DNA sequences. *Harvard Papers in Botany* 6:481–499.
765 Blackmore S, Cannon MJ. 1983. Palynology and systematics of Morinaceae. *Rev*
766 *Palaeobot Palynol* 40: 207–266.

767 Bolger AM, Lohse, M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina
768 Sequence Data. *Bioinformatics* (*Oxford, England*) 30: 2114–2120.
769 Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut
770 A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary
771 analysis. *PLoS Computational Biology* 10: e1003537.
772 Caputo G, Cozzolino S.1994. A cladistic analysis of Dipsacaceae (Dipsacales). *Plant*
773 *Systematics and Evolution* 189: 41–61.
774 Carlson SE, Mayer V, Donoghue MJ. 2009. Phylogenetic relationships, taxonomy, and
775 morphological evolution in Dipsacaceae (Dipsacales) inferred by DNA sequence
776 data. *Taxon* 58: 1075–1091.
777 Cavender-Bares J, Gonzalez-Rodriguez A, Eaton DAR, Hipp AAL, Beulke
778 A, Manos PS. 2015. Phylogeny and biogeography of the American live oaks
779 (*Quercus* subsection *Virentes*): A genomic and population genetics approach.
780 *Molecular Ecology* 24: 3668–3687.
781 Chamala S, García, N, Godden GT, Krishakumar V, Jordon-Thaden IE, Smet RD et al.
782 2015. MarkerMiner 1.0: A new application for phylogenetic marker development
783 using angiosperm transcriptomes. *Applications in Plant Sciences* 3: 1400115.
784 Copetti D, Búrquez A, Bustamante E, Charboneau JLM, Childs KL, Eguiarte LE, Lee
785 S, Liu TL, McMahon MM, Whiteman NK, Wing RA, Wojciechowski MF,
786 Sanderson MJ. 2017. Extensive gene tree discordance and hemiplasy shaped the
787 genomes of North American columnar cacti. *Proceedings of the National*
788 *Academy of Sciences* 114: 12003-12007.
789 Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and
790 the multispecies coalescent. *Trends in Ecology and Evolution* 24: 332–340.
791 Ding WN, Ree RH, Spicer RA, Xing YW. 2020. Ancient orogenic and monsoon-driven
792 assembly of the world's richest temperate alpine flora. *Science* 369: 578–581.
793 Donoghue MJ, Bell CD, Winkworth RC. 2003. The evolution of reproductive
794 characters in Dipsacales. *International Journal of Plant Sciences* 164:
795 S453–S464.

796 Donoghue MJ, Eriksson T, Reeves PA, Olmstead RG. 2001. Phylogeny and
797 phylogenetic taxonomy of Dipsacales, with special reference to *Sinadoxa* and
798 *Tetradoxa* (Adoxaceae). *Harvard Papers in Botany* 6: 459–479.
799 Donoghue MJ, Olmstead RG, Smith JF, Palmer JD. 1992. Phylogenetic relationships
800 of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden*
801 79: 333–345.
802 Doyle JJ, Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh
803 leaf tissue. *Phytochemical Bulletin* 19: 11–15.
804 Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with
805 BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.
806 Edwards SV. 2009. Is a new and general theory of molecular systematics emerging?
807 *Evolution* 63: 1–19.
808 Feng Y, Peter Comes H, Qiu YX. 2020. Phylogenomic insights into the
809 temporal-spatial divergence history, evolution of leaf habit and hybridization in
810 *Stachyurus* (Stachyuraceae). *Molecular Phylogenetics and Evolution* doi:
811 https://doi.org/10.1016/j.ympev.2020.106878.
812 Friis EM. 1985. Angiosperm fruits and seeds from the Middle Miocene of Jutland
813 (Denmark). *Biologiske Skrifter* 24:1–165.
814 Folk RA, Mandel JR, Freudenstein JV. 2017. Targeted enrichment of intronic sequence
815 markers for recent radiations: a phylogenomic example from *Heuchera*
816 (Saxifragaceae). *Applications in Plant Sciences* 3: 1500039.
817 Galtier N, Daubin V. 2008. Dealing with incongruence in phylogenomic analyses.
818 *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:
819 4023–4029.
820 Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss
821 SJ, Osborne EJ, Sreedharan VT. et al. 2011. Multiple reference genomes and
822 transcriptomes for *Arabidopsis thaliana*. *Nature* 477: 419–423.
823 García N, Folk RA, Meerow AW, Chamala S, Gitzendanner MS, de Oliveira RS, Soltis
824 DE, Soltis PS. 2017. Deep reticulation and incomplete lineage sorting obscure the

825 diploid phylogeny of rain-lilies and allies (Amaryllidaceae tribe Hippeastreae).
826 *Molecular Phylogenetics and Evolution* 111: 231–247.
827 Gould KR, Donoghue MJ. 2000. Phylogeny and biogeography of *Triosteum*
828 (Caprifoliaceae). *Harvard Papers in Botany* 5:157–166.
829 Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes
830 directly from genomic next-generation sequencing reads-a baiting and iterative
831 mapping approach. Nucleic Acids Research 41, e129.
832 Hahn MW, Nakhleh L. 2017. Irrational exuberance for resolved species trees.
833 *Evolution* 70:7–17.
834 Hara H. 1983. A Revision of the Caprifoliaceae of Japan with Reference to Allied
835 Plants in Other Districts and the Adoxaceae. *Academia Scientific Books, Tokyo*.
836 Jacobs B, Geuten K, Pyck N, Huysmans S, Jansen S, Smets E. 2011. Unraveling
837 the phylogeny of *Heptacodium* and *Zabelia* (Caprifoliaceae): an interdisciplinary
838 approach. *Systematic Botany* 36: 231–252.
839 Jacobs B, Pyck N, Smets E. 2010. Phylogeny of the *Linnaea* clade: are *Abelia* and
840 *Zabelia* closely related? *Molecular Phylogenetics and Evolution* 57: 741–752.
841 Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega
842 NJC, Wickett NJ. 2016. HybPiper: extracting coding sequence and introns for
843 phylogenetics from high-throughput sequencing reads using target enrichment.
844 *Applications in Plant Sciences* 4: 1600016.
845 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software Version
846 7: improvements in performance and usability. *Molecular Biology and Evolution*
847 30: 772–780.
848 Kim T. 1998. Phylogenetic studies of tribe Linnaeeae (Caprifoliaceae). *Ph. D Thesis,*
849 *Chonbuk National University.*
850 Kim TJ, Sun BY, Suh Y. 2001. Palynology and cytotaxonomy of the genus *Abelia s.l.*,
851 Caprifoliaceae. *Korean Journal of Plant Taxonomy* 31: 91–106.
852 Koenen EJM, Ojeda DI, Bakker FT, Wieringa JJ, Catherine K, Hardy OJ, Pennington
853 RT, Herendeen PS, Bruneau A, Hughes CE. 2020. The origin of the legumes is a

854 complex paleopolyploid phylogenomic tangle closely associated with the
855 cretaceous-paleogene (K-Pg) mass extinction event. *Systematic Biology* syaa041.
856 Konowalik K, Wagner F, Tomasello S, Vogt R, Oberprieler C. 2015. Detecting
857 reticulate relationships among diploid Leucanthemum Mill. (Compositae,
858 Anthemideae) taxa using multilocus species tree reconstruction methods and
859 AFLP fingerprinting. *Molecular Phylogenetics and Evolution* 92: 308–328.
860 Landrein S. 2010. *Abelia* and its relatives. *Plantsman, NS* 9: 40–47.
861 Landrein S, Prenner G. 2013. Unequal twins? Inflorescence evolution in the twinflower
862 tribe Linnaeeae (Caprifoliaceae *s.l.*). *International Journal of Plant Sciences* 174:
863 200–233.
864 Landrein S, Prenner G. 2016. Structure, ultrastructure and evolution of floral nectaries
865 in the twinflower tribe Linnaeeae and related taxa (Caprifoliaceae). *Botanical*
866 *Journal of Linnean Society* 181: 37–69.
867 Landrein S, Prenner G, Chase MW, Clarkson JJ. 2012. *Abelia* and relatives:
868 Phylogenetics of Linnaeeae (Dipsacales Caprifoliaceae *s.l.*) and a new
869 interpretation of their inflorescence morphology. *Botanical Journal of Linnean*
870 *Society* 169: 692–713.
871 Lańcucka-Rodoniowa M. 1967. Two new genera: *Hemiptelea* Planch. and *Weigela*
872 Thumb. in the younger Tertiary of Poland. *Acta Palaeobotanica* 8: 1–17.
873 Lee-Yaw JA, Grassa CJ, Joly S, Andrew RL, Rieseberg LH. 2019. An evaluation of
874 alternative explanations for widespread cytonuclear discordance in annual
875 sunflowers (*Helianthus*). *New Phytologist* 221: 515–526.
876 Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete
877 morphological character. *Systematic Biology* 50: 913–925.
878 Li HT, Yi TS, Gao LM, Ma PF, Zhang T, Yang JB, Gitzendanner MA, Fritsch PW, Cai
879 J, Luo Y, Wang H, Bank MV, Zhang SD, Wang QF, Wang J, Zhang ZR, Fu CN,
880 Yang J, Hollingsworth PM, Chase MW, Soltis DE, Soltis PS, Li DZ. 2019. Origin
881 of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* 5: 461–470.
882 Liu L, Xi Z, Wu S, Davis CC, Edwards SV. 2015. Estimating phylogenetic trees from
883 genome-scale data. *Annals of the New York Academy of Sciences* 1360: 36–53.

884 Lin HY, Hao YJ, Li JH, Fu CX, Soltis PS, Soltis DE, Zhao YP. 2019. Phylogenomic
885 conflict resulting from ancient introgression following species
886 diversification in *Stewartia s.l.* (Theaceae). *Molecular Phylogenetics and*
887 *Evolution* 135: 1–11.
888 Linder HP, Naciri Y. 2015. Species delimitation and relationships: the dance of the
889 seven veils. *Taxon* 64: 3–16.
890 Long C, Kubatko L. 2018. The Effect of Gene Flow on Coalescent-based
891 Species-Tree Inference. *Systemmatic biology* 67: 770-785.
892 Lu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL, Sun M, Miller JT, Mathews S, Hu HH,
893 Niu YT, Peng DX, Chen YH, Smith SA, Chen M, Xiang KL, Le CT, Dang VC, Lu
894 AM, Soltis PS, Soltis DE, Li JH, Chen ZD. 2018. Evolutionary history of the
895 angiosperm flora of china. *Nature* 554: 234–238.
896 Mallet J. 2007. Hybrid speciation. *Nature* 466: 279–283.
897 Maddison WP, Maddison DR. 2018. Mesquite: a modular system for evolutionary
898 analysis. Version 3.51 http://www.mesquiteproject.org.
899 Manchester SR. 2000. Late Eocene fossil plants of the John Day Formation, Wheeler
900 County, Oregon. *Oregon Geology* 62: 51–63.
901 Manchester SR, Donoghue MJ. 1995. Winged fruits of Linnaeeae (Caprifoliaceae) in
902 the Tertiary of western North America: *Diplodipelta* gen. nov. *International*
903 *Journal of Plant Science* 156: 709–722.
904 Moore BR, Donoghue MJ. 2007. Correlates of diversification in the plant clade
905 Dipsacales: Geographic movement and evolutionary innovations. *American*
906 *Naturalist* 170: S28–S55.
907 Morales-Briones DF, Liston A, Tank DC. 2018. Phylogenomic analyses reveal a deep
908 history of hybridization and polyploidy in the Neotropical genus *Lachemilla*
909 (Rosaceae). *New Phytologist* 218: 1668–1684.
910 Morales-Briones DF, Gehrke B, Huang CH, Liston A, Ma H, Marx HE, Tank DC,
911 Yang Y. 2020. Analysis of paralogs in target enrichment data pinpoints multiple
912 ancient polyploidy events in *Alchemilla s.l.* (Rosaceae). *BioRxiv*.
913 *https://doi.org/10.1101/2020.08.21.261925*.

914 Morales-Briones DF, Kadereit G,Tefarikis DT, Moore MJ, Smith SA,
915 Brockington SF, Timoneda A, et al. 2021. Disentangling sources of gene tree
916 discordance in phylotranscriptomic datasets: a case study from Amaranthaceae *s.l.*
917 *Systematic Biology* 70: 219-235.
918 Pease JB, Brown JW, Walker JF, Hinchliff CE, Smith SA. 2018.
919 Quartet Sampling distinguishes lack of support from conflicting support
920 in the green plant tree of life. *American Journal of Botany* 105: 385–403.
921 Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and*
922 *Evolution* 25:1253–1256.
923 Pyck N. 2001. Phylogenetic relationships within Dipsacales: a combined molecular and
924 morphological approach. PhD thesis. Katholieke University, Louvain, Belgium.
925 Robinson TJ, Ruiz-Herrera A, Avise JC. 2008. Hemiplasy and homoplasy in the
926 karyotypic phylogenies of mammals. *Proceedings of the National Academy of*
927 *Sciences* 105: 14477-14481.
928 Salichos L, Stamatakis A, Rokas A. 2014. Novel information theory-based measures
929 for quantifying incongruence among phylogenetic trees. *Molecular Biology and*
930 *Evolution* 31: 1261–1271.
931 Sayyari E, Mirarab S. 2016. Fast coalescent-based computation of local branch support
932 from quartet frequencies. *Molecular Biology and Evolution* 33: 1654–1668.
933 Schwarz G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:
934 461–464.
935 Schulte P, Alegret L, Arenillas I, Arz JA, Barton PJ, Bown PR, Bralower TJ, Christeson
936 GL, Claeys P, Cockell CS, Collins GS, Deutsch A, Goldin TJ, Goto K,
937 Grajales-Nishimura M, et al. 2010. The Chicxulub asteroid impact and mass
938 extinction at the Cretaceous-Paleogene boundary. *Science* 327: 1214–1218.
939 Smith SA, Beaulieu JM, Donoghue MJ. 2010. An uncorrelated relaxed clock analysis
940 suggests an earlier origin for flowering plants. *Proceedings of the National*
941 *Academy of Sciences of the United States of America* 107: 5897–5902.
942 Smith SA, Dunn CW. 2008. Phyutility: a phyloinformatics tool for trees,
943 alignments and molecular data. *Bioinformatics* 24: 715–716.

944 Smith SA, Moore MJ, Brown JW, Yang Y. 2015. Analysis of phylogenomic datasets
945 reveals conflict, concordance, and gene duplications with examples from animals
946 and plants. *BMC Evolutionary Biology* 15: 150.
947 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post
948 analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
949 Stevens PF. 2001 onwards. Angiosperm Phylogeny Website. Version Apr. 15, 2019
950 http://www.mobot.org/MOBOT/research/APweb/.
951 Solís-Lemus C, Ané C. 2016. Inferring phylogenetic networks with maximum
952 pseudolikelihood under incomplete lineage sorting. *PLOS Genetics* 12:
953 e1005896.
954 Solís-Lemus C, Yang MY, Ané C. 2016. Inconsistency of Species Tree Methods under
955 Gene Flow. *Systematic Biology* 65: 843-851.
956 Stull GW, Soltis PS, Soltis DE, Gitzendanner MA, Smith SA. 2020. Nuclear
957 phylogenomic analyses of asterids conflict with plastome trees and support novel
958 relationships among major lineages. *American Journal of Botany* 107: 1–16.
959 Sugiura N. 1978. Further analysis of the data by Akaike's information criterion and
960 the finite corrections. *Communications in Statistics - Theory and Methods A* 7:
961 13–26.
962 Sukumaran J, Holder MT. 2010. DendroPy: A Python library for phylogenetic
963 computing. *Bioinformatics* 26: 1569–1571.
964 Sun M, Soltis DE, Soltis PS, Zhu X, Burleigh JG, Chen Z. 2015. Deep phylogenetic
965 incongruence in the angiosperm clade Rosidae. *Molecular Phylogenetics and*
966 *Evolution* 83: 156–166.
967 Szöllősi GJ, Tannier E, Daubin V, Boussau B. 2015. The inference of gene trees with
968 species trees. *Systematic Biology* 64: e42–e62.
969 Tang YC, Lu AM. 2005. Paraphyletic group, PhyloCode and phylogenetic species
970 – the current debate and a preliminary commentary. *Acta Phytotaxonomica*
971 *Sinica* 43: 403–419.

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

Than C, Ruths D, Nakhleh L. 2008. A software package for analyzing and reconstructing reticulate evolutionary histories. *BMC Bioinformatics* 9: 322

Traiser C, Klotz S, Uhl D, Mosbrugger V. 2005. Environmental signals from leaves a physiognomic analysis of European vegetation. *New Phytologist* 166: 465–484.

Vargas OM, Heuertz M, Smith SA, Dick CW. 2019.Target sequence capture in the Brazil nut family (Lecythidaceae): Marker selection and in silico capture from genome skimming data. *Molecular Phylogenetics and Evolution* 135: 98–104.

Vargas OM, Ortiz EM, Simpson BB. 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytologist* 214: 1736–1750.

Verlaque R. 1983. Contribution l'étude du genre *Morina* L. *Pollen ET Spores* 25: 143–162.

Wang HF, Landrein S, Dong WP, Nie ZL, Kondo K, Funamoto T, Wen J, Zhou SL. 2015. Molecular phylogeny and biogeographic diversification of Linnaeoideae (Carprifoliaceae *s.l.*) disjunctively distributed in Eurasia, North American and Mexico. *PLoS One* 10: e0116485.

Wang HX, Liu H, Moore MJ, Landrein S, Liu B, Zhu ZX, Wang HF. 2020. Plastid phylogenomic insights into the evolution of the Caprifoliaceae *s.l.* (Dipsacales). *Molecular Phylogenetics and Evolution* 142: 106641.

Wang YX, Cao Z, Ogilvie HA, Nakhleh L. 2020. Phylogenomic assessment of the role of hybridization and introgression in trait evolution. BioRxiv preprint doi: https://doi.org/10.1101/2020.09.16.300343.

Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A. 2014. Hyb-Seq: combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences* 2: 9.

Wen D, Yu Y, Zhu J, Nakhleh L. 2018. Inferring phylogenetic networks using PhyloNet. *Systematic Biology* 67: 735–740.

1002 Widhelm TJ, Grewe F, Huang JP, Mercado-Díaz JA, Gofnet B, Lücking R, Moncada
1003 B, Mason-Gamer R, Lumbsch HT. 2019. Multiple historical processes obscure
1004 phylogenetic relationships in a taxonomically difficult group (Lobariaceae,
1005 Ascomycota). *Scientific Reports* 9: 8968.
1006 Xiang CL, Dong HJ, Landrein S, Zhao F,Yu WB, Soltis DE, Soltis PS, Backlund A,
1007 Wang HF, Li DZ, Peng H. 2020. Revisiting the phylogeny of Dipsacales: new
1008 insights from phylogenomic analyses of complete chloroplast genome sequences.
1009 *Journal of Systematics and Evolution* 58: 103–117.
1010 Xiang YZ, Huang CH, Hu YH, Wen J, Li SS, Yi TS, Chen HY, Xiang J, Ma H. 2016.
1011 Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of
1012 geological times and genome duplication. *Molecular Biology and Evolution* 34:
1013 262-281.
1014 Yakimowski SB, Rieseberg LH. 2014. The role of homoploid hybridization in
1015 evolution: a century of studies synthesizing genetics and ecology. *American*
1016 *Journal of Botany* 101: 1247–1258.
1017 Yang QR, Landrein S. 2011. Caprifoliaceae. In: Wu Z, Hong D, Raven PH eds. *Flora*
1018 *of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 19:
1019 617–618.
1020 Yang Y, Smith SA. 2014. Orthology inference in nonmodel organisms
1021 using transcriptomes and low-coverage genomes: improving accuracy and
1022 matrix occupancy for phylogenomics. *Molecular Biology and Evolution* 31:
1023 3081–3092.
1024 Yu Y, Degnan JH, Nakhleh L. 2012. The probability of a gene tree topology within a
1025 phylogenetic network with applications to hybridization detection. *PLoS Genetics*
1026 8: e1002660.
1027 Zhai W, Duan X, Zhang R, Guo C, Li L, Xu G, Shan H, Kong H, Ren Y. 2019.
1028 Chloroplast Genomic Data Provide New and Robust Insights into the Phylogeny
1029 and Evolution of the Ranunculaceae. *Molecular Phylogenetics and Evolution*
1030 135: 12–21.

1031 Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time
1032 species tree reconstruction from partially resolved gene trees. *BMC*
1033 *Bioinformatics*. 19: 523.
1034 Zhang WH, Chen ZD, Li JH, Chen HB, Tang YC. 2003. Phylogeny of the Dipsacales
1035 *s.l.* based on chloroplast *trnL-F* and *ndhF* sequences. *Molecular Phylogenetics*
1036 *and Evolution* 26: 176–189.
1037
1038

Fig. 1. Floral diversity of Dipsacales. (A) *Kolkwitzia amabilis*; (B) *Zabelia integrifolia*; (C) *Scabiosa comosa* (D) *Valeriana flaccidissima*; (E) *Acanthocalyx nepalensis* subsp. *delavayi*; (F) *Lonicera fragrantissima* var. *lancifolia*; (G) *Weigela coraeensis;* (H) *Viburnum opulus* subsp. *calvescens.*

Fig. 2. Alternative relationships for the Caprifoliaceae backbone based on previous analyses. (A) Donoghue et al. (2001); parsimony analyses based on chloroplast *rbcL* sequences and morphological characteristics; (B) Bell et al. (2001); maximum likelihood tree from the combined chloroplast DNA data; (C) Zhang et al. (2003); maximum likelihood tree based on chloroplast *trnL-F* and ndhF sequences; (D) Jacobs et al. (2010); maximum parsimony Dipsacales phylogeny based on nuclear and chloroplast sequence data; (E) Wang et al.(2020); maximum likelihood tree based on 68 complete plastomes. (F) This study; species tree based on nuclear concatenated data set.

Fig. 3. Species tree of the nuclear dataset inferred with ASTRAL-□. Local posterior probabilities and internode certainty all (ICA) scores are shown above branches for main clades. Pie charts next to the nodes present the proportion of congruent gene trees that supports that clade (blue), the proportion of discordant gene trees of the main alternative topology for that clade (green), the proportion of discordant trees for the remaining alternative topologies (red), dark gray represents the proportion of uninformative gene trees (bootstrap support < 50%), and light gray is the proportion of missing data. Numbers above branches indicate (LPP) / ICA score / number concordant gene trees / number of all discordant gene trees. Major taxonomic groups or main clades in the family as currently recognized are indicated by branch colors as a visual reference to relationships.

Fig. 4. Tanglegram of the nuclear concatenated (left) and plastid (right) phylogenies. Dotted lines connect taxa between the two phylogenies. Maximum likelihood bootstrap support values are shown above branches. The asterisks indicate maximum likelihood bootstrap support of 100%. Major taxonomic groups or main clades in the family as currently recognized are indicated by branch colors as a visual reference to relationships.

Fig. 5. Best supported species networks inferred with Phylonet for the (a) 11-taxa, (b) 9-taxa, and (c) 7-taxa data sets. Numbers next to the inferred hybrid branches indicate inheritance probabilities. Red lines represent minor hybrid edges (edges with an inheritance contribution < 0.50)

Fig. 6. BEAST analysis of divergence times based on the nuclear data set, under Analysis I. Calibration points are indicated by A, B. and C. Numbers 1–11 represent major divergence events in Caprifoliaceae; mean divergence times and 95% highest posterior densities are provided for each nodes of interests.

Fig. 7. Maximum likelihood inference of character evolution in Caprifoliaceae based on the plastid matrix. Left, Number of stamens; Right, Style exertion.

Fig. 8. Maximum likelihood inference of character evolution in Caprifoliaceae based on the plastid matrix. Left, fruit type; Right, Number of carpels.

Fig. 9. Maximum likelihood inference of character evolution in Caprifoliaceae based on the plastid matrix. Left, number of seeds; Right, epicalyx presence/absence.

Fig. S1. Simplified ML tree generated from the nuclear gene data showing the distribution of selected character states. The asterisks indicate maximum likelihood bootstrap support of 100%.

Figure. S2. Heatmaps showing gene recovery efficiency for the nuclear gene in 46 species of Caprifoliaceae. Columns represent genes, and each row is one sample. Shading indicates the percentage of the reference locus length coverage.

Fig. S3. ASTRAL-III species tree. Numbers above branches indicate the number of gene trees concordant/conflicting with that node in the species tree. Numbers below the branches are the Internode Certainty All (ICA) score. Pie charts next to the nodes present the proportion of gene trees that supports that clade (blue), the proportion that supports the main alternative for that clade (green), the proportion that supports the remaining alternatives (red), light gray means missing data, and dark gray mean uninformative (BS < 50%).

Fig. S4. Results of the Quartet Sampling of the ASTRAL tree. Node labels indicate QC/Quartet Differential (QD)/Quarte Informativeness (QI) scores.

Fig. S5. Results of the Quartet Sampling of the nuclear concatenated RAxML tree. Node labels indicate QC/Quartet Differential (QD)/Quarte Informativeness (QI) scores.

Fig. S6. Maximum likelihood cladogram of Caprifoliaceae inferred from RAxML analysis of the concatenated 713-nuclear gene supermatrix. Numbers above branches indicate the number of gene trees concordant/conflicting with that node in the species tree. Numbers below the branches are the Internode Certainty All (ICA) score. Pie charts next to the nodes present the proportion of gene trees that supports that clade (blue), the proportion that supports the main alternative for that clade (green), the proportion that supports the remaining alternatives (red), light gray means missing data, and dark gray mean uninformative (BS < 50%).

Fig. S7. Results of the Quartet Sampling of the chloroplast tree. Node labels indicate QC/Quartet Differential (QD)/Quarte Informativeness (QI) scores.

Fig. S8. Phylogeny of the plastid DNA dataset; numbers above branches represent clade frequencies of the simulated gene trees.

Fig. S9. Best supported species networks inferred with Phylonet for the (a) 11-taxa, (b) 9-taxa, and (c) 7-taxa data sets. Blue branches connect the hybrid nodes. Numbers next to the hybrid branches indicate inheritance probabilities.

Fig. S10. BEAST analysis of divergence times based on the nuclear data. Calibration points are indicated by A, B. and C. Numbers 1–11 represent major divergence events in Caprifoliaceae; mean divergence times and 95% highest posterior densities are provided for each. □, □ and □ indicate the three analyses that varied in the placement of the *Diplodipelta* fossil constraint.

Fig. S11. BEAST analysis of divergence times based on the cpDNA data. Calibration points are indicated by A, B, and C. Numbers 1–10 represent major divergence events in Caprifoliaceae; mean divergence times and 95% highest posterior densities are provided for each. □, □ and □ indicate the three analyses that varied in the placement of the *Diplodipelta* fossil constraint.

Fig. S12. Maximum likelihood inference of character evolution in Caprifoliaceae based on the nuclear matrix. Left, Number of stamens; Right, Style exertion.

Fig. S13. Maximum likelihood inference of character evolution in Caprifoliaceae based on the nuclear matrix. Left, Style of fruit; Right, Number of carpels.

Fig. S14. Maximum likelihood inference of character evolution in Caprifoliaceae based on the nuclear matrix. Left, number of seeds; Right, epicalyx presence/absence.

Table 1 Dataset statistics, including the number of taxa, number of characters, number of PI characters, missing data.

Alignment	No. of	No. sites	Missing data (%)	No. of variable/ Parsimony	ML analysis

	taxa			informative sites	
Nuclear	46	343,609	34	144,517/96,479	GTR+G
cpDNA	46	208,607	25	55,059/32,960	GTR+G+I

Table 2 Model selection of different species networks and bifurcating trees.

Topology	lnL	Parameters	Loci	Number of hybridizations	Information criteria		
					AIC	AICc	BIC
11 taxa							
nuclear ASTRAL	-2834.940114	19	178	N/A	5707.880227	5712.690354	5768.334115
nuclear RAxML	-2848.731017	19	178	N/A	5735.462034	5740.272161	5795.915922
cpDNA	-3004.815841	19	178	N/A	6047.631682	6052.441809	6108.08557
Network 1	-2745.311471	21	178	1	5532.622943	5538.54602	5599.440397
Network 2	-2639.120484	23	178	2	5324.240969	5331.4098	5397.421991
Network 3	-2628.091812	25	178	3	5306.183624	5314.736256	5385.728213
Network 4	**-2571.727153**	27	**178**	**4**	**5197.454305**	**5207.534305**	**5283.362461**
Network 5	-2626.829726	27	178	4	5307.659452	5317.739452	5393.567608
9 taxa							
nuclear ASTRAL	-3436.346996	15	328	N/A	6902.693992	6904.232453	6959.589196
nuclear RAxML	-3438.060346	15	328	N/A	6906.120691	6907.659153	6963.015895
cpDNA	-3489.910056	15	328	N/A	7009.820111	7011.358573	7066.715316
Network1	-3324.554786	17	328	1	6683.109571	6685.083765	6747.590803
Network2	-3321.944353	19	328	2	6681.888706	6684.356239	6753.955965
Network3	-3161.006574	21	328	3	6364.013147	6367.032755	6443.666433
Network4	-3189.59271	21	328	3	6421.185421	6424.205028	6500.838706
Network 5	**-3037.728627**	**21**	**328**	**3**	**6117.457254**	**6120.476862**	**6197.11054**
7 taxa							
nuclear ASTRAL	-2756.695803	11	496	N/A	5535.391606	5535.937061	5581.663942
nuclear RAxML	-2762.789144	11	496	N/A	5547.578288	5548.123743	5593.850623

cpDNA	-2801.967642	11	496	N/A	5625.935285	5626.480739	5672.20762
Network1	-2750.416642	13	496	1	5526.833285	5527.588472	5581.518772
Network2	-2750.618801	13	496	1	5527.237602	5527.992789	5581.923089
Network3	-2746.173709	17	496	3	5526.347418	5527.627753	5597.859209
Network4	**-2601.483732**	**17**	**496**	**3**	**5236.967465**	**5238.247799**	**5308.479255**
Network 5	-2671.907806	17	496	3	5377.815611	5379.095946	5449.327402

Table S1. List of species and vouchers used in this study.

	Family	Subfamily	Taxon	Locality	Voucher specimen number	Serial number	SRA accession
1	Caprifoliaceae	Linnaeoideae	*Dipelta floribunda* Maximowicz Bull.	Longnan, Gansu, China.	HUTB, E57	E57	SRR13705796.
2	Caprifoliaceae	Linnaeoideae	*Dipelta floribunda* Maximowicz Bull.	Longnan, Gansu, China.	HUTB, E56	E56	SRR13705795
3	Caprifoliaceae	Linnaeoideae	*Dipelta floribunda* Maximowicz Bull.	Longnan, Gansu, China.	HUTB, E55	E55	SRR13705794
4	Caprifoliaceae	Linnaeoideae	*Diabelia serrata* (Siebold et Zucc.) Landrein	Anan, Tokushima, Japan	HUTB, E123	E123	SRR13705758
5	Caprifoliaceae	Linnaeoideae	*Diabelia ionostachya* var. *stenophylla* (Siebold & Zucc.) Landrein	Tanyama, Nagano, Japan	HUTB, E306	E306	SRR13705756
6	Caprifoliaceae	Linnaeoideae	*Diabelia sanguinea* (Siebold & Zucc.) Landrein	Sendai, Miyagi, Japan	HUTB, E301	E301	SRR13705755
7	Caprifoliaceae	Linnaeoideae	*Diabelia spathulata* var. *spathulata* (H. Hara) Landrein	Shiga, Japan	HUTB, E198	E198	SRR13705757
8	Caprifoliaceae	Linnaeoideae	*Diabelia ionostachya* var. *wenzhouensis* (Siebold & Zucc.) Landrein	Wenzhou, Zhejiang, China	HUTB, E204	E204	SRR13705754
9	Caprifoliaceae	Linnaeoideae	Kolkwitzia amabilis Graebn.	Weinan, Shaanxi, China	HUTB, E9	E9	SRR13705797
10	Caprifoliaceae	Linnaeoideae	*Abelia macrotera* (Graebn. et Buchw.) Rehd.	Nanchuan, Chongqing, China	HUTB, E110	E110	SRR13705799
11	Caprifoliaceae	Linnaeoideae	*Abelia uniflora* R. Brown	Wuyishan, Fujian, China	HUTB, E51	E51	SRR13705798
12	Caprifoliaceae	Linnaeoideae	*Abelia chinensis* R. Brown	Jujiang, Jiangxi, China	HUTB, E206	E206	SRR13705765
13	Caprifoliaceae	Linnaeoideae	*Abelia chinensis* var. *ionandra* (André) Rehd.	Yilan, Taiwan, China	HUTB, E30	E30	SRR13705776
14	Caprifoliaceae	Linnaeoideae	*Abelia forrestii* (Diels) W. W. Smith	Nujiang, Yunnan, China	HUTB, E37	E37	SRR13705787
15	Caprifoliaceae	Linnaeoideae	*Vesalea occidentalis* (Villarreal) Wang, H.F. & Landrein	Durango, Mexico	HUTB, E96	E96	SRR13705793

16	Caprifoliaceae	Linnaeoideae	*Vesalea coriacea* (Hemsl.) T.Kim & B.Sun ex Landrein	San Luis Potosi, Mexico	HUTB, E284	E284	SRR13705790
17	Caprifoliaceae	Linnaeoideae	*Vesalea coriacea* (Hemsl.) T.Kim & B.Sun ex Landrein	San Luis Potosi, Mexico	HUTB, E89	E89	SRR13705789
18	Caprifoliaceae	Linnaeoideae	*Vesalea mexicana* Villarreal	San Luis Potosi, Mexico	HUTB, E93	E93	SRR13705792
19	Caprifoliaceae	Linnaeoideae	*Vesalea floribunda* M.Martens & Galeotti	Oaxaca, Mexico	HUTB, E92	E92	SRR13705791
20	Caprifoliaceae	Linnaeoideae	*Linnaea borealis* Linn.	Yili, Xinjiang, China	HUTB, E59	E59	SRR13705788
21	Caprifoliaceae	Linnaeoideae	*Linnaea borealis* Linn.	Yili, Xinjiang, China	HUTB, E14	E14	SRR13705786
22	Caprifoliaceae	*Zabelia*	*Zabelia biflora* Turcz.	Dushanbe, Tajikistan	HUTB, E100	E100	SRR13705775
23	Caprifoliaceae	*Zabelia*	*Zabelia integrifolia* Koidz	Fukuoka, Kyushu, Japan	HUTB, E15	E15	SRR13705777
24	Caprifoliaceae	*Zabelia*	*Zabelia dielsii* (Graebn.) Makino	Ganzi, Sichuan, China	HUTB, E286	E286	SRR13705778
25	Caprifoliaceae	*Zabelia*	*Zabelia dielsii* (Graebn.) Makino	Ganzi, Sichuan, China	HUTB, E108	E108	SRR13705774
26	Caprifoliaceae	*Zabelia*	*Zabelia triflora* R.Br. ex Wall.	Bangalore, India	HUTB, E276	E276	SRR13705773
27	Caprifoliaceae	Monrinoideae	*Morina longifolia* Wall. ex DC.	Dushanbe, Tajikistan	HUTB, E20	E20	SRR13705771
28	Caprifoliaceae	Monrinoideae	*Acanthocalyx alba* (Hand.-Mazz.) M. Connon	Nujiang, Yunnan, China	HUTB, E19	E19	SRR13705772
29	Caprifoliaceae	Valerianoideae	*Centranthus ruber* DC.	San Francisco, California, United States	US 2998828	E220	SRR13705784
30	Caprifoliaceae	Valerianoideae	*Valerianella dentata* (L.) Pollich	Smyrna,Tennessee, United States	US 2998828	E217	SRR13705782
31	Caprifoliaceae	Valerianoideae	*Valeriana urticifolia* Kunth	Autopista Orizaba-Puebla, Vera Cruz, Mexico	US 3714857	E219	SRR13705785
32	Caprifoliaceae	Valerianoideae	*Valeriana officinalis* Linn.	Baoding, Hebei, China	HUTB, E27	E27	SRR13705783
33	Caprifoliaceae	Dipsacoideae	*Scabiosa canescens* Waldst. & Kit.	Athus, Espe, Belgium	US 1273936	E227	SRR13705781
34	Caprifoliaceae	Dipsacoideae	*Scabiosa tschiliensis* Grüning	Yanqing, Beijing, China	HUTB, E21	E21	SRR13705780
35	Caprifoliaceae	Dipsacoideae	*Dipsacus japonicus* Miq.	Beijing, China	HUTB, E23	E23	SRR13705779
36	Caprifoliaceae	Caprifolioideae	*Lonicera arizonnica* Hemsl.	Miyun, Beijing, China	HUTB, E269	E269	SRR13705770

37	Caprifoliaceae	Caprifolioideae	*Lonicera ligustrina* Wall.	Shennongjia, Hubei, China	HUTB, E74	E74	SRR13705769
38	Caprifoliaceae	Caprifolioideae	*Lonicera confusa* D.C.	Haikou, Hainan, China	HUTB, E193	E193	SRR13705768
39	Caprifoliaceae	Caprifolioideae	*Lonicera korolkowii* Stapf.	Baoding, Hebei, China	HUTB, E212	E212	SRR13705767
40	Caprifoliaceae	Caprifolioideae	*Symphoricarpos orbiculatus* (L.) Macm.	Lynchburg, Virginia, United States	US 2099602	E237	SRR13705766
41	Caprifoliaceae	Caprifolioideae	*Heptacodium miconioides* Rehder	Hangzhou, Zhejiang, China	HUTB, E28	E28	SRR13705764
42	Caprifoliaceae	Divervilloideae	*Diervilla lonicera* Mill.	Lambton, Ontario, United States	US 3666234	E331	SRR13705763
43	Caprifoliaceae	Divervilloideae	*Weigela florida* (Bunge) A. DC.	Kangwon, Korea	HUTB, E99	E99	SRR13705762
44	Adoxaceae-Outgroup	Adoxoideae	*Sambucus williamsii* Hance.	Haidian, Beijing, China	HUTB, E208	E208	SRR13705760
45	Adoxaceae-Outgroup	Adoxoideae	*Sambucus nigra* Linn.	Yanqing, Beijing, China	HUTB, E207	E207	SRR13705761
46	Adoxaceae-Outgroup	Opuloideae	*Viburnum opulus* Linn.	Qingdao, Shandong, China	HUTB, E162	E162	SRR13705759

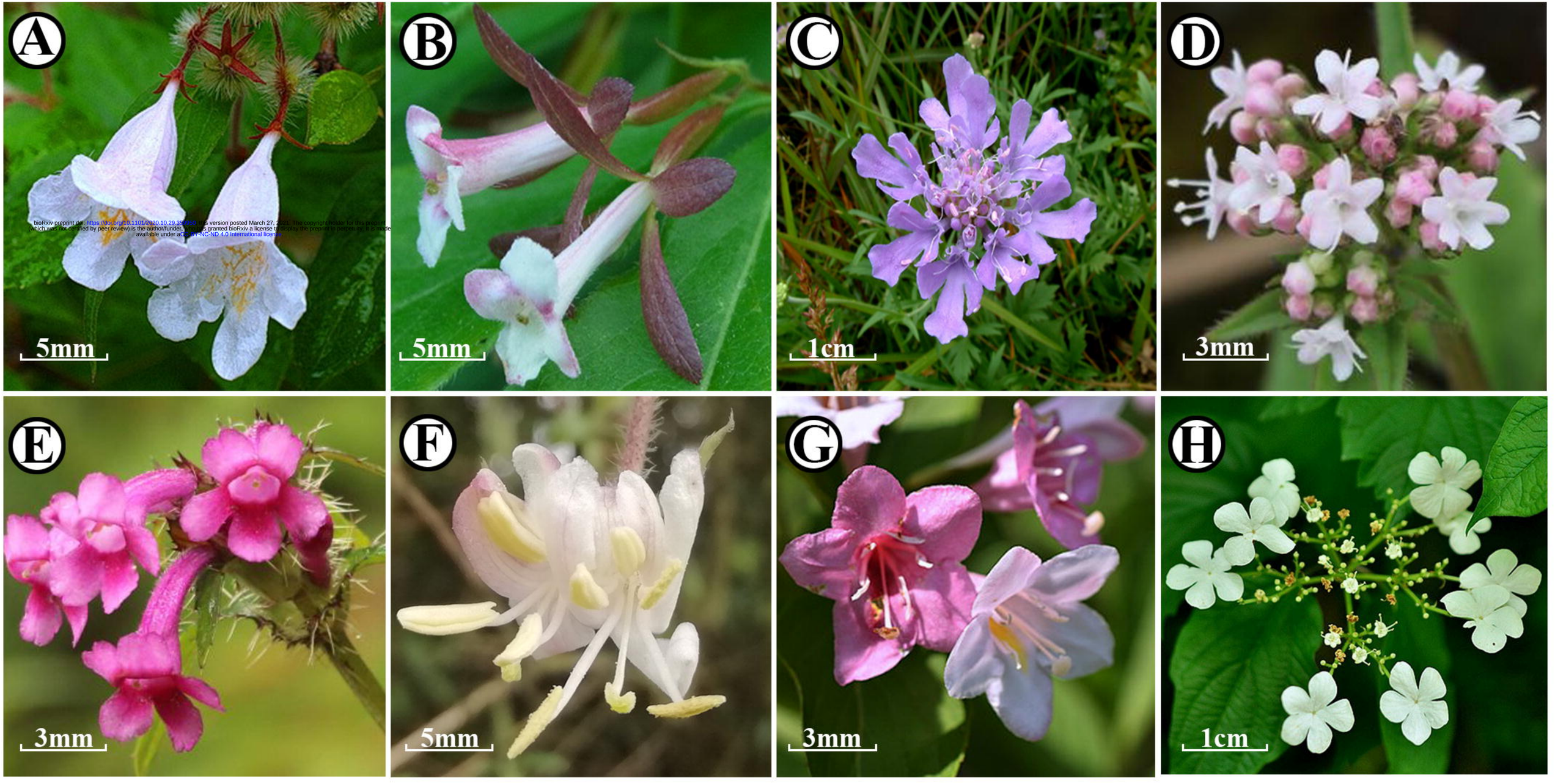
Table S2. HybPiper assembly statistics

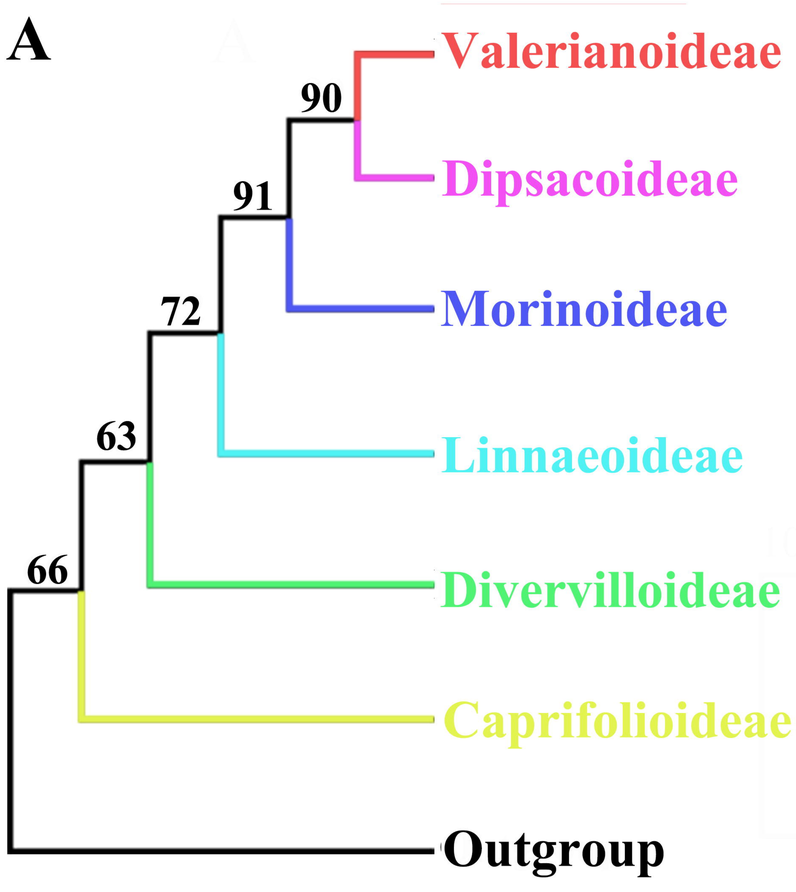
Species	Number of reads	Number of reads on target	Percent reads on target	Number of exons with reads	Number of exons with contigs	Number of exons with sequences	Number of exons with sequences > 25% of the target length	Number of exons with sequences > 50% of the target length	Number of exons with sequences > 75% of the target length	Number of exons with sequences > 150% of the target length	Number of exons with paralog warnings
*Abelia_chinensis*_E206	9076240	3538761	0.39	1118	1021	994	993	979	930	1	162
Abelia_chinensis var. i*onandra*_E30	8512137	2058453	0.242	1054	969	945	942	917	835	1	72
*Abelia_forrestii*_E37	9666409	2890168	0.299	1114	1005	977	975	961	921	0	178
*Abelia_macrotera*_E110	9134947	2888390	0.316	1072	988	960	959	940	901	1	84
*Abelia_uniflora*_E51	18769946	3613529	0.193	1086	991	967	966	946	892	2	102
*Acanthocalyx_alba*_E19	6118686	1489979	0.244	1060	948	927	922	891	821	1	339
*Centranthus_ruber*_E220	6873833	2738073	0.398	1021	822	779	769	715	632	0	324
Diabelia_ionostachya var. *stenophylla*_E306	9313351	1888267	0.203	1119	1007	977	974	960	889	0	79
Diabelia_ionostachya var. *wenzhouensis*_E204	6642696	2077347	0.313	1103	1006	983	982	962	908	1	76
*Diabelia_sanguinea*_E301	12028399	2724711	0.227	1151	1062	1032	1032	1021	989	3	619
*Diabelia_serrata*_E123	21639680	7331755	0.339	1120	1033	998	996	980	933	1	136
Diabelia_spathulata var. *spathulata*_E198	9267804	3459777	0.373	1130	1022	993	992	981	928	1	92
*Diervilla_lonicera*_E331	5850914	1589800	0.272	1077	965	939	938	915	857	0	105
*Dipelta_floribunda*_E55	4636783	105360	0.023	924	660	642	639	601	507	0	27
*Dipelta_floribunda*_E56	6580640	248258	0.038	992	853	833	830	793	680	0	43
*Dipelta_floribunda*_E57	3074426	54170	0.018	866	547	530	518	461	367	0	14

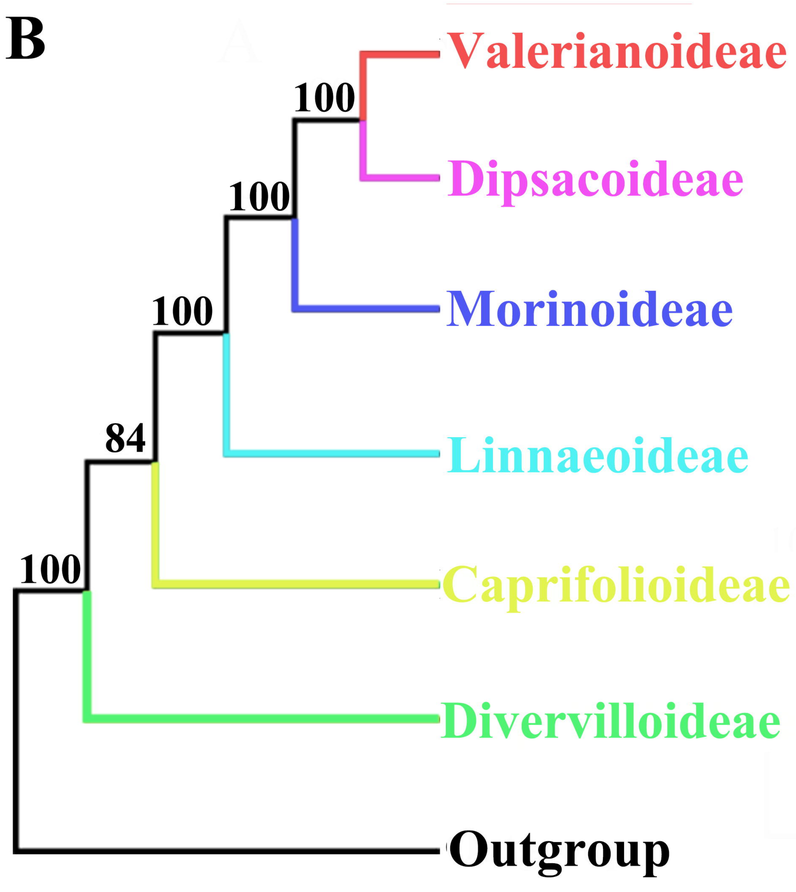
*Dipsacus_japonicus*_E23	2017716	174475	0.086	910	552	525	501	417	284	1	27
*Heptacodium_miconioides*_E28	36252001	9788612	0.27	1125	1005	978	975	953	922	0	221
*Kolkwitzia_amabilis*_E9	9779050	2641147	0.27	1097	1005	984	983	967	916	0	75
*Linnaea_borealis*_E14	9523146	1355451	0.142	1069	981	962	961	946	902	1	73
*Linnaea_borealis*_E59	3831881	51899	0.014	867	536	516	507	443	333	0	11
*Lonicera_arizonica*_E269	7277244	1732355	0.238	1049	927	901	894	852	772	0	83
*Lonicera_confusa*_E193	9914689	3445473	0.348	1090	950	920	918	896	850	0	12
*Lonicera_korolkowii*_E212	10617680	3453855	0.325	1075	962	929	926	911	872	0	11
Lonicera_ligustrina var. *pileata*_E74	4988272	157960	0.032	958	740	724	724	702	653	1	15
*Morina_longifolia*_E20	23680514	8543790	0.361	1104	1004	980	978	951	895	1	56
*Sambucus_nigra*_E207	5805790	460514	0.079	964	701	678	665	631	575	0	108
*Sambucus_williamsii*_E208	5071638	402725	0.079	931	705	674	667	635	573	0	108
*Scabiosa_canescens*_E223	2521156	272582	0.108	948	624	589	555	450	302	0	27
*Scabiosa_tschiliensis*_E21	21530335	3744566	0.174	1060	918	890	887	859	799	2	303
*Symphoricarpos_orbiculatus*_E237	13570556	6350425	0.468	1114	996	961	958	937	879	0	101
*Valeriana_officinalis*_E27	2004483	203244	0.101	880	673	651	636	572	489	1	41
Valeriana_urticifolia var. *scorpioides*_E219	7471262	2066511	0.277	1004	784	751	744	706	637	1	33
*Valerianella_dentata*_E217	6930308	2552865	0.368	1024	785	741	732	688	605	0	21
*Vesalea_coriacea*_E284	7770516	2129447	0.274	1099	1013	977	974	952	883	0	61
*Vesalea_coriacea*_E89	6728245	436748	0.065	1107	951	926	925	910	856	0	276
*Vesalea_floribunda*_E92	3249704	22487	0.007	718	311	287	267	205	130	0	1
*Vesalea_mexicana*_E93	4712564	111119	0.024	938	692	670	667	625	530	1	12
*Vesalea_occidentalis*_E96	6434742	142056	0.022	974	730	712	703	650	559	0	23
*Viburnum_opulus var. americanum*_E162	4690628	682398	0.145	961	811	785	781	751	691	0	105
*Weigela_florida*_E99	13576145	1103870	0.081	1062	942	916	914	888	850	0	116

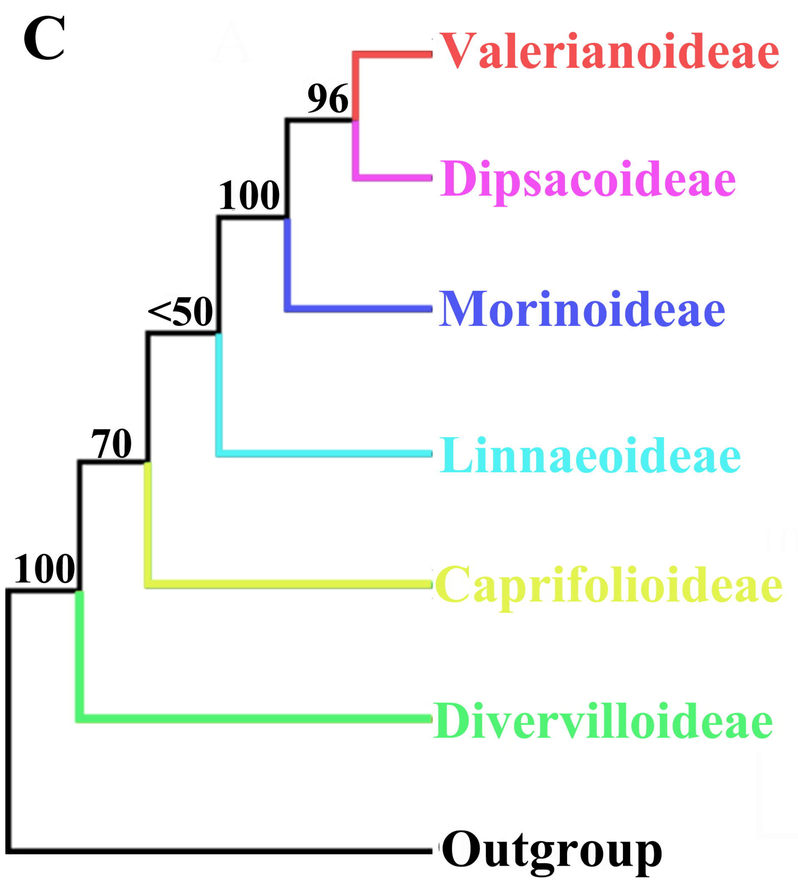
*Zabelia_biflora*_E100	15232347	4165064	0.273	1108	1031	1009	1006	983	943	3	567
*Zabelia_dielsii*_E108	3134211	287330	0.092	1029	907	885	881	856	785	3	340
*Zabelia_dielsii*_E286	4085869	86244	0.021	923	625	611	587	493	348	0	45
*Zabelia_integrifolia*_E15	2031719	222622	0.11	1012	872	855	848	815	733	2	270
*Zabelia_triflora*_E276	6552409	721374	0.11	1090	988	960	954	912	798	1	321

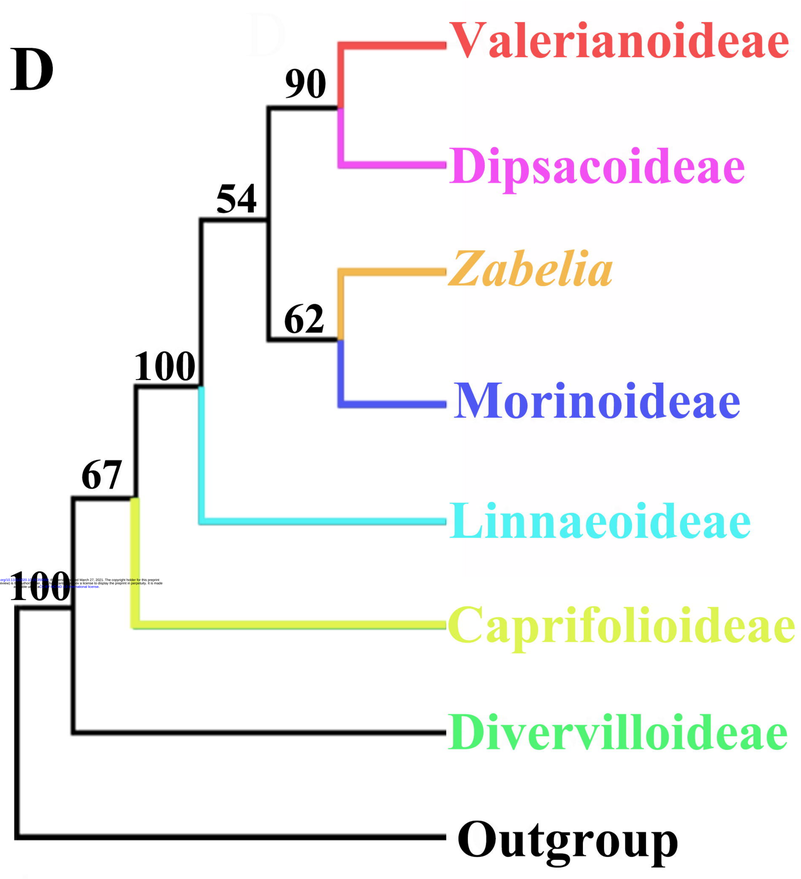
1209

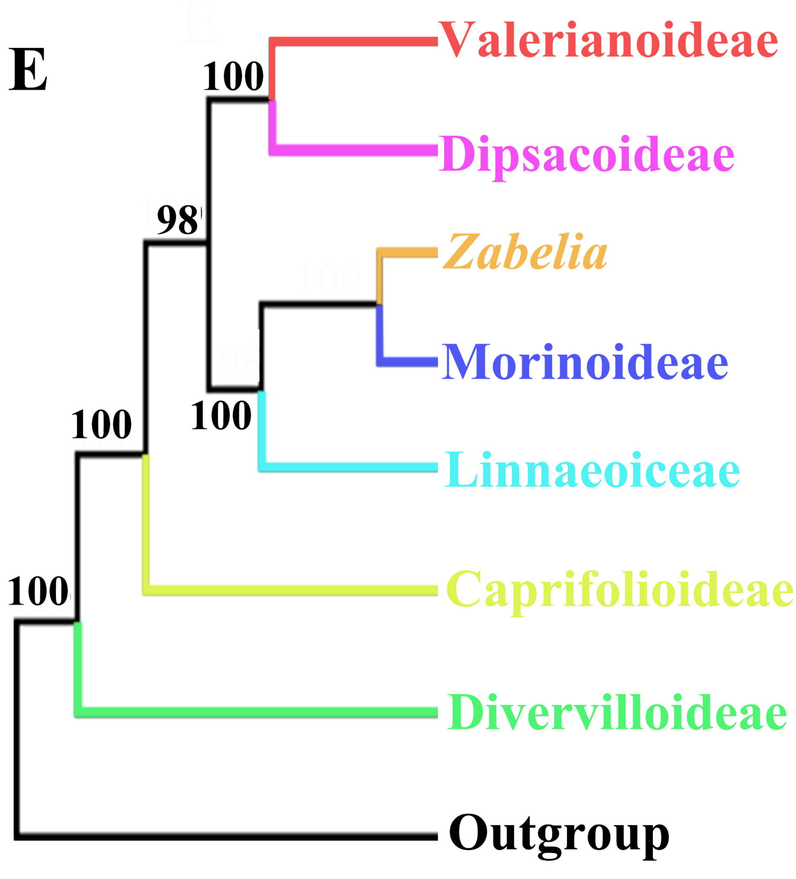


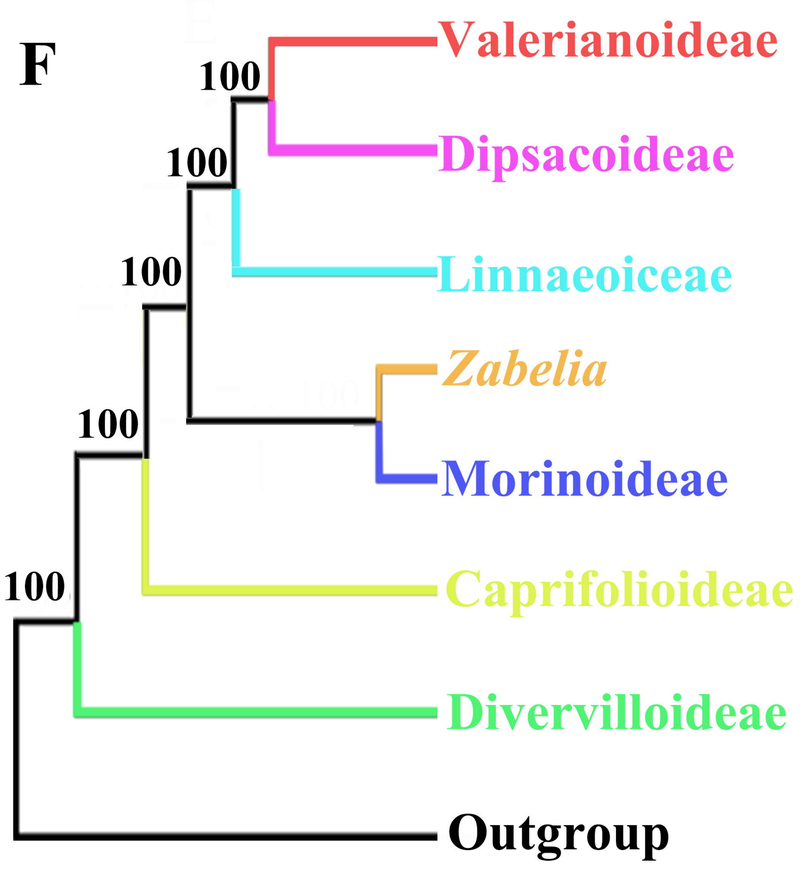

A
Valerianoideae
Dipsacoideae
Morinoideae
Linnaeoideae
Divervilloideae
Caprifolioideae
Outgroup
90
91
72
63
66


B
Valerianoideae
Dipsacoideae
Morinoideae
Linnaeoideae
Caprifolioideae
Divervilloideae
Outgroup
100
100
100
84
100


C
Valerianoideae
Dipsacoideae
Morinoideae
Linnaeoideae
Caprifolioideae
Divervilloideae
Outgroup
96
100
<50
70
100


D
Valerianoideae
Dipsacoideae
Zabelia
Morinoideae
Linnaeoideae
Caprifolioideae
Divervilloideae
Outgroup
90
54
62
100
67
100


E
Valerianoideae
Dipsacoideae
Zabelia
Morinoideae
Linnaeoiceae
Caprifolioideae
Divervilloideae
Outgroup
100
98
100
100
100


F
Valerianoideae
Dipsacoideae
Linnaeoiceae
Zabelia
Morinoideae
Caprifolioideae
Divervilloideae
Outgroup
100
100
100
100
100

Abelia uniflora E51
A. macrotera E110
A. forrestii E37
A. chinensis var. *ionandra* E30
A. chinensis E206
Diabelia ionostachya var. *stenophylla* E306
D. sanguinea E301
D. serrata E123
D. ionostachya var. *wenzhouensis* E204
D. spathulata var. *spathulata* E198
Kolkwitzia amabilis E9
Dipelta floribunda E57
D. floribunda E56
D. floribunda E55
Vesalea coriacea E284
V. coriacea E89
V. occidentails E96
V. mexicana E93
V. floribunda E92
Linnaea borealis E59
L. borealis E14
Dipsacus japonicus E23
Valeriana urticifolia E219
V. offcinalis E27
Centranthus ruber E220
Valerianella dentata E217
Scabiosa canescens E223
S. tschiliensis E21
Zabelia biflora E100
Z. dielsii E108
Z. integrifolia E15
Z. dielsii E286
Z. triflora E276
Acanthocalyx alba E19
Morina longifolia E20
Lonicera korolkowii E212
L. ligustrina E74
L. arizonica E269
L. confusa E193
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Diervilla lonicera E331
Weigela florida E99
Sambucus nigra E207
S. williamsii E208
Viburnum opulus E162

0.94/0.08/54/295
1/0.07/45/349
1/0.2/149/196
0.71/-0.06/15/396
1/0.06/20/393
0.95/0.69/276/50
1/0.08/106/401
1/0.33/101/122
1/0.20/128/215
1/0.08 /48/420
1/0.88/301/11
1/0.09/86/375
0.99/0.58/314/44
1/0.11/63/307
0.99/0.37/223/85
1/0.32/417/181
1/0.18/86/268
1/0.81/453/31
1/0.55/306/92
1/0.72/409/39
1/0.73/266/23
1/0.23/82/114
1/0.50/161/55
1/0.79/294/31

Linnaeoideae
Dipsacoideae
Valerianoideae
Dipsacoideae
Zabelia
Morinoideae
Caprifolioideae
Diervilloideae
Outgroup

0.8

Nuclear concatenated

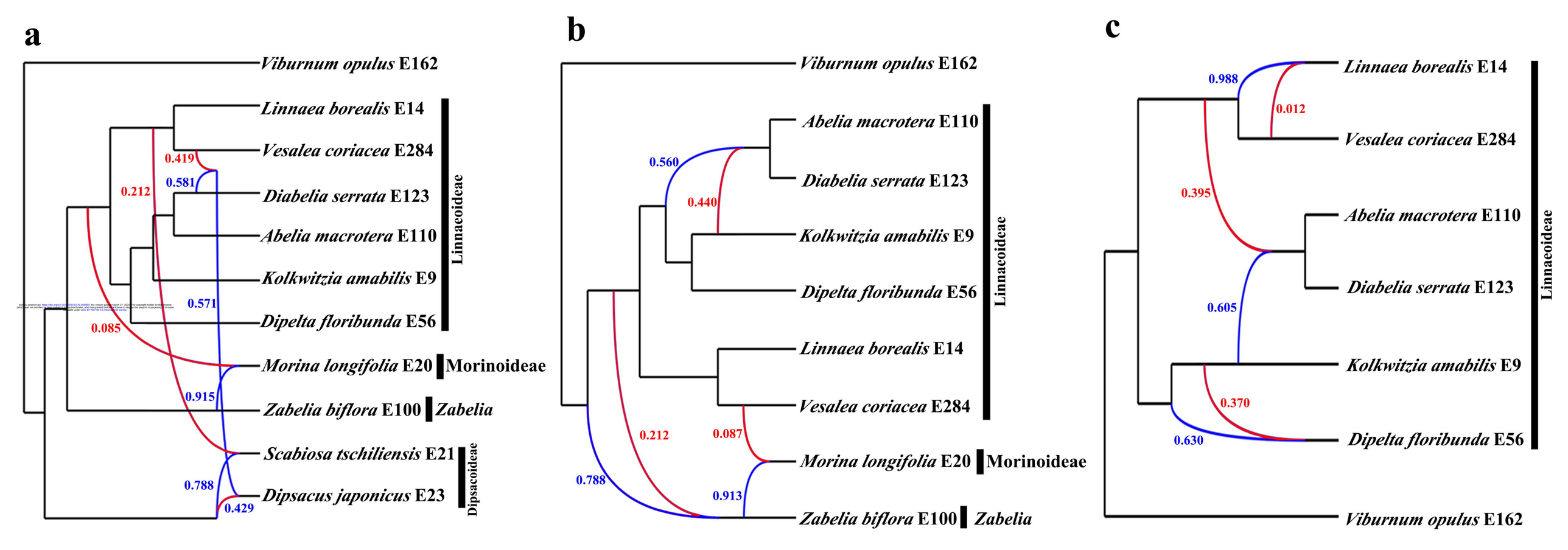
Plastome

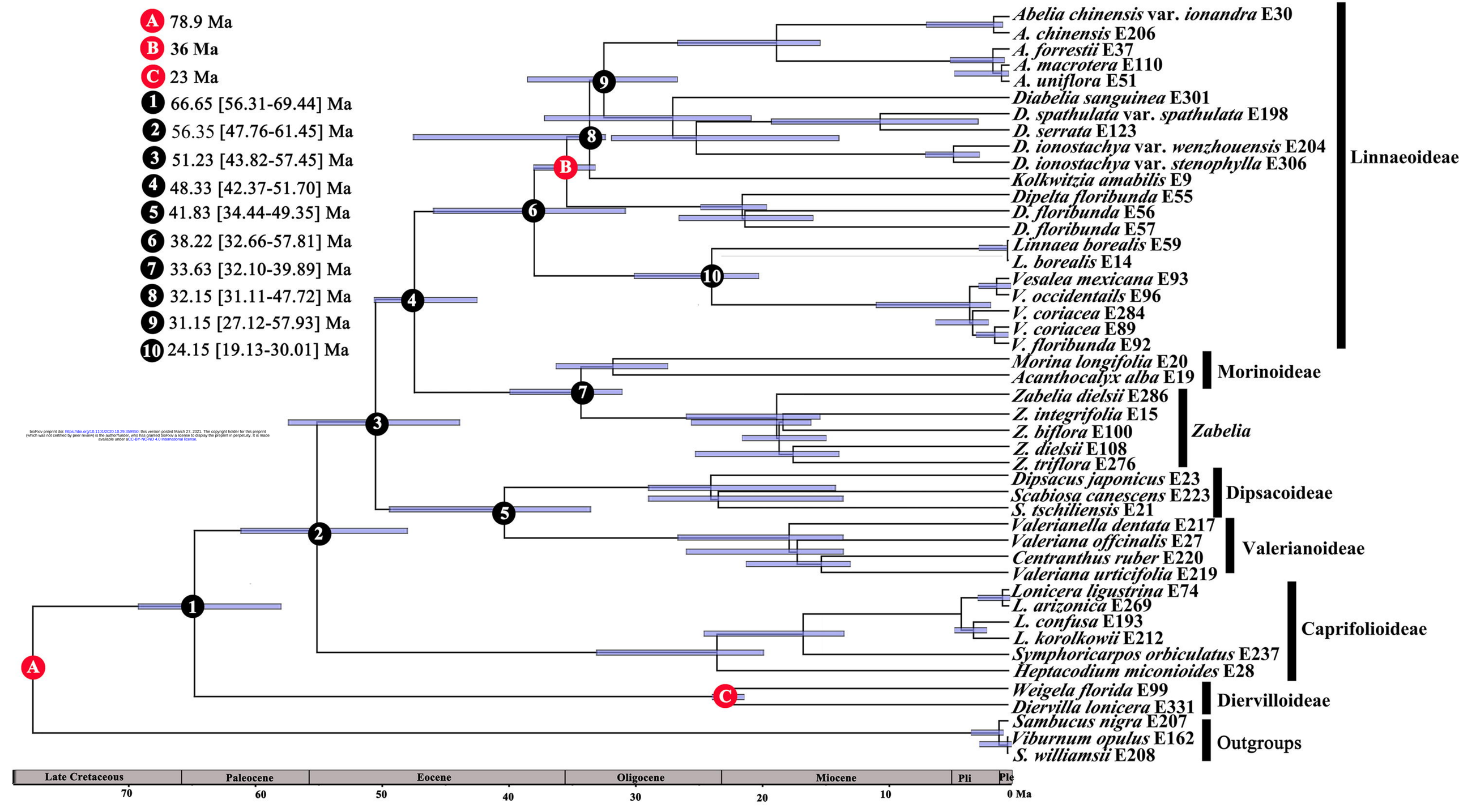
Abelia uniflora
A. macrotera
A. forrestii
A. chinensis var. *ionandra*
A. chinensis
Diabelia serrata
D. ionostachya var. *wenzhouensis*
D. ionostachya var. *stenophylla*
D. sanguinea
D. spathulata var. *spathulata*
Dipelta floribunda
D. floribunda
D. floribunda
Kolkwitzia amabilis
Vesalea occidentails
V. mexicana
V. floribunda
V. coriacea
V. coriacea
Linnaea borealis
L. borealis
Valeriana urticifolia
Centranthus ruber
Valeriana offcinalis
Valerianella dentata
Scabiosa canescens
S. tschiliensis
Dipsacus japonicus
Zabelia dielsii
Z. integrifolia
Z. biflora
Z. dielsii
Z. triflora
Acanthocalyx alba
Moina longifolia
Lonicera arizonica
L. ligustrina
L. confusa
L. korolkowii
Symphoricarpos orbiculatus
Heptacodium miconioides
Diervilla lonicera
Weigela florida
Sambucus nigra
S. williamsii
Viburnum opulus

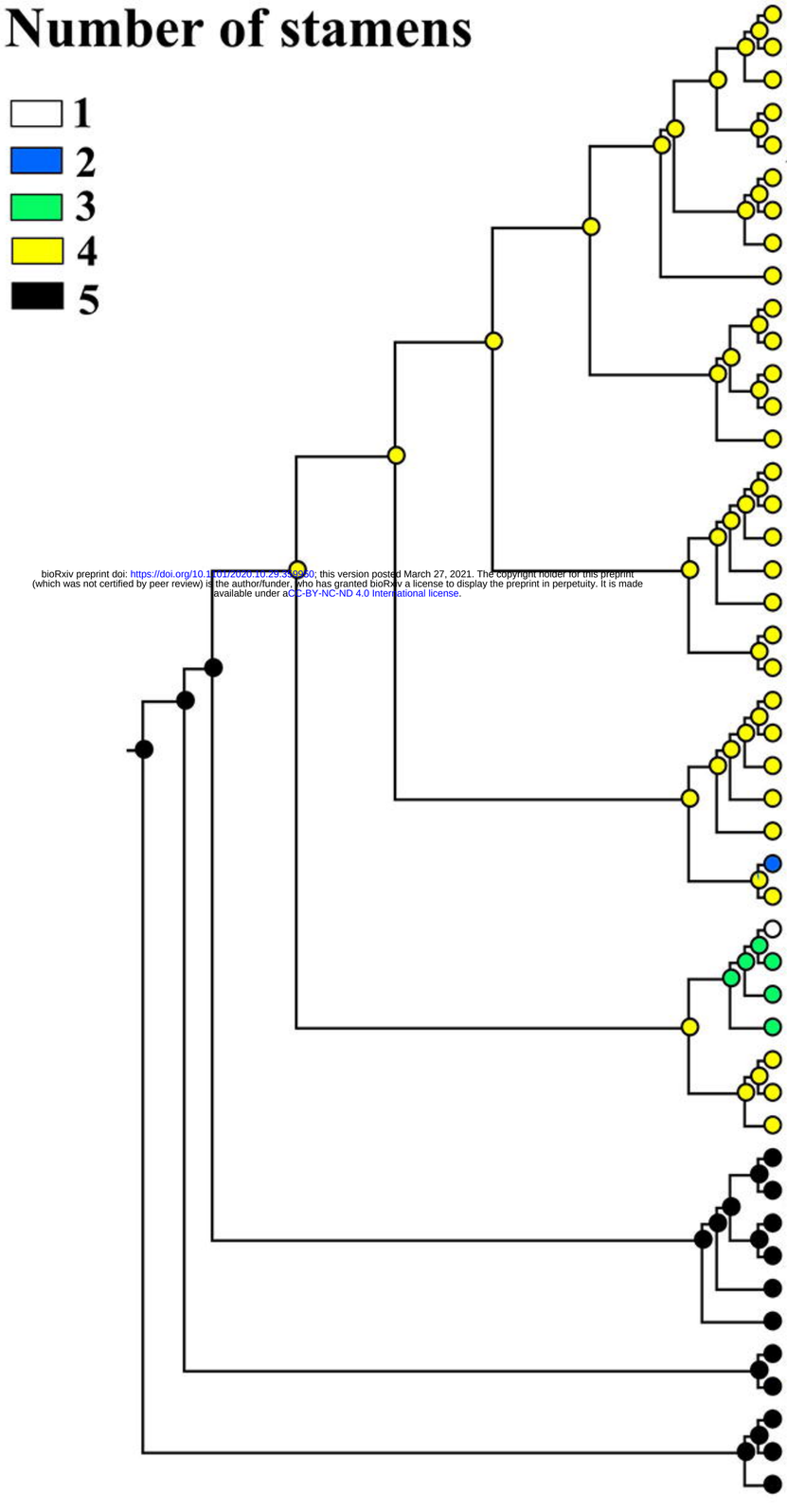
85
89
75
79

Linnaeoideae
Valerianoideae
Dipsacoideae
Zabelia
Morinoideae
Caprifolioideae
Diervilloideae
Outgroups

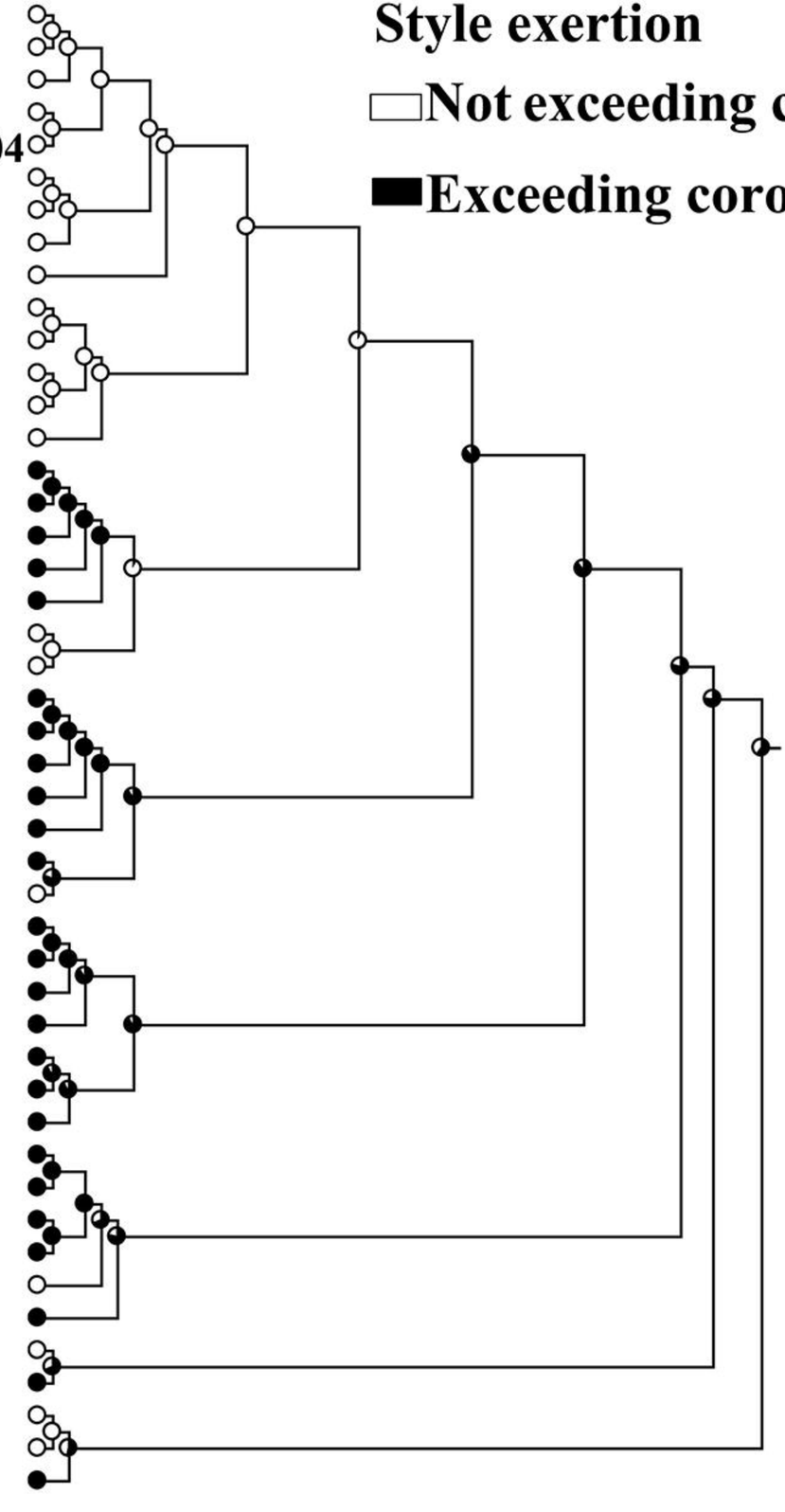
0.04
0.0


a
Viburnum opulus E162
Linnaea borealis E14
Vesalea coriacea E284
Diabelia serrata E123
Abelia macrotera E110
Kolkwitzia amabilis E9
Dipelta floribunda E56
Morina longifolia E20
Zabelia biflora E100
Scabiosa tschiliensis E21
Dipsacus japonicus E23
Linnaeoideae
Morinoideae
Zabelia
Dipsacoideae
0.419
0.581
0.212
0.571
0.085
0.915
0.788
0.429
b
Viburnum opulus E162
Abelia macrotera E110
Diabelia serrata E123
Kolkwitzia amabilis E9
Dipelta floribunda E56
Linnaea borealis E14
Vesalea coriacea E284
Morina longifolia E20
Zabelia biflora E100
Linnaeoideae
Morinoideae
Zabelia
0.560
0.440
0.212
0.087
0.788
0.913
c
Linnaea borealis E14
Vesalea coriacea E284
Abelia macrotera E110
Diabelia serrata E123
Kolkwitzia amabilis E9
Dipelta floribunda E56
Viburnum opulus E162
Linnaeoideae
0.988
0.012
0.395
0.605
0.370
0.630





Diabelia serrata E123
D. ionostachya var. *stenophylla* E306
D. sanguinea E301
D. spathulata var. *spathulata* E198
D. ionostachya var. *wenzhouensis* E204
Dipelta floribunda E57
D. floribunda E56
D. floribunda E55
Kolkwitzia amabilis E9
Abelia macrotera E110
A. uniflora E51
A. chinensis E206
A. chinensis var. *ionandra* E30
A. forrestii E37
Vesalea coriacea E89
V. coriacea E284
V. floribunda E92
V. mexicana E93
V. occidentails E96
Linnaea borealis E59
L. borealis E14
Zabelia dielsii E286
Z. biflora E100
Z. integrifolia E15
Z. dielsii E108
Z. triflora E276
Moina longifolia E20
Acanthocalyx alba E19
Centranthus ruber E220
Valerianella dentata E217
Valeriana urticifolia E219
Valeriana offcinalis E27
Scabiosa canescens E223
S. tschiliensis E21
Dipsacus japonicus E23
Lonicera korolkowii E212
L. confusa E193
L. ligustrina E74
L. arizonica E269
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Weigela florida E99
Diervilla lonicera E331
Sambucus nigra E207
S. williamsii E208
Viburnum opulus E162



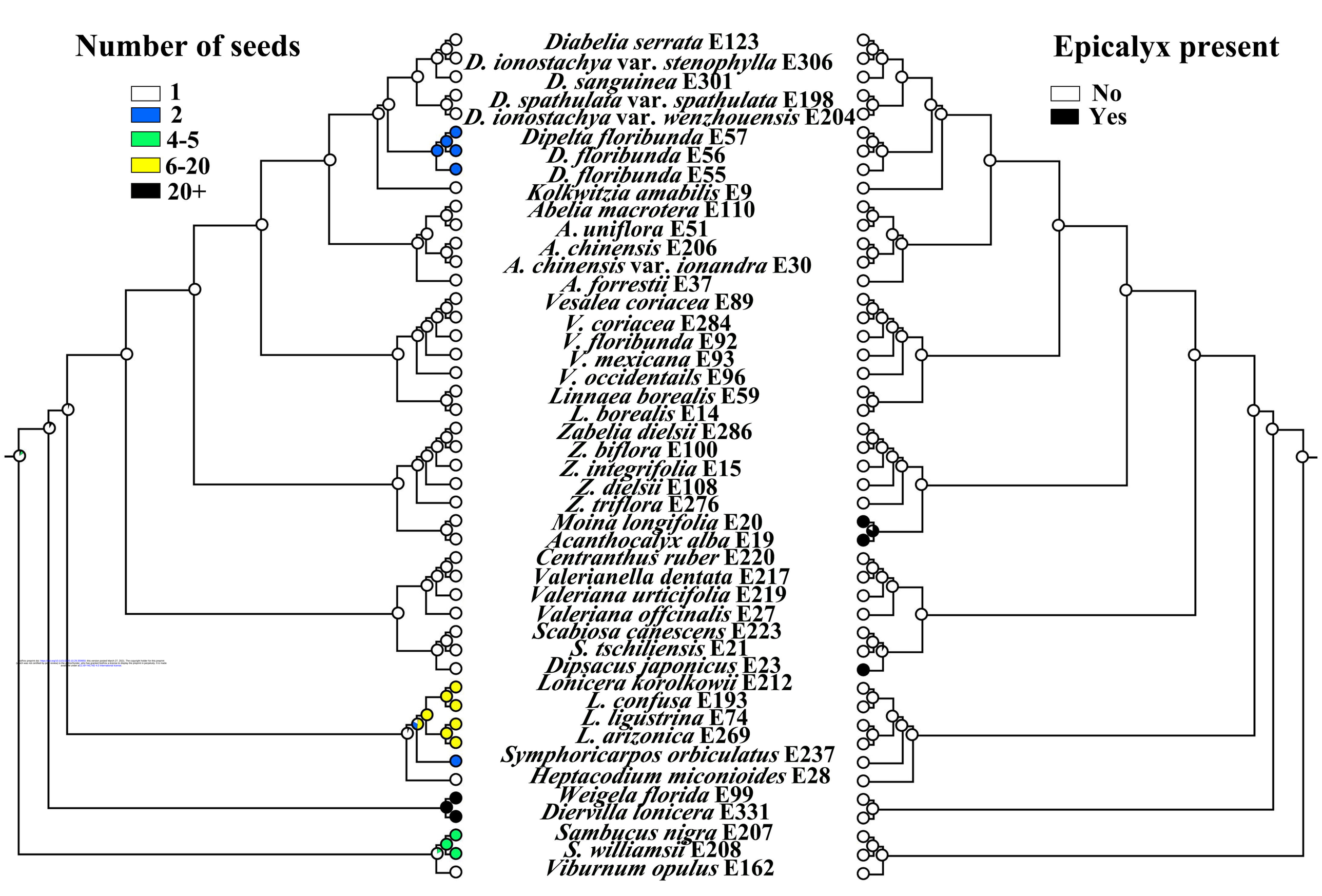
Fruit stype

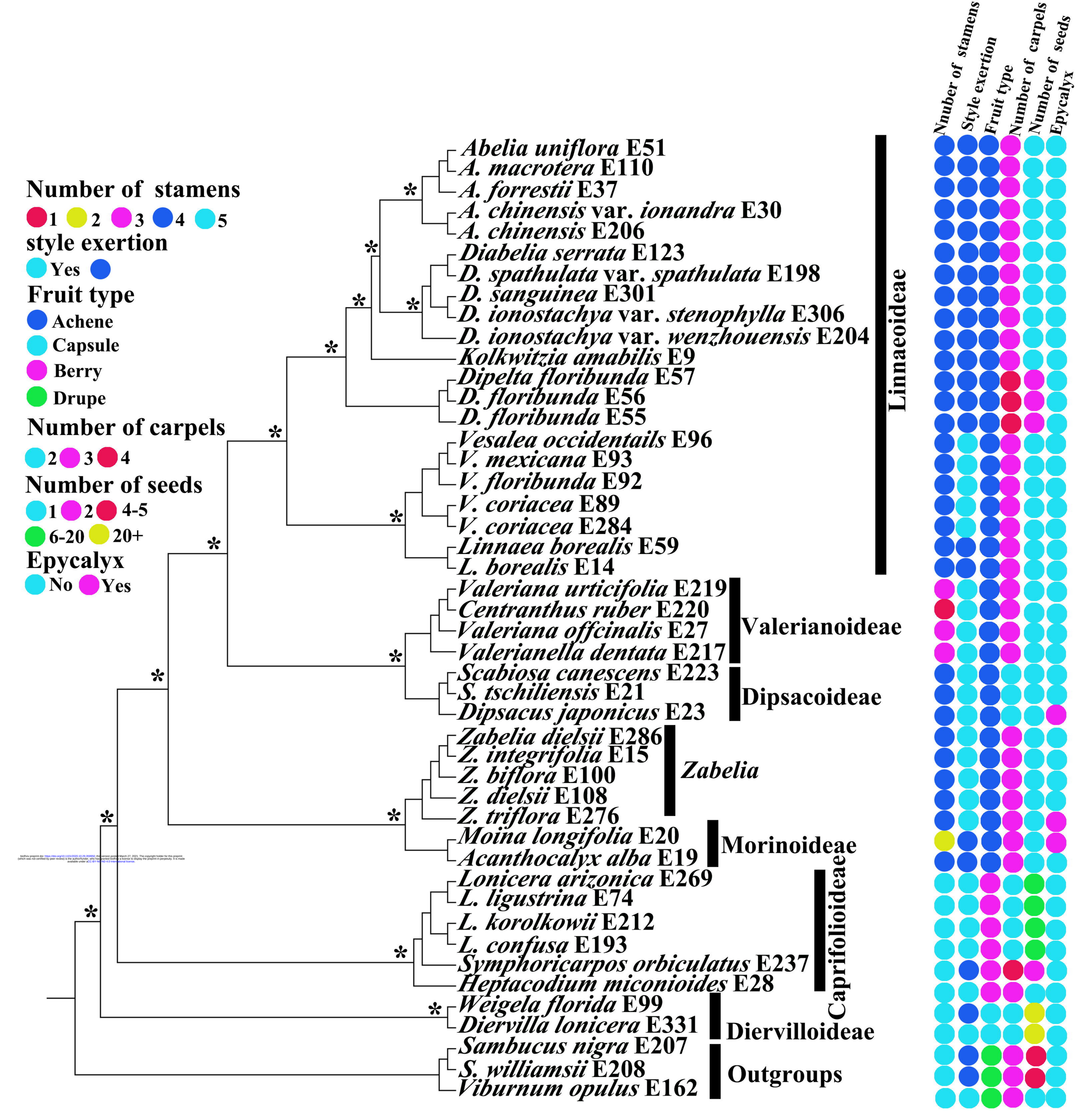
- Achene
- Capsule
- Berry
- Drupe

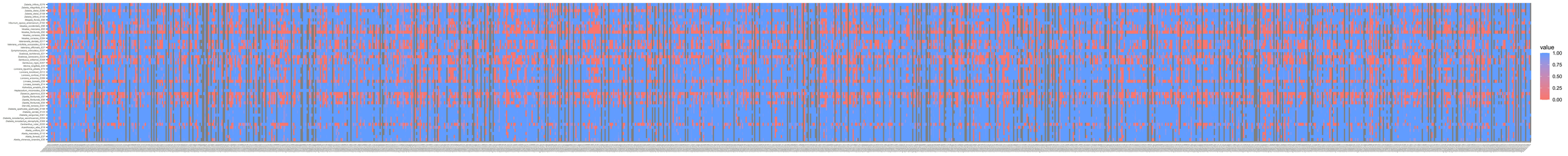
Diabelia serrata E123
D. ionostachya var. *stenophylla* E306
D. sanguinea E301
D. spathulata var. *spathulata* E198
D. ionostachya var. *wenzhouensis* E204
Dipelta floribunda E57
D. floribunda E56
D. floribunda E55
Kolkwitzia amabilis E9
Abelia macrotera E110
A. uniflora E51
A. chinensis E206
A. chinensis var. *ionandra* E30
A. forrestii E37
Vesalea coriacea E89
V. coriacea E284
V. floribunda E92
V. mexicana E93
V. occidentails E96
Linnaea borealis E59
L. borealis E14
Zabelia dielsii E286
Z. biflora E100
Z. integrifolia E15
Z. dielsii E108
Z. triflora E276
Moina longifolia E20
Acanthocalyx alba E19
Centranthus ruber E220
Valerianella dentata E217
Valeriana urticifolia E219
Valeriana offcinalis E27
Scabiosa canescens E223
S. tschiliensis E21
Dipsacus japonicus E23
Lonicera korolkowii E212
L. confusa E193
L. ligustrina E74
L. arizonica E269
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Weigela florida E99
Diervilla lonicera E331
Sambucus nigra E207
S. williamsii E208
Viburnum opulus E162

Number of carpels

- 2
- 3
- 4



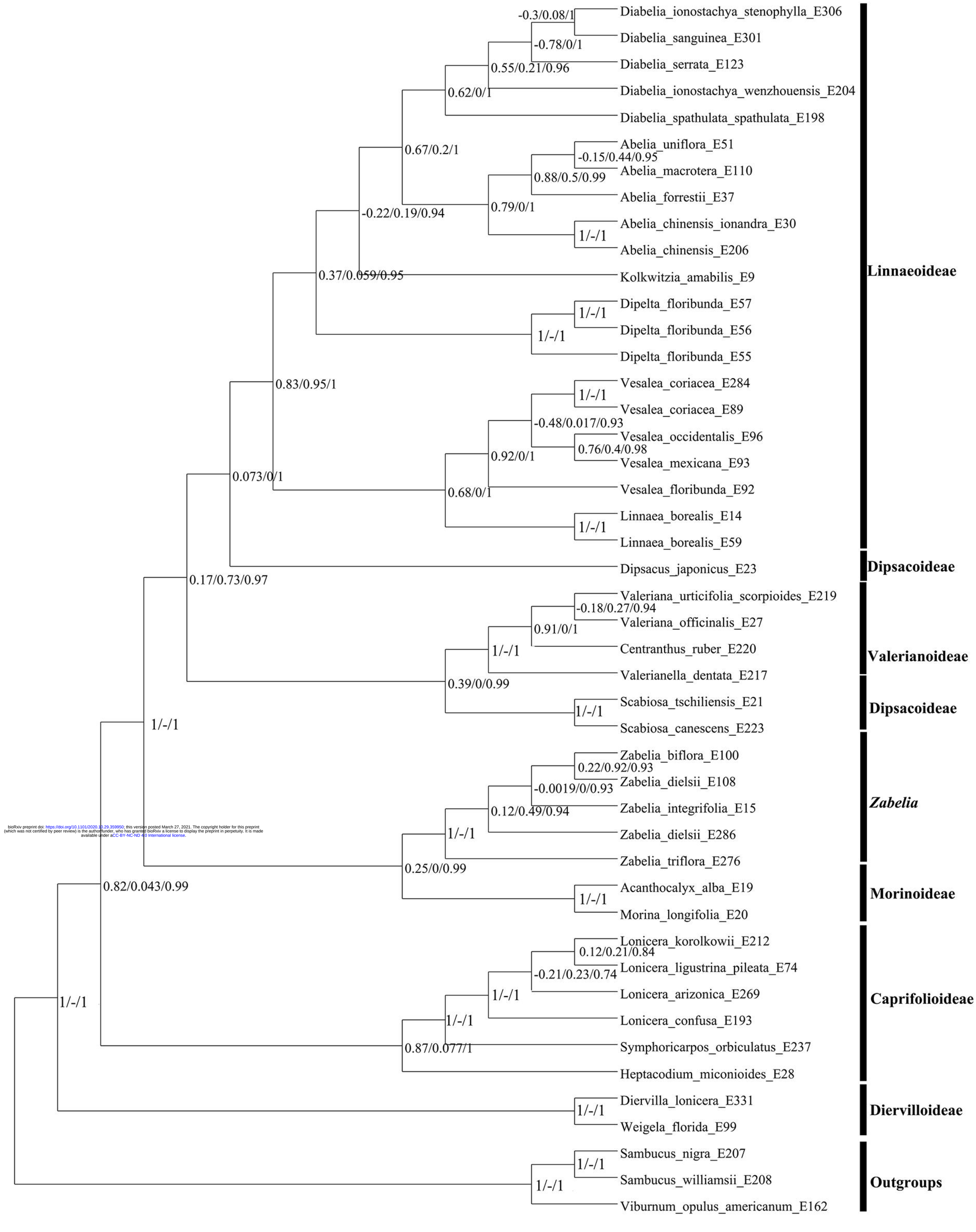


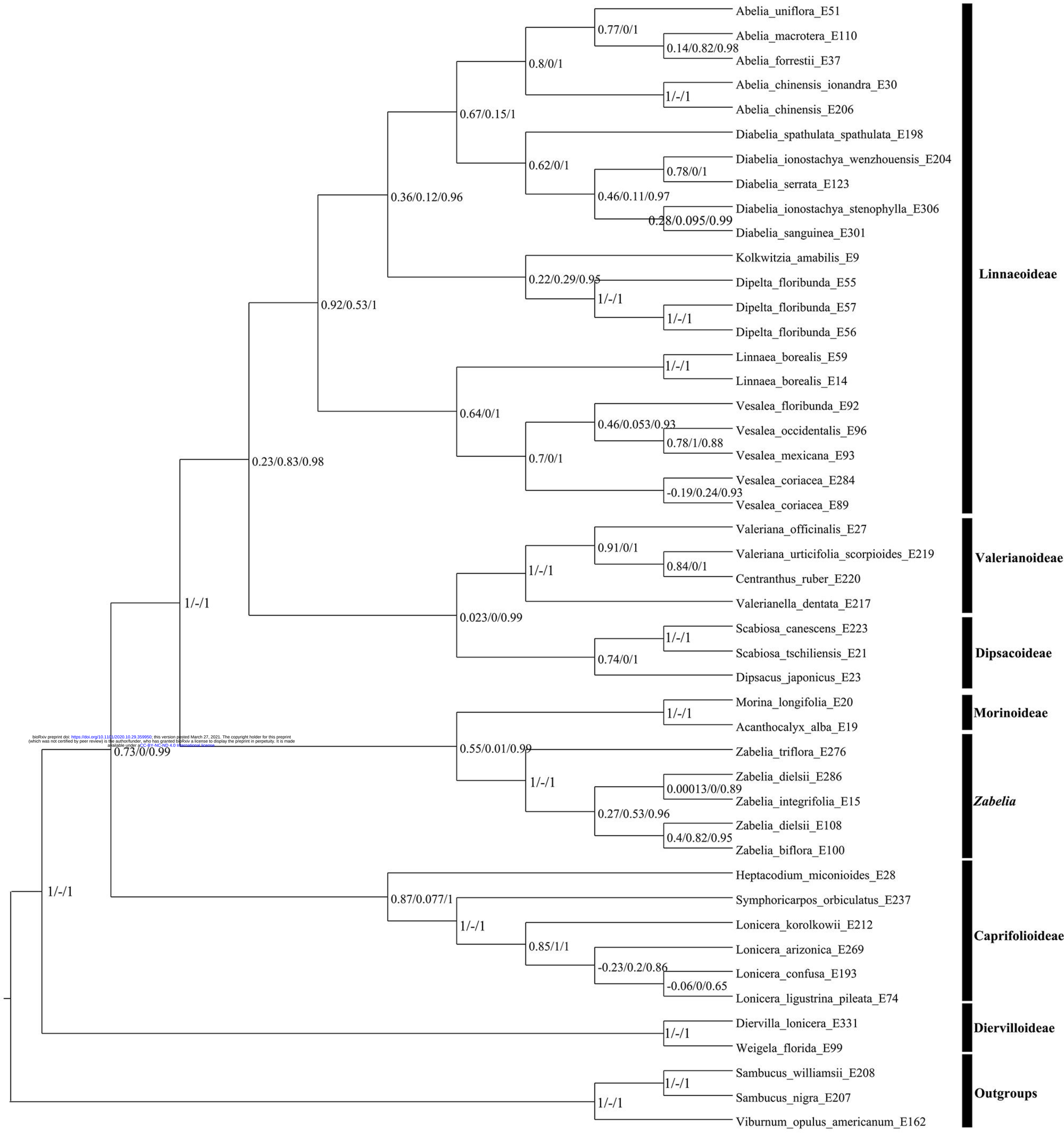

Dahlia_imperialis_ERR
Dahlia_merckii_ERR
Dahlia_pinnata_ERR
Dahlia_rosea_ERR
Dahlia_variabilis_ERR
Magnolia_floribunda_ERR
Viburnum_dilatatum_accession_ERR
Weigela_coraeensis_ERR
Weigela_florida_ERR
Weigela_hortensis_ERR
Weigela_middendorffiana_ERR
Abelia_chinensis_ERR
Abelia_dielsii_ERR
Abelia_mosanensis_ERR
Valeriana_officinalis_accession_ERR
Symphoricarpos_albus_ERR
Symphoricarpos_orbiculatus_ERR
Diabelia_hortensis_ERR
Diabelia_serrata_ERR
Sambucus_chinensis_ERR
Sambucus_nigra_ERR
Viburnum_dilatatum_ERR
Lonicera_japonica_dilatata_ERR
Lonicera_japonica_ERR
Lonicera_praeflorens_ERR
Lonicera_subsessilis_ERR
Lonicera_xylosteum_ERR
Linnaea_borealis_ERR
Abelia_x_grandiflora_ERR
Heptacodium_miconioides_ERR
Diabelia_hortensis_ERR
Diabelia_spathulata_ERR
Diabelia_tschonoskii_ERR
Diabelia_serrata_ERR
Diabelia_spathulata_ERR
Diabelia_serrata_ERR
Diabelia_tschonoskii_ERR
Diabelia_sinodavidiana_stenophylla_ERR
Dipelta_floribunda_ERR
Kolkwitzia_amabilis_ERR
Abelia_chinensis_accession_ERR
Abelia_dielsii_ERR
Abelia_mosanensis_ERR
Abelia_x_grandiflora_ERR
value
1.00
0.75
0.50
0.25
0.00

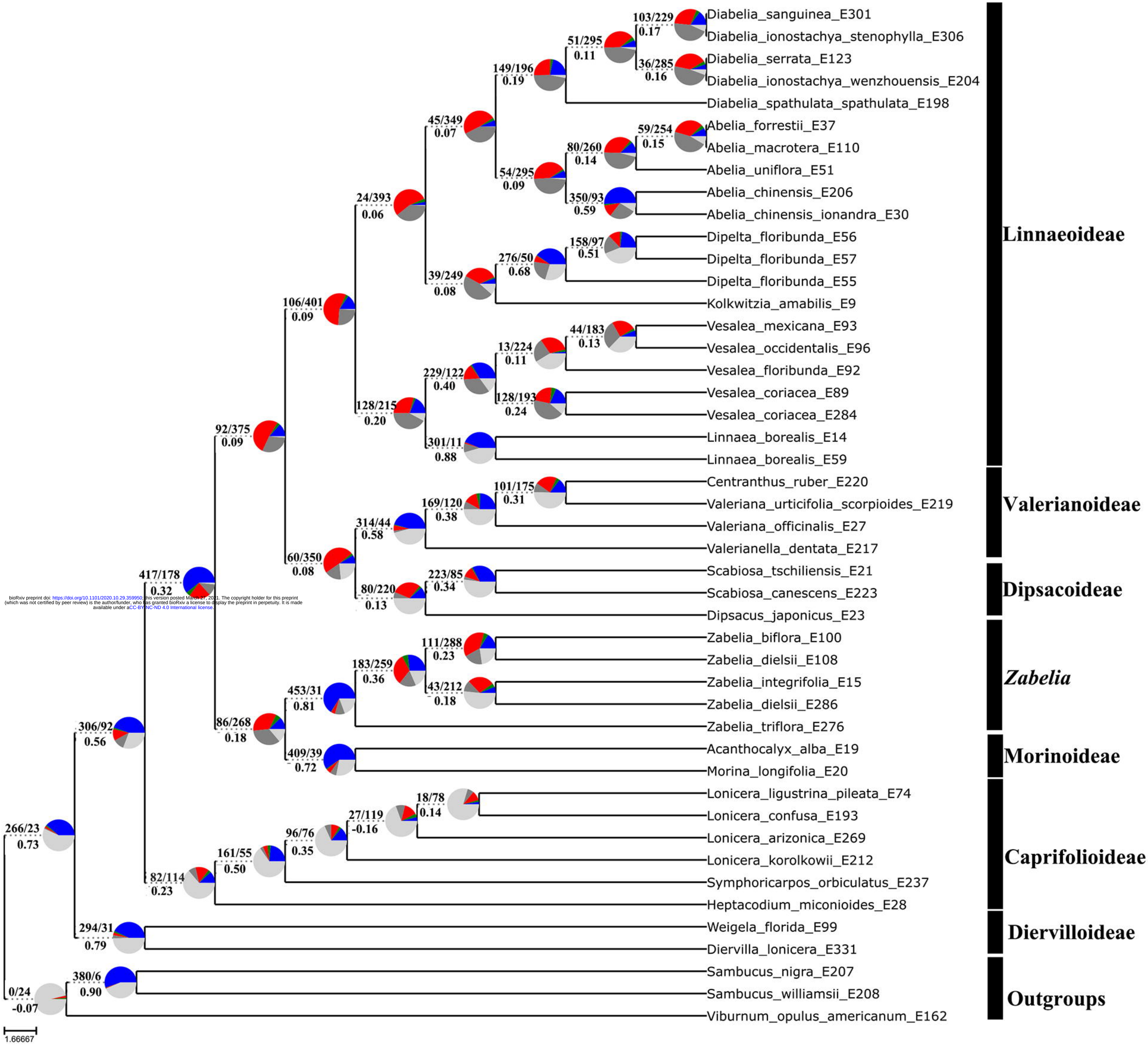
Abelia_macrotera_E110
Abelia_uniflora_E51
Abelia_forrestii_E37
Abelia_chinensis_E206
Abelia_chinensis_ionandra_E30
Diabelia_sanguinea_E301
Diabelia_ionostachya_stenophylla_E306
Diabelia_serrata_E123
Diabelia_ionostachya_wenzhouensis_E204
Diabelia_spathulata_spathulata_E198
Kolkwitzia_amabilis_E9
Dipelta_floribunda_E56
Dipelta_floribunda_E57
Dipelta_floribunda_E55
Vesalea_mexicana_E93
Vesalea_occidentalis_E96
Vesalea_coriacea_E89
Vesalea_coriacea_E284
Vesalea_floribunda_E92
Linnaea_borealis_E59
Linnaea_borealis_E14
Dipsacus_japonicus_E23
Valeriana_officinalis_E27
Valeriana_urticifolia_scorpioides_E219
Centranthus_ruber_E220
Valerianella_dentata_E217
Scabiosa_canescens_E223
Scabiosa_tschiliensis_E21
Zabelia_dielsii_E108
Zabelia_biflora_E100
Zabelia_integrifolia_E15
Zabelia_dielsii_E286
Zabelia_triflora_E276
Morina_longifolia_E20
Acanthocalyx_alba_E19
Lonicera_ligustrina_pileata_E74
Lonicera_korolkowii_E212
Lonicera_arizonica_E269
Lonicera_confusa_E193
Symphoricarpos_orbiculatus_E237
Heptacodium_miconioides_E28
Weigela_florida_E99
Diervilla_lonicera_E331
Sambucus_williamsii_E208
Sambucus_nigra_E207
Viburnum_opulus_americanum_E162

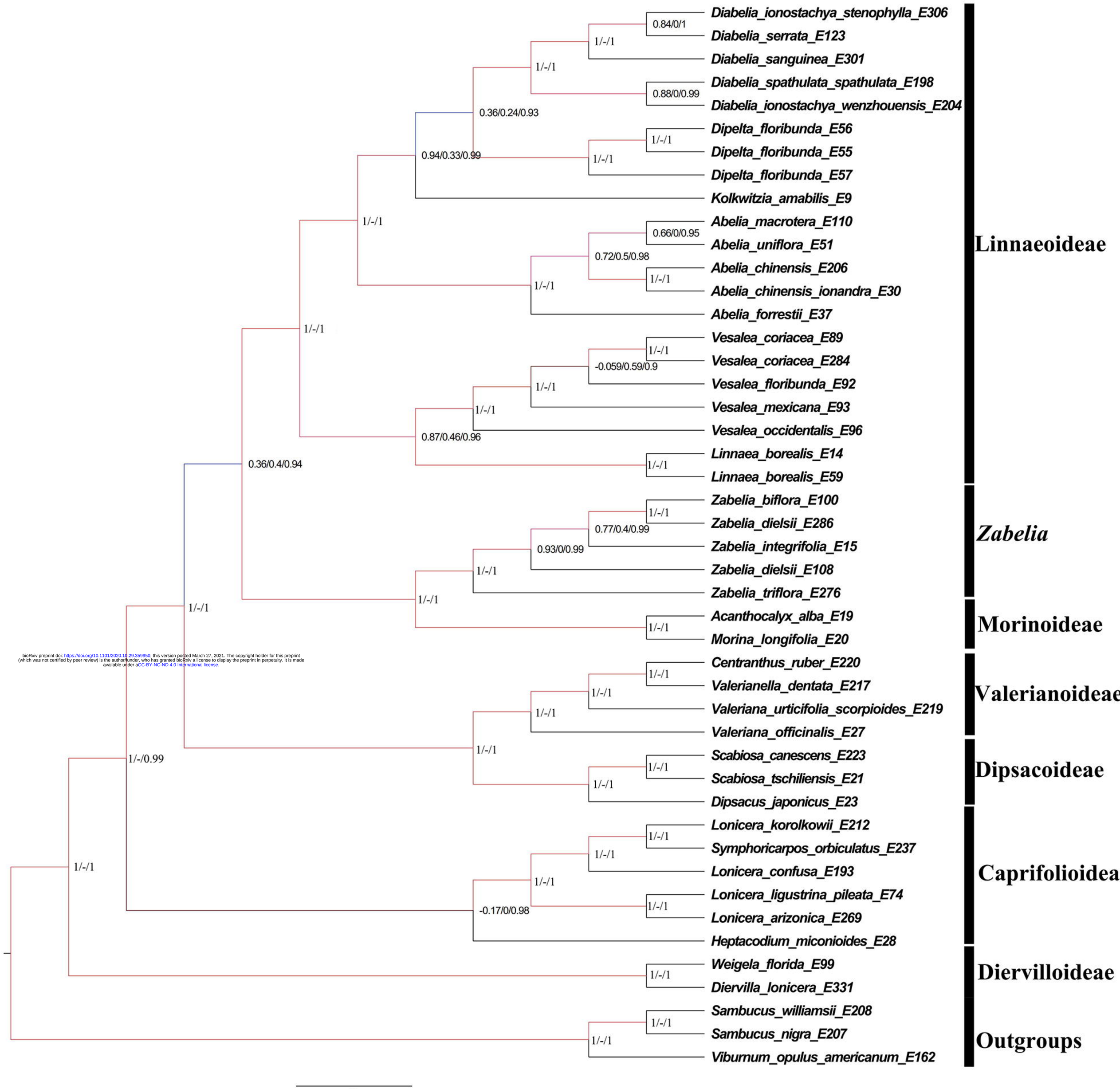
Linnaeoideae
Dipsacoideae
Valerianoideae
Dipsacoideae
Zabelia
Morinoideae
Caprifolioideae
Diervilloideae
Outgroups

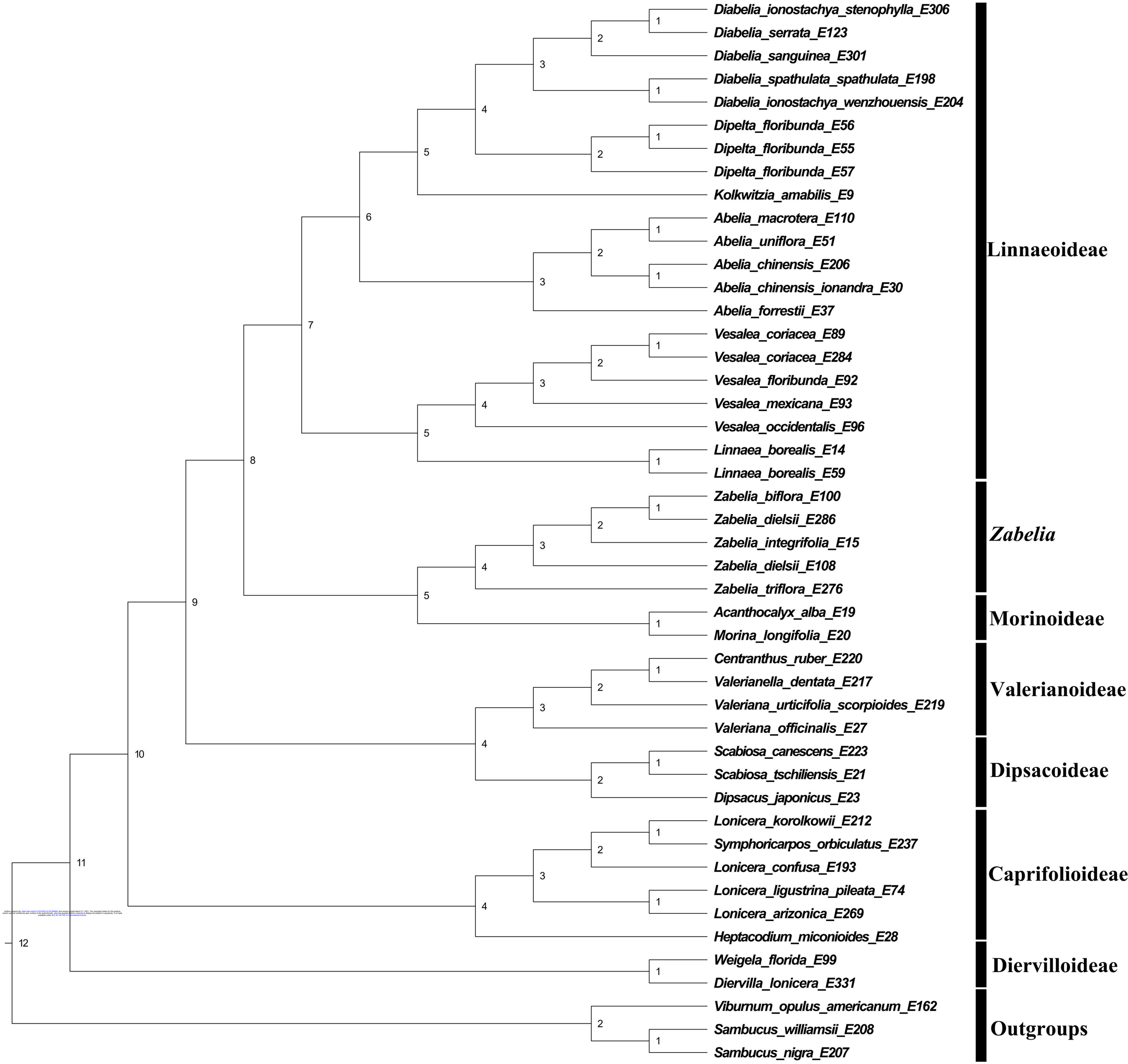
1.66667

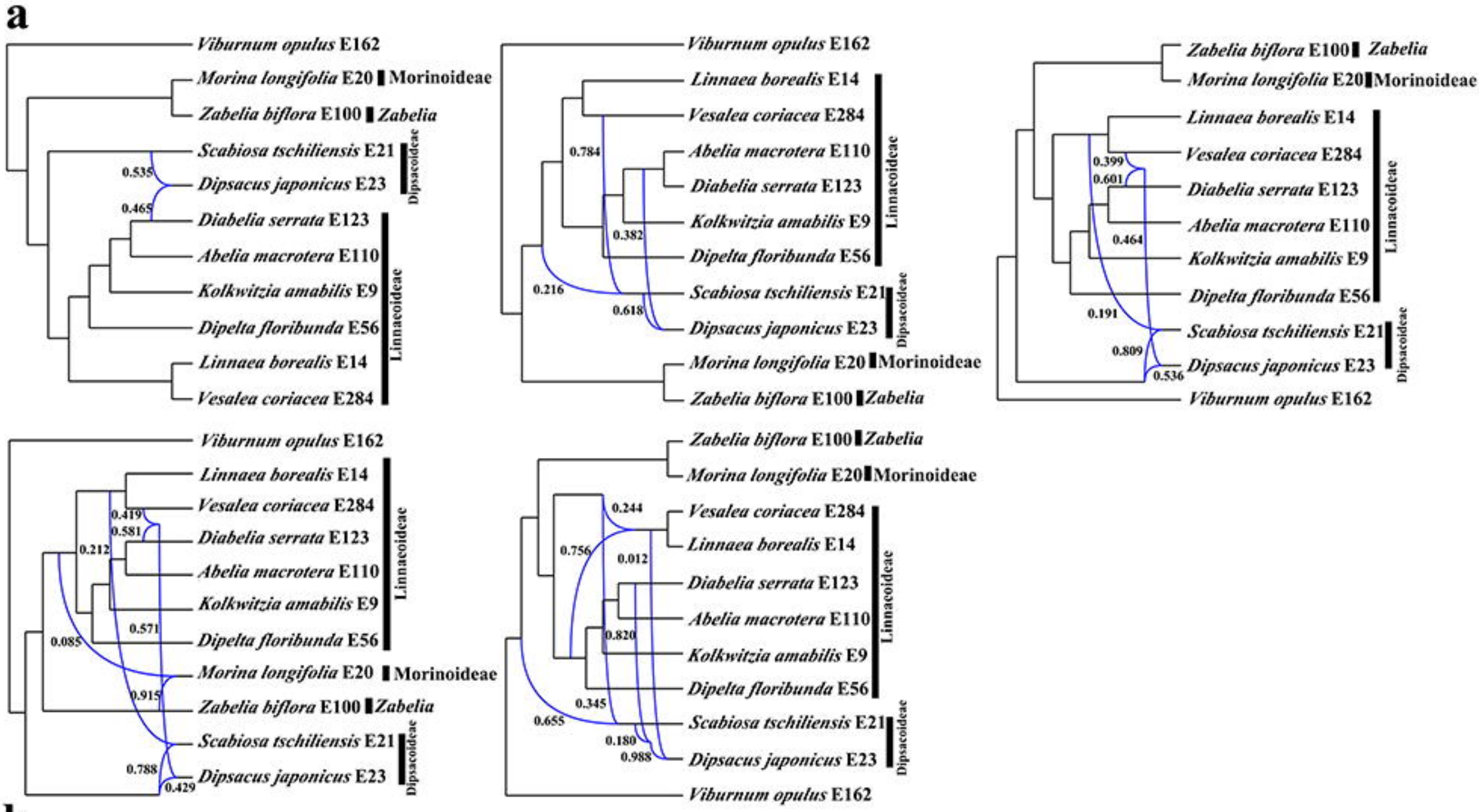


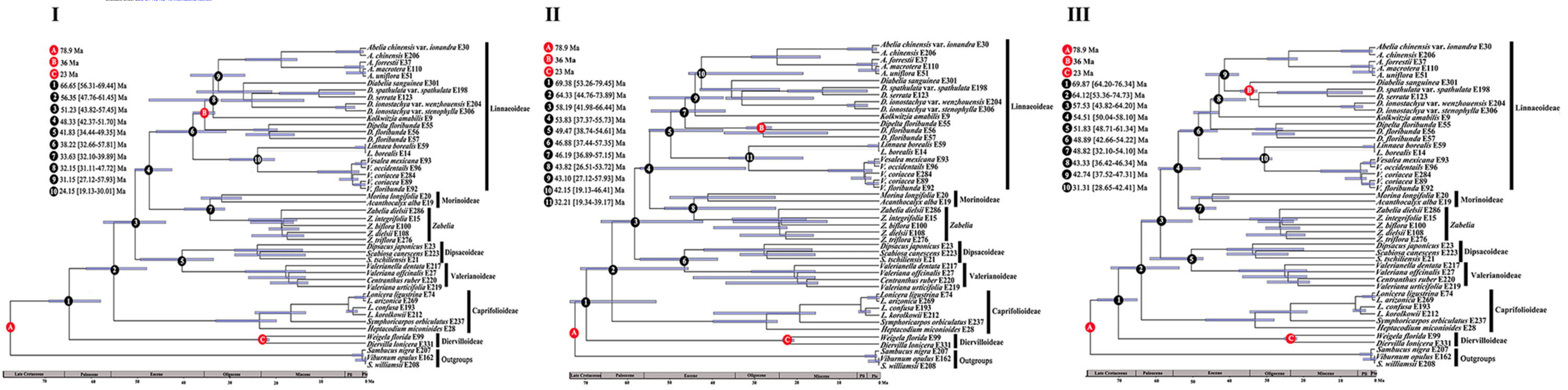







Diabelia_ionostachya_stenophylla_E306
Diabelia_serrata_E123
Diabelia_sanguinea_E301
Diabelia_spathulata_spathulata_E198
Diabelia_ionostachya_wenzhouensis_E204
Dipelta_floribunda_E56
Dipelta_floribunda_E55
Dipelta_floribunda_E57
Kolkwitzia_amabilis_E9
Abelia_macrotera_E110
Abelia_uniflora_E51
Abelia_chinensis_E206
Abelia_chinensis_ionandra_E30
Abelia_forrestii_E37
Vesalea_coriacea_E89
Vesalea_coriacea_E284
Vesalea_floribunda_E92
Vesalea_mexicana_E93
Vesalea_occidentalis_E96
Linnaea_borealis_E14
Linnaea_borealis_E59
Zabelia_biflora_E100
Zabelia_dielsii_E286
Zabelia_integrifolia_E15
Zabelia_dielsii_E108
Zabelia_triflora_E276
Acanthocalyx_alba_E19
Morina_longifolia_E20
Centranthus_ruber_E220
Valerianella_dentata_E217
Valeriana_urticifolia_scorpioides_E219
Valeriana_officinalis_E27
Scabiosa_canescens_E223
Scabiosa_tschiliensis_E21
Dipsacus_japonicus_E23
Lonicera_korolkowii_E212
Symphoricarpos_orbiculatus_E237
Lonicera_confusa_E193
Lonicera_ligustrina_pileata_E74
Lonicera_arizonica_E269
Heptacodium_miconioides_E28
Weigela_florida_E99
Diervilla_lonicera_E331
Viburnum_opulus_americanum_E162
Sambucus_williamsii_E208
Sambucus_nigra_E207
Linnaeoideae
Zabelia
Morinoideae
Valerianoideae
Dipsacoideae
Caprifolioideae
Diervilloideae
Outgroups
0.02




I
A 78.9 Ma
B 36 Ma
C 23 Ma
1 66.65 [56.31-69.44] Ma
2 56.35 [47.76-61.45] Ma
3 51.23 [43.82-57.45] Ma
4 48.33 [42.37-51.70] Ma
5 41.83 [34.44-49.35] Ma
6 38.22 [32.66-57.81] Ma
7 33.63 [32.10-39.89] Ma
8 32.15 [31.11-47.72] Ma
9 31.15 [27.12-57.93] Ma
10 24.15 [19.13-30.01] Ma
Abelia chinensis var. ionandra E30
A. chinensis E206
A. forrestii E37
A. macrotera E110
A. uniflora E51
Diabelia sanguinea E301
D. spathulata var. spathulata E198
D. serrata E123
D. ionostachya var. wenzhouensis E204
D. ionostachya var. stenophylla E306
Kolkwitzia amabilis E9
Dipelta floribunda E55
D. floribunda E56
D. floribunda E57
Linnaea borealis E59
L. borealis E14
Vesalea mexicana E93
V. occidentails E96
V. coriacea E284
V. coriacea E89
V. floribunda E92
Linnaeoideae
Morina longifolia E20
Acanthocalyx alba E19
Morinoideae
Zabelia dielsii E286
Z. integrifolia E15
Z. biflora E100
Z. dielsii E108
Z. triflora E276
Zabelia
Dipsacus japonicus E23
Scabiosa canescens E223
Dipsacoideae
S. tschiliensis E21
Valerianella dentata E217
Valeriana officinalis E27
Centranthus ruber E220
Valeriana urticifolia E219
Valerianoideae
Lonicera ligustrina E74
L. arizonica E269
L. confusa E193
L. korolkowii E212
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Caprifolioideae
Weigela florida E99
Diervilla lonicera E331
Diervilloideae
Sambucus nigra E207
Viburnum opulus E162
S. williamsii E208
Outgroups
Late Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Pli
Plc
70
60
50
40
30
20
10
0 Ma
II
A 78.9 Ma
B 36 Ma
C 23 Ma
1 69.38 [53.26-79.45] Ma
2 64.33 [44.76-73.89] Ma
3 58.19 [41.98-66.44] Ma
4 53.83 [37.37-55.73] Ma
5 49.47 [38.74-54.61] Ma
6 46.88 [37.44-57.35] Ma
7 46.19 [36.89-57.15] Ma
8 43.82 [26.51-53.72] Ma
9 43.10 [27.12-57.93] Ma
10 42.15 [19.13-46.41] Ma
11 32.21 [19.34-39.17] Ma
Abelia chinensis var. ionandra E30
A. chinensis E206
A. forrestii E37
A. macrotera E110
A. uniflora E51
Diabelia sanguinea E301
D. spathulata var. spathulata E198
D. serrata E123
D. ionostachya var. wenzhouensis E204
D. ionostachya var. stenophylla E306
Kolkwitzia amabilis E9
Dipelta floribunda E55
D. floribunda E56
D. floribunda E57
Linnaea borealis E59
L. borealis E14
Vesalea mexicana E93
V. occidentails E96
V. coriacea E284
V. coriacea E89
V. floribunda E92
Linnaeoideae
Morina longifolia E20
Acanthocalyx alba E19
Morinoideae
Zabelia dielsii E286
Z. integrifolia E15
Z. biflora E100
Z. dielsii E108
Z. triflora E276
Zabelia
Dipsacus japonicus E23
Scabiosa canescens E223
Dipsacoideae
S. tschiliensis E21
Valerianella dentata E217
Valeriana officinalis E27
Centranthus ruber E220
Valeriana urticifolia E219
Valerianoideae
Lonicera ligustrina E74
L. arizonica E269
L. confusa E193
L. korolkowii E212
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Caprifolioideae
Weigela florida E99
Diervilla lonicera E331
Diervilloideae
Sambucus nigra E207
Viburnum opulus E162
S. williamsii E208
Outgroups
Late Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Pli
Plc
70
60
50
40
30
20
10
0 Ma
III
A 78.9 Ma
B 36 Ma
C 23 Ma
1 69.87 [64.20-76.34] Ma
2 64.12[53.36-74.73] Ma
3 57.53 [43.82-64.20] Ma
4 54.51 [50.04-58.10] Ma
5 51.83 [48.71-61.34] Ma
6 48.89 [42.66-54.22] Ma
7 48.82 [32.10-54.10] Ma
8 43.33 [36.42-46.34] Ma
9 42.74 [37.52-47.31] Ma
10 31.31 [28.65-42.41] Ma
Abelia chinensis var. ionandra E30
A. chinensis E206
A. forrestii E37
A. macrotera E110
A. uniflora E51
Diabelia sanguinea E301
D. spathulata var. spathulata E198
D. serrata E123
D. ionostachya var. wenzhouensis E204
D. ionostachya var. stenophylla E306
Kolkwitzia amabilis E9
Dipelta floribunda E55
D. floribunda E56
D. floribunda E57
Linnaea borealis E59
L. borealis E14
Vesalea mexicana E93
V. occidentails E96
V. coriacea E284
V. coriacea E89
V. floribunda E92
Linnaeoideae
Morina longifolia E20
Acanthocalyx alba E19
Morinoideae
Zabelia dielsii E286
Z. integrifolia E15
Z. biflora E100
Z. dielsii E108
Z. triflora E276
Zabelia
Dipsacus japonicus E23
Scabiosa canescens E223
Dipsacoideae
S. tschiliensis E21
Valerianella dentata E217
Valeriana officinalis E27
Centranthus ruber E220
Valeriana urticifolia E219
Valerianoideae
Lonicera ligustrina E74
L. arizonica E269
L. confusa E193
L. korolkowii E212
Symphoricarpos orbiculatus E237
Heptacodium miconioides E28
Caprifolioideae
Weigela florida E99
Diervilla lonicera E331
Diervilloideae
Sambucus nigra E207
Viburnum opulus E162
S. williamsii E208
Outgroups
Late Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Pli
Plc
70
60
50
40
30
20
10
0 Ma