

Phylogeny, historical biogeography, and diversification of angiosperm order Ericales suggest ancient Neotropical and East Asian connections

Jeffrey P. Rose^{a,*}, Thomas J. Kleist^b, Stefan D. Löffstrand^c, Bryan T. Drew^d, Jürg Schönenberger^e, Kenneth J. Sytsma^a

^a Department of Botany, University of Wisconsin-Madison, 430 Lincoln Dr., Madison, WI 53706, USA

^b Department of Plant Biology, Carnegie Institution for Science, 260 Panama St., Stanford, CA 94305, USA

^c Department of Ecology, Environment and Botany, Stockholm University, SE-106 91 Stockholm Sweden

^d Department of Biology, University of Nebraska-Kearney, Kearney, NE 68849, USA

^e Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, AT-1030, Vienna, Austria

ARTICLE INFO

Keywords:

Ericaceae
Ericales
Long distance dispersal
Supermatrix
Theaceae
Vicariance

ABSTRACT

Inferring interfamilial relationships within the eudicot order Ericales has remained one of the more recalcitrant problems in angiosperm phylogenetics, likely due to a rapid, ancient radiation. As a result, no comprehensive time-calibrated tree or biogeographical analysis of the order has been published. Here, we elucidate phylogenetic relationships within the order and then conduct time-dependent biogeographical and diversification analyses by using a taxon and locus-rich supermatrix approach on one-third of the extant species diversity calibrated with 23 macrofossils and two secondary calibration points. Our results corroborate previous studies and also suggest several new but poorly supported relationships. Newly suggested relationships are: (1) holoparasitic Mitrastemonaceae is sister to Lecythidaceae, (2) the clade formed by Mitrastemonaceae + Lecythidaceae is sister to Ericales excluding balsaminoids, (3) Theaceae is sister to the styracoids + sarracenioids + ericoids, and (4) subfamilial relationships with Ericaceae suggest that Arbutioideae is sister to Monotropeoideae and Pyroloideae is sister to all subfamilies excluding Arbutioideae, Enkianthoideae, and Monotropeoideae. Our results indicate Ericales began to diversify 110 Mya, within Indo-Malaysia and the Neotropics, with exchange between the two areas and expansion out of Indo-Malaysia becoming an important area in shaping the extant diversity of many families. Rapid cladogenesis occurred along the backbone of the order between 104 and 106 Mya. Jump dispersal is important within the order in the last 30 My, but vicariance is the most important cladogenetic driver of disjunctions at deeper levels of the phylogeny. We detect between 69 and 81 shifts in speciation rate throughout the order, the vast majority of which occurred within the last 30 My. We propose that range shifting may be responsible for older shifts in speciation rate, but more recent shifts may be better explained by morphological innovation.

1. Introduction

The relative contributions of taxonomic versus character sampling to improved estimation of phylogenetic histories has long been debated. On the one hand, increasing character sampling allows for more data to infer relationships, and is especially useful in the case of rapid radiations where branch lengths are short. On the other hand, increasing taxonomic sampling may “discover” phylogenetically informative characters and may also break up spurious, long branches (Wiens, 2006). Recently, “supermatrix” approaches that infer the phylogenetic history of large numbers of taxa in the presence of high levels of missing

data have been used (de Queiroz and Gatesy, 2007; Janies et al., 2013). Such studies have generated novel insights into the phylogeny, biogeography, diversification, and character evolution in land plant groups where phylogenetic inference based on few taxa and relatively few characters has been problematic (Givnish et al., 2015; Hinchliff and Roalson, 2013; Rose et al., 2016; Smith et al., 2011; Soltis et al., 2013; Testo and Sundue, 2016; Wurdack and Davis, 2009). We present here a supermatrix analysis of the speciose asterid order Ericales and use the resulting phylogenetic framework to evaluate species diversification in the context of biogeographical vicariance and dispersal across continents.

* Corresponding author.

E-mail addresses: jrose3@wisc.edu (J.P. Rose), tkleist@carnegiescience.edu (T.J. Kleist), stefan.lofstrand@su.se (S.D. Löffstrand), bdrewfb@yahoo.com (B.T. Drew), juerg.schoenenberger@univie.ac.at (J. Schönenberger), kjsytsma@wisc.edu (K.J. Sytsma).

<https://doi.org/10.1016/j.ympev.2018.01.014>

Received 25 August 2017; Received in revised form 8 January 2018; Accepted 18 January 2018

Available online 02 February 2018

1055-7903/ © 2018 Elsevier Inc. All rights reserved.

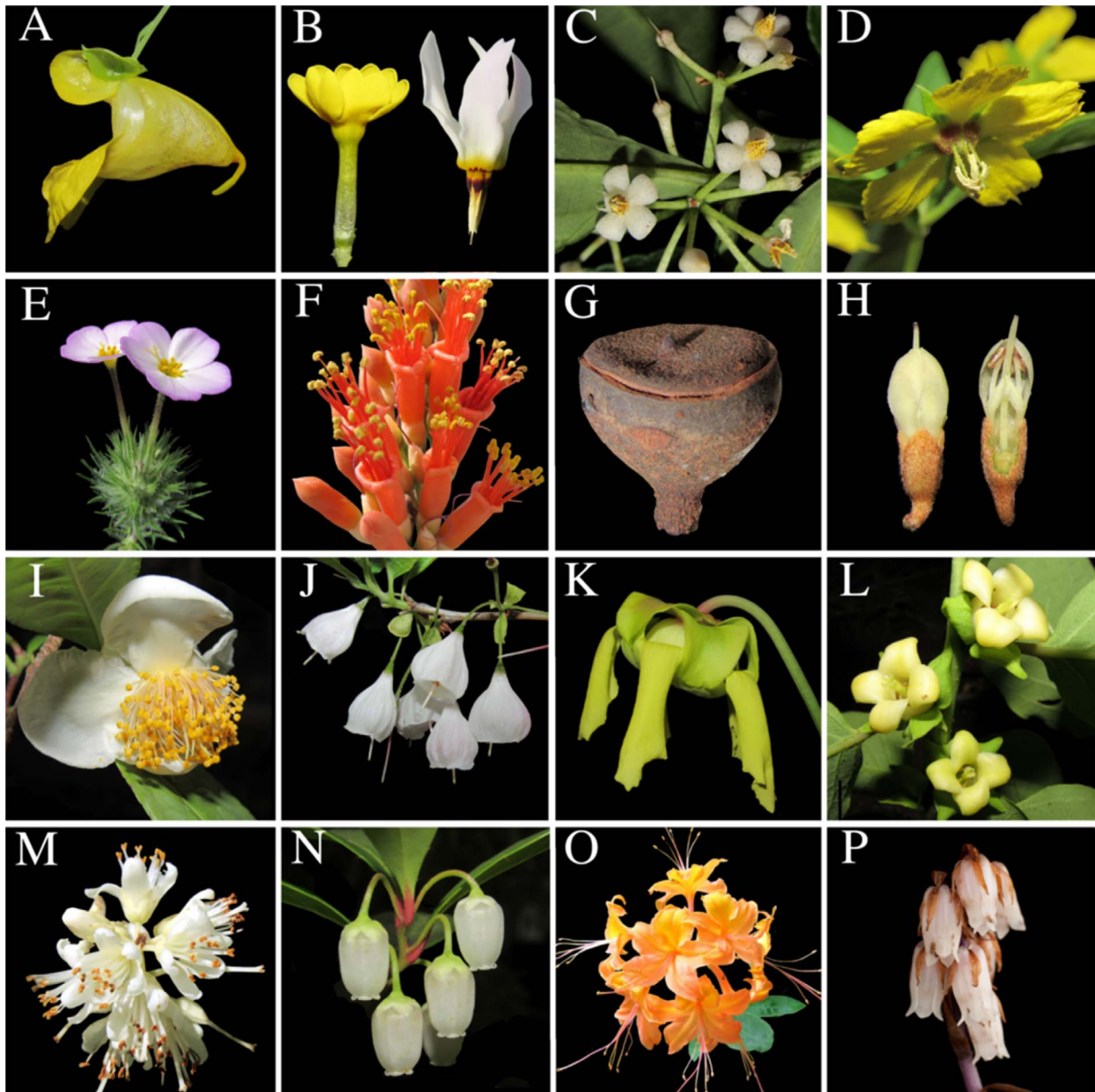


Fig. 1. Representative reproductive diversity within the order Ericales. (A) *Impatiens pallida* (Balsaminaceae) (B) *Primula veris* (left) and *P. meadia* (right) (Primulaceae) (C) *Ardisia crenata* (Primulaceae) (D) *Lysimachia lancifolia* (Primulaceae) (E) *Leptosiphon montanus* (Polemoniaceae) (F) *Fouquieria splendens* (Fouquieriaceae) (G) *Eschweilera sagotiana* (Lecythidaceae) (H) *Synsepalum dulciferum* (Sapotaceae) (I) *Camellia sinensis* (Theaceae) (J) *Halesia carolina* (Styracaceae) (K) *Sarracenia flava* (Sarraceniaceae) (L) *Diospyros virginiana* (Ebenaceae) (M) *Clethra alnifolia* (Clethraceae) (N) *Gaultheria procumbens* (Ericaceae) (O) *Rhododendron calendulaceum* (Ericaceae) (P) *Monotropsis reynoldsiae* (Ericaceae).

1.1. Families and known relationships within Ericales

Ericales is a largely woody, economically important order of ~12,000 species distributed in 21 or 22 families (APG IV, 2016), depending on if Sladeniaceae is recognized as distinct from Pentaphragmaceae. Families now recognized in the order exhibit a wide range of floral diversity, particularly in the androecium, but also in perianth fusion and various other floral characters (Chartier et al., 2017; Schönenberger et al., 2005; Fig. 1). Therefore, it is not surprising that based on morphology alone, members belonging to this order had been thought to be distantly related (Cronquist, 1981). However, the monophyly of the order has been clearly demonstrated (Albach et al., 2001; Anderberg et al., 2002; Bremer et al., 2002; Chase et al., 1993; Ruhfel et al., 2014; Soltis et al., 2000, 2011; Schönenberger et al., 2005). Within the order, monophyly of families has generally been substantiated with the exception of Theaceae (Anderberg et al., 2002; now largely broken up into Theaceae and Pentaphragmaceae) but

inferring relationships among families in the order has been difficult (Anderberg et al., 2002; Geuten et al., 2004; Lens et al., 2007; Schönenberger et al., 2005). As in other higher-level clades of problematic angiosperm groups (e.g., rosids, placement of Caryophyllales, Malphigiales, Saxifragales), this difficulty has been attributed to a rapid radiation in the span of several million years (Davis et al., 2005; Soltis et al., 2010; 2013) which may explain incongruence between genomic regions observed by Geuten et al. (2004). In addition, in those clades of Ericales where the relationships among families have been well-resolved, the relationships often are at odds from a morphological perspective and necessitate comparative anatomical studies to determine both putative synapomorphies and facilitate an understanding of floral evolution in the order (Löfstrand and Schönenberger, 2015a; Schönenberger, 2009; Schönenberger et al., 2010; von Balthazar and Schönenberger, 2013).

The most recent molecular phylogeny of the order is that of Schönenberger et al. (2005) which used 11 molecular markers (two

ribosomal, two mitochondrial, and seven chloroplast) from 59 exemplars of the order. While providing some further resolution along the backbone of the phylogeny from the prior study of [Anderberg et al. \(2002\)](#), several relationships were still unresolved. Nevertheless, [Schönenberger et al. \(2005\)](#) found support for seven major supra-familial clades (informally named here and referred to throughout the text): (1) “balsaminoids” (Balsaminaceae, Marcgraviaceae, Tetrameristaceae), (2) “polemonioids” (Fouquieriaceae, Polemoniaceae), (3) “primuloids” (Ebenaceae, Primulaceae, Sapotaceae), (4) “styracoids” (Diapensiaceae, Styracaceae, Symplocaceae), (5) a clade comprised of Actinidiaceae, Clethraceae, Cyrillaceae, Ericaceae, Roridulaceae, Sarraceniaceae, which has since been divided into the “ericoids” (Clethraceae, Cyrillaceae, Ericaceae) and “sarracenioids” (Actinidiaceae, Roridulaceae, Sarraceniaceae) ([Löfstrand and Schönenberger, 2015a,b](#); [Schönenberger et al., 2012](#)), (6) “theoids” (Theaceae, ericoids, sarracenioids, and styracoids), and (7) “core Ericales” (all clades excluding balsaminoids, polemonioids, and Lecythidaceae)

The holoparasitic family Mitrastemonaceae, consisting of two species in the genus *Mitrastemon* ([Meijer and Veldkamp, 1993](#); [Reveal, 2010](#)), was not included by [Schönenberger et al. \(2005\)](#), despite a clear affinity to the order based on mitochondrial *matR*, *atp1*, and *cox1* sequence data ([Barkman et al., 2004, 2007](#)). Unfortunately, studies including *Mitrastemon* have lacked enough sampling within Ericales to confidently hypothesize its exact placement (APG IV, 2016). The polystemonous condition with connate filaments ([Meijer and Veldkamp, 1993](#)) and unusual placentation suggest potential affinities to Lecythidaceae ([Mori et al., 2007](#); [Mori et al., 2015](#)), although the family appears fairly remote from other Ericales based on overall floral form ([Chartier et al., 2017](#)).

The recent emphasis on whole-plastome sequencing has provided additional information on phylogenetic relationships in Ericales. Unfortunately, only a handful of plastomes from few ericacean families have been sequenced and these families belong to the well-supported clades of [Schönenberger et al. \(2005\)](#). Nevertheless, phylogenetic reconstructions based on coding genes ([Jo et al., 2016](#); [Yan et al., 2016](#)) recapitulate the findings of [Schönenberger et al. \(2005\)](#) but also provide support for a relationship of Theaceae as sister to the ericoids, sarracenioids, and styracoids.

1.2. Biogeographical distributions of Ericales

Biogeographically, Ericales is widespread with both pantropical (Lecythidaceae, Sapotaceae) and cosmopolitan (Ericaceae, Primulaceae) clades. Because classifications based on morphology differ so dramatically compared to the circumscription of the order as currently recognized, no explicit hypotheses about the biogeographic history of the order have been proposed. Biogeographical studies on individual families or subfamilies using more complex models than simple character mapping have been conducted on Diapensiaceae ([Hou et al., 2016](#)), Ebenaceae ([Duangjai et al., 2009](#)), Ericaceae ([Kron and Luteyn, 2005](#)), Sarraceniaceae ([Ellison et al., 2012](#)), Sapotaceae subf. Chrysophylloideae ([Bartish et al., 2011](#); [Swenson et al., 2014](#)), Styracaceae ([Fritsch et al., 2001](#)), and Symplocaceae ([Fritsch et al., 2015](#)), but the order remains to be treated as a whole. These studies have demonstrated a Northern Hemisphere (especially Eurasian) origin of Diapensiaceae, Ericaceae, and Symplocaceae, African and South American or Eurasian and South American origin of Ebenaceae, Southeast Asian or Australian origin of Sapotaceae subf. Chrysophylloideae, and a South American origin of sarracenioids. As most of these families are closely related, no basal nodes of the order have been represented in biogeographical analyses.

Based on the present-day distributions of the grade of families leading to the core families of Ericales, which at least partially occupy Indo-Malaysia and/or the Neotropics (balsaminoids, Lecythidaceae, polemonioids), one might hypothesize the origin of the order in either of these regions. However, the largely Afrotropical distribution of

subfamilies of Lecythidaceae, with the Neotropics as apparently an “apomorphic” area complicates this hypothesis ([Mori et al., 2007](#)). In addition, the pantropical or amphi-Pacific distribution patterns of many families (Lecythidaceae, Primulaceae subf. Myrsinoideae, Sapotaceae) beg the question of whether long distance dispersal or vicariance has played a role in shaping their current distribution patterns, and, if both, what the relative contributions of each has been. Specifically, with regard to widespread tropical clades, the question has been whether their distributions are explained by vicariance following the breakup of Gondwana, boreotropical land bridges, or long distance dispersal ([Givnish and Renner, 2004](#); [Nie et al., 2012](#); [Pennington et al., 2006](#); [Thomas et al., 2015](#)). In their analysis of the amphi-Pacific Symplocaceae, [Fritsch et al. \(2015\)](#) suggested boreotropical migration for the family with an origin in Southeast Asia, spread to North America, and then later spread to South America. Overall, studies on angiosperms have largely suggested that, at least at the family level, clades are generally too young to have their distributions explained by the breakup of Gondwana but may be explained by vicariance through land bridges (either northern or southern hemisphere) or long distance dispersal ([Berger et al., 2016](#); [Cardinal-McTeague et al., 2016](#); [Ruhfel et al., 2016](#); [Thomas et al., 2015](#)). As many families within Ericales have a pantropical or amphi-Pacific distribution, simultaneous analyses of these groups allows for concurrent testing to see if the processes that led to the present-day distributions are similar and/or occurred within the same timeframe.

The recent explosion in the number of discovered Cretaceous fossils of ericacean affinity ([Crepet et al., 2013](#); [Martínez et al., 2016](#); [Schönenberger and Friis, 2001](#); [Schönenberger et al., 2012](#)) from Europe and North America indicates a north temperate origin of at least the ericoid clade of [Schönenberger et al. \(2005\)](#). Not only does this recently enriched fossil record provide clues as to the origin of the order or certain clades within the order (which can be recapitulated with biogeographical analyses of extant tips), but also provides numerous fossils to calibrate a molecular phylogeny.

1.3. Diversification shifts within Ericales

Within Ericales, the occurrence of several large genera (*Impatiens* (Balsaminaceae), ~1,000 spp.; *Rhododendron* (Ericaceae), ~1,000 spp.; *Diospyros* (Ebenaceae), ~700 spp.; *Erica* (Ericaceae), ~700 spp.; *Primula* (Primulaceae) ~400 spp.; *Symplocos* (Symplocaceae), ~300 spp.) suggests the possibility of multiple, temporally and geographically widespread shifts in diversification. Evidence for shifts in these groups has been presented for *Erica* and *Rhododendron* ([Bouchenak-Khelladi et al., 2015](#); [Schwery et al., 2015](#); putatively associated with the rise of montane areas) and *Primula* ([de Vos et al., 2014](#); putatively associated with the evolution of heterostyly). However, it remains to be seen if these putative shifts are significant in the context of the order as a whole.

In this study, we present an updated time-calibrated phylogeny of Ericales using a supermatrix approach, incorporating a rich sample of both taxa and genes. Using this data set, we attempt to (1) resolve backbone relationships in the order, (2) place the holoparasitic family Mitrastemonaceae, (3) estimate when and where major clades of the order originated, (4) examine the biogeographic processes responsible for current day distribution patterns, testing hypotheses related to pantropical distributions (specifically vicariance through polar migrations vs. long distance dispersal), and (5) examine the location and timing of shifts in diversification rates and provide hypotheses to describe their distribution in phylogenetic space.

2. Materials and methods

2.1. Phylogenetic data

We utilized a supermatrix approach of increasing sampling of taxa

and characters, an approach that produces accurate phylogenies (even in the presence of highly incomplete data for some taxa) suitable for the analysis of macroevolutionary patterns (Rose et al., 2016; Soltis et al., 2013; Testo and Sundue, 2016). We gathered sequence data from 4,943 species. These included members of Ericales (4,531 species), Cornales (411 species), and Aquifoliales (one species), with Cornales as the ultimate outgroup.

Genetic data totaled 25 loci representing the chloroplast (*ndhF*, *psbA-trnH*, *rbcL*, *rpl16*, *rpoC1*, *rps16*, *trnL-trnL-trnF*, *trnS-trnG*, *psbB*, *rps2*, *rps4*, *atpB*, *matK*), mitochondrial (*atp1*, *ccmB*, *cob*, *cox1*, *matR*, *mttB*, *nad4-nad5*, *nad5*, *rps3*), and ribosomal (ITS, 18S, 26S) genomes. Most sequence data were derived from GenBank but we also added 470 unpublished sequences, mostly from the mitochondrial genome. (PCR methods follow Schönerberger, et al. (2005); loci follow Soltis et al., 2011). Complete GenBank voucher information and new sequences are listed in Supplementary Table 1 of Appendix A and taxonomic coverage per locus is indicated in Supplementary Table 2 of Appendix A. Individual loci were aligned in MAFFT (Katoh and Standley, 2013) using default parameters. Sequence direction was adjusted according to that of the first sequence to account for potentially reversed sequences in GenBank. Alignments for each locus were then concatenated into a single matrix.

Geuten et al. (2004), and to a lesser extent Schönerberger, et al. (2005) reported higher-level topological conflict between genomic regions within the order. However, most topological conflicts presented in Fig. 3 of Geuten et al. (2004) are not well supported except for relationships between balsaminoid families and a Primulaceae + Pentaphragaceae relationship supported in a plastid-coding dataset.

Our concatenated data matrix was subjected to maximum likelihood (ML) analysis in RaxML (Stamatakis, 2014). To account for different substitution rates among genomes, we used PartitionFinder2 (Lanfear et al., 2017) to test for the best partitioning scheme among three possible *a priori* partitions corresponding to the three genomic regions explored in this study (ribosomal, chloroplast, and mitochondrial). Branch lengths were set to linked. We limited testing for the best fitting nucleotide substitution model to the GTR models implemented in RaxML with AICc selection between models. The best fitting partitioning scheme was found using the “rcluster” search scheme (Lanfear et al., 2014). PartitionFinder2 suggested all three *a priori* partitions should be retained and suggested GTR + I + G as the best fitting nucleotide substitution model for each. Because the RaxML manual recommends against adding invariant sites as a free parameter, we used the slightly less parameterized GTR + G model in all phylogenetic analyses. ML analysis was conducted in RaxML under the GTRGAMMA model of sequence evolution. All other parameters were set to their default values. The best tree was found using 10 separate search replicates. Branch support was assessed through two runs of 250 rapid bootstrap replicates. ML bootstrap samples were summarized on the best tree using SumTrees v. 4.2.0 in the Python package DendroPy v. 4.2.0 (Sukumaran and Holder, 2010).

As the backbone relationships of Ericales have been controversial, we also wanted to assess topological uncertainty given the assumptions in our ML analysis. We therefore conducted a Bayesian analysis (BI) on a subset of 191 species. These species were selected based on taxonomic representativeness and genetic completeness (with the fewest missing nucleotide sequences from the 25 targeted loci). Because the motivation for including a Bayesian analysis in this study was to account for topological uncertainty given the parameters used in our ML analysis, we analyzed this reduced matrix using MrBayes version 3.2.6 (Ronquist et al., 2012) under the same model of sequence evolution and the same three partitions mentioned above. Six separate analyses of three runs of four chains each were run for 10 million generations with sampling every 1,000 generations. All other parameters were set to their default values. We examined parameter convergence in all 18 runs using Tracer v. 1.6.0 (Rambaut et al., 2014). All parameters for this dataset converged within 2 million generations. We discarded the first 5,000 trees

(equivalent to 5 million generations) as a burn-in. The posterior distribution of trees was summarized on a maximum clade credibility (MCC) tree and on a pruned version of the best ML tree using SumTrees. The alignment file and partitions are available via [Mendeley Data](#).

2.2. Fossil calibrations

As the size of the final data matrix precluded dating in BEAST (Drummond et al., 2012), we utilized penalized likelihood as implemented in treePL (Smith and O’Meara, 2012) on our best ML tree. This program has been shown to estimate divergence dates similar to those of BEAST (Lagomarsino et al., 2016). In total, we used 25 calibration points (23 macrofossils and 2 secondary calibrations). We restricted fossil calibrations to macrofossils (flowers, fruits, seeds, leaves) of both Ericales and Cornales that could be confidently assigned to crown groups, as treePL only permits placing minimum and maximum dates around nodes. Several good fossils such as *Glandulocalyx upatoiensis* (Schönerberger et al., 2012) are assignable to the same crown groups as different but older fossils and were therefore not used in our dating analysis. For some important fossils assignable to stems, we placed them on the crown group preceding that stem. In addition to fossils, we restricted the crown ages of asterids, and Ericales + Aquifoliales using secondary calibration points from Magallón et al. (2015) (Table 1, Supplementary Text 1 of Appendix A). For our treePL analysis, we conducted a thorough search using default parameters except we specified 8 threads, random cross-validation, and a smoothing parameter of 100. The treePL configuration file is available via [Mendeley Data](#).

2.3. Biogeographical analyses

To both facilitate ease of coding as well as to remove potential bias in coding, we opted for an automatic means of coding tips. To do this, we first downloaded all databased records of Ericales in the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>). We limited our search to only specimen records and to specimens with no known coordinate issues following established procedures (Maldonado et al., 2015; Spalink et al., 2016a,b). Using this set of coordinates, we then determined which biogeographic realm these specimens belonged to by extracting the “realm” raster layer from the “Terrestrial Ecoregions of the World” shape file (Olson et al., 2001) available from the World Wildlife Fund website (<http://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world>) using the R package *raster* (Hijmans and van Etten, 2016). Realms were then condensed into six areas as follows: (1) Afrotropics, (2) Australasia + Oceania (including Hawaii), (3) Indo-Malaysia, (4) Nearctic, (5) Neotropics, and (6) Palearctic. Each tip in our phylogeny was either present or absent in each area. After initial coding, the range assigned to each tip was checked a second time to ensure that realm assignment was accurate (i.e., the sign of latitudinal coordinates in GBIF was not inverted, cultivated specimens were not included). Additionally, important taxa not present in the GBIF data were scored by hand based on specimen data available on Tropicos (<http://www.tropicos.org/>).

In general, the distribution of each tip is representative of the geographic range of that species. However, for some tips, the tip is representative of the distribution of a higher-level clade as follows. (1) *Mitrastemon yamamotoi* was used as a placeholder for Mitrastemonaceae, which also occurs in Central America. (2) *Pentamerista neotropica* was used as a placeholder for this genus and its sister genus, *Tetramerista*, which occurs in Southeast Asia. (3) *Gordonia singaporeana* was used as a placeholder for the Australasian *G. papuana*, as *G. papuana* has been suggested to have as a close relationship to Indonesian and Philippine taxa (Barker, 1980). (4) *Impatiens capensis* was used as a placeholder for Central American *I. turrialbana*, as its floral morphology is similar to North American species (Donnell-Smith,

Table 1

Macrofossil calibrations used in this study with the minimum and maximum ages and most recent common ancestors of the nodes constrained. * indicates secondary dates ** indicates fossils best assigned to stem lineages instead placed on the next youngest crown.

Fossil (clade)	Age range (My)	Crown group	Reference
<i>Actinocalyx bohrii</i>	84.0–80.6	<i>Diapensia</i> + <i>Styrax</i> **	Fritsch et al. (2015)
<i>Alangium</i> sp.	73–66	<i>Alangium</i> + <i>Cornus</i>	Xiang et al. (2011)
<i>Androsace</i> sp.	5.2–23.2	<i>Androsace</i> + <i>Douglasia</i>	Zhang et al. (2001)
Aquifoliales + Ericales [*]	107.4–117.1	Aquifoliales + Ericales	Magallón et al. (2015)
asterids [*]	110.4–118.9	root	Magallón et al. (2015)
<i>Austrodiospyros cryptostoma</i>	38–47.8	<i>Diospyros</i> + <i>Euclea</i>	Basinger and Christophel (1985)
<i>Dichroa bornensis</i>	29.4–27.4	<i>Hydrangea macrophylla</i> + <i>Dichroa</i>	Schenk and Hufford (2010)
<i>Empetrum</i> sp.	16–11.6	<i>Empetrum eamesii</i> + <i>E. sibiricum</i>	Schwery et al. (2015)
<i>Gilisenium hueberi</i>	42–46.2	<i>Ipomopsis</i> + <i>Phlox</i>	Lott et al. (1998)
<i>Gordonia</i> / <i>Gordoniopsis</i> / <i>Andrewsiocarpon</i>	38–47.8	<i>Camellia</i> + <i>Polyspora</i>	Grote and Dilcher (1992)
<i>Hironoia fusiformis</i>	80–90	<i>Nasa</i> + <i>Nyssa</i>	Schenk and Hufford (2010); Xiang et al. (2011)
<i>Klaprothiopsis dyscrita</i>	45–15	<i>Kissenia</i> + <i>Nasa</i>	Poinar et al. (2015)
<i>Lyonia</i> sp.	16–11.6	<i>Lyonia fruticosa</i> + <i>L. ovalifolia</i>	Schwery et al. (2015)
<i>Lysimachia angulata</i>	33.9–28.1	<i>Trientalis</i> + <i>Lysimachia</i>	Boucher et al. (2016)
<i>Paleoenkianthus sayrevillensis</i>	89.8–93.9	<i>Cyrilla</i> + <i>Pyrola</i> **	Nixon and Crepet (1993)
<i>Parasaurauia allonensis</i>	72.1–83.6	<i>Actinidia</i> + <i>Roridula</i> **	Keller et al. (1996)
<i>Pouteria costata</i>	20.4–28.1	<i>Sarcosperma</i> + <i>Planchonella</i>	Bartish et al. (2011)
<i>Primula rosiae</i>	16–11.6	<i>Primula maedia</i> + <i>P. sieboldii</i>	Boucher et al. (2016)
<i>Raritaniflora</i> spp.	89.8–93.9	<i>Arctous</i> + <i>Clethra</i>	Crepet et al. (2013)
<i>Rehderodendron stonoi</i>	56.0–47.8	<i>Rehderodendron</i> + <i>Styrax</i>	Fritsch et al. (2015)
<i>Rhododendron newberryianum</i>	55–56	<i>Rhododendron groenlandicum</i> + <i>R. redowskianum</i> **	Collinson and Crane (1978)
<i>Richeaphyllum waimumuensis</i>	28.1–16	<i>Dracophyllum</i> + <i>Sphenotoma</i>	Jordan et al. (2010)
<i>Schima nanlinensis</i>	11.6–5.3	<i>Schima sinensis</i> + <i>S. wallichii</i>	Li et al. (2013a,b)
<i>Symplocos pseudogregaria</i>	37.8–33.9	<i>Symplocos zizyphoides</i> + <i>S. povedae</i>	Fritsch et al. (2015)
<i>Symplocos quadrilocularis</i>	56.0–47.8	<i>Cordyloblaste</i> + <i>Symplocos</i>	Fritsch et al. (2015)

1897).

One unresolved issue in coding was the absence of molecular data for several clades of Pentaphragmaceae. These include the African genus *Balthasaria*, African species of *Ternstroemia*, and the endemic New Guinean *Archboldiodendron* and *Adinandra forbesii* (Barker, 1980). In the case of *Ternstroemia*, only two species are present in Africa and were hypothesized by Kobuski (1961) to be closely related to the numerous Asian and South American species in the genus with no further hypotheses on their relationships. The DNA of *Balthasaria* (*Melchiora*) has never been sequenced. The phylogeny of Luna and Ochoterena (2004) suggests strong topological conflict between molecular and morphological data and thus *Balthasaria* cannot be confidently placed based on morphological data. Barker (1980) suggested a close relationship of *Adinandra forbesii* to Indonesian species with glabrous ovaries. However, these species are absent from the phylogeny. Based on morphology, Luna and Ochoterena (2004) found *Archboldiodendron* to be sister to *Adinandra*. However, *Adinandra* is not monophyletic (Wu et al., 2007). Because it is also possible that *Archboldiodendron* may be nested in *Adinandra*, it is hard to place this genus in the phylogeny without molecular data. Therefore, the range of areas represented in Pentaphragmaceae is not completely representative, but it is likely that these missing clades are recent in origin.

We performed biogeographical ancestral range (AR) estimation (ARE) analysis on our best ML tree using the R package BioGeoBEARS (Matzke, 2012, 2013). As Ericales is sister to a clade of cosmopolitan asterids, we excluded all outgroups for the biogeographical analysis. We performed a stratified analysis using the same dispersal probability assumptions and time periods as in Berger et al. (2016), allowing ancestors to occupy up to four areas. Specifically, we employed the Dispersal, Extinction and Cladogenesis model (DEC) (Ree and Smith, 2008) with and without jump dispersal (j parameter in the BioGeoBEARS supermodel; Matzke, 2012, 2014). The j parameter allows for a daughter lineage to occupy a new area that is different from the parental lineage. As we were also interested in the location and probability of particular events on each edge of the phylogeny, we conducted 100 replicates of biogeographical stochastic mapping in BioGeoBEARS (Matzke, 2016). All input files for BioGeoBEARS are available via Mendeley Data.

2.4. Diversification analyses

We utilized BAMM version 2.5.0 (Rabosky, 2014; Rabosky et al., 2017) on our ML tree to examine the location and timing of shifts in Ericales. Because of the large size of the asterids sister to Ericales, we discarded all outgroup taxa. As our ML tree only contains about a third of the species diversity in the order, we provided BAMM with a sampling fraction file. We assigned tips to groups in the fraction file following a thorough survey of the taxonomic literature with a particular focus on monographic treatments and paying close attention to synonymy at the generic level and the probable placement of genera for which no genetic data exists (Supplementary Table 3 of Appendix A). Group assignment in the fraction file was as detailed as possible (generally genus level) and we used the following procedure: (1) We examined the topology of our ML tree to determine where genera or subgenera were monophyletic, paraphyletic or polyphyletic. (2) We referenced the taxonomic literature to find monographic treatments for each genus or subgenus. (3) In cases where the genus is monophyletic, assignment was straightforward. (4) If a genus was paraphyletic or not broadly polyphyletic, we assigned a number to that genus, plus any genera embedded within it. (5) In cases of broad polyphyly of a genus or in cases where a family or subfamily was missing genera, we assigned a species number to that entire clade. Following our thorough search of monographic treatments for individual families and genera, a total of 12,520 species of Ericales are recognized, although specific numbers for some families or genera are estimates due to taxonomic uncertainty (e.g., Sapotaceae, Govaerts et al. (2001); *Impatiens*, Yu et al., 2015).

Priors for BAMM were generated using the R package BAMMtools v. 2.1.6 (Rabosky et al., 2014). We ran four MCMC chains for 100,000,000 generations with parameters logged every 2,000 generations using the generated BAMM priors and 1 expected shift. All other priors were set to their default values. Chain convergence was assessed using the R package coda v. 0.19-1 (Plummer et al., 2006). After removing 25% of generations as a burn-in, we utilized BAMMtools to compute the 95% credible speciation rate shift configuration using a Bayes factor (BF) criterion of 6 with the expected number of shifts set to 1. A large BF indicates strong evidence for a shift on a given branch. We estimated

the speciation rate shift configuration with the highest posterior probability after excluding all non-core shifts. In addition, we examined the diversification metrics of speciation rate, extinction rate, and net diversification rate for individual families as well as subfamilies in Ericaceae and Primulaceae using the “getCladeRates” function in BAMMtools. All input files for BAMM are available via [Mendeley Data](#).

To examine if there has been a tree-wide shift in speciation and/or extinction rate through time, we also used the CoMET model as implemented in the R package TESS (Höhna et al., 2016; May et al., 2016). The MCMC was conducted for a maximum of 20,000,000 iterations and a burn-in of 5,000,000 iterations. Because our tree is not fully sampled, we set the sampling probability to 0.363 (4,543/12,520 spp.). All other priors were set to their default values. After the analysis finished, parameter convergence was assessed using coda. As with the BAMM analysis, BF that exceeded 6 were interpreted as support for a tree-wide shift in diversification rates.

3. Results

3.1. Phylogenetic inference and divergence dates

The best ML tree from the analysis of our supermatrix of 49,435 aligned nucleotides (87.6% missing data, varying from 26.8% to 99.5% per locus; Supplementary Table 2 of Appendix A) with three genomic partitions resulted in a fully resolved tree of the order (Fig. 2, Supplementary data 1), albeit with very poor support for some relationships. Our results largely recapitulate those of Schönenberger et al. (2005). The topology of the MCC tree from our BI result (Supplementary data 2) is identical to the ML tree for all major nodes, with the only topological differences being found within families, especially where branch lengths are small. These topological differences are the placement of *Cyclamen* within Primulaceae subf. Myrsinoideae, relationships within Sapotaceae subf. Chrysophylloideae, placement of *Franklinia* relative to *Gordonia* and *Schima* in Theaceae, and placement of *Bryanthus* in Ericaceae. Though the topology is identical between the two optimality criteria, low support values for these branches indicate that topological uncertainty around several nodes persists, but for fewer nodes than in previous studies.

We recover balsaminoids as the clade (BS = 99, PP = 1.0) sister to the rest of Ericales (BS = 95, PP = 0.99). Relationships within balsaminoids remain unclear, with either Marcgraviaceae or Tetrameristaceae sister to Balsaminaceae. Both our BI and ML analyses suggest the placement of Marcgraviaceae as sister to the remaining balsaminoid families (BS = 53, PP = 0.29). The next clade sister to the remaining Ericales is comprised of Lecythidaceae + Mitrastemonaceae (BS = 62, PP = 0.54), with the monophyly of the rest of the order moderately to weakly supported (BS = 58, PP = 0.90). Monophyly of Lecythidaceae receives fairly weak support in the ML analysis (BS = 77, PP = 1.0). The next clade we recover is the polemonioids (BS = 98, PP = 1.0), which is sister to a monophyletic but weakly supported core Ericales (BS = 50, PP = 0.69). Within core Ericales, the primuloids are recovered as monophyletic but with weak support (BS = 59, PP = 0.69) and are sister to the remainder of the core Ericales (BS = 34, PP = 0.69). Within this remaining core Ericales, monophyletic Pentaphragaceae (BS = 93, PP = 1.0) and Theaceae (BS = 99, PP = 1.0) form a grade with the remainder of the core Ericales (BS = 75, PP = 0.69 and BS = 42, PP = 0.69, respectively), with Theaceae more closely related to the styracoids, sarracenioids, and ericoids.

Styracoids are monophyletic (BS = 83, PP = 0.98) and sister to sarracenioids and ericoids (BS = 93, PP = 1.0). Likewise, sarracenioids are monophyletic (BS = 78, PP = 1.0) and sister to ericoids (BS = 75, PP = 1.0). Within ericoids, Clethraceae and Cyrillaceae form a grade sister to Ericaceae, with Cyrillaceae more closely related to Ericaceae. However, monophyly of Clethraceae (including *Purdiaea*) is questionable (BS = 71, PP = 0.67). Within Ericaceae, sub-familial relationships

suggest Arbutoideae + Monotropoideae is sister to the remainder of the family, excluding Enkianthoideae (BS = 98, PP = 1.0), with the Pyroloideae weakly supported as comprising a clade with the remaining subfamilies (BS = 65, PP = 0.62). Vaccinioideae and Styphelioideae form a clade (BS = 87, PP = 1.0) sister to Harrimanelloideae. These three subfamilies are in turn sister to Cassiopoideae and Ericoideae. As mentioned above, within Ericoideae the placement of *Bryanthus* is topologically discordant between the ML and BI results. In the ML analysis it is sister to *Ledothamnus* (BS = 72), while in the BI results it is sister to *Daboecia* + *Calluna* + *Erica* but with no support (PP = 0.0035).

Our treePL chronogram (Fig. 3, Supplementary data 3) places the age of the crown of Ericales at ~110 Mya. Rapid cladogenesis along the backbone of the order occurred between 106 and 104 Mya. By ~83 Mya stem lineages of all extant families had originated, but familial crown ages are highly variable, ranging from ~95 Mya (Pentaphragaceae) to ~11 Mya (Fouquieriaceae). Rapid cladogenesis is also found within balsaminoids with a crown age of ~82 Mya and divergence date for Tetrameristaceae and Balsaminaceae at ~81 Mya. Lecythidaceae and Mitrastemonaceae diverged ~104 Mya. Crown polemonioids dates to ~99 Mya, with crown Polemoniaceae much older than crown Fouquieriaceae (~54 Mya and ~15 Mya, respectively). Within Polemoniaceae, subf. Acanthogillioideae is recovered as sister to the remainder of the family (BS = 84, PP = 1.0). Crown primuloids dates to ~102 Mya with a Primulaceae/Ebenaceae divergence date of ~80 Mya. Core Ericales began to diverge ~104 Mya followed quickly by the divergence of crown theoids at ~102.2 Mya and crown styracoids + sarracenioids + ericoids at ~101.6 Mya. Crown styracoids date to ~99 Mya with the divergence date of Styracaceae/Diapensiaceae at ~93 Mya. Crown sarracenioids date to ~93 Mya with an Actinidiaceae/Roridulaceae divergence date of ~80 Mya. Divergence between ericoid families occurred between ~98 and 91 Mya, with crown Ericaceae dating to ~91 Mya.

3.2. Historical biogeography

We were able to automatically and confidently score geographic ranges for 3,875 tips, with an additional 333 tips scored manually. Based on a likelihood ratio test, the BioGeoBEARS DECj model of ancestral ranges (lnL = -5709.69) was a significantly better fit for the data compared to DEC (lnL = -5772.72) based on a likelihood ratio test (D = 126.06, P = 0.0). Since jump dispersal is an important facet of the historical biogeography, subsequent discussion will only involve the model with this parameter (Fig. 4; Supplementary Fig. 1 of Appendix A). Parameters of the DECj model include anagenetic dispersal rate $d = 9.77 \times 10^{-3}$, extinction rate $e = 3.64 \times 10^{-5}$, and cladogenetic dispersal rate $j = 5.91 \times 10^{-3}$. Biogeographical stochastic mapping, given the parameters of our DECj model, indicates that of 4,207 cladogenetic events, 84% are sympatry, 10% are subset sympatry, 3.5% are vicariance, and 2.7% are founder-event dispersals. However, the relative contribution of vicariance and dispersal to the biogeographic history of the order is not uniformly distributed through time, with vicariance playing a much more important role early on in the history of the order, with dispersal only becoming important in the last 30 My. Within the last 30 My, however, both processes have contributed to the historical biogeography of the family (Fig. 5).

While probabilities for individual areas of combinations of areas were low for the basal-most nodes of Ericales, many of these nodes have relatively high probabilities for an ARE containing the Neotropics and/or Indo-Malaysia. Crown Ericales originated in a shared Neotropical + Indo-Malaysian AR (P = 0.29) with all other possible ranges having a small probability. The backbone nodes of Ericales are reconstructed as remaining in this shared AR or switching to solely to a Neotropical AR. Balsaminoids originated in this shared Neotropical + Indo-Malaysian AR, (P = 0.55) and shifted to an Indo-Malaysian AR along stem Balsaminaceae. Mitrastemonaceae + Lecythidaceae most likely

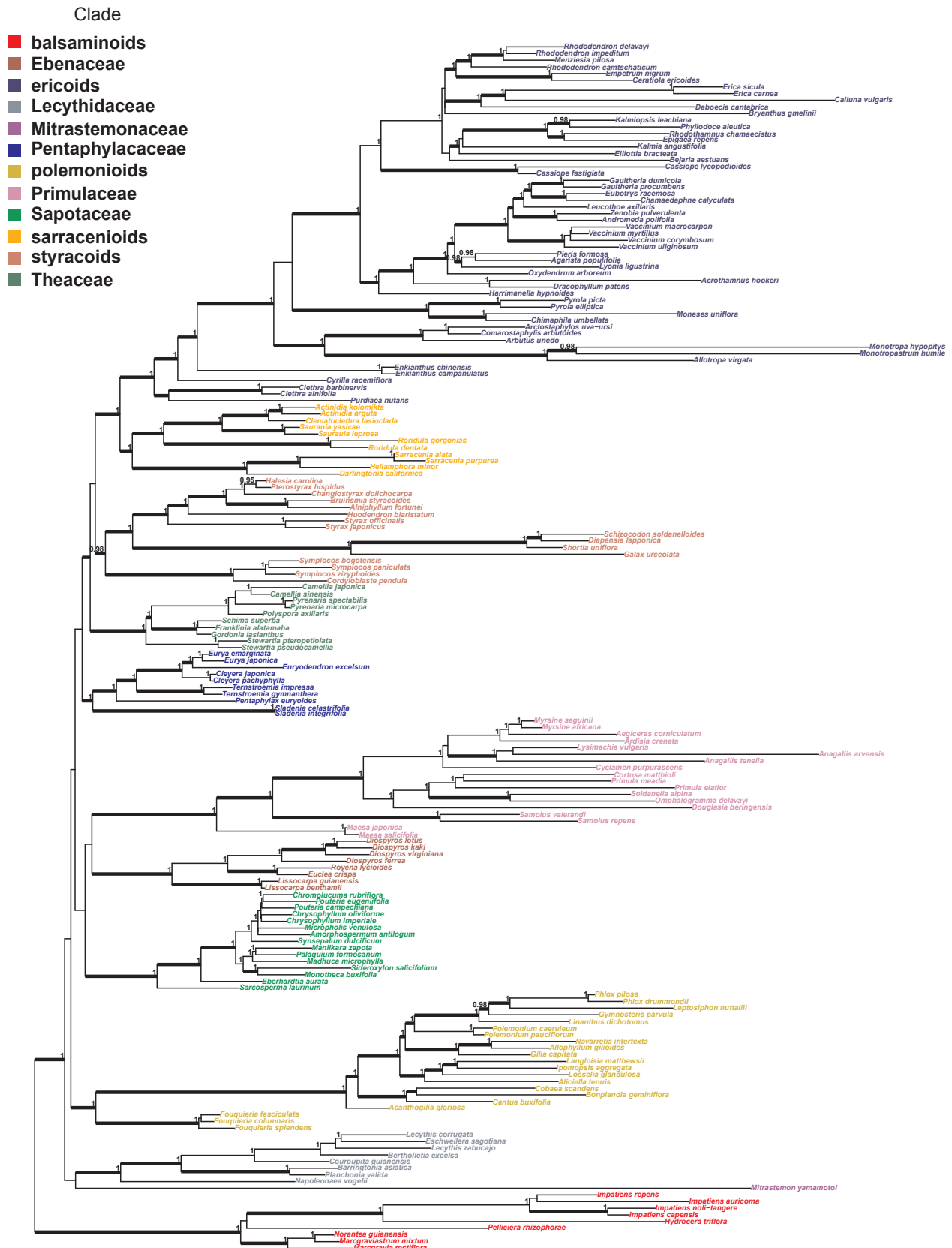


Fig. 2. Best maximum likelihood (ML) phylogram of Ericales pruned to match the taxonomic sampling of the Bayesian analysis with major clades colored. Thickened branches indicate ML bootstrap values > 70. Bayesian posterior probabilities > 0.95 are indicated numerically above branches. Note that toward the tips ML bootstrap values may be low because of pruning. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

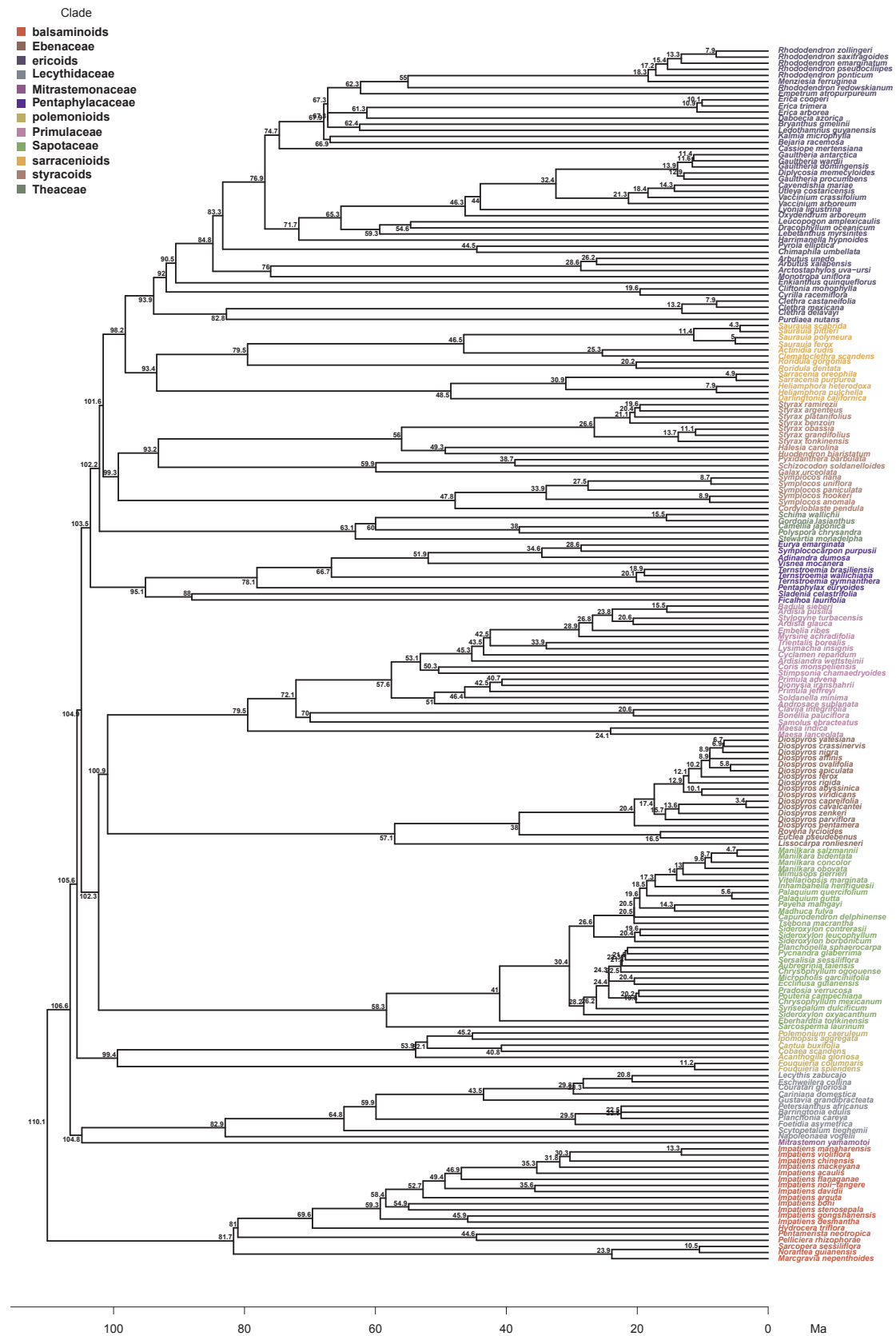


Fig. 3. Chronogram of Ericales. Numbers at nodes indicate node ages for each clade rounded to the nearest tenth of a million years. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

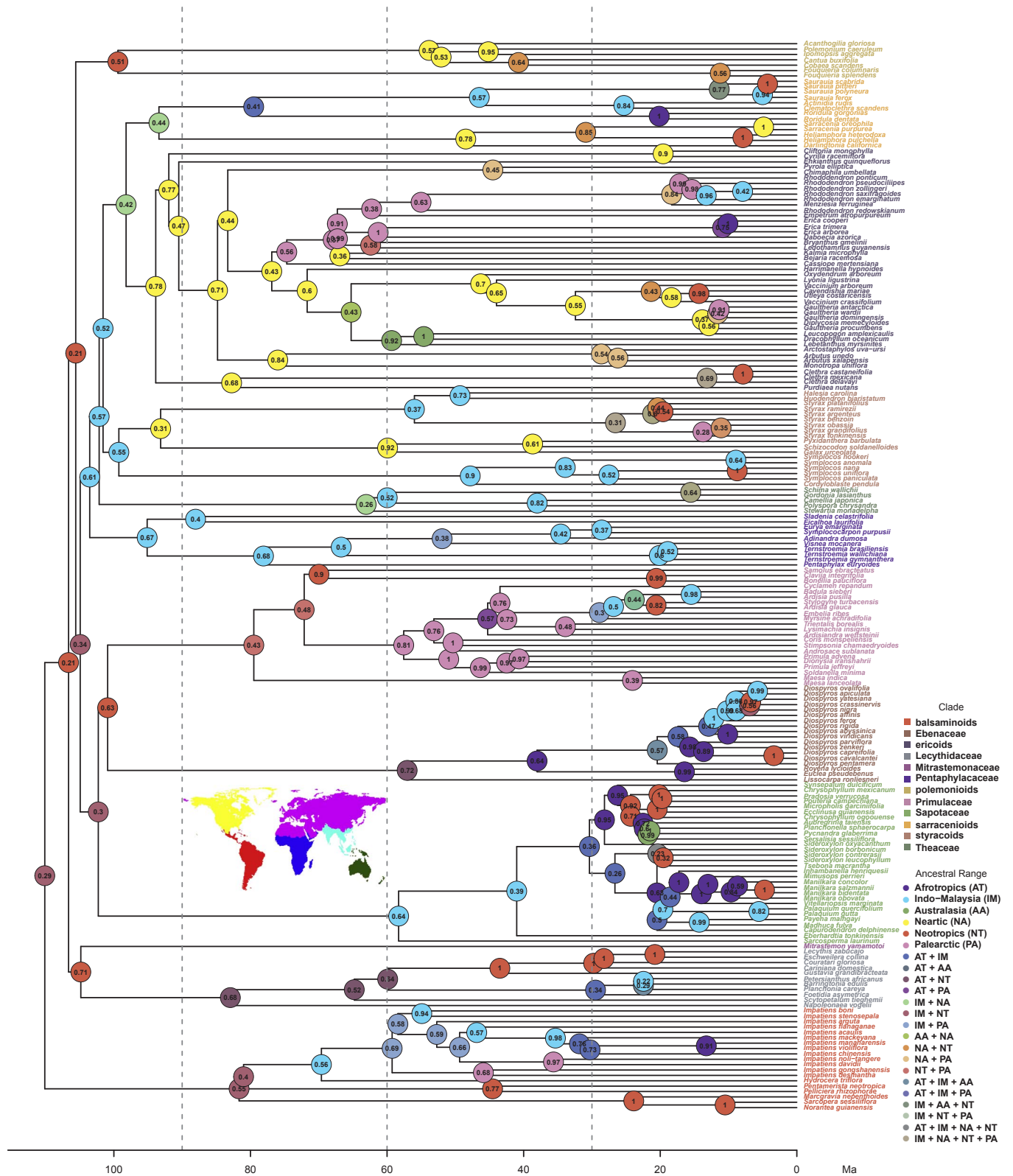


Fig. 4. Most likely ancestral range for major inter-familial and intra-familial clades of Ericales based on our DECj BioGeoBEARS model. Geographic areas as defined in this study are indicated on the map. Circles at nodes indicate the most likely ancestral range of the most recent common ancestor of that clade, with the probability of occupying this ancestral range indicated within each circle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

originated in the Neotropics ($P = 0.71$) with dispersal to Indo-Malaysia along stem Mitrastemonaceae. Crown Lecythidaceae originated in a shared Afrotropical + Neotropical AR ($P = 0.58$). Polemonioids likely originated in the Neotropics ($P = 0.51$), although a Nearctic origin is

also likely. Primuloids originated in a shared Neotropical + Indo-Malaysian AR ($P = 0.30$), although a Neotropical origin is also possible. Stem Sapotaceae originated in Indo-Malaysia, an AR that the crown of the family also occupied ($P = 0.64$). Crown Ebenaceae + Primulaceae

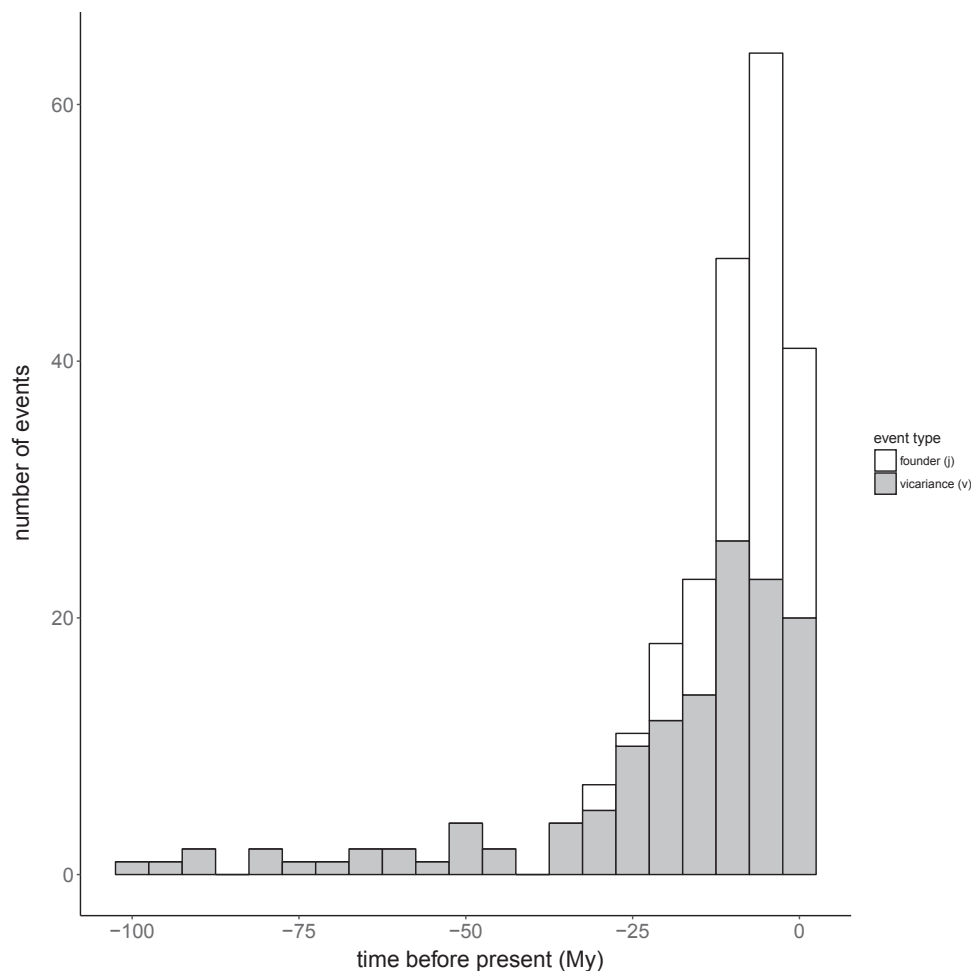


Fig. 5. Histogram of the temporal contribution of founder and vicariance events to the historical biogeography of Ericales as indicated based on 100 stochastic maps on our DECj BioGeoBEARS model. Each bin corresponds to a 5 My interval. Note that all jump dispersal events have occurred within the last 30 Mya.

occupied the Neotropics ($P = 0.63$). Core Ericales, crown Pentaphragmaceae, and crown styracoids all originated in Indo-Malaysia ($P = 0.61, 0.67, \text{ and } 0.55$, respectively).

Starting with the origin of sarracenioids, styracoids, and ericoids, the Nearctic became a much more important geographic region in the historical biogeography of Ericales. Diapensiaceae + Styracaceae originated in the Nearctic ($P = 0.31$). Crown ericoids + sarracenioids and crown sarracenioids originated in a shared Nearctic/Indo-Malaysian range ($P = 0.42 \text{ and } 0.44$, respectively). Backbone nodes of the ericoids as well as crown Ericaceae originated in the Nearctic. While crown Ericaceae most likely originated in the Nearctic ($P = 0.47$), an origin in a shared Nearctic + Palearctic range is also possible.

3.3. Diversification within Ericales

Our BAMM analysis detected between 60 and 105 shifts in speciation rate in Ericales. The 95% credible set of speciation rate shifts included 477 distinct configurations. The configurations with the nine highest posterior probabilities contain between 69 and 81 shifts. The best shift configuration of our BAMM analysis (Supplementary Fig. 2 of Appendix A) suggested that there have been 70 shifts in speciation rate throughout the order (excluding the root shift), only three of which have been decreases. Of these shifts, 56 have occurred since the Oligocene, ~ 33 Mya (Fig. 6). On a phylogeny-wide scale, two peaks in the number of speciation events are evident, one 50–45 Mya and another 15–10 Mya. Examination of individual plots of net diversification, speciation, and extinction rates (Fig. 7) suggests that speciation rate has been increasing exponentially since ~ 60 Mya. Consistent with this, speciation rate has also increased markedly since this time period,

though speciation rate has increased at a more uniform rate through time, with a particularly sharp increase in the last ~ 10 Mya. Unlike speciation rate, extinction rate in the order increased markedly at ~ 65 Mya with an additional increase ~ 20 Mya. Extinction rate also increased sharply at this time, but decreased between ~ 35 and 20 Mya. BAMM analyses support a diversity-dependent speciation process with a mean speciation rate of 0.25 spp./My and a mean extinction rate of 0.09 spp./My. Within individual clades (families or subfamilies), diversification rates vary dramatically. The highest net diversification rates are found in Ericaceae subf. Ericoideae (0.26 spp./My), subf. Vaccinioideae (0.25 spp./My), and Ebenaceae (0.25 spp./My) (Table 2). The lowest net diversification rates are in Ericaceae subf. Cassiopoideae (0.036 spp./My), Diapensiaceae (0.037 spp./My), Tetrameristaceae (0.037 spp./My), and Ericaceae subf. Monotropeoideae (0.037 spp./My) (Table 2). Of the 70 shifts in the best shift configuration, all are confined to within families

Our CoMET analysis (Supplementary Fig. 3 of Appendix A) suggests there is weak evidence (BF between 2 and 6) for tree-wide shifts in speciation and extinction rates in Ericales. The first potential tree-wide shift occurred at ~ 90 Mya (BF of 2.0–2.5 for a decrease in speciation rate and an increase in extinction rate) and a second tree-wide shift at ~ 15 Mya (BF of 3.1–4.4 for an increase in speciation rate and an decrease in extinction rate).

4. Discussion

4.1. Phylogenetic relationships in Ericales

This study provides a new, more fully-resolved and comprehensive

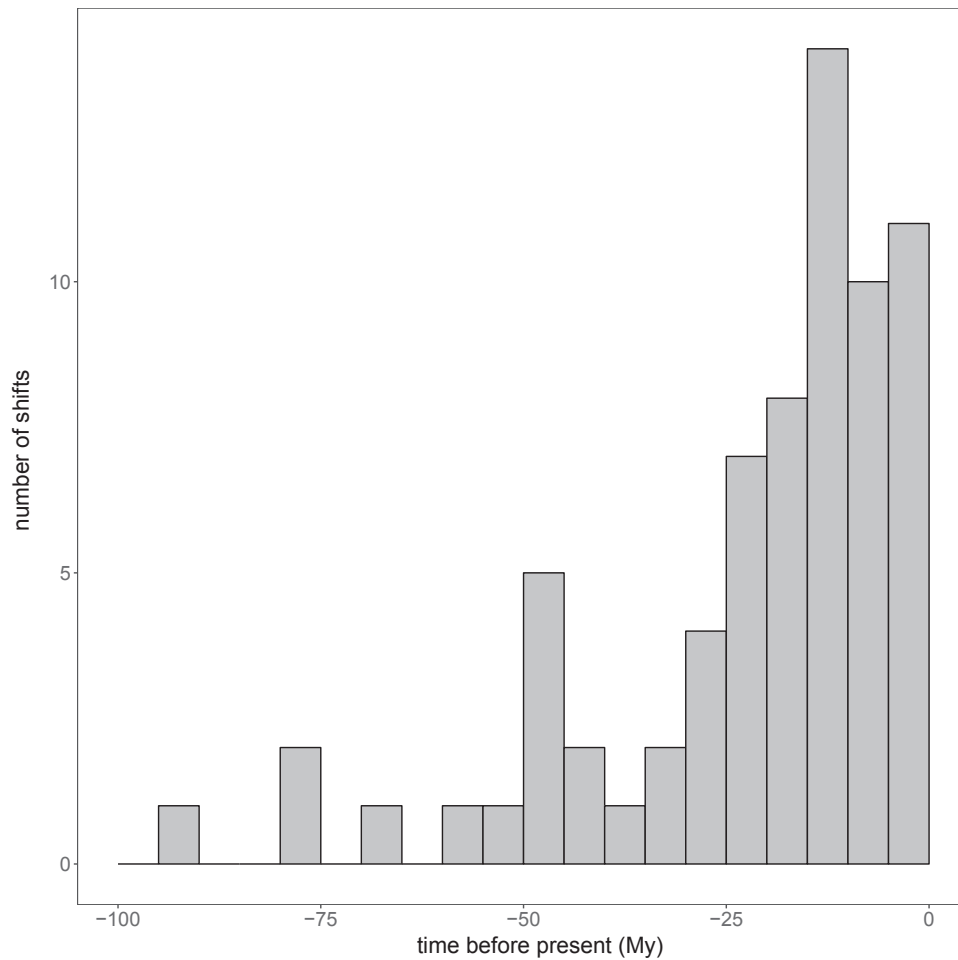


Fig. 6. Histogram of the temporal occurrence of the 70 shifts in speciation rate within Ericales detected in the best BMM configuration. Each bin corresponds to a 5 My interval. Note the sharp increase in number of shifts between 50 and 45 Mya and the second rapid accumulation since the last 30 Mya, especially in the last 15–10 Mya.

phylogeny of the order Ericales, although most relationships along the backbone remain unclear. Specifically, four new hypotheses of relationships are presented here. (1) Mitrastemonaceae is sister to Lecythidaceae. (2) This Mitrastemonaceae + Lecythidaceae clade is

sister to the Ericales excluding balsaminoids. (3) Theaceae is sister to sarracenoids + styracoids + ericoids. This finding is consistent with relationships as suggested by protein-coding genes from whole plastome sequences (Jo et al., 2016; Yan et al., 2016) but is perhaps not

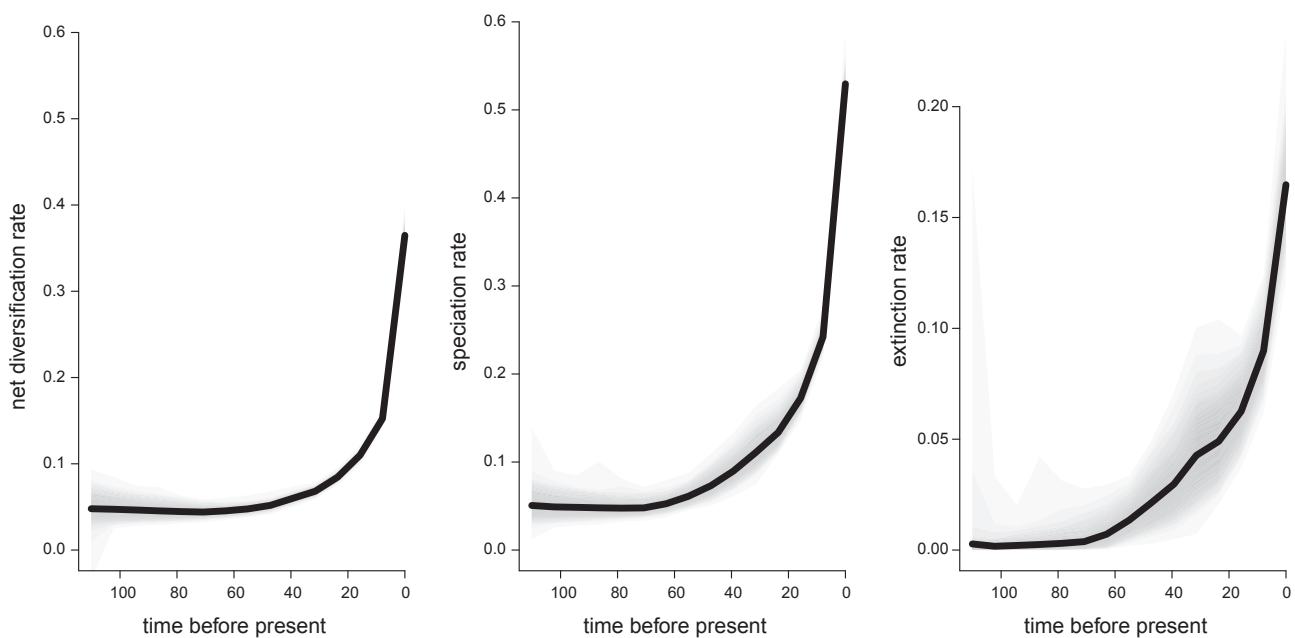


Fig. 7. BMM diversification rates through time in Ericales: net diversification (left), speciation (center), and extinction (right).

Table 2

Summary of crown ages, speciation (λ), extinction (μ), and net diversification (r) rates for major clades of Ericales. The number of shifts in speciation rate in each clade as suggested in the best BAMM shift configuration as well as notable clades in which these shifts occur are also indicated.

Clade	Crown age (Mya)	λ (mean)	μ (mean)	r ($\lambda - \mu$)	Number of descendant shifts	Notable speciation rate shifts
Actinidiaceae	46.5	0.42	0.25	0.17	3	<i>Saurauia</i> , <i>Actinidia</i> + <i>Clematoclethra</i>
Balsaminaceae	69.6	0.27	0.17	0.10	2	<i>Impatiens</i>
Clethraceae	82.8	0.20	0.080	0.12	1	<i>Clethra</i>
Cyrillaceae	19.8	0.041	0.0032	0.038	–	–
Diapensiaceae	59.9	0.038	0.0015	0.037	–	–
Ebenaceae	57.1	0.28	0.036	0.25	5	<i>Euclea</i> + <i>Royena</i> ; New Caledonia <i>Diospyros</i>
Ericaceae: Arbutoideae	28.6	0.23	0.047	0.18	3	<i>Arctostaphylos</i>
Ericaceae: Cassiopoideae	37.2	0.037	0.0014	0.036	–	–
Ericaceae: Enkianthoideae	11.3	0.16	0.10	0.058	–	crown shift
Ericaceae: Ericoideae	67.9	0.49	0.24	0.26	10	<i>Empetrum</i> , <i>Erica</i> , <i>Rhododendron</i>
Ericaceae: Harrimanelloideae	51.4	0.039	0.0019	0.038	–	–
Ericaceae: Monotropeoideae	65.8	0.038	0.0015	0.037	–	–
Ericaceae: Pyroloideae	44.5	0.12	0.036	0.080	1	<i>Pyrola</i>
Ericaceae: Styphelioideae	59.3	0.16	0.030	0.13	4	<i>Dracophyllum</i> ; <i>Leucopogon</i> (in part)
Ericaceae: Vaccinioideae	46.2	0.37	0.12	0.25	4	Vaccinieae; <i>Gaultheria</i> ; <i>Cavendishia</i>
Fouquieriaceae	11.2	0.080	0.027	0.053	–	–
Lecythidaceae	82.9	0.13	0.015	0.11	1	–
Marcgraviaceae	23.9	0.32	0.17	0.15	1	crown shift
Pentaphragaceae	95.1	0.30	0.17	0.13	4	<i>Adinandra</i> ; <i>Eurya</i> ; <i>Ternstroemia</i>
Polemoniaceae	53.9	0.15	0.027	0.12	3	Gillioideae (in part); <i>Phlox</i> ; <i>Polemonium</i>
Primulaceae: Maesoideae	24.1	0.35	0.21	0.14	1	<i>Maesa</i>
Primulaceae: Myrsinoideae	53.2	0.30	0.087	0.21	6	<i>Ardisia</i> , <i>Myrsine</i>
Primulaceae: Primuloideae	46.4	0.22	0.027	0.19	4	<i>Androsace</i> ; <i>Primula</i> (in part)
Primulaceae: Theophrastoideae	70.0	0.12	0.022	0.097	1	<i>Bonellia</i> + <i>Clavija</i> + <i>Jacquinia</i>
Roridulaceae	20.2	0.039	0.0014	0.038	–	–
Sapotaceae	58.3	0.22	0.040	0.18	6	<i>Manilkara</i> ; <i>Palaquium</i> (in part); <i>Pycnanthus</i>
Sarraceniaceae	48.5	0.17	0.11	0.067	1	<i>Heliamphora</i> + <i>Sarracenia</i>
Styracaceae	56.0	0.14	0.033	0.10	1	<i>Styrax</i>
Symplocaceae	47.8	0.45	0.26	0.19	2	<i>Symplocos</i>
Tetrameristaceae	44.6	0.041	0.0043	0.037	–	–
Theaceae	63.1	0.27	0.082	0.19	2	<i>Camellia</i>

surprising given the large contribution of the chloroplast genome to our supermatrix. (4) Within Ericaceae, sub-familial relationships place Arbutoideae + Monotropeoideae as sister to the remainder of the family, excluding Enkianthoideae, with the Pyroloideae in turn sister to the remaining subfamilies.

Taking the order as a whole, some relationships are still uncertain as evidenced by poor bootstrap support values, and warrant further investigation. Nonetheless, many of the clades with low ML bootstrap support values have high posterior probabilities in the BI analysis. Taxa in the BI analysis were carefully selected to include both well-sequenced and taxonomically representative species, so it is unclear to what extent missing data, taxonomic representation, and optimality criterion selection effects support values, though taxon sampling may have a large effect (de la Torre-Bárcena et al., 2009). Nevertheless, as mentioned above, the topologies based on both the full ML and reduced BI supermatrices are nearly identical, and the nodes at which they differ are on short branches within families. While backbone relationships are identical between the ML and BI analyses, support for some of these relationships remains low. While we provide evidence for new backbone relationships here, support is less than optimal. As a result, relationships among major clades in the order continue to remain recalcitrant in the same places as in previous studies. Specifically, these are the placement of Mitrastemonaceae, Lecythidaceae, polemonioids, primuloids, Pentaphragaceae, and theoids relative to each other. Based on our chronogram, these branches are short and are also the result of multiple, rapid cladogenetic events in the Late Cretaceous. Consequently, there may be little phylogenetic information to help resolve them. Geuten et al. (2004) concluded that incongruence between genomic regions was responsible for difficulties in inferring relationships along the backbone of Ericales, and attributed the discordance to either different biological process operating on different genes or erroneous tree construction. Based on the ancient, rapid radiation (cladogenesis within the span of three million years 100 Mya), we suggest

that there might not be enough information preserved in the genomes to resolve this issue with full confidence, especially with Sanger sequence data. However, future genomic-level data from many low copy nuclear genes, as well as more whole-plastome sequences from throughout the order may provide robust support for the proposed relationships along the backbone. It is encouraging that the strongly supported relationships uncovered thus far using whole chloroplast genomes (Jo et al., 2016; Yan et al., 2016) are also found in this study.

In this study, we highlight a better placement of Mitrastemonaceae within Ericales as compared to initial, conflicting placement with different taxonomic and mitochondrial genome sampling (Barkman et al., 2004; Barkman et al., 2007). Our hypothesized placement of Mitrastemonaceae as sister to Lecythidaceae should be more closely examined in future studies to see if this placement makes sense in light of morphological data. Furthermore, comprehensive studies of morphological evolution in Ericales beyond what was examined by Schönenberger et al. (2005) should be undertaken to better understand character evolution in the order, specifically if traits found in either or both of Mitrastemonaceae or Lecythidaceae are symplesiomorphic with the balsaminoids.

Relationships among balsaminoids families have remained problematic. This is clearly a result of another rapid radiation in the Late Cretaceous ~81 Mya. Schönenberger et al. (2005) depicted these relationships as a polytomy, while Janssens et al. (2009) reported Marcgraviaceae as sister to Balsaminaceae with 98% bootstrap support based on an *atpB-rbcL* phylogeny. In this study, our hypothesis of Tetrameristaceae as sister to Balsaminaceae, while not statistically supported, fits well with data from floral anatomy (von Balthazar and Schönenberger, 2013). Our phylogenetic reconstruction of *Impatiens* differs slightly from the relationships consistently recovered within the genus (e.g., Yu et al., 2015). In this study, subgenus *Clavicarpa* is recovered as nested within subgenus *Impatiens* as opposed to being sister to it. The chloroplast marker *atpB-rbcL* is commonly used in

phylogenetic reconstructions in *Impatiens* but is not included in this analysis. The second widely used marker, ITS does not appear to contain much phylogenetic information at deeper levels (cf. Yuan et al., 2004), so it is unclear how much influence the variation in *atpB-rbcL* and gap coding has on phylogenetic reconstruction within the family.

We find surprisingly low support for monophyly of primuloids, especially in light of the strong support for the existence of this clade in Schönberger et al. (2005). We note that many species of Sapotaceae are only represented by barcoding markers *rbcL* and *matK*. We posit that missing data within Sapotaceae may be degrading support for the monophyly of this clade.

Within families, resolution is generally similar to that of previous studies, with expected poor support in the larger clades of Sapotaceae (Swenson and Anderberg, 2005) and Symplocaceae (Fritsch et al., 2008). This problem is further compounded by both the breakup of branches with increased taxonomic sampling in Ebenaceae, Sapotaceae, and Symplocaceae and the rapid radiations of these families, as evidenced by both backbone divergence times within these families as well as our BAMM results.

The phylogenetic relationships within Primulaceae subf. Myrsinoideae have never been studied as a whole (Bone et al., 2012), although tribe Lysimachiaeae has been targeted (Anderberg et al., 2007). Here, we demonstrate that relationships in the large, taxonomically complex, mostly woody Myrsinoideae (tribes Ardisieae and Myrsineae), while resolved, are unsupported. In the future, this clade should be targeted for detailed phylogenetic study, especially with inclusion of Hawaiian *Myrsine*. As noted above, our BI and ML analyses differ in the placement of *Cyclamen* within Primulaceae subf. Myrsinoideae. Previous studies have found little support for backbone relationships of this subfamily (Källersjö et al., 2000; Martins et al., 2003), so this result may be expected. Because of the strong signal for a Palearctic origin for nearly the entire backbone of Primulaceae, in our ARE, the placement of this subfamily does not have much effect on the results of our downstream analyses.

The relationships at the basal nodes of Ericaceae have been difficult to determine. Kron et al. (2002), while recovering subf. Enkianthoideae as sister to the remainder of the family, were not able to conclusively determine the relationships between Arbutoideae, Monotropeoideae, and Pyroloideae. Schwery et al. (2015), based solely on *matK* and *rbcL*, placed all three subfamilies in a clade sister to the remaining Ericaceae excluding Enkianthoideae, with Pyroloideae sister to Arbutoideae + Monotropeoideae. Using a larger set of molecular markers from all genomes, Freudenstein et al. (2016) corroborated this result, though they found support for this relationship only with nuclear ribosomal genome data, while organellar DNA supported a closer relationship of the Pyroloideae to core Ericaceae. However, in all cases, support is suboptimal. Here, we suggest that the relationship of the Pyroloideae as more closely related to core Ericaceae than Arbutoideae + Monotropeoideae is the better hypothesis in both our ML and BI analyses. As noted above, the position of *Bryanthus* within subf. Ericoideae differs between our ML and BI results. In the ML dataset, *Bryanthus* is sister to *Ledothamnus*. However, *Ledothamnus* was excluded from our BI analysis. The placement of these two genera (as a clade) within the Ericoideae is difficult to convincingly resolve with the markers available to us (cf. Gillespie and Kron, 2010).

4.2. Early historical biogeography of Ericales

Inference of the biogeographical origins of Ericales is confounded by the widespread distribution of the order, by the many pantropical or amphipacific families, and by the phylogenetic uncertainty surrounding the early, rapid cladogenesis of the family. Our analysis does not shed much more light on the origin of the most basal nodes of the order, given low probabilities of possible ARs ($P = 0.20\text{--}0.30$, Fig. 4). These low probabilities may be due in part to the assumptions of our dispersal matrix, which assigns a low probability to movement between

these geographically disparate areas. Nevertheless, the greatest AR probability is of an ancestor found in a combined Neotropical/Indo-Malaysian AR. Signal for this combined area or a strictly Neotropical AR is evident in both the basal-most nodes of the balsaminoids and those of remainder of the order (Fig. 4). Alternative AREs are all of very low probability and involve various combinations of geographically widespread ancestors. Furthermore, movement to or from one or both of these ARs is clearly involved in the formation of major clades such as all balsaminoids families (both), Lecythidaceae + Mitrastemonaceae (Neotropics), polemonioids (Neotropics), primuloids (both), Sapotaceae (Indo-Malaysia), Ebenaceae + Primulaceae (Neotropics), and core ericoids (Indo-Malaysia). Furthermore, dispersal from Indo-Malaysia has resulted in the re-colonization of the New World in the sarracenoids and ericoids, perhaps via movement across the Pacific Rim (Fig. 4).

4.2.1. Historical biogeography of balsaminoids

Balsaminoids began diversifying ~83 Mya in a shared Indo-Malaysian/Neotropical AR. Ancestors along the backbone of the clade continued to occupy this region until the crown of Balsaminaceae, with two speciation events in the Neotropics leading to Tetrameristaceae and Marcgraviaceae. Stem Balsaminaceae originated Indo-Malaysia following the split from Tetrameristaceae ~81 Mya. This divergence date is much older than a divergence estimate of 48.1 Mya by Janssens et al. (2009). This latter study also recovered a much younger crown age for balsaminoids as a whole (which effects the age of Balsaminaceae), and may be the result of Janssens et al. (2009) using mostly non-Ericales fossils as calibration points. Within Balsaminaceae, the crown of the family either occupied Indo-Malaysia (highest probability, $P = 0.56$) or the combined area of Indo-Malaysia + Palearctic. The crown of *Impatiens* occupied Indo-Malaysia + Palearctic ~59 Mya ($P = 0.69$). Grey-Wilson (1980) proposed a west Gondwanan origin for Balsaminaceae ~50 Mya, with movement to Southeast Asia and subsequent radiation from that region. With a better understanding of phylogenetic relationships, Yuan et al. (2004), while only using a simplistic parsimony optimization and lacking a timeframe, suggested a clear Southeast Asian origin for the genus with three later arrivals in Africa or Madagascar. Consistent with the study of Yuan et al. (2004), we suggest the genus diversified in Indo-Malaysia and Indo-Malaysia + Palearctic AR with movement from the Palearctic to the Nearctic ~14 Mya. In addition, we suggest three movements of the genus from Indo-Malaysia to the Afrotropics and five subsequent shifts back to Indo-Malaysia, with the most dramatic movement between these regions occurring 31–35 Mya, well after the regions were separated. The conflicting phylogenetic position of section *Clavicarpa*, as noted above, appears to not have an influence on the reconstruction of the historical biogeography of the genus, as it is restricted to Indo-Malaysia (Janssens et al., 2009).

4.2.2. Historical biogeography of Lecythidaceae

Stem and crown Lecythidaceae originated in a shared Afrotropics/Neotropics AR ~105 and ~83 Mya, respectively. Vicariant events gave rise to all subfamilies: Napoleonaeoideae (Afrotropics), Planchonioideae (Afrotropics), Lecythidoideae (Neotropics), and Scyttopetalaloideae (Neotropics). Stem Scyttopetalaloideae subsequently recolonized the Afrotropics to occupy a combined amphipacific distribution ~24 Mya where another vicariant event gave rise to the ancestor of monotypic *Asteranthos brasiliensis* in the Neotropics and the remainder of the subfamily in the Afrotropics. Stem and crown of the largest clade of Lecythidaceae, Neotropical subfamily Lecythidoideae arose ~59 Mya and ~43 Mya, respectively. Within subf. Planchonioideae, especially within *Barringtonia*, there has been extensive interchange between Indo-Malaysia and Australasia in the last ~29 Mya, generally via jump dispersal.

4.2.3. Historical biogeography of polemonioids

A Neotropical AR of stem crown polemonioids is only slightly more

likely than a Nearctic origin. This reconstruction is ambiguous as a result of the close proximity of these two regions and the occurrence of species of *Fouquieria* at the boundary between them. Grant (1959) suggested a central Mexican origin of Polemoniaceae, albeit based on spurious assumptions about “primitive” and “derived” characters in the family, and without the knowledge of a close relationship of the family to Fouquieriaceae. Our results do not shed much light on this problem based on our coarse regional coding, but a strictly Nearctic origin is preferred ($P = 0.57$) over a Nearctic + Neotropical origin. The crown of subfamilies Cobaeoideae + Polemonioidae is Nearctic + Neotropical in origin ($P = 0.53$), though the presence of *Bonplandia* and *Cobaea* in southern Mexico and Central America suggests an origin in Mesoamerica instead of in the Andes (*Cantua*). Our estimate of the age of crown Polemoniaceae at ~54 Mya is much younger than the Late Cretaceous origin of the family estimated by Porter et al. (2010) and differences in age estimates extend to all other lineages in the family. Given the extant distributions of the Cobaeoideae (central Mexico to Bolivia) and Acanthogilioideae (Baja, Mexico), the origin of the family is likely central Mexico.

4.2.4. Historical biogeography of Sapotaceae

The biogeographical history of this large, pantropical family is complex but with clear geographic signal in most clades. The stem and crown of Sapotaceae date to ~102 and ~58 Mya, respectively, both with an AR of Indo-Malaysia. Our divergence time estimates differ markedly from those of Bartish et al. (2011) and Swenson et al. (2014). For example, Swenson et al. (2014) date the divergence of *Planchonella* and *Pycnantha* to 51.2 Mya, (range of 38.5–60.7), whereas we suggest the age of this divergence is much younger, at 21.5 Mya. This appears to be the result of Swenson et al. (2014) using ambiguous fossils (leaf fossils merely assigned to “Sapotaceae” used to constrain inter-familial nodes) and employing geologically-enforced constraints. From Indo-Malaysia, Sapotaceae subsequently moved to Afrotropics + Indo-Malaysia by 30 Mya. Subsequent diversification of the family has largely relied on the Afrotropics as a source realm, with this range as the origin of Clades 2 and 3 of Swenson and Anderberg (2005).

Subfamily Sideroxyloideae (excluding *Sideroxylon discolor* and *S. oxyacanthum*) originated in the Afrotropics or a shared Afrotropics/Indo-Malaysian AR ~27 Mya, with movement to the Palearctic in *Argania* + *Monothea* and African *Sideroxylon* (shared Afrotropical + Palearctic AR). New Caledonian *Nesoluma* is deeply embedded in Neotropical/Nearctic *Sideroxylon* and diverged ~11 Mya, though it is unclear if this distribution is due to vicariance or long distance dispersal.

As with *Impatiens*, tribe Sapoteae shows a complex history of pan-Indian connections of the Afrotropics, Indo-Malaysia and Australasia beginning ~18 Mya and extending to the present. Neotropical *Manilkara* originated from a long distance dispersal event from the Afrotropics. The recently described Malagasy *Bemangidia lowryi* was originally placed in a clade with sympatric *Tsebona macrantha* and *Capurodendron* (Gautier et al., 2013). However, with different taxonomic sampling and no gap coding, we instead suggest that this genus is nested in *Madhuca* and is the result of a recent vicariant event between the Afrotropics and Indo-Malaysia.

Our ARE for subf. Chrysophylloideae is in near complete agreement with the parsimony inference of historical biogeography of Bartish et al. (2011) despite increased taxon sampling and much younger divergence times. In general, we support their conclusion that long-distance dispersal is the main driver of the clade distributions. It is inferred that the majority of the Australasian Chrysophylloideae originated from a long-distance dispersal event from the Afrotropics ~21.5 Mya. Within the Chrysophylloideae, more recent dispersal events from the Afrotropics to the Neotropics and vice versa, as well as from the Afrotropics to Australasia in the last 10 My suggest continued dispersal routes.

4.2.5. Historical biogeography of Ebenaceae

Crown Ebenaceae originated in a combined Neotropics/Africa AR ~57 Mya, A vicariant event formed the stem lineage of *Lissocarpa* in the Neotropics and the stem lineage of *Diospyros s.l.* (including *Euclea* and *Royena*) in Africa. The backbone of *Diospyros* indicates that complex interactions across the Indian Ocean between the Afrotropics, Australasia, and Indo-Malaysia contributed to the diversification of Ebenaceae between ~25 Mya and the present. This exchange has generally been through dispersal.

Our interpretation of the historical biogeography of *Diospyros* contrasts greatly with the study of Duangjai et al. (2009), who stressed importance of the Northern Hemisphere, particularly Eurasia, in the historical biogeography of the family. Their contrasting interpretation may be because they lumped large swathes of Southeast Asia into their concept of Eurasia and also because their biogeographical analysis lacked a time component.

Basinger and Christophel (1985) reported middle-Eocene fossils of *Diospyros* from Australia (calibration point used in this study). Movements across the Indian Ocean during this time period along the stem of clades II-IX of Duangjai et al. (2009) from the Afrotropics to a widespread Indian Ocean ancestor, suggest that Ebenaceae could indeed have reached Australasia during the Eocene. However, extant Australasian lineages of *Diospyros* date to no older than 20 Mya with two major arrivals in the area resulting in Clade II of Duangjai et al. (2009) from a pan-Indian ancestor and a New Caledonian/Pacific Island Clade nested in Clade III of Duangjai et al. (2009) from the Afrotropics. Additional movements to Australasia from Indo-Malaysia have also occurred, but usually in singular instances or from range expansion of species found in both regions.

Diospyros has reached the New World four times, including members of Clades IV and XI of Duangjai et al. (2009) from the Afrotropics and Indo-Malaysia, respectively, and two more times in *D. virginiana* and *D. texana* (both from Indo-Malaysia). However, our topology differs from that of Duangjai et al. (2009) in that their study placed *D. texana* as closely related to Neotropical members of Clade XI with a Bayesian posterior probability of 1.0.

4.2.6. Historical biogeography of Primulaceae

Crown Primulaceae originated from a Neotropical ancestor and was found in a combined Neotropical + Palearctic AR ~80 Mya. The nearly wholly Neotropical subf. Theophrastoideae (Källersjö and Ståhl, 2003) originated in the Neotropics with relatively recent movements outside of that region in the aquatic genus *Samolus*. The split between Theophrastoideae and Myrsinoideae + Primuloideae represents a vicariant event between the Neotropics and Palearctic ~72 Mya with the latter two subfamilies occupying a Palearctic AR. Within Myrsinoideae and especially *Lysimachia*, much of the backbone is of Palearctic origin. The endemic Hawaiian *Lysimachia* (Oh et al., 2013) apparently arrived in the region from a Eurasian ancestor 13.3 Mya, though it is on a long stem, as the crown of the clade is only 1.9 Mya. This clade is not the closest relative of the widespread Indian and Pacific Ocean *L. mauritiana*, as might be inferred by geography, but instead the Iberian *L. ephemereum*. Relationships within this supported clade have been difficult to infer (Anderberg et al., 2007; Hao et al., 2004) but the Hawaiian clade is clearly Palearctic in origin.

Within the woody Myrsinoideae (tribes Ardisieae and Myrsineae), the biogeographical history is unclear, apparently due to short branch lengths, low support, and Neotropical species found throughout this larger clade, especially in Myrsineae. However, a large proportion of Neotropical *Ardisia* appears to be a clade. This clade is sister to a grade of Australasian taxa, *Elingamita jonsonii* from New Zealand and *Ardisia pachyrrhachis* from Australia, and appears to have arisen as a result of a dispersal event from Australasia.

Old World *Ardisia* (with the exception of Philippino *A. glauca* and Javan *A. speciosa* which are nested in the large Neotropical clade) originated from a vicariant event between a combined

Australasian + Indo-Malaysian AR ~24 Mya. The East Indian Ocean genera *Badula* and *Oncostemum* (Bone et al., 2012) likely dispersed from Indo-Malaysia ~9.8 Mya. Overall, a clearer picture of the biogeography of tribes Ardisieae and Myrsineae awaits a better taxonomically sampled and better-supported phylogeny. Divergence time estimates within Primuloideae are in general agreement with those of de Vos et al. (2014), but that study contained wide confidence intervals for the age estimates. At the resolution scale of geographic areas used in this study, the biogeography of Primuloideae is uniform, with a clear origin and persistence in the Palearctic until the last ~16 Mya when the subfamily began to move into the Nearctic.

4.2.7. Historical biogeography of Pentaphragaceae

Crown Pentaphragaceae originated in Indo-Malaysia at ~95 Mya. The historical biogeography of the backbone of the family indicates that it has persisted in Indo-Malaysia throughout its history, with at least three movements to the New World in *Ternstroemia* (from the Palearctic or Indo-Malaysia), *Euryodendron* (from a joint Indo-Malaysian + Nearctic distribution) and *Frezeria* (from Indo-Malaysia).

4.2.8. Historical biogeography of Theaceae

Stem Theaceae originated in Indo-Malaysia ~102 Mya and subsequently moved to an Indo-Malaysian or Indo-Malaysian + Nearctic distribution at the crown of the family ~63 Mya. This crown age estimate is markedly different from the estimate of 49 Mya by Li et al. (2013a) and 42.4 Mya by Zhang et al. (2014). In addition to topological differences relative to Li et al. (2013a) regarding relationships among tribes Theaceae, Gordoniaceae, and Stewartiaceae, there are also problems with chronogram estimation in these studies. The calibrations of Li et al. (2013b) are too strictly enforced, and Zhang et al. (2014) do not include the MRCA of all *Camellia* in their tree despite attempting to place a calibration point at the crown of the genus. In both cases, allowing a less strict age in their calibration would result in an older crown age for the family. Our estimate for the topology of Theaceae and crown age for the family is nearly in agreement with the estimate of 53–60 Mya, and well within the 95% HPD of Yu et al. (2017).

The biogeographical reconstruction of the origin of the family in Li et al. (2013b) is comparable to this study. The probability of a shared Indo-Malaysian/Nearctic distribution is clearly the result of North American *Stewartia* sister to the remainder of the genus. *Stewartia* itself has been a “classic” example of a vicariance pattern between North America and Eastern Asia. However, molecular data has challenged this view (Prince, 2002). Our reconstruction is ambiguous as to the exact biogeographic scenario of *Stewartia*. However, two vicariant events, one giving rise to *S. rostrata* + *S. malacodendron* and another giving rise to *S. ovata*, are much more likely than dispersal, though support for relationships within this clade are low.

The backbone of Theaceae is clearly Indo-Malaysian in origin. Our ARE suggests the distribution of *Laplacea* in the Neotropics is the result of long distance dispersal from Indo-Malaysia ~20.5 Mya. The origin of tribe Theaceae in Indo-Malaysia appears inconsistent with Eocene fruits of *Gordoniopsis polysperma* from North America, broadly assigned to tribe Theaceae in this study (Grote and Dilcher, 1992). However, given the multiple arrivals of Theaceae into the New World, it seems possible that Theaceae may have arrived in North America at yet another time and that *Gordoniopsis polysperma* represents an extinct lineage.

4.2.9. Historical biogeography of styracoids

The styracoids began to diversify ~99 Mya in Indo-Malaysia. This date is consistent with the estimate of 80–94 Mya in Fritsch et al. (2015). The AR of stem and crown Diapensiaceae + Styracaceae may have originated in the Nearctic (highest probability), Indo-Malaysian, or a combined area. However, both stem and crown Diapensiaceae clearly occupied the Nearctic, with the crown of the family dating to ~60 Mya, slightly younger than the ~65 Mya estimate of Hou et al. (2016), but well within the 95% HPD (though their tree topology is

problematic). The distribution of the eastern North American endemic *Shortia galacifolia* appears to be the result of a vicariant event from a Nearctic + Palearctic ancestor ~12 Mya.

While the ARE for stem Styracaceae is clearly Indo-Malaysian, the AR of the crown of *Styrax* is problematic. Fritsch et al. (2001), with limited taxonomic sampling and no temporal context, suggested a Eurasian origin for the family. It is unclear what concept of Eurasia Fritsch et al. (2001) used, so our Indo-Malaysian origin may or may not be consistent with this result. Inference of the AR of *Styrax* itself seems to be confounded by low support for infra-generic relationships (Fritsch, 2001; Fritsch et al., 2001) and by geographic heterogeneity between the Old and New World within these two clades. This geographic heterogeneity results in our ARE invoking vicariant events from widespread ancestors to explain the present-day distribution of *Styrax*.

Stem and crown Symplocaceae both occupied Indo-Malaysia. Our inferred crown age of 48 Mya is consistent with the 48–57 Mya divergence time estimate of Fritsch et al. (2015). Crown *Symplocos* originated in Indo-Malaysia ~39 Mya, with an age estimate again almost exactly identical to the estimate of Fritsch et al. (2015). As in Fritsch et al. (2015), our estimate of the AR for the origin of *Symplocos* is in the Old World, although we parse this area out more finely. Fritsch et al. (2015) inferred movement from Eurasia to North America (including Central America) along stem *Symplocos*. However, we find no evidence for a shift to the New World along the stem. As in Fritsch et al. (2015), *Symplocos* (excluding sect. *Palura*) can be broadly delimited into “New World” (sect. *Hopea* and sect. *Ciponima*) and “Old World” (sect. *Lodhra*) clades. However, our interpretation of the historical biogeography of this distribution differs markedly. Fritsch et al. (2015) argued for a boreotropical origin of the Neotropical species. However, their analysis did not include the Australasian *S. paucistamenia*, which appears to be nested in a largely New World clade of sect. *Hopea* and sect. *Ciponima*. Consequently, our DECj model suggests an Australasian origin of the stem of this clade, with subsequent anagenetic movement into a shared Australasian + Neotropical range 27–18 Mya, perhaps via Antarctica

4.2.10. Historical biogeography of sarracenioids

The Nearctic appears to be an important area in the early diversification of sarracenioids, as both the stem and crown of the group have the highest probabilities involving this area and Indo-Malaysia. Our estimate of the crown age of sarracenioids at ~93 Mya is much older than the divergence estimate of 48.62 Mya by Ellison et al. (2012). Sarracenioids most likely evolved in a shared Nearctic + Indo-Malaysian AR, greatly contrasting with a South American origin proposed by Ellison et al. (2012). Vicariance between Indo-Malaysia and the Nearctic is responsible for the split between new world Sarraceniaceae and Actinidiaceae + Roridulaceae. Crown Sarraceniaceae dates to ~48 Mya, again much older than the estimate of 34.91 Mya of Ellison et al. (2012). Nevertheless, our estimate for the divergence of *Sarracenia* and *Heliophora* is reasonably similar at ~31 Mya as opposed to 22.76 Mya (range 14–32) in Ellison et al. (2012). Ellison et al. (2012) concluded that vicariance is responsible for the extant distribution of Sarraceniaceae. Likewise, we find support for the hypothesis that vicariance is responsible for the Nearctic/Neotropical disjunction between these genera.

Stem Actinidiaceae + Roridulaceae originated in Indo-Malaysia and subsequently dispersed to the Afrotropics, with a vicariance event ~80 Mya leading to the isolation of Roridulaceae in South Africa. Again, this is consistent with Ellison et al. (2012) although they interpreted it as a South American/South African disjunction as opposed to an initial Afrotropics/Indo-Malaysian disjunction as proposed in this study.

Crown Roridulaceae is only ~20 Mya, much younger than the recently described 35–47 Mya roridulaceous fossil from Russia (Sadovskii et al., 2015). While our biogeographical inference based on distribution of the extant Roridulaceae provides no evidence of a Palearctic distribution along the stem of the family, it is certainly possible that the

family may have extended into more northern areas during the Eocene, especially given the northern hemisphere distribution of portions of Actiniaceae and strong connections between these two landmasses in other ericacean groups such as *Erica* and Ebenaceae.

Actiniaceae arose in Indo-Malaysia ~48 Mya, with *Actinidia* + *Clematoclethra* remaining in the ancestral range with several slight northward extensions into the Palearctic within *Actinidia*. Stem *Saurauia* originated in Indo-Malaysia but had dispersed to a pan-south Pacific distribution by the crown of the genus ~15 Mya. The present day amph-Pacific distribution of the genus appears to be a result of vicariance with an Indo-Malaysian + Australasian and a Neotropical clade, so movement via an Antarctic land bridge is possible.

4.2.11. Historical biogeography of ericoids

Crown ericoids and stems and crowns of all three families in this clade (Clethraceae, Cyrillaceae, Ericaceae) clearly originated in the Nearctic. The emergence of the stems of all families occurred within a short time between 98 and 90 Mya. This is consistent with the large number of Late Cretaceous fossils in North America assigned to the crown or stem groups of this clade (Crepet et al., 2013; Martínez et al., 2016; Nixon and Crepet, 1993; Schönenberger et al., 2012). The biogeographical history of *Clethra*, especially in the Old World, appears complex. Lack of phylogenetic support due to incomplete molecular data likely contributes to this uncertainty (Fior et al., 2003). Specifically, the placement of *C. ferruginea* as sister to the remainder of the genus instead of within larger clade of Neotropical section *Cuellaria* complicates the ARE of the genus. Nevertheless, it appears that the present day distribution of *Clethra* is due to vicariance, with uncertainty as to how widespread the ancestor of the genus was. The entire backbone of section *Clethra* apparently was represented by a widespread ancestor and all major clades in this section are the result of vicariant events, consistent with the widespread fossil record of the genus in the Miocene of Europe (Fior et al., 2003). The Madeiran endemic *C. arborea* and the eastern North America endemics *C. acuminata* and *C. alnifolia* all appear to have originated through independent vicariant events from widespread northern hemisphere ancestors within the last 10 My. Crown Cyrillaceae retains the Nearctic AR of the ericoids, with an age of ~20 Mya. The range of *Cyrilla racemiflora* has subsequently expanded southward into the northern Neotropics.

Ericaceae represents the largest extant member of the ericoids, and indeed the largest family in the entire order. A more detailed survey of the biogeography of the family is underway (Rose et al., in prep.), so only the biogeographic history of the major clades will be discussed here. Kron and Luteyn (2005) suggested a Northern Hemisphere origin for the family, likely a joint area of Asia and North America. Particularly problematic for reconstruction of the AR of Ericaceae has been a combination of uncertainty about the relationships at the basal nodes of the family (discussed above) and the relationship of the East Asian subfamily Enkianthoideae as sister to the remainder of the family. Unfortunately, clarification of the relationships within Ericaceae in this study, combined with thorough sampling of the closest relatives of the family, has not helped resolve this problem. Stem Ericaceae clearly originated in the Nearctic, but the crown of the family may have originated in the Nearctic or a joint Nearctic + Palearctic area, with a slightly higher probability to the former scenario. The backbone of Ericaceae shows a clear Nearctic ancestral range, as in Kron and Luteyn (2005).

Hileman et al. (2001) suggested that the New World/Mediterranean distribution of Arbutoideae is the result of a vicariant event at the Paleogene/Neogene boundary (~23 Mya) and supports the Madrean-Tethyan hypothesis. Here, we corroborate that divergence of Old World *Arbutus* is the result of a vicariant event between the Nearctic and Palearctic, but suggest a more complex scenario. The ancestor of the entire subfamily occupied this wide range, with New World *Arbutus* originating from an additional vicariant event. The vicariant event giving rise to Old World *Arbutus* occurred around the Paleogene/Neogene boundary at ~26 Mya.

Stem Vaccinioideae + Styphelioideae originated in the Nearctic and dispersed to Australasia to occupy a shared distribution at ~65 Mya; the divergence of these two subfamilies is due to vicariance. The Australasian + Patagonian distribution of Styphelioideae, with the single New World species *Lebetanthus myrsinoides*, is the result of a vicariant event ~23 Mya, and perhaps invoking an initial Antarctic land bridge. Within the Ericoideae, *Erica* arrived in South Africa ~14–11 Mya from an Afrotropical + Palearctic ancestor, much as *E. arborea*, the species sister to South African *Erica*, is distributed today. The phylogenetic position of Northeast Asian *Bryanthus* as sister to Venezuelan *Ledothamnus* represents an interesting biogeographical scenario. This divergence is old (~62 Mya) and appears to represent a vicariant event from anagenetic dispersal from the Palearctic to Nearctic (Gillespie and Kron, 2010, 2012). The ARE of the crown of *Rhododendron*, excluding section *Therorhodium* occupied a shared Nearctic + Palearctic range and represents anagenetic dispersal from a Palearctic ancestor.

4.3. Cladogenetic drivers of disjunct tropical distributions

This study within Ericales has demonstrated complex patterns of distributions between areas. Importantly, within a given landmass, species of a given clade are often not each other's closest relatives and are usually more closely related to geographically disparate species. This complicates interpretation of disjunct patterns, as multiple processes may have led to the current within-region assemblage of species and movements to a given area have not occurred at a single time. However, in cases where the divergence time between species showing this disjunct pattern is > 30 Mya, vicariance is always the preferred explanation for their disjunct distribution (Fig. 5; Fig. 8). Tropical disjunctions include pantropical, amph-Atlantic (Afrotropics-Neotropics), amph-Pacific (Neotropics and Australasia and/or Indo-Malaysia), trans-Indian (Afrotropics and Australasia and/or Indo-Malaysia), and Wallacean (Australasia and Indo-Malaysia). Examination of the temporal patterns of tropical distributions highlights that drivers of these disjunct tropical distributions are heterogeneous within Ericales (Fig. 8). Nevertheless, several trends emerge: (1) vicariance has been a driver of tropical disjunctions in deeper time but vicariant events have continued to be important in shaping extant diversity well into more recent times. (2) Pantropical disjunctions are usually the result of founder events (jump dispersal). (3) Wallacean disjunctions are always the result of vicariance, not jump dispersal.

4.4. Shifts in speciation rate within Ericales

Shifts detected in this study largely corroborate smaller-scale studies on diversification in the order, but also refine the location of these shifts and put them in an ordinal-level context. Previously, de Vos et al. (2014) detected three shifts in speciation rate in Primulaceae subfamily Primuloideae: one in European/American *Androsace* + *Douglasia* (also detected in Roquet et al., 2013) and two in *Primula s.l.*: one at the crown of the genus and one in *Soldanella*. We also detect the shift in European/American *Androsace*, but suggest the shift in diversification within *Primula* is not as spectacular when the diversity in subfamily Myrsinoideae is more accurately accounted for by increasing taxonomic sampling. While we do detect shifts in *Primula* (see below), we place a shift in speciation rate close to the crown of Myrsinoideae + Primuloideae. Within Ericaceae, Schwery et al. (2015) detected six increases in speciation rate, though they noted uncertainty in placing some of them based on limited taxonomic sampling. As with Primulaceae, we again detect these shifts but more accurately place them. Moreover, we detect additional shifts within clades. Specifically, we note that (1) within *Erica*, while a shift does occur along the stem of the genus, the most dramatic shift occurred after movement into Africa. Several subsequent shifts occurred within South Africa (as in Pirie et al., 2016) and (2) three additional shifts are noted in subfamily Styphelioideae other than

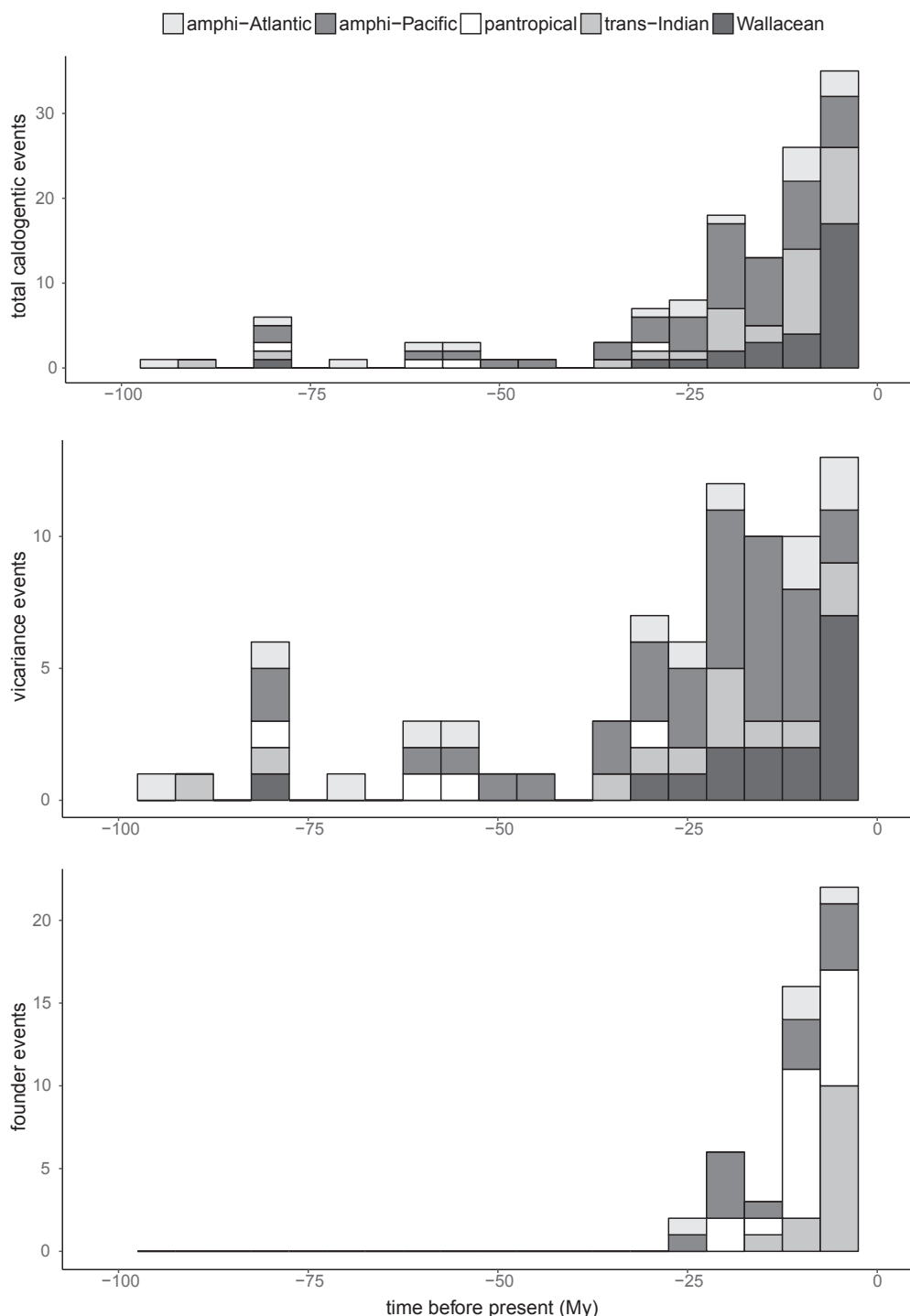


Fig. 8. Temporal distribution of tropical disjunctions driven by cladogenic events based on 100 stochastic maps in BioGeoBEARS. Panels correspond to all cladogenic events (top), vicariance only (middle) and founder events (jump dispersals) only (bottom).

within tribe Richeeae.

Our diversification rate through time plots (Fig. 7), in which an increase in speciation rate at the level of the entire order 70–60 Mya is evident, further highlights the general trend of a post-Cretaceous increase in speciation rates within angiosperms (Antonelli et al., 2015). One of the most important takeaways from our diversification analysis is that most of the shifts in speciation rate responsible for the extant diversity in Ericales are relatively recent. Of the 70 detected shifts in speciation rate in the best shift configuration, only 18 are older than the Eocene/Oligocene boundary (~34 Mya). In particular, we emphasize the sharp spike in branch-specific speciation rates in the Early-Middle Miocene as detected in our BAMM analysis (Fig. 6) and also on a tree-

wide scale in our CoMET analysis, although evidence for this is weak in our CoMET analysis. One limitation of CoMET analyses is that only a tree-wide sampling probability for extant tips can be provided. The analysis therefore assumes random sampling of extant tips, an assumption that is undoubtedly violated in our phylogeny. BAMM, on the other hand, does not assume this and allows the user to specify a clade-specific sampling fraction. Therefore, our CoMET analysis may be understating the strength of evidence for a tree-wide shift during the Early-Middle Miocene.

Ericales appears to have a relatively low rate of species formation during its early history, despite short branch lengths. This is reflected in the early branching pattern of many extant families (e.g. Ebenaceae,

Lecythidaceae, Sapotaceae, Symplocaceae) where major clades are on long branches. Furthermore, this basal divergence is represented by species-poor clades sister to species-rich clades. In addition, branch lengths along the species-rich clade are also oftentimes short. Therefore, despite ancient, short branches along the backbone of the order, our BAMM analysis suggests that the most rapid speciation rates have occurred later in the history of the order, after ~50 Mya, at the time of formation of most higher-level interfamilial clades.

Within the balsaminoids, Marcgraviaceae and Balsaminaceae both show increases in speciation rate. Marcgraviaceae is generally hummingbird and bat pollinated (Dressler and Tschapka, 2002), which may facilitate high rates of speciation via reproductive isolation as a result of specialized reproductive strategies. *Impatiens* is one of the larger genera of flowering plants with ca. 1,000 species and greatly exceeds the species diversity of its closest relatives Tetrameristaceae (3 spp.) and *Hydrocera* (1 sp.). It is therefore not surprising that we detect a shift in speciation rate along stem *Impatiens*. In addition, we detect an additional shift in sect. *Uniflorae*. Janssens et al. (2009) similarly detected increased rates of net diversification within the genus. Janssens et al. (2009) ascribed this change to climatic changes during the Pliocene. As mentioned above, their divergence dates within Balsaminaceae are much younger than ours and both of our detected shifts in speciation rate are pre-Pliocene. This suggests that speciation as a result of climatic changes during the Pliocene is not the best hypothesis for explaining extant diversity in the genus.

Shifts along the stem of Lecythidaceae excluding subf. Napoleoneaeoideae, stem Polemoniaceae, stem Sapotaceae excluding *Sarcosperma* and *Eberhartia*, and two separate but temporally similar increases in *Diospyros*, seem to suggest a repeating pattern of small, species-poor clades on long branches sister to larger clades. It is unclear if any morphological aspects may have triggered these changes in speciation rate.

Within Sapotaceae, we detected five shifts in speciation rate. A shift in the New Caledonian *Pycnandra* clade appears to be associated in part with geographic endemism. However, the diversifications of the *Madhuca* and *Payena* clades in Indo-Malaysia and the *Manilkara* clade in the Afrotropics appear unrelated to geography.

Within *Diospyros* (Ebenaceae), geography appears to have played a large role in diversification. Of the five shifts in speciation rate in the genus, two appear correlated with diversification in insular systems: one in a clade of species largely endemic to Mauritius, and two in clades of New Caledonian taxa, including the clade recently studied by Turner et al. (2013) and Paun et al. (2016). Paun et al. (2016) suggested this clade adaptively radiated as a result of edaphic factors.

Within Primulaceae, the increases in speciation rate in *Maesa* may be due to taxonomic confusion more than anything else, with most authors giving very approximate estimates of species diversity (cf. Utteridge and Saunders, 2004). Furthermore, with genetic sampling of only ten (out of 150–200) species, it is unclear if the detected shift is appropriately placed within the genus. *Maesa* has at least partially diverged in sexual system with both unisexual and hermaphroditic species (Utteridge and Saunders, 2001). While it is also currently unclear where unisexuality has evolved within the genus, it is possible with more taxonomic sampling that the shift in speciation rate in this genus will coincide with the evolution of unisexuality. Within the remainder of Primulaceae, we detected a shift at the crown of Myrsinoideae + Primuloideae. Within Myrsinoideae the species-poor clade formed by monotypic *Coris* and *Stimpsonia* shows a decrease in speciation rate. The highest rates of speciation within Primulaceae occur within Myrsinoideae tribes Ardisieae and Myrsineae (Källersjö et al., 2000). Specifically, these shifts occur in *Myrsine* s.l. and in Old World *Ardisia*. However, as with *Maesa*, there is a great deal of taxonomic uncertainty as to the number of species in these clades, so the location of this shift might change with a better sense of taxonomic limits in these tribes.

As mentioned above, we fail to detect the shift in speciation rate

along the stem lineage of *Primula*. In an earlier examination of diversification rates within *Primula*, de Vos et al. (2014) detected a shift in net diversification near crown *Primula* and suggested that this shift was associated with the evolution of heterostyly. We do detect shifts in speciation rate within *Primula*, but in two deeply nested clades. These correspond to Clade V and New World members of Clade II in Mast et al. (2001). A significant number of species in these clades are homostylous (Mast et al., 2001), so it may be that loss of heterostyly in these clades is driving diversification. Another possibility is that ecological opportunities afforded by expansion into the Nearctic has facilitated speciation in some lineages of this clade.

Of five increases in speciation rate within Pentaphylacaceae, the oldest increase occurs along the stem of *Ternstroemia*. This is not associated with any change in ancestral range, but Luna and Ochoterena (2004) suggested that this clade is unique in the family in having punctate leaves and seeds with a sarcotesta. It is possible that the evolution of a sarcotesta facilitated within-realm seed dispersal and seedling establishment. One of the other major diversification events within Pentaphylacaceae, although recent, occurs in the *Eurya* + *Freziera* clade. Luna and Ochoterena (2004) did not recover these genera as monophyletic based on morphological data, but examination of apomorphic character states for each genus suggests that both are characterized by the evolution of unisexual flowers, which may have driven increased speciation in this group.

One of the rare cases of a detected shift along the stem lineage of an entire ericalean family occurs in Theaceae. However, this shift may or may not involve a shift in geographic range, and comparative morphological studies are needed to determine if any morphological characters may have spurred diversification in this group. Within Theaceae, we detected a shift within tribe Theaeae along the stem *Camellia* + *Pyrenaria* clade, which contains most of the species diversity in the tribe and family. Yu et al. (2017) also detected a shift along the stem to this lineage. Prince and Parks (2001) discussed several potential reasons for this increase in species richness. One possible synapomorphy for these two genera is the presence of semi-fleshy capsules. The evolution of this trait may have led to increased dispersal ability via biotic vectors, as with *Ternstroemia*, and is a possible reason for the increased species richness in this group.

One shift in speciation rate occurs within Styracaceae along stem *Styrax*. Fritsch (1999) discovered eight potential synapomorphies for *Styrax*. Most involved gynoeceal traits, but they also included the dispersal-related traits such as the presence of fleshy mesocarp in the fruit and a thickened seed coat. These traits, in addition to the drupaceous fruit, suggest adaptations to biotic dispersal may have spurred increased rates of speciation in the genus.

We detect two increases in speciation rate in Symplocaceae at the crowns of both large clades of *Symplocos*: sect. *Hopea* + sect. *Ciponima* and sect. *Lodhra*. Again, this is a case of a small clades (2–3 species in *Cordyloblaste* + *Symplocos sonoharuae*) sister to a large radiation of ~300 species. Neither of these shifts appears associated with geographical range shifting. Given the morphological variability present in *Symplocos* in terms of androecium arrangement, sexual condition, and pollinator reward (Fritsch et al., 2008), more detailed studies are needed to determine what morphological features, if any, are responsible for these two increases in speciation rate. Within sarracenioids, we detect four increases in speciation rate. The oldest increase is in Sarraceniaceae at the crown of *Heliamphora* + *Sarracenia*. As noted above, this corresponds to a range expansion into the Neotropics. Within Actinidiaceae the oldest shifts occur separately along the stem lineages of *Actinidia* + *Clematoclethra* and *Saurauia*. None of the mapped characters in Löfstrand and Schönenberger (2015b) are synapomorphies for either clade, so these characters cannot be implicated as reasons for the increase in speciation rate in these clades. One important factor to consider is the issue of species delimitation in these genera, especially *Saurauia*, of which the vast majority of species in the genus (ca. 170/250) occur in Southeast Asia. These Old World species of *Saurauia* have

no comprehensive monographic treatment (Hunter, 1966; Soejarto, 1969). *Actinidia* and *Saurauia* both exhibit complex sexual systems (Soejarto, 1969; Chat et al., 2004) and also have fleshy fruits, though it is unclear if dioecy is homologous or homoplastic in the family. In addition, these genera are most abundant in high-elevation areas (Li et al., 2007; Soejarto, 1969) where geographically induced reproductive isolation may occur. All or some of these traits may be responsible for the increased rates of speciation that we detect. A plurality of detected shifts in speciation rate occurs within Ericaceae, which contains about one-third of the entire species diversity in the order. A total of 22 shifts in speciation occur in this family including in the hyper-diverse genera *Erica* (six shifts) and *Rhododendron* (two shifts). One of the more ancient radiations exhibiting a shift in speciation rate occurs in Vaccinioideae excluding *Oxydendrum*. However, it is unclear what, if any morphological traits may have spurred diversification on these branches, as few traits distinguish Vaccinioideae (Kron et al., 2002). The much more numerous recent radiations in Ericaceae may coincide with diversification of floral traits (Kriebel et al., in prep.). Overall, our BAMM, and to some extent CoMET results highlight the importance of post-Cretaceous, and especially Early-Middle Miocene speciation events on shaping extant diversity in the order. In addition, our results lay a framework for exploring how morphological or geographic features contributed to the diversification of lineages, and here provide some hypotheses to examine, particularly relating to the evolution of atypical sexual systems or means of dispersal via biotic factors.

5. Conclusions

Here, we provide a new, comprehensive time-calibrated phylogeny of the order Ericales based on a summary of the extant molecular data for the order. Our results corroborate previous findings and resolve previously ambiguous relationships, but they also underscore the difficulty of confidently inferring the phylogeny of the order along its ancient backbone. To what extent taxonomically representative sequencing of both plastomes and multiple nuclear loci will help elucidate these currently obscured relationships remains an open question. However, this phylogenetic and biogeographic study provides a framework for formulating and testing hypotheses regarding morphological evolution in this phenotypically diverse order. This framework will be useful for investigating hypotheses at both the whole-order and familial level and, importantly, this study provides the phylogenetic context needed to more systematically investigate driving factors behind observed increases in speciation rate within this clade.

Funding

This work was supported in part through funding from a NSF doctoral dissertation improvement grant (award 1501867) to KJS and JPR and the ON and EK Allen Fellowship awarded to JPR through the University of Wisconsin-Madison Department of Botany.

Acknowledgments

This study was conducted as part of the doctoral dissertation of JPR while at the University of Wisconsin-Madison. We thank previous Ericales researchers for providing invaluable sequence data for this broad synthesis, the associate editor and two anonymous reviewers for helpful comments to greatly improve the manuscript, and the University of Wisconsin-Madison Department of Botany for financial support.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.ymppev.2018.01.014>.

References

- Albach, D.C., Soltis, P.S., Soltis, D.E., Olmstead, R.G., 2001. Phylogenetic analysis of asterids based on sequences of four genes. *Ann. Missouri Bot. Gard.* 88, 163–212.
- Anderberg, A.A., Rydin, C., Källersjö, M., 2002. Phylogenetic relationships in the order Ericales s.l.: analyses of molecular data from five genes from the plastid and mitochondrial genomes. *Am. J. Bot.* 89, 677–687.
- Anderberg, A.A., Manns, U., Källersjö, M., 2007. Phylogeny and floral evolution of the Lysimachieae (Ericales: Myrsinaceae): evidence from ndhF sequence data. *Willdenowia*. 37, 407–421.
- Antonelli, A., Zizka, A., Silvestro, D., Scharn, R., Cascales-Miñana, B., Bacon, C.D., 2015. An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Front. Genet.* 6, 130.
- APG, I.V., 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linnean Soc.* 181, 1–20.
- Barker, W.R., 1980. Taxonomic revisions in Theaceae in Papuaia. 1. *Gordonia*, *Terstroemia*, *Adinandra* and *Archboldiodendron*. *Brunonia* 3, 1–60.
- Barkman, T.J., Lim, S., Salleh, K.M., Nais, J., 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. *Proc. Natl. Acad. Sci. U.S.A.* 101, 787–792.
- Barkman, T.J., McNeal, J.R., Lim, S., Coat, G., Croom, H.B., Young, N.D., dePamphilis, C.W., 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evol. Biol.* 7, 248.
- Bartish, I.V., Antonelli, A., Richardson, J.E., Swenson, U., 2011. Vicariance or long-distance dispersal: historical biogeography of the pantropical subfamily Chrysophylloideae (Sapotaceae). *J. Biogeogr.* 38, 177–190.
- Basinger, J.F., Christophel, D.C., 1985. Fossil flowers and leaves of Ebenaceae from the Eocene of Southern Australia. *Can. J. Bot.* 63, 1825–1843.
- Berger, B.A., Kriebel, R., Spalink, D., Sytsma, K.J., 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Mol. Phylogenet. Evol.* 95, 116–136.
- Bone, R.E., Strijk, J.S., Fritsch, P.W., Buerki, S., Strasberg, D., Thebaud, C., Hodkinson, T.R., 2012. Phylogenetic inference of *Badula* (Primulaceae), a rare and threatened genus endemic to the Mascarene Archipelago. *Bot. J. Linnean Soc.* 169, 284–296.
- Bouchenak-Khelladi, Y.N., Onstein, R.E., Xing, Y., Schwery, O., Linder, H.P., 2015. On the complexity of triggering evolutionary radiations. *New Phytol.* 207, 313–326.
- Boucher, F.C., Zimmermann, N.E., Conti, E., 2016. Allopatric speciation with little niche divergence is common among Alpine Primulaceae. *J. Biogeogr.* 43, 591–602.
- Bremer, B., Bremer, K., Heidari, N., Erixon, P., Olmstead, R.G., Anderberg, A.A., Källersjö, M., Barkhordarian, E., 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Mol. Phylogenet. Evol.* 24, 274–301.
- Cardinal-McTeague, W.M., Sytsma, K.J., Hall, J.C., 2016. Biogeography and diversification of Brassicales: a 103million year tale. *Mol. Phylogenet. Evol.* 99, 204–224.
- Chartier, M., Löfstrand, S., von Balthazar, M., Gerber, S., Jabbour, F., Sauquet, H., Schönenberger, J., 2017. How (much) do flowers vary? Unbalanced disparity among flower functional modules and a mosaic pattern of morphospace occupation in the order Ericales. *Proc. R. Soc. Lond. (Biol.)* 284, 20170066.
- Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qui, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrén, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Fournier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Leary, G.H., Graham, S.W., Barrett, S.C.H., Dayanandan, S., Albert, V., 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcl*. *Ann. Missouri Bot. Gard.* 80, 528–580.
- Chat, J., Jauregui, B., Petit, R.J., Nadot, S., 2004. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. *Am. J. Bot.* 91, 736–747.
- Collinson, M.E., Crane, P.R., 1978. *Rhododendron* seed from the Palaeocene of southern England. *Bot. J. Linnean Soc.* 76, 195–205.
- Crepet, W.L., Nixon, K.C., Daghljan, C.P., 2013. Fossil Ericales from the Upper Cretaceous of New Jersey. *Int. J. Plant Sci.* 174, 572–584.
- Cronquist, A., 1981. An Integrated System of Classification of Flowering Plants. Columbia University Press, New York.
- Davis, C.C., Webb, C.O., Wurdack, K.J., Jaramillo, C.A., Donoghue, M.J., 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of tropical rain forests. *Am. Nat.* 165, E36–E65.
- de Queiroz, A., Gates, J., 2007. The supermatrix approach to systematics. *Trends Ecol. Evol.* 22, 34–41.
- Donnell-Smith, J., 1897. Undescribed plants from Guatemala and other Central American republics XVIII. *Bot. Gaz.* 23, 235–251.
- Dressler, S., Tschapka, M., 2002. Bird versus bat pollination in the genus *Marcgravia* and the description of a new species. *Curtis's Botanical Mag.* 19, 104–109.
- Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–2197.
- Duangjai, S., Samuel, R., Munzinger, J., Forest, F., Wallnöfer, B., Barfuss, M.A., Fischer, G., Chase, M.W., 2009. A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. *Mol. Phylogenet. Evol.* 52, 602–620.
- Ellison, A.M., Butler, E.D., Hicks, E.J., Naczi, R.F.C., Calie, P.J., Bell, C.D., Davis, C.C., 2012. Phylogeny and biogeography of the carnivorous plant family Sarracenaceae. *PLoS ONE*. 7, e39291.
- Fior, S., Karis, P.O., Anderberg, A.A., 2003. Phylogeny, Taxonomy, and Systematic

- Position of *Clethra* (Clethraceae, Ericales) with Notes on Biogeography: Evidence from Plastid and Nuclear DNA Sequences. *Int. J. Plant Sci.* 164, 997–1006.
- Freudenstein, J.V., Broe, M.B., Feldenkris, E.R., 2016. Phylogenetic relationships at the base of Ericaceae: implications for vegetative and mycorrhizal evolution. *Taxon*. 65, 794–804.
- Fritsch, P.W., 1999. Phylogeny of *Styrax* based on morphological characters, with implications for biogeography and infragenetic classification. *Syst. Bot.* 24, 356–378.
- Fritsch, P.W., 2001. Phylogeny and Biogeography of the flowering Plant Genus *Styrax* (Styracaceae) based on chloroplast DNA restriction sites and DNA sequences of the internal transcribed spacer region. *Mol. Phylogenet. Evol.* 19, 387–408.
- Fritsch, P.W., Morton, C.M., Chen, T., Meldrum, C., 2001. Phylogeny and biogeography of the styracaceae. *Int. J. Plant Sci.* 162, S95–S116.
- Fritsch, P.W., Kelly, L.M., Wang, Y., Almeda, F., Kriebel, R., 2008. Revised infrafamilial classification of Symplacaceae based on phylogenetic data from DNA sequences and morphology. *Taxon* 57, 823–852.
- Fritsch, P.W., Manchester, S.R., Stone, R.D., Cruz, B.C., Almeda, F., 2015. Northern Hemisphere origins of the amphipacific tropical plant family Symplacaceae. *J. Biogeogr.* 42, 891–901.
- Gautier, L., Naciri, Y., Anderberg, A.A., Swenson, U., 2013. A new species, genus and tribe of Sapotaceae, endemic to Madagascar. *Taxon* 62, 972–983.
- Geuten, K., Smets, E., Schols, P., Yuan, Y., Janssens, S., Küpfer, P., Pyck, N., 2004. Conflicting phylogenies of balsaminoid families and the polytomy in Ericales: combining data in a Bayesian framework. *Mol. Phylogenet. Evol.* 31, 711–729.
- Gillespie, E.L., Kron, K.A., 2010. Molecular phylogenetic relationships and a revised classification of the subfamily Ericoideae (Ericaceae). *Mol. Phylogenet. Evol.* 56, 343–354.
- Gillespie, E.L., Kron, K.A., 2012. A new tribe, Bryanthaeae (Ericoideae: Ericaceae), composed of disjunct genera from South America and Japan. *Brittonia* 64, 73–75.
- Givnish, T.J., Renner, S.S., 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *Int. J. Plant Sci.* 65, S1–S6.
- Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J.D., Clements, M.A., Arroyo, M.T.K., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., Cameron, K.M., 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. R. Soc. Lond. (Biol.)* 282 DOI: 10.1098/rspb.2015.1553.
- Govaerts, R., Frodium, D.G., Pennington, T.D., 2001. World Checklist and Bibliography of Sapotaceae. Surrey, UK, Royal Botanic Gardens, Kew.
- Grant, V., 1959. Natural History of the Phlox Family. Martinus Nijhoff, The Hague, Netherlands.
- Grey-Wilson, C., 1980. *Impatiens of Africa. Morphology, Pollination, and Pollinators, Ecology, Phytogeography, Hybridization, Keys and a Systematic Treatment of all African Species with a Note on Collecting and Cultivation.* A.A. Balkema, Rotterdam.
- Grote, P.J., Dilcher, D.L., 1992. Fruits and seeds of tribe gordoniaceae (Theaceae) from the eocene of North America. *Am. J. Bot.* 79, 744–753.
- Hao, G., Yuan, Y., Hu, C., Ge, X., Zhao, N., 2004. Molecular phylogeny of *Lysimachia* (Myrsinaceae) based on chloroplast trnL-trnF and nuclear ribosomal ITS sequences. *Mol. Phylogenet. Evol.* 31, 323–339.
- Hijmans, R.J., van Etten J., 2016. Raster: Geographic Analysis and Modeling with Raster Data. R package version 2.5-8. <http://CRAN.R-project.org/package=raster>.
- Hileman, L.C., Vasey, M.C., Parker, V.T., 2001. Phylogeny and biogeography of the Arbutioideae (Ericaceae): implications for the Madrean-Tethyan hypothesis. *Syst. Bot.* 26, 131–143.
- Hinchliff, C.E., Roalson, E.H., 2013. Using supermatrices for phylogenetic inquiry: an example using the sedges. *Syst. Biol.* 62, 205–219.
- Höhna, S., May, M.R., Moore, B.R., 2016. TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics* 32, 789–791.
- Hou, Y., Björå, C.S., Ikeda, H., Brochmann, C., Popp, M., 2016. From the north into the Himalayan-Hengduan Mountains: fossil-calibrated phylogenetic and biogeographical inference in the arctic-alpine genus *Diapensia* (Diapensiaceae). *J. Biogeogr.* 43, 1502–1513.
- Hunter, G.E., 1966. Revision of Mexican and Central American *Saurauia* (Dilleniaceae). *Ann. Missouri Bot. Gard.* 53, 47–89.
- Janies, D.A., Studer, J., Handelman, S.K., Linchangco, G., 2013. A comparison of supermatrix and supertree methods for multilocus phylogenetics using organismal datasets. *Cladistics* 29, 560–566.
- Janssens, S.B., Knox, E.B., Huysmans, S., Smets, E.F., Merckx, V.S., 2009. Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. *Mol. Phylogenet. Evol.* 52, 806–824.
- Jo, S., Kim, H., Kim, Y., Cheon, S., Kim, K., 2016. The first complete plastome sequence from the family Sapotaceae, *Pouteria campechiana* (Kunth) Baehni, Mitochondrial DNA B. Resour. 1, 734–736.
- Jordan, G.J., Bannister, J.M., Mildenhall, D.C., Zetter, R., Lee, D.E., 2010. Fossil Ericaceae from New Zealand: deconstructing the use of fossil evidence in historical biogeography. *Am. J. Bot.* 97, 59–70.
- Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30, 772–780.
- Källersjö, M., Bergqvist, G., Anderberg, A.A., 2000. Generic realignment in primuloid families of the Ericales s.l.: a phylogenetic analysis based on DNA sequences from three chloroplast genes and morphology. *Am. J. Bot.* 87, 1325–1341.
- Källersjö, M., Ståhl, B., 2003. Phylogeny of Theophrastaceae (Ericales s. lat.). *Int. Plant Sci.* 164, 579–591.
- Keller, J.A., Herendeen, P.S., Crane, P.R., 1996. Fossil flowers and fruits of the Actinidiaceae from the Campanian (Late Cretaceous) of Georgia. *Am. J. Bot.* 83, 528–541.
- Kobuski, C.E., 1961. Studies in the Theaceae, XXX the African Species of *Ternstroemia*. *J. Arnold Arb.* 42, 81–86.
- Kron, K.A., Judd, W.S., Stevens, P.F., Crayn, D.M., Anderberg, A.A., Gadek, P.A., Quinn, C.J., Luteyn, J.L., 2002. Phylogenetic classification of Ericaceae: molecular and morphological evidence. *Bot. Rev.* 68, 335–423.
- Kron, K.A., Luteyn, J.L., 2005. Origin and biogeographic patterns in Ericaceae: New insights from recent phylogenetic analyses. *Biol. Skr.* 55, 479–500.
- Lagomarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., Davis, C.C., 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210, 1430–1442.
- Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* 14, 82.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773.
- Lens, F., Schönenberger, J., Baas, P., Jansen, S., Smets, E., 2007. The role of wood anatomy in phylogeny reconstruction of Ericales. *Cladistics* 23, 229–254.
- Li, J., Li, X., Soejarto, D.D., 2007. Actinidiaceae. In: Wu, Z.Y., Raven, P.H., Hong, D.Y. (Eds.), *Flora of China*. Vol. 12 (Hippocastanaceae through Theaceae). Missouri Botanical Garden Press, St. Louis.
- Li, M., Li, J., del Tredici, P., Corajod, J., Fu, C., 2013a. Phylogenetics and biogeography of Theaceae based on sequences of plastid genes. *J. Syst. Evol.* 51, 396–404.
- Li, Y., Awasthi, N., Yang, J., Li, C., 2013b. Fruits of *Schima* (Theaceae) and seeds of *Toddalia* (Rutaceae) from the Miocene of Yunnan Province, China. *Rev. Palaeobot. Palynol.* 193, 119–127.
- Löfstrand, S.D., Schönenberger, J., 2015a. Comparative floral structure and systematics in the sarracenioid clade (Actinidiaceae, Roridulaceae and Sarraceniaceae) of Ericales. *Bot. J. Linnean Soc.* 178, 1–46.
- Löfstrand, S.D., Schönenberger, J., 2015b. Molecular phylogenetics and floral evolution in the sarracenioid clade (Actinidiaceae, Roridulaceae and Sarraceniaceae) of Ericales. *Taxon* 64, 1209–1224.
- Lott, T.A., Manchester, S.R., Dilcher, D.L., 1998. A unique and complete polemoniaceous plant from the middle Eocene of Utah USA. *Rev. Palaeobot. Palyn.* 104, 39–49.
- Luna, I., Ochoterena, H., 2004. Phylogenetic relationships of the genera of Theaceae based on morphology. *Cladistics* 20, 223–270.
- Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L., Hernández-Hernández, T., 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytol.* 207, 437–453.
- Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J., Chilquillo, E., Ronsted, N., Antonelli, A., 2015. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? *Glob. Ecol. Biogeogr.* 24, 973–984. <http://dx.doi.org/10.1111/geb.12326>.
- Martinez, C., Choo, T.Y.S., Allevato, D., Nixon, K.C., Crepet, W.L., Harbert, R.S., Daghlian, C.P., 2016. *Rariglandia jerseyensis*, a new ericalean fossil flower from the Late Cretaceous of New Jersey. *Botany* 94, 747–758.
- Martins, L., Oberprieler, C., Hellwig, F.H., 2003. A phylogenetic analysis of Primulaceae s.l. based on internal transcribed spacer (ITS) DNA sequence data. *Plant Syst. Evol.* 237, 75–85.
- Mast, A.R., Kelso, S., Richards, A.J., Lang, D.J., Feller, D.M.S., Conti, E., 2001. Phylogenetic Relationships in *Primula* L. and Related Genera (Primulaceae) based on noncoding chloroplast DNA. *Int. J. Plant Sci.* 162, 1381–1400.
- Matzke, N.J., 2012. Founder-event speciation in BioGeoBEARS package dramatically improves likelihoods and alters parameter inference in Dispersal-Extinction-Cladogenesis (DEC) analyses. *Front. Biogeogr.* 4, 210.
- Matzke, N.J., 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts, CRAN: The Comprehensive R Archive Network. Austria, Vienna Website <http://cran.r-project.org/package=BioGeoBEARS>.
- Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63, 951–970.
- Matzke, N. J. 2016. Stochastic mapping under biogeographical models. Website http://phylo.wikidot.com/biogeobears#stochastic_mapping.
- May, M.R., Höhna, S., Moore, B.R., 2016. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods Ecol. Evol.* 7, 947–959.
- Meijer, W., Veldkamp, J.F., 1993. A revision of *Mitrastema* (Rafflesiaceae). *Blumea* 38, 221–229.
- Mori, S.A., Tsou, C.H., Wu, C.C., Cronholm, B., Anderberg, A.A., 2007. Evolution of Lecythidaceae with an emphasis on the circumscription of Neotropical genera: information from combined *ndhF* and *trnL-F* sequence data. *Am. J. Bot.* 94, 289–301.
- Mori, S.A., Matos, C.C., Huang, Y., Smith, N.P., de Moraes Potaschke, C., 2015. The utility of placentalization in the circumscription of genera of New World Lecythidaceae (Brazil nut family). *Phytoneuron* 13, 1–46.
- Nie, Z.L., Sun, H., Manchester, S.R., Meng, Y., Luke, Q., Wen, J., 2012. Evolution of the intercontinental disjunctions in six continents in the *Ampelopsis* clade of the grape family (Vitaceae). *BMC Evol. Biol.* 12, 17.
- Nixon, K.C., Crepet, W.L., 1993. Late Cretaceous Fossil Flowers of Ericalean Affinity. *Am. J. Bot.* 80, 616–623.
- Oh, I., Schönenberger, J., Motley, T.J., Myrenäs, M., Anderberg, A.A., 2013. Phylogenetic relationships among endemic Hawaiian *Lysimachia* (Ericales: Primulaceae): Insights from nuclear and chloroplast DNA sequence data. *Pac. Sci.* 67, 237–251.
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* 51, 933–938.
- Paun, O., Turner, B., Trucchi, E., Munzinger, J., Chase, M.W., Samuel, R., 2016. Processes

- driving the adaptive radiation of a tropical tree (*Diospyros*, Ebenaceae) in New Caledonia, a biodiversity hotspot. *Syst. Biol.* 65, 212–227.
- Pennington, R.T., Richardson, J.E., Lavin, M., 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytol.* 172, 605–616.
- Pirie, M.D., Oliver, E.G.H., de Kuppler, A.M., Gehrke, B., Le Maitre, N.C., Kandziora, M., Bellstedt, D.U., 2016. The biodiversity hotspot as evolutionary hot-bed: spectacular radiation of *Erica* in the Cape Floristic Region. *BMC Evol. Biol.* 16, 190. <http://dx.doi.org/10.1186/s12862-016-0764-3>.
- Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: Convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.
- Poinar Jr., G.O., Weigend, M., Henning, T., 2015. *Klaprothiopsis dyscrita* gen et sp. nov. (Loasaceae) in mid-Tertiary Dominican Amber. *J. Bot. Res. Inst. Texas* 9, 369–379.
- Prince, L.M., 2002. Circumscription and biogeographic patterns in the Eastern North American–East Asian genus *Stewartia* (Theaceae: Stewartieae): insight from chloroplast and nuclear DNA sequence data. *Castanea* 67, 290–301.
- Prince, L.M., Parks, C.R., 2001. Phylogenetic relationships of Theaceae inferred from chloroplast DNA data. *Am. J. Bot.* 88, 2309–2320.
- Rabosky, D.L., 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9, e89543.
- Rabosky, D.L., Grudler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., Huang, H., Larson, J.G., 2014. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Meth. Ecol. Evol.* 5, 701–707.
- Rabosky, D.L., Mitchell, J.S., Chang, J., 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* DOI: 10.1093/sysbio/syx037.
- Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6, Website <http://tree.bio.ed.ac.uk/software/tracer/>.
- Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57, 4–14.
- Reveal, J.L., 2010. Proposal to conserve the name *Mitrastemon* (Rafflesiaceae) with that spelling. *Taxon* 59, 299–300.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- Roquet, C., Boucher, F.C., Thuiller, W., Lavergne, S., 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886.
- Rose, J.P., Kriebel, R., Sytsma, K.J., 2016. Shape analysis of moss (Bryophyta) sporophytes: insights into land plant evolution. *Am. J. Bot.* 103, 652–662.
- Ruhfel, B.R., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E., Burleigh, J.G., 2014. From algae to angiosperms—inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evol. Biol.* 14, 23.
- Ruhfel, B.R., Bove, C.P., Philbrick, C.T., Davis, C.C., 2016. Dispersal largely explains the Gondwanan distribution of the ancient tropical clusioid plant clade. *Am. J. Bot.* 103, 1117–1128.
- Sadowski, E., Seyfullah, L.J., Sadowski, F., Fleischmann, A., Behling, H., Schmidt, A.R., 2015. Carnivorous leaves from Baltic amber. *Proc. Natl. Acad. Sci. U.S.A.* 112, 190–195.
- Schenk, J.J., Hufford, L., 2010. Effects of substitution models on divergence time estimates: simulations and an empirical study of model uncertainty using cornales. *Syst. Bot.* 35, 578–592.
- Schönenberger, J., 2009. Comparative floral structure and systematics of Fouquieriaceae and Polemoniaceae (Ericales). *Int. J. Plant Sci.* 170, 1132–1167.
- Schönenberger, J., Friis, E.M., 2001. Fossil flowers of ericalean affinity from the late cretaceous of Southern Sweden. *Am. J. Bot.* 88, 467–480.
- Schönenberger, J., Anderberg, A.A., Sytsma, K.J., 2005. Molecular phylogenetics and patterns of floral evolution in the Ericales. *Int. J. Plant Sci.* 166, 265–288.
- Schönenberger, J., von Balthazar, M., Sytsma, K.J., 2010. Diversity and evolution of floral structure among early diverging lineages in the Ericales. *Proc. R. Soc. Lond. (Biol.)* 365, 437–448.
- Schönenberger, J., von Balthazar, M., Takahashi, M., Xiao, X., Crane, P.R., Herendeen, P.S., 2012. *Glandulocalyx upatoiensis*, a fossil flower of Ericales (Actinidiaceae/Clethraceae) from the Late Cretaceous (Santonian) of Georgia USA. *Ann. Bot.* 109, 921–936.
- Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J., Linder, H.P., 2015. As old as the mountains: the radiations of the Ericaceae. *New Phytol.* 207, 355–367.
- Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
- Smith, S.A., Beaulieu, J.M., Stamatakis, A., Donoghue, M.J., 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *Am. J. Bot.* 98, 404–414.
- Soejarto, D.D.M., 1969. Aspects of reproduction in *Saurauia*. *J. Arnold Arbor.* 50, 180–196.
- Soltis, D.E., Moore, M.J., Burleigh, J.G., Bell, C.D., Soltis, P.S., 2010. Assembling the angiosperm tree of life: progress and future prospects. *Ann. Missouri Bot. Gard.* 97, 514–526.
- Soltis, D.E., Mort, M.E., Latvis, M., Mavrodiev, E.V., O'Meara, B.C., Soltis, P.S., Burleigh, J.G., de Casas, R.R., 2013. Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *Am. J. Bot.* 100, 916–929.
- Soltis, D.E., Smith, S., Cellinese, N., Wurdack, K.J., Tank, D., Brockington, S.F., Refulio-Rodriguez, N.F., Moore, M.J., Carlswald, B., Bell, C.D., Latvis, M., Crawley, S., Black, C., Diouf, D., Xi, Z., Gitzendanner, M.A., Sytsma, K.J., Qiu, Y.-L., Hill, K.W., Manchester, S.R., Davis, C.C., Sanderson, M.J., Olmstead, R., Judd, W.S., Donoghue, M., Soltis, P.S., 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *Am. J. Bot.* 98, 704–730.
- Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *Bot. J. Linnean Soc.* 133, 381–461.
- Spalink, D., Drew, B.T., Pace, M.C., Zaborsky, J.G., Li, P., Cameron, K.M., Givnish, T.J., Sytsma, K.J., 2016a. Evolution of geographical place and niche space: Patterns of diversification in the North American sedge (Cyperaceae) flora. *Mol. Phylogenet. Evol.* 95, 183–195. <http://dx.doi.org/10.1016/j.ympev.2015.09.028>.
- Spalink, D., Drew, B.T., Pace, M.C., Zaborsky, J.G., Starr, J.R., Cameron, K.M., Givnish, T.J., Sytsma, K.J., 2016b. Biogeography of the cosmopolitan sedges (Cyperaceae) and the area–richness correlation in plants. *J. Biogeogr.* 43, 1893–1904. <http://dx.doi.org/10.1111/jbi.12802>.
- Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
- Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26, 1569–1571.
- Swenson, U., Anderberg, A.A., 2005. Phylogeny, character evolution, and classification of Sapotaceae (Ericales). *Cladistics* 21, 101–130.
- Swenson, U., Nylinder, S., Munzinger, J., 2014. Sapotaceae biogeography supports New Caledonia being an old Darwinian island. *J. Biogeogr.* 41, 797–809.
- Testo, W., Sundue, M., 2016. A 4000-species dataset provides new insights into the evolution of ferns. *Mol. Phylogenet. Evol.* 105, 200–211.
- Thomas, D.C., Chatrou, L.W., Stull, G.W., Johnson, D.M., Harris, D.J., Thongpairaj, U., Saunders, R.M.K., 2015. The historical origins of palaeotropical intercontinental disjunctions in the pantropical flowering plant family Annonaceae. *Perspect. Plant Ecol. Evol. Syst.* 17, 1–16.
- de la Torre-Bárcena, J.E., Kolokotronis, S., Lee, E.K., Stevenson, D.W., Brenner, E.D., Katari, M.S., Coruzzi, G.M., DeSalle, R., 2009. The impact of outgroup choice and missing data on major seed plant phylogenetics using genome-wide EST data. *PLoS One*. <http://dx.doi.org/10.1371/journal.pone.0005764>.
- Turner, B., Munzinger, J., Duangjai, S., Tensch, E.M., Stockenhuber, Barfuss, M.H.J., Chase, M.W., Samuel, R., 2013. Molecular phylogenetics of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers. *Mol. Phylogenet. Evol.* 69, 740–763.
- Utteridge, T.M.A., Saunders, R.M.K., 2001. Sexual dimorphism and functional dioecy in *Maesa perlarium* and *M. japonica* (Maesaceae/Myrsinaceae). *Biotropica* 33, 368–374.
- Utteridge, T.M.A., Saunders, R.M.K., 2004. The genus *Maesa* (Maesaceae) in the Philippines. *Bot. J. Linnean Soc.* 145, 17–43.
- de Vos, J.M., Hughes, C.E., Schneeweiss, G.M., Moore, B.R., Conti, E., 2014. Heterostyly accelerates diversification via reduced extinction in primroses. *Proc. R. Soc. Lond. (Biol.)* 81, 20140075 doi: 10.1098/rspb.2014.0075.
- Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. *J. Biomed. Inform.* 39, 34–42.
- Wu, C., Hsu, Z.F., Tsou, C.H., 2007. Phylogeny and taxonomy of *Eurya* (Ternstroemiaceae) from Taiwan, as inferred from ITS sequence data. *Bot. Stud.* 48, 97–116.
- Wurdack, K.J., Davis, C.C., 2009. Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *Am. J. Bot.* 96, 1551–1570.
- Xiang, Q.Y., Thomas, D.T., Xiang, Q.P., 2011. Resolving and dating the phylogeny of Cornales—Effects of taxon sampling, data partitions, and fossil calibrations. *Mol. Phylogenet. Evol.* 59, 123–138.
- Yan, M., Moore, M.J., Meng, A., Yao, X., Wang, H., 2016. The first complete plastome sequence of the basal asterid family Styracaceae (Ericales) reveals a large inversion. *Plant Syst. Evol.* <http://dx.doi.org/10.1007/s00606-016-1352-0>.
- Yu, S., Janssens, S.B., Zhu, X., Lidén, M., Gao, T., Wang, W., 2015. Phylogeny of *Impatiens* (Balsaminaceae): integrating molecular and morphological evidence into a new classification. *Cladistics* 32, 179–197.
- Yu, X.Q., Gao, L.M., Soltis, D.E., Soltis, P.S., Yang, J.B., Fang, L., Yang, S.X., Li, D.Z., 2017. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol.* 215, 1235–1248.
- Yuan, Y., Song, Y., Geuten, K., Rahelivololona, E., Wohlhauser, S., Fischer, E., Smets, E., Küpfer, P., 2004. Phylogeny and biogeography of balsaminaceae inferred from ITS sequences. *Taxon* 53, 391–403.
- von Balthazar, M., Schönenberger, J., 2013. Comparative floral structure and systematics in the balsaminoid clade including Balsaminaceae, Marcgraviaceae and Tetrameristaceae (Ericales). *Bot. J. Linnean Soc.* 173, 325–386.
- Zhang, L., Comes, H.P., Kadereit, J.W., 2001. Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *Am. J. Bot.* 88, 2331–2345.
- Zhang, W., Kann, S., Zhao, H., Li, Z., Wang, X., 2014. Molecular phylogeny of tribe Theeae (Theaceae s.s.) and its implications for generic delimitation. *PLoS One* 9, e98133.