

Chapter 29

Diversification History of Neotropical Lecythidaceae, an Ecologically Dominant Tree Family of Amazon Rain Forest



Oscar M. Vargas and Christopher W. Dick

Abstract The Neotropical subfamily of Lecythidaceae (Lecythidoideae) is a clade of 10 genera with an estimated number of 232 species. Lecythidaceae is the third most abundant family of trees in Amazon forests, and its most diverse genus, *Eschweilera* (ca. 100 species) is the most abundant genus of Amazon trees. In this chapter we explore the diversification history of the Lecythidoideae through space and time in the Neotropics. We inferred a time-calibrated phylogeny of 118 species, which we used to reconstruct the biogeographic origins of Lecythidoideae and its main clades. To test for significant changes of speciation rates in the subfamily, we performed a diversification analysis. Our analysis dated the crown clade of Lecythidoideae at 46 Ma (95% CI = 36.5–55.9 Ma) and the stem age at 62.7 Ma (95% CI = 56.7–68.9 Ma), suggesting dispersal from the paleotropics long after the Gondwana breakup. Most major crown clades in the Lecythidoideae (*Grias*, *Gustavia*, *Eschweilera*, *Couroupita*, *Couratari*, and all *Lecythis* and *Eschweilera* subclades) differentiated during the Miocene (ca. 5.3–23 Ma). The Guayana floristic region (Guiana Shield + north-central Amazon) is the inferred ancestral range for 8 out of the 18 Lecythidoideae clades (129 species, ~55%), highlighting the region's evolutionary importance, especially for the species-rich *Bertholletia* clade, which includes the genera *Eschweilera*, *Lecythis*, *Corythophora* and *Bertholletia*. Our results indicate that the *Bertholletia* clade colonized the Trans-Andean region at least three times in the last 10 Ma. We found no significant changes in the rate of diversification inside Lecythidoideae over the Cenozoic, and found no evidence of increased speciation during the Pleistocene. Lecythidoideae has diversified not in

O. M. Vargas (✉)

Department of Ecology and Evolutionary Biology and Herbarium, University of Michigan, Ann Arbor, MI, USA

Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA

C. W. Dick (✉)

Department of Ecology and Evolutionary Biology and Herbarium, University of Michigan, Ann Arbor, MI, USA

Smithsonian Tropical Research Institute, Panama City, Republic of Panama

e-mail: [cwick@umich.edu](mailto:cwdick@umich.edu)

pulses, but in a pattern of steady accumulation, akin to a museum model of diversification.

Keywords Ericales · Brazil nut · Speciation · Boreotropics · Long-distance dispersal · Phylogeny

1 Introduction

Lecythidaceae is an important family of woody plants in the Amazon forests, both in terms of species richness and abundance. As such, its diversification history can greatly illuminate our understanding of the assembly of biological communities in the Neotropics. Lecythidaceae is a pantropical family of trees in the order Ericales often referred to include either five (APG IV 2016) or three subfamilies (Huang et al. 2015). In the latter classification, Napoleonaeoideae and Scytopetaloidae are excluded from Lecythidaceae, leaving three core Lecythidaceae subfamilies Foetidioideae (in Madagascar), Planchonioideae (in Asia and Africa), and the Lecythidoideae, which is restricted to the Neotropics (Mori et al. 2017). The Lecythidoideae contains ca. 232 (Mori 2017) of the ca. 278 known species in the family (Mori et al. 2007, 2017; Huang et al. 2015; Mori 2017). An enigmatic species, *Asteranthos brasiliensis* Desf., is the single Neotropical representative of the West African family Scytopetalaceae, which is sometimes included in the Lecythidaceae (subfamily Scytopetaloidae; Mori et al. 2017).

Neotropical Lecythidaceae (excluding *A. brasiliensis*, and henceforth called “Lecythidoideae”) are understory, canopy, or emergent trees with distinctive floral morphology, fibrous bark, and woody fruit capsules. Although the subfamily is distributed from Mexico to Southern Brazil, its epicenter of species diversity is Amazonia. Lecythidaceae is the third most abundant family of trees in Amazon forests, ranking only behind Fabaceae and Sapotaceae (ter Steege et al. 2013). The most species-rich genus, *Eschweilera* with ca. 100 species (Mori 2017), is also the most abundant Amazonian tree genus (ter Steege et al. 2013). The canopy species *Eschweilera coriacea* is the most common tree in much of Amazonia, contributing substantially to the biomass and carbon sequestration of Amazon forests (ter Steege et al. 2013). Other notable Lecythidaceae include the Brazil nut tree, *Bertholletia excelsa*, which has been a keystone food source for Amazon peoples for millennia (Shepherd and Ramirez 2011); *Cariniana micrantha*, the monkey nut tree (tauari), which is the oldest documented Amazon tree (carbon dated at >1400 years; Chambers et al. 1998); and *Couroupita guianensis* (cannonball tree), a cauliflorous tree with cannonball-sized fruits that is grown in tropical botanical gardens around the world. Lecythidaceae includes important timber species, such as *Cariniana legalis*, known from Brazil’s Atlantic Forest. Most of the species are found in lowland rain forest, although some species are adapted to tropical montane forest, flooded forest (igapó and várzea), and savanna. Bees are the main pollinators for Lecythidoideae, with some taxa showing specialization toward carpenter and euglossine bees (Huang

2010) and even bats (Prance and Mori 1979). Fruits in Lecythidoideae are woody and can be dehiscent or indehiscent. Although there is much to be learned about seed dispersal in the family, preliminary studies and field observations have suggested that water (e.g. some species of *Allantoma*), wind (*Cariniana* and *Couratari*), parrots and macaws (some *Eschweilera* spp.), fish (*Gustavia* spp.), rodents (including agoutis, which gnaw through the Brazil nut fruit wall), monkeys (some *Lecythis* spp.), and bats (*Lecythis* spp.) act as dispersal agents (Prance and Mori 1979; Mori and Prance 1990).

Phylogenetic analyses of Lecythidaceae, using chloroplast markers and the internal transcribed spacer (ITS), identified the Neotropical subfamily Lecythidoideae as monophyletic and sister to the Paleotropical clade Foetidioideae + Planchonioideae (Mori et al. 2007). These analyses also identified the ecologically important and diverse Bertholletia clade, comprised of four genera with zygomorphic flowers (*Eschweilera*, *Lecythis*, *Corythophora*, and *Bertholletia*) (Mori et al. 2007; Huang et al. 2015). The Bertholletia clade is nested within a paraphyletic grade of actinomorphic-flowered (*Grias*, *Gustavia*, and *Allantoma*) and zygomorphic-flowered genera (*Cariniana*, *Couroupita*, and *Couratari*; Huang et al. 2015; Mori et al. 2017). The two most species-rich genera in the Bertholletia clade, *Lecythis* and *Eschweilera*, are not monophyletic in these phylogenetic reconstructions. These groups may be paraphyletic. Alternatively, the lack of monophyly may result from poor resolution of deeper nodes using ITS and a handful of chloroplast markers.

Given that Lecythidoideae is nested in a grade comprising predominantly Paleotropical clades with actinomorphic flowers (Foetidioideae, Planchonioideae, Napoleonaceae and Scytopetalaceae), the leading hypotheses to explain the dispersal of Lecythidoideae from the Paleotropics are (1) a vicariance event after the breakup of Gondwana, (2) migration using a high latitude land bridge(s) (e.g. North Atlantic land bridge or Beringia, Graham 2018), and (3) a long-distance dispersal event across the Atlantic (Mori et al. 2017). In a phylogenetic analysis of Ericales, the crown and stem ages of Lecythidoideae have been estimated to be 43 Ma and 53 Ma, respectively, with Lecythidoideae sister to Southeast Asian Planchonioideae (Rose et al. 2018). This time frame post-dates the breakup of Gondwana (90–100 Ma) by tens of millions of years, supporting a land-bridge migration or a long-distance dispersal.

After its establishment in the Neotropics, Lecythidoideae evolved into a speciose clade with ca. 232 species. This species richness has been partially attributed to Pleistocene refugia (Haffer 1969) in *Cariniana*, *Couratari*, and *Eschweilera* (Prance 1973, 1978), but a holistic investigation about Lecythidoideae diversification is missing. Considering the Eocene arrival of the family to the Neotropics (Rose et al. 2018), it is also possible that Lecythidoideae have accumulated species steadily akin to a museum model (Stebbins 1974).

In this chapter, we present the first chronogram and ancestral range reconstructions for the Lecythidoideae. Our phylogenetic inference combines published data (Mori et al. 2007; Huang et al. 2015) and DNA sequences of 13 highly informative regions extracted from 26 plastomes (Thomson et al. 2018; Yu et al. 2017). Our main goals were to (1) identify the geographic origins and divergence times of the primary

clades of Lecythidoideae, (2) address the hypotheses (land-bridge migration *vs.* oceanic dispersal) for the origin of Lecythidaceae in the Neotropics, and (3) test for heterogeneity in diversification rates across the tree to address the Pleistocene refuge *vs.* the museum hypotheses.

2 Methods

2.1 Tree Building

Lecythidaceae genetic data were obtained from GenBank using PYPHLAWD (Smith and Brown 2018). We focused on the Lecythidoideae and the outgroup genus *Barringtonia*, which was identified as a close relative (Mori et al. 2007). We retained DNA markers represented in >50 species (i.e. ITS, *ndhF*, *psbA-trnH*, and *trnL-F*) and species with least two DNA regions (we noted inconsistencies in the placement of species represented by only one marker). We combined the PYPHLAWD-generated matrix with a matrix containing 13 plastome regions identified by Thomson et al. (2018) as phylogenetically informative for the Lecythidoideae. The 13-region matrix was extracted from an alignment of 26 plastomes that contained all Lecythidoideae genera and three *Barringtonia* species (Thomson et al. 2018; Yu et al. 2017). Our strategy of combining a species-rich matrix (PYPHLAWD-generated) with a sequence-rich one (13-region) aimed to create a comprehensive Lecythidoideae tree with a strong backbone. The combined matrix contained a total of 118 species (including 8 outgroups) and 16 markers (PYPHLAWD-mined: ITS, *ndhF*, and *trnL-F*; 13 highly-informative plastome regions: *petN-trnD*, *psaJ-rps18*, *psbM-trnD*, *psbZ-trnfM*, *rps15-ycf1*, *trnE-trnT*, *trnfM-psbA*, *trnK-rps16*, *trnT-psbD*, *trnV-atpE*, *ycf1(1)*, *ycf1(2)*; *psbA-trnH* was shared by both marker-sets). We used MAFFT v7.310 (Katoh and Standley 2013) to align individual regions before concatenation and searched for the best maximum likelihood (ML) tree in RAxML v. 8.2.11 (Stamatakis 2014) using the option “-f a”, which performs rapid bootstraps (200) and then searches for the best-scoring ML topology. Every region was treated by RAxML as an independent partition using a GTRGAMMA model of molecular evolution (Stamatakis 2015). The best-scoring ML tree was subsequently time-calibrated using BEAST v.2.5 (Bouckaert et al. 2014). We employed two fossils to calibrate our phylogeny: (1) a seed fossil, *Lecythidospermum* (Pons and Rica 1983), assigned to the crown node containing *Allantoma*, *Cariniana*, *Couratari*, and the Bertholletia clade, with a lognormal prior distribution with a median of 24.3 Ma and a 95% confidence interval (CI) of 20.0–29.2 Ma; (2) a wood fossil, *Barringtonioxylon deccanense* (Shallom 1960; Srivastava et al. 2009), to the stem node of *Barringtonia*, with a lognormal prior distribution with a median of 66.0 Ma and a 95% CI of 60.3–72.0 Ma. We ran three independent BEAST analyses of ten million generation sampling every 4000; these results were combined to calculate a chronogram using a 0.25 burnin fraction in LogCombiner v.2.5 and TreeAnnotator v.2.5 (Bouckaert et al. 2014). We checked

for a minimum effective sample size >200 for the estimated parameters with Tracer v.1.6 (Rambaut et al. 2018).

2.2 *Ancestral Range Reconstruction*

To infer the geographic history of Lecythidoideae in the Neotropics and identify areas of importance for its diversification, we performed an ancestral range reconstruction on the chronogram, using a dispersal-extinction-cladogenesis model (DEC, Ree and Smith 2008) in BioGeoBEARS (Matzke 2013). Given a phylogeny and the presence or absence of extant taxa in a set of geographical regions, DEC infers the geographical ranges of internal nodes by modeling dispersal, extinction, and cladogenesis (speciation). We did not include a founder speciation event parameter “J” (Matzke 2014), given the caveats noted by Ree and Sanmartín (2018). Although the models DIVA (Ronquist 1997) and BayArea (Landis et al. 2013) are also implemented in BioGeoBEARS and typically run in conjunction with DEC, we opted not employ them in our analysis because DIVA does not model some cladogenetic processes (i.g. parapatric speciation; Kodandaramaiah 2010)—a process expected to have occurred in the Lecythidoideae given the numerous species found in Amazon forests—and BayArea is optimized for taxa distributed on numerous geographic regions like islands systems (Landis et al. 2013)—Lecythidoideae are found in contiguous low and middle elevations forests throughout tropical Central and South America.

We divided the Neotropical region in eight areas, based on a previous bioregionalization (Morrone 2014; Löwenberg-Neto 2014) modified in QGIS v2.18 (QGIS Development Team 2005). Presence of species in our regions was codified from a database of specimens that included data from GBIF (<https://www.gbif.org>) and the New York Botanical Garden. Duplicates and unlikely data points (e.g. maritime areas) were filtered out.

2.3 *Diversification Analyses*

We performed an analysis of diversification on our chronogram using BAMM v2.5 (Rabosky 2014). BAMM uses reversible-jump Markov chain Monte Carlo to infer diversification processes using phylogenies. We ran BAMM with four chains for ten million generations, sampling every 1000, with a sampling correction that considers missing taxa in the phylogeny, and based on the number of species sampled in each one of the main Lecythidoideae clades described in Mori et al. (2017) (Table 29.1). We enforced an effective sample size >200 and used BAMMtools (Rabosky et al. 2014) to overlay speciation rates onto our phylogeny and to graph average speciation rates over time.

Table 29.1 Crown clade ages, with 95% confidence interval (CI), most probable origin with its probability, speciation rates (Lambda), and percentage sampled for main Lecythidoideae clades

Clade/genus	Age	95% CI	Most probable origin	Prob.	Species	Lambda	Sampled (%)
Lecythidoideae	46.1	36.5–55.9	W Amazonia	0.036	232	0.159	47
<i>Grias</i>	7.0	2.4–12.8	Transandean, Andean, N&W Amazonia	0.239	11	0.156	27
<i>Gustavia</i>	20.4	11.2–30.9	Transandean	0.109	45	0.163	20
<i>Couroupita</i>	8.8	3.9–14.2	Transandean, N&W Amazonia, Guayana	0.106	3	0.149	100
<i>Allantoma</i>	4.8	2.1–8.0	Amazonia	0.302	8	0.151	38
<i>Cariniana</i>	23.2	15.3–30.2	W Amazonia	0.710	9	0.151	33
Echinata (<i>Couratari</i>)	17.2	9.8–25.7	Transandean, W Amazonia	0.187	6	0.152	50
Guianensis (<i>Couratari</i>)	13.3	6.8–20.5	Guayana	0.319	13	0.151	31
Bertholletia clade	28.6	22.4–34.7	Guayana	0.822	136	0.163	62
Chartacea (<i>Lecythis</i>)	16.3	11.8–21.5	Guayana	0.956	14	0.154	71
Integrifolia (<i>Eschweilera</i>)	12.2	7.8–16.5	Guayana	0.165	22	0.156	77
Tetrapetala (<i>Eschweilera</i>)	11.4	6.3–17.2	Cerrado & Caatinga	0.337	8	0.152	38
Ollaria (<i>Lecythis</i>)	13.8	8.5–19.3	Transandean, N Amazonia	0.464	3	0.153	100
Poiteaui (<i>Lecythis</i>)	17.0	11.2–22.8	Guayana	0.799	6	0.153	83
Pisonis (<i>Lecythis</i>)	9.0	5.2–13.2	Amazonia, Cerrado & Caatinga, Mata Atlantica	0.106	5	0.154	80
<i>Corythophora</i>	12.5	7.8–17.5	Guayana	0.990	4	0.154	100
Corrugata (<i>Lecythis</i>)	8.2	4.6–12.2	Guayana	0.856	5	0.164	100
Parvifolia (<i>Eschweilera</i>)	13.6	10.2–17.1	Guayana	0.600	65	0.186	46

All GenBank accessions, DNA alignments, locality data, bioregionalization map, control files, scripts, and additional results and figures used for or generated by this study can be found at https://bitbucket.org/oscarvargash/lecythidaceae_diversification

3 Results

Our matrix comprised 118 species and 16 DNA regions, containing 12,726 aligned nucleotides. Overall matrix gene-occupancy was 39%, with ITS = 84%, *ndhF* = 97%, *psbA-trnH* = 83%, *trnL-F* = 95%, and the remaining 12 highly-informative regions = 22%. Nearly half (48%) of the nodes in our ML phylogeny had high ($\geq 80\%$) bootstrap support (BS), while 14% had moderate BS (60–79) and 38% had low BS (< 60). Our tree had more supported nodes than the most recently published comprehensive molecular tree for the Lecythidoideae (Huang et al. 2015) which contained 24% nodes with high BS. Our topology largely agrees with that of Huang et al. (2015) and recovers all of the clades detected by their study: *Bertholletia* clade (*Eschweilera* + *Lecythis* + *Corythophora* + *Bertholletia*); the *Lecythis* clades Ollaria, Pisonis, Corrugata, Poiteaui, and Chartacea; the *Eschweilera* clades Integrifolia, Tetrapetala, and Parvifolia; *Corythophora*, and the species *Bertholletia excelsa* (Fig. 29.1). There are three main differences in our tree when compared against previous phylogenies (Mori et al. 2007; Huang et al. 2015). First, our phylogeny recovers *Couratari* as polyphyletic (with moderate BS), while previous studies recovered it as monophyletic. Second, our phylogeny suggests that *Bertholletia excelsa* is sister to the Chartacea clade and the Integrifolia clade, while the phylogeny of Huang et al. (2015) placed it as sister of the Integrifolia clade. Finally, in our tree, *Eschweilera amazoniciformis* is sister to a clade comprising the Pisonis clade, *Corythophora*, the Corrugata clade, and the Parvifolia clade, while in previous phylogenies this species was inferred to be nested within the Integrifolia clade (Huang et al. 2015). The positions of *Bertholletia excelsa* and *Eschweilera amazoniciformis*, however, have low support.

When comparing prior vs. posterior distributions for the nodes calibrated with our Bayesian analysis, we found no difference for the node calibrated with *Barringtonioxylum*, while the node dated with *Lecythidospermum* showed a posterior distribution slightly shifted towards older ages. The latter suggests that the sequence data used in this study provide information for the estimation of a posterior estimate of *Lecythidospermum*. The chronogram dates the crown clade for the Lecythidoideae at 46 Ma (95% CI = 36.5–56 Ma) (Fig. 29.1, Table 29.1) and its stem age at 62.71 Ma (95% CI = 56.7–68.92 Ma). *Bertholletia* clade's stem and crown ages date back to the Oligocene. In the *Bertholletia* clade, stem and crown ages of main clades (e.g. Parvifolia, Chartaceae, Huang et al. 2015) fall within the Miocene, with only the stem age of the Tetrapetala clade falling within the Oligocene.

Our biogeographic analysis shows a preliminary historical range reconstruction for Lecythidoideae (Fig. 29.2, Table 29.1). While many of the range reconstructions for early ancestors in the subfamily are ambiguous (e.g. the ancestor for the Lecythidoideae), the reconstruction for many ancestors in the *Bertholletia* clade is dominated by the Guayana area. Furthermore, Guayana is reconstructed as the most probable ancestral region for eight diverse clades (comprising ~129 species), suggesting it as the most important geographic region for neotropical diversification in the subfamily. Despite ambiguous reconstructions for ancestors of *Grias* and

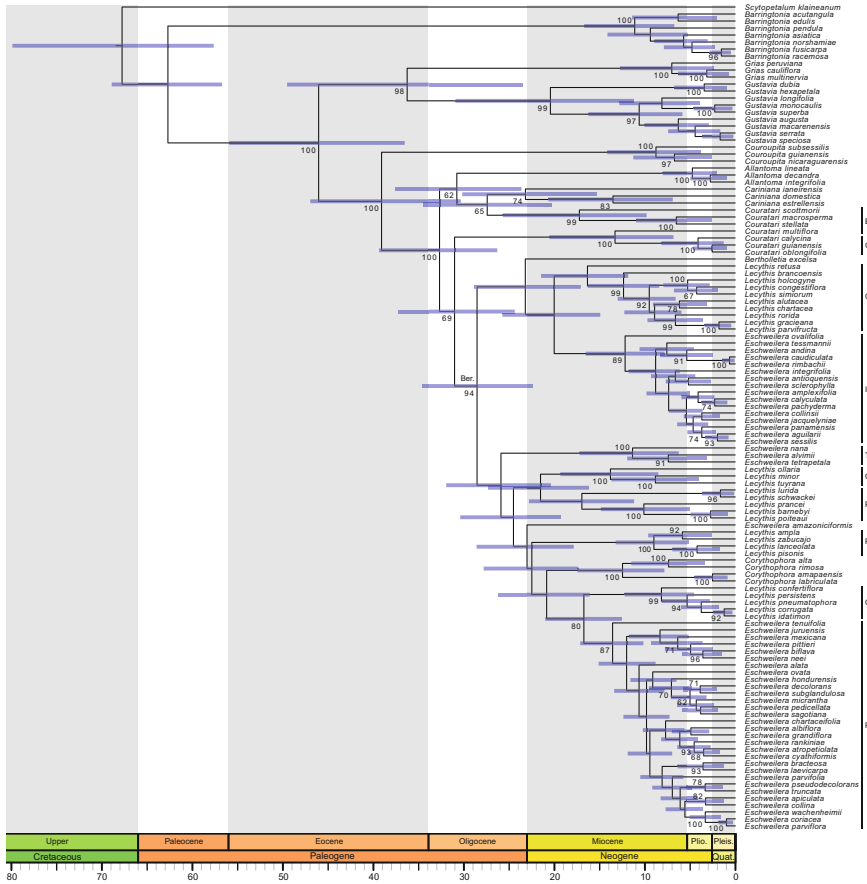


Fig. 29.1 Lecythidoideae chronogram. Blue bars at nodes indicate the 95% confidence intervals for the age of a given node. Numbers at nodes indicate bootstrap support, not shown for values <60. Main clades are indicated as: Ber. = *Bertholletia*, Ech. = *Echinata*, Gui. = *Guianensis*, Cha. = *Chartacea*, Int. = *Integrifolia*, Tet. = *Tetrapetala*, Oll. = *Ollaria*, Poi. = *Poiteau*, Pis. = *Pisonis*, Cor. = *Corrugata*, Par. = *Parvifolia*. Stars indicate newly defined clades in this study

Gustavia, the Trans-Andean region seems central to their diversification (Fig. 29.2, Table 29.1). Similarly, the Western Amazon area seems to have played an important role as the setting for diversification of the clade comprised by *Allantoma* + *Echinata* (*Couratari*) clade + *Cariniana*. Finally, our reconstruction also suggests at least three instances of dispersal into the Trans-Andean region within the *Bertholletia* clade, whose center of diversification appears to be the Guayana region, in the last 10 Ma.

We found no evidence for shifts in the diversification dynamics of the Lecythidoideae through the Cenozoic: a scenario of no significant changes in diversification rates was found to have the highest posterior probability (0.62). The

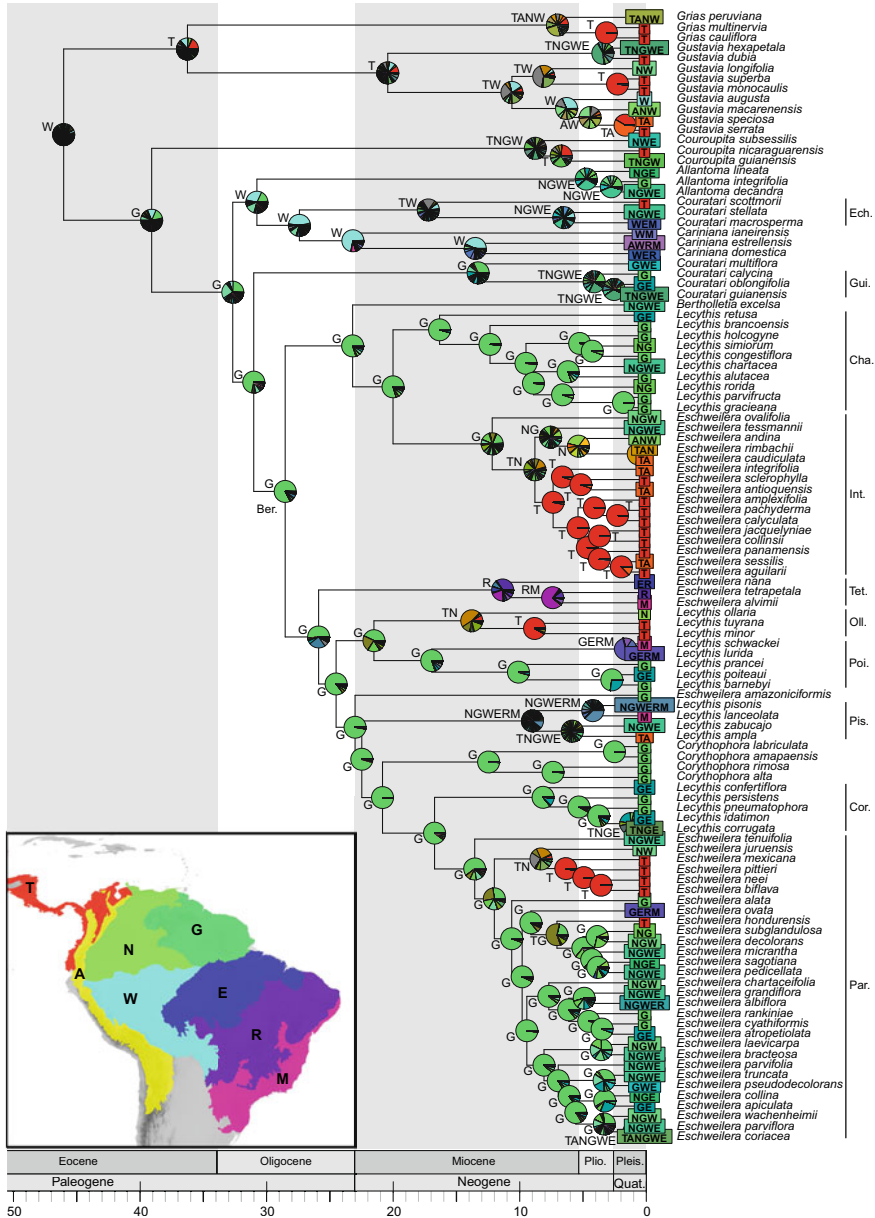


Fig. 29.2 Biogeographic ancestral reconstruction for Lecythidoideae. Percent probabilities are shown as pie charts, the most probable range is indicated on top of the pie chart. Notice that ancestral ranges comprising multiple areas are represented by a combination of colors/letters of the individual areas. Main clades are indicated as in Fig. 29.1

next probable scenario is that of one significant change, yet with a posterior probability of 0.27; scenarios with two or more changes have a posterior probability of <0.09 (see additional BAMM results in the on-line repository). Furthermore, our estimations of speciation rates (λ) suggest stability through time with a slight increase towards the present (Fig. 29.3); overall mean λ is 0.156 for the Lecythidoideae with the Parvifolia clade (*Eschweilera*) having the highest λ of 0.186 (this higher rate is not statistically different), and *Couroupita* having the lowest λ of 0.149 (Table 29.1).

4 Discussion

4.1 Systematics of Lecythidoideae

We present the most comprehensive phylogeny for the Lecythidoideae to date. Our phylogeny, which contains 111 species for the subfamily (representing 47% of its species), largely agrees with that of Huang et al. (2015), recovering all the main clades described by them (Fig. 29.1, Table 29.1), and supporting the polyphyly of *Eschweilera* and *Lecythis*. Our topology suggests, however, with moderate support, that *Couratari* is not monophyletic; instead, its species appear to belong to two clades. To facilitate the discussion and future communication among botanists, we named the two clades of *Couratari* as Echinata and Guianensis, with the caveat that we only sampled seven species of the genus (37%). Echinata comprises the species of section Echinata, and Guianensis comprises species of the sections *Couratari* and *Microcarpa* (Mori and Prance 1990). The inferred positions of *Bertholletia excelsa* and *Eschweilera amazoniciformis* are also in conflict with the phylogeny of Huang et al. (2015), but they are poorly supported in both phylogenies. The placement of these two taxa is problematic because they both bear long branches and seem to be of old age (both dating from the Oligocene/Miocene boundary).

While our phylogeny adds more evidence to the polyphyly of *Eschweilera* and *Lecythis* and suggests for the first time the non-monophyly of *Couratari*, we maintain a conservative position about re-circumscribing these genera until a larger number of unlinked nuclear markers are included in the analysis. Our matrix is primarily composed of plastome markers, and preliminary phylogenomic evidence for Lecythidoideae suggest deep incongruence between the plastome phylogeny and a species tree inferred with >300 nuclear genes (Vargas et al. 2019).

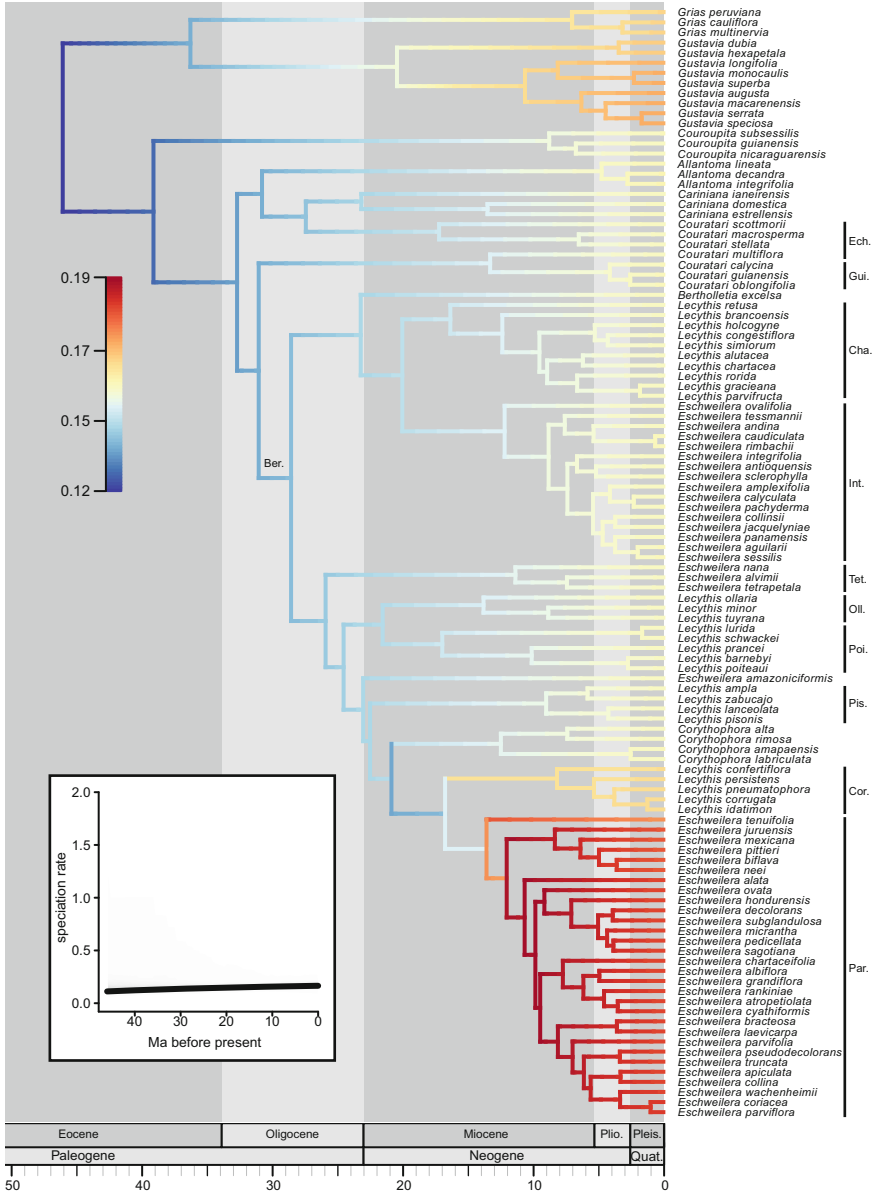


Fig. 29.3 Model-averaged speciation rates plotted on the Lecythidoideae phylogeny and speciation rates through time (inset). Main clades are indicated as in Fig. 29.1

4.2 *Spatiotemporal Patterns in the Lecythidoideae*

4.2.1 Dispersal of Lecythidaceae to the New World

Our chronogram suggests that stem (62.7 Ma 95% CI = 56.7–68.9 Ma) and crown (46 Ma 95% CI = 36.5–55.9 Ma) clade ages for the Lecythidoideae date to the Paleocene and Eocene, respectively (Fig. 29.1), firmly within the Cenozoic (i.e. within past 66 Ma). Because Lecythidoideae is nested within Neotropical clades (Foetidioideae, Planchonioideae, Napoleonaceae and Scytopetalaceae; Mori et al. 2017), we can infer that Neotropical Lecythidaceae has its roots in the Paleotropics. Given that the Gondwana breakup occurred in the mid-Cretaceous (e.g. 80–100 Ma), our age estimates rule out a Gondwana vicariance explanation for the pantropical disjunction of the Lecythidaceae, supporting an oceanic dispersal (Fig. 29.4a) or a land-bridge migration in the early Cenozoic (Fig. 29.4b).

The land-bridge scenario (Fig. 29.4b) for the Neotropical distribution of Lecythidaceae proposes that early members crossed one of the high latitude land bridges (e.g. Bering Land Bridge, or North Atlantic Land Bridge) early in the Cenozoic, when Earth surface temperatures were higher, and cold-intolerant thermophilic (“tropical”) lineages extended as far north as 50 degrees (Graham 2011). During the Paleocene-Eocene thermal maximum (PETM), the warmest period of the Cenozoic, some plant taxa of tropical affinities were widely distributed in the Northern hemisphere (Eurasia and North America) creating a boreotropical distribution—palms and figs were left as fossils in New England and London. Lavin and Luckow (1993) proposed a phylogenetic test of such high latitude (boreotropical) dispersal for Neotropical lineages. It requires (1) that the Neotropical lineage be nested within a tropical Asian or African clade, and (2) existence of high latitude fossil records. Several important Neotropical tree clades are shown to meet these criteria, including Melastomataceae (Renner et al. 2001), Burseraceae (Weeks et al. 2005), Meliaceae (Muellner et al. 2006), and some clades of Lauraceae (Chanderbali and van der Werff 2001; Huang et al. 2016) and Annonaceae (Couvreur et al. 2011). Land-bridge migration can also occur from the Neotropics to the Paleotropics: there is evidence that Malpighiaceae (Davis et al. 2002) originated in the Neotropics but obtained its pantropical distribution through boreotropical or “Laurasian” dispersal.

Because there are no fossils of either the distinctive wood or woody fruits of Lecythidaceae in any high latitude site, we propose the oceanic dispersal hypothesis as the most plausible to explain the Lecythidaceae arrival to the Neotropics (Fig. 29.4a). As with the disjunct Amazonian *Asteranthos*, which is nested within west African Scytopetalaceae, the ancestor of many Amazon tree lineages likely arrived by oceanic dispersal (Pennington and Dick 2004). This conclusion stands in contrast to that of Rose et al. (2018), who estimated similar crown/stem ages for Lecythidoideae but accepted a Gondwana vicariance explanation. We believe that the disagreements between our study and that of Rose et al. (2018) are based on the evidence and analyses taken into consideration to draw interpretations. While we take fossil evidence (absence of Lecythidaceae in North America) and

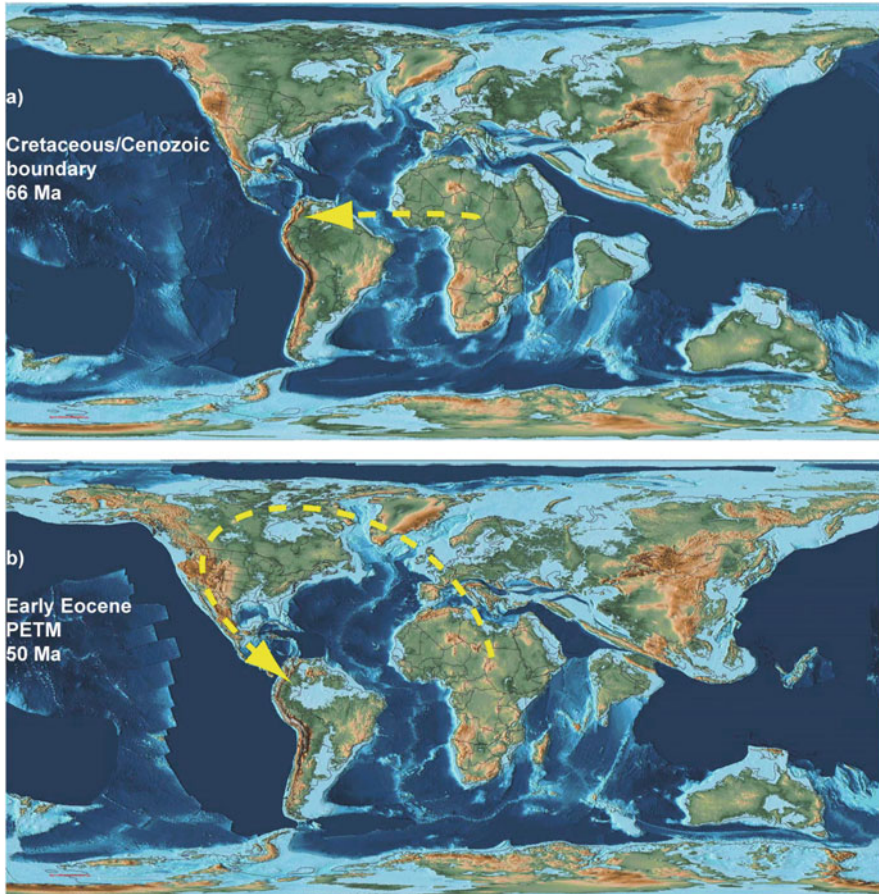


Fig. 29.4 Competing hypothesis about the origin of the Lecythidoideae. Arrows represent colonization routes. **(a)** Long dispersal hypothesis (supported by the present study) represented on the Cretaceous-Tertiary boundary earth tectonic reconstruction 66 Ma, which is close to the Lecythidoideae stem age 62.7 Ma (95% confidence interval = 56.7–68.9 Ma). **(b)** Boreotropical colonization hypothesis represented on the Earth tectonic reconstruction during the Paleocene-Eocene thermal maximum 50 Ma. Figures modified from Scotese (2001)

paleogeography into consideration (stem and crown Lecythidoideae ages postdate the break up Gondwana), the conclusions of Rose et al. appear to be based solely on an ancestral biogeographic reconstruction that suggests a vicariance event, in which the parental node of the Neotropical Lecythidoideae, Paleotropical Foetidoideae, and Paleotropical Planchonioideae is reconstructed as being distributed in both the Neotropics and the Afrotropics with a moderate probability of 0.61 (Rose et al. 2018, p. 67).

4.2.2 Diversification Within the Neotropics

Mori et al. (2017) suggested that the first Lecythydaceae arrived in Central America and first diverged into the *Gustavia* + *Grias* clade. While our phylogeny supports the *Gustavia* + *Grias* clade as sister to the rest of Lecythydoideae, we cannot pinpoint the arrival location to the Neotropics from the Paleotropics due to the ambiguity of our reconstruction for the deepest nodes in the subfamily (Fig. 29.2).

Our analyses suggest that the Bertholletia clade diversified in the transition of the Oligocene to the Miocene at 28.6 Ma (95%CI = 22.4–34.7), with a high probability of having originated in the Guayana region (0.82) (Fig. 29.1, Table 29.1). All the clades that make up the Bertholletia clade *sensu* Huang et al. (2015) (e.g. Chartacea, Parvifolia) appear to have diverged during the Miocene, with six clades having Guayana as the most probable center of origin (Chartacea, Integrifolia, Poiteaui, Corythophora, Guianensis, and Parvifolia). These results suggest that the Guayana region has been a center of diversification for the Bertholletia clade.

We hypothesize that the elevated terrain of the Guayana shield, which predates the origin of the Lecythydaceae, has provided constant non-flooded *terra firme* for the family to diversify over time. In contrast to Guayana, Northern Amazonia and Western Amazonia (following our bioregionalization, Fig. 29.2) experienced marine incursions and wetland systems during the Miocene (Hoorn et al. 2010), making these areas improbable for Lecythydaceae speciation. It has been documented that Lecythydoideae is more abundant in *terra-firme*, with near 70% of the Neotropical species inhabiting non-flooded forest (Mori et al. 2017). The preference for *terra-firme* predicts that other areas, i.e. the proto-Andean Cordillera, which could have sustained *terra-firme* forest in the Eocene, could also be a center of diversification for the family. Estimates indicate that the proto-Andean cordilleras began their uplift during the Cretaceous (Gregory-Wodzicki 2000), likely providing low elevation mountains (<1000 m) but with enough elevation for a *terra-firme* forest during the Eocene—when the Lecythydoideae started to diversify. In fact, current distributions of *Grias*, *Gustavia*, and the Integrifolia clade show higher concentration of species in the hills or in the surroundings of the Andes Cordillera (Mori et al. 2017). There is little support for the proto-Andean hypothesis in our biogeographic analysis (Fig. 29.2, Table 29.1), but the absence of the Andean region in our ancestral reconstruction could be an artifact of the (perhaps too stringent) threshold used in our analysis to score Andean species (>1000 m), and our biogeographic modeling failing to include Andean paleoelevation estimates that could allow for an appropriate inference of the distribution of ancestors in the early Andean Cordillera.

4.3 Diversification Rates

Our estimations of diversification rates in the Lecythydoideae shows that the subfamily has diversified relatively constantly over time, with a slight tendency to

increase speciation rates towards the present (Fig. 29.3). Our analysis does not show significant increases of speciation rates for the Pleistocene (inset Fig. 29.3), as in a refuge hypothesis (Haffer 1969), which was previously a favored explanation for the sympatric occurrence of closely related *Eschweilera* species in central Amazon forests (Prance 1978). Instead, our results favor a museum hypothesis (Stebbins 1974) with most nodes (stem and crown) leading to the main Lecythidoideae crown clades being older than 10 Ma. Nevertheless, there are species that diverged recently during the Pleistocene (e.g. some species of *Lecythis* and *Eschweilera*) and our phylogeny is missing 53% of Lecythidoideae species. Beyond these classical hypotheses to explain the Amazon biodiversity, our results show that Lecythidaceae arrived to the Amazon forests after its formation (Graham 2011) and that Lecythidoideae speciated at a steady rate until the present. While our results support the museum hypothesis, we believe there is much to learn from Amazon speciation beyond binary categorizations (i.e. cradle vs. museum). In particular, we believe that causes for speciation are pivotal for understanding biological diversification in the Amazon region. Lineage splitting processes remain a mystery and should be studied to, for example, explain the presence of numerous closely-related Lecythidaceae species in the Guayana floristic region.

4.4 Caveats of Our Study

Despite the fact that we presented the most comprehensive phylogeny of Lecythidoideae to date, we sampled only 47% of its current diversity, with the lowest sampled genera being *Grias* and *Gustavia* with 27% and 20%, respectively. The gaps in our sampling, specifically in *Grias* and *Gustavia*, affected our biogeographic reconstruction (inconclusive for some of the ancestors of these two sister taxa) and diversification analysis; a denser taxon sampling may increase diversification rates in main clades especially towards the Pleistocene. Sampling efforts are unequal across the Amazon basin as major Lecythidaceae collectors, Mori and Prance, focused primarily in the central Amazon around Manaus and the Guayana region. It is possible that these collections biased our sampling and our estimation of distributional ranges for our biogeographic analysis.

Our dating analysis is based solely on two fossils, with only one of them, *Lecythidospermum*, located in the ingroup. Our diversification analysis is based solely on tree shape and no fossils were used to estimate diversification rates. While additional New World fossils have been reported (Mori et al. 2017), these need to be revised in terms of age or taxonomic identity for accurate utilization in the Lecythidoideae phylogeny (C. Martinez, pers. comm.).

Finally, our model for reconstructing ancestral ranges does not consider historical geology and climate, meaning that our reconstruction assumes no change in landscape and forest coverage through time. Northwestern South America was considerably different before the middle Miocene because major uplift of the Central and Northern Andes occurred in the last 10 Ma (Gregory-Wodzicki 2000), meaning that

Andean paleoelevation during the Eocene was much lower. Additionally, the isthmus of Panama remained opened until at least 15 Ma (Montes et al. 2015) and it is known that wetlands existed during the Miocene in western Amazonia. These historical landscape changes, not accounted in this study, most certainly affected the coverage of Amazon forests and hence the past distribution of the Lecythidoideae.

4.5 Conclusions and Future Directions

Based on the most comprehensive Lecythidoideae phylogeny to date, we inferred spatiotemporal diversification patterns of this subfamily during the Cenozoic. Our results support an oceanic dispersal event from Africa as early as the late Cretaceous and early divergence of the Lecythidoideae during the Eocene. We found that its most important center of diversification is the Guayana region, which was likely the origin of the Bertholletia clade that includes the most speciose genera in the family, *Eschweilera* and *Lecythis*. Since its arrival to the Neotropics, Lecythidoideae has speciated at a steady rate, with a slight but non-significant increase in speciation rates towards the present. To corroborate some of new hypotheses stated in this study it will be necessary to increase the number of taxa included in the phylogenetic reconstruction of the subfamily, especially for the non-Bertholletia genera *Allantoma*, *Couratari*, *Grias*, and *Gustavia*. The inclusion of nuclear data will be pivotal in obtaining a robust species tree in the future, and a comprehensive study and reassessment of the fossil record should provide further insight into the history of the family. Comparative phylogenetic approaches, like the reconstruction of historical niches, have the promise of complementing the preliminary spatiotemporal inferences presented here.

Acknowledgements The National Science Foundation (DEB 1240869 and FESD Type I 1338694 to CWD and Stephen Smith) and the University of Michigan (Associate Professor Award to CWD) provided financial support for this work. We thank the members of the Stephen Smith lab, Scott Mori, Ana Carnaval, Lucia Lohmann, João Capurucho, and Giorgia Auteri for useful comments and support. Camila Martinez kindly provided expert advice about fossil utilization in the dating analysis. We are grateful to Scott Mori for sharing many insights into the family, and for providing access to botanical collections and DNA samples through the New York Botanical Garden.

References

- APG IV (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc* 181:1–20. <https://doi.org/10.1111/boj.12385>
- Bouckaert R, Heled J, Kühnert D et al (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10:1–6. <https://doi.org/10.1371/journal.pcbi.1003537>

- Chambers JQ, Higuchi N, Schimel JP (1998) Ancient trees in Amazonia. *Nature* 391:135–136. <https://doi.org/10.1038/34325>
- Chanderbali PAS, van der Werff H (2001) Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. *Ann Missouri Bot Gard* 88:104–134
- Couvreur TLP, Pirie MD, Chatrou LW et al (2011) Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. *J Biogeogr* 38:664–680. <https://doi.org/10.1111/j.1365-2699.2010.02434.x>
- Davis CC, Bell CD, Mathews S, Donoghue MJ (2002) Laurasia migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *Proc Natl Acad Sci U S A* 99:6833–6837. <https://doi.org/10.1073/pnas.102175899>
- Graham A (2011) A natural history of the New World, the ecology and evolution of plants in the Americas. University of Chicago Press, Chicago
- Graham A (2018) Land bridges: ancient environments, plant migrations, and new world connections. University of Chicago Press, Chicago
- Gregory-Wodzicki KM (2000) Uplift history of the Central and Northern Andes: a review. *Geol Soc Am Bull* 112:1091–1105. [https://doi.org/10.1130/0016-7606\(2000\)112<1091](https://doi.org/10.1130/0016-7606(2000)112<1091)
- Haffer J (1969) Speciation in Amazonian forest birds. *Science* 165:131–137
- Hoorn C, Wesselingh FP, ter Steege H et al (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330:927–931. <https://doi.org/10.1126/science.1194585>
- Huang YY (2010) Systematics of Lecythidoideae (Lecythidaceae) with emphasis on *Bertholletia*, *Corythophora*, *Eschweilera*, and *Lecythis*. A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York
- Huang YY, Mori SA, Kelly LM (2015) Toward a phylogenetic-based generic classification of neotropical Lecythidaceae—I. Status of *Bertholletia*, *Corythophora*, *Eschweilera* and *Lecythis*. *Phytotaxa* 203:85–121. <https://doi.org/10.11646/phytotaxa.203.2.1>
- Huang JF, Li L, van der Werff H et al (2016) Origins and evolution of cinnamon and camphor: a phylogenetic and historical biogeographical analysis of the *Cinnamomum* group (Lauraceae). *Mol Phylogenet Evol* 96:33–44. <https://doi.org/10.1016/j.ympev.2015.12.007>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Kodandaramaiah U (2010) Use of dispersal – vicariance analysis in biogeography – a critique. *J Biogeogr* 37:3–11. <https://doi.org/10.1111/j.1365-2699.2009.02221.x>
- Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP (2013) Bayesian analysis of biogeography when the number of areas is large. *Syst Biol* 62:789–804. <https://doi.org/10.1093/sysbio/syt040>
- Lavin M, Luckow M (1993) Origins and relationships of tropical North America in the context of the Boreotropics hypothesis. *Am J Bot* 80:1–14
- Löwenberg-Neto P (2014) Neotropical region: a shapefile of Morrone’s (2014) biogeographical regionalisation. *Zootaxa* 3802:300. <https://doi.org/10.11646/zootaxa.3802.2.12>
- Matzke NJ (2013) BioGeoBEARS: Biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts, CRAN: The Comprehensive R Archive Network, Vienna, Austria. <http://cran.r-project.org/package=BioGeoBEARS>
- Matzke NJ (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst Biol* 63:951–970. <https://doi.org/10.1093/sysbio/syu056>
- Montes C, Cardona A, Jaramillo C et al (2015) Middle Miocene closure of the Central American Seaway. *Science* 348:226–229. <https://doi.org/10.1126/science.aaa2815>
- Mori SA (2017) The Lecythidaceae pages. <http://sweetgum.nybg.org/science/projects/lp/>
- Mori SA, Prance GT (1990) Lecythidaceae. Part II. The zygomorphic-flowered New World genera (*Couropita*, *Corythophora*, *Bertholletia*, *Couratari*, *Eschweilera*, & *Lecythis*). *Flora Neotrop Monogr* 21:1–376

- Mori SA, Tsou CH, Wu CC et al (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. <https://doi.org/10.3732/ajb.94.3.289>
- Mori SA, Kiernan EA, Smith NP et al (2017) Observations on the phylogeography of the Lecythidaceae clade (Brazil nut family). *Phytoneuron* 30:1–85
- Morrone JJ (2014) Biogeographical regionalisation of the neotropical region. *Zootaxa* 3782:1–110. <https://doi.org/10.11646/zootaxa.3782.1.1>
- Muellner AN, Savolainen V, Samuel R, Chase MW (2006) The mahogany family “out-of-Africa”: divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. *Mol Phylogenet Evol* 40:236–250. <https://doi.org/10.1016/j.ympev.2006.03.001>
- Pennington RT, Dick CW (2004) The role of immigrants in the assembly of the South American rainforest tree flora. *Philos Trans R Soc Lond B Biol Sci* 359:1611–1622. <https://doi.org/10.1098/rstb.2004.1532>
- Pons D, Rica P (1983) *Lecythidospermum bolivarensis* (Berry) nov. comb., graine fossile de Lecythidaceae (Angiospermae) du Tertiaire de la Colombie. *Ann Paléontologie* 69:1–12
- Prance GT (1973) Phylogeographic support the theory of Pleistocene forest refuges in the Amazon Basin, based on evidence from distribution patterns in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. *Acta Amaz* 3:5–28
- Prance GT (1978) The origin and evolution of the Amazon flora. *Interciencia* 3:207–222
- Prance GT, Mori SA (1979) Lecythidaceae. Part I. The actinomorphic-flowered New World Lecythidaceae (*Asteranthos*, *Gustavia*, *Grias*, *Allantoma*, & *Cariniana*). *Flora Neotrop Monogr* 21:1–270
- QGIS Development Team (2005) QGIS geographic information system. Open source geospatial foundation project. <http://qgis.osgeo.org>
- Rabosky DL (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9:e89543. <https://doi.org/10.1371/journal.pone.0089543>
- Rabosky DL, Grudler M, Anderson C et al (2014) BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol Evol* 5:701–707. <https://doi.org/10.1111/2041-210X.12199>
- Rambaut A, Drummond AJ, Xie D et al (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ree RH, Sanmartín I (2018) Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *J Biogeogr* 45:741–749. <https://doi.org/10.1111/jbi.13173>
- Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst Biol* 57:4–14. <https://doi.org/10.1080/10635150701883881>
- Renner SS, Clausing G, Meyer K (2001) Historical biogeography of Melastomataceae: the roles of tertiary migration and long-distance dispersal. *Am J Bot* 88:1290–1300
- Ronquist F (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst Biol* 46:195–203. <https://doi.org/10.1093/sysbio/46.1.195>
- Rose JP, Kleist TJ, Löfstrand SD et al (2018) Phylogeny, historical biogeography, and diversification of angiosperm order Ericales suggest ancient Neotropical and East Asian connections. *Mol Phylogenet Evol* 122:59–79. <https://doi.org/10.1016/j.ympev.2018.01.014>
- Scotese CR (2001) Atlas of earth history, vol 1, Paleogeography, PALEOMAP Project. Arlington, TX
- Shallom LJ (1960) Fossil dicotyledonous wood of Lecythidaceae from the Deccan Intertrappean beds of Mahurzari. *J Ind Bot Soc* 39:198–203
- Shepard GH, Ramirez H (2011) “Made in Brazil”: human dispersal of the Brazil nut (*Bertholletia excelsa*, Lecythidaceae) in ancient Amazonia. *Econ Bot* 65:44–65. <https://doi.org/10.1007/s12231-011-9151-6>

- Smith SA, Brown JW (2018) Constructing a broadly inclusive seed plant phylogeny. *Am J Bot* 105:1–13. <https://doi.org/10.1002/ajb2.1019>
- Srivastava R, Kapgate DK, Chatterjee S (2009) Permineralized fungal remains in the fossil wood of *Barringtonia* from the Deccan Intertrappean sediments of Yavatmal District, Maharashtra, India. *The Palaeobotanist* 58:11–19
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stamatakis A (2015) The RAxML v8.1.X manual. <http://sco.h-its.org/exelixis/web/software/raxml/index.html>
- Stebbins GL (1974) Flowering plants: evolution above the species level. Belknap Press, Cambridge
- ter Steege H, Pitman NCA, Sabatier D et al (2013) Hyperdominance in the Amazonian tree flora. *Science* 342:325–342. <https://doi.org/10.1126/science.1243092>
- Thomson AM, Vargas OM, Dick CW (2018) Complete plastome sequences from *Bertholletia excelsa* and 23 related species yield informative markers for Lecythidaceae. *Appl Plant Sci* 6: e1151. <https://doi.org/10.1002/aps3.1151>
- Vargas OM, Heuertz M, Smith SA, Dick CW (2019) Target sequence capture in the Brazil nut family (Lecythidaceae): marker selection and in silico capture from genome skimming data. *Mol Phylogenet Evol* 135:98–104. <https://doi.org/10.1016/j.ympcv.2019.02.020>
- Weeks A, Daly DC, Simpson BB (2005) The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. *Mol Phylogenet Evol* 35:85–101. <https://doi.org/10.1016/j.ympcv.2004.12.021>
- Yu XQ, Gao LM, Soltis DE et al (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. <https://doi.org/10.1111/nph.14683>