Botanical Journal of the Linnean Society, 2020, **XX**, 1–19. With 5 figures.

Rapid diversification rates in Amazonian Chrysobalanaceae inferred from plastid genome phylogenetics

JEROME CHAVE[1,*], CYNTHIA SOTHERS[2], AMAIA IRIBAR[1], UXUE SUESCUN[1], MARK W. CHASE[2,3] and GHILLEAN T. PRANCE[2]

[1]*Laboratoire Evolution et Diversité Biologique UMR 5174 CNRS, IRD, Université Paul Sabatier 31062 Toulouse, France*
[2]*Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK*
[3]*Department of Environment and Agriculture, Curtin University, Perth, Western Australia, Australia*

Received 27 February 2020; revised 9 June 2020; accepted for publication 13 June 2020

We studied the evolutionary history of Chrysobalanaceae with phylogenetic analyses of complete plastid genomes from 156 species to assess the tempo of diversification in the Neotropics and help to unravel the causes of Amazonian plant diversification. These plastid genomes had a mean length of 162 204 base pairs, and the nearly complete DNA sequence matrix, with reliable fossils, was used to estimate a phylogenetic tree. Chrysobalanaceae diversified from 38.9 Mya (95% highest posterior density, 95% HPD: 34.2–43.9 Mya). A single clade containing almost all Neotropical species arose after a single dispersal event from the Palaeotropics into the Amazonian biome *c.* 29.1 Mya (95% HPD: 25.5–32.6 Mya), with subsequent dispersals into other Neotropical biomes. All Neotropical genera diversified from 10 to 14 Mya, lending clear support to the role of Andean orogeny as a major cause of diversification in Chrysobalanaceae. In particular, the understory genus *Hirtella* diversified extremely rapidly, producing > 100 species in the last 6 Myr (95% HPD: 4.9–7.4 Myr). Our study suggests that a large fraction of the Amazonian tree flora has been assembled *in situ* in the last 15 Myr.

ADDITIONAL KEYWORDS: Amazonia – Malpighiales – molecular dating – phylogenetic inference – tropical forest.

INTRODUCTION

How the Neotropical rainforest biome has been assembled is a fascinating question for biogeography and evolutionary biology (Gentry, 1982). The Neotropics harbour no fewer than 90 000 species of plants, more than the rest of the tropics combined (Antonelli & Sanmartín, 2011), and this outstanding diversity appears to be the result of a confluence of factors, including large areas with stable, favourable environmental conditions for clade persistence (Wallace, 1878) and long periods of continental isolation (Raven & Axelrod, 1974). Also, major geological events, especially the uplift of the Andes (Gregory-Wodzicki, 2000; Hoorn *et al.*, 2010) are thought to have contributed to the hydrological remodelling of the region now covered by Amazon forest (Hughes *et al.*, 2012; Hoorn *et al.*, 2017; Jaramillo *et al.*, 2017). In the late Miocene, the demise of the Pebas wetland in western Amazonia (Figueiredo *et al.*, 2009; Sacek, 2014) created new habitats such as white sand forests that may have promoted edaphic specialization (Fine *et al.*, 2010) and contributed to allopatric speciation by creating dispersal barriers (Coyne & Orr, 2004; Smith *et al.*, 2014). Plio-Pleistocene climatic fluctuations could also explain some recent rates of increased diversification (Prance, 1982; Haffer, 2008). These repeated drier and cooler episodes (Wang *et al.*, 2017) could have spurred diversification as they would have caused geographical isolation of wet-forest clades.

Dated molecular phylogenetic trees have proved essential in advancing our understanding on the diversification of plants in the tropical rainforest biome, and they shed light on episodes of intercontinental migration and pace of *in situ* diversification (Eiserhardt, Couvreur & Baker, 2017).

*Corresponding author. E-mail: jerome.chave@univ-tlse3.fr

© 2020 The Linnean Society of London,

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

A common feature in the history of Neotropical lowland plant clades is that they cannot be understood without accounting for intercontinental dispersal during the Neogene (Pennington & Dick, 2004). Combined with reliable fossil constraints, studies have identified key migrations across biomes (Donoghue & Edwards, 2014; Fine, Zapata & Daly, 2014; Donoghue & Sanderson, 2015), and a large number of Neotropical forest clades turn out to have originated outside the lowland Neotropical forests, as is the case for palms in Arecoideae (Arecaceae), which result from a single dispersal event into the Neotropics (Baker & Couvreur, 2013). The complex environmental history of South America has left an imprint on *in situ* diversification for many woody plant clades that today characterize the Amazonian forest (Antonelli & Sanmartín, 2011). Miocene onsets of diversification have been detected in Arecaceae (Roncal *et al.*, 2013; Bacon *et al.*, 2018), Burseraceae (Fine *et al.*, 2014), Annonaceae (Pirie *et al.*, 2018), Fabaceae (Schley *et al.*, 2018), Orchidaceae (Pérez-Escobar *et al.*, 2017) and Meliaceae (Koenen *et al.*, 2015). However, a complete picture of diversification for Amazonia should ideally be based on a comprehensive sampling of its tree flora, and here we provide insights from a floristically important plant family in lowland Amazonian habitats.

Chrysobalanaceae, the coco plum family, with 545 species, are a mid-sized pantropical family and a notable component of the Neotropical tree flora (Prance, 1972), with *c.* 80% of species found only in the Neotropics (Prance & Sothers, 2003). All Chrysobalanaceae are woody plants, ranging in height from 10 to > 40 m, with a uniform vegetative architecture (Prance & White, 1988). Across Amazonian rainforests, an analysis of 1170 tree inventory plots (≥ 10 cm in trunk diameter) reveals that Chrysobalanaceae rank seventh in tree dominance behind Fabaceae, Lecythidaceae, Sapotaceae, Malvaceae, Moraceae and Burseraceae (ter Steege *et al.*, 2013). Their centre of diversity is Amazonia, with 251 species restricted to lowland Amazonian forests, but they are found in virtually all Neotropical biomes, notably the Brazilian Atlantic forest (47 species), forests of Central America, Colombian Chocó and the Caribbean (47 species), dry habitats (29 species: 15 in the Brazilian cerrado, 11 in South American savannas and three in seasonally dry forests) and at high elevation (29 species: 18 in the Andes and 11 restricted to the Guyana highlands; Prance & Sothers, 2003).

Generic delimitation has been challenging in Chrysobalanaceae (for a historical account, see Prance & White, 1988). There are currently 27 genera recognized, 11 Palaeotropical, 12 Neotropical and three amphi-Atlantic (*Chrysobalanus* L., *Parinari* Aubl. and *Maranthes* Blume). The two non-Neotropical species of *Hirtella* L. from eastern Africa and Madagascar should be reassigned to the old genus name from Madagascar, *Thelira* Thouars. Eight genera are found in Oceania and Southeast Asia, and ten occur in Africa and Madagascar. Recent changes to generic delimitations in Chrysobalanaceae include the resurrection of *Angelesia* Korth. for three Southeast Asian species previously included in *Licania* Aubl. (Sothers & Prance, 2014), a revision of *Couepia* Aubl., with four species transferred to other genera, and a new genus, *Gaulettia* Sothers & Prance, created to accommodate species of the former *parillo* clade of *Couepia* (Sothers *et al.*, 2014), and splitting of the large polyphyletic Neotropical genus *Licania* into eight genera (Sothers, Prance & Chase, 2016). With clarified generic delimitation, an improved interpretation of the diversification and biogeography is now possible for the family.

In this contribution, we provide a phylogenetic reconstruction based on an expanded taxon sampling, aimed at all major clades of Chrysobalanaceae, but with a focus on the Neotropical clades. In a previous study, Bardon *et al.* (2016) sequenced plastid genomes for 51 species of Chrysobalanaceae, complemented with limited DNA sequences of 88 additional species, and concluded that the family diversified in the Palaeotropics in the early Oligocene (33 Mya) and subsequently dispersed once to the Neotropics. The hypothesis of a Palaeotropical origin for the family was motivated by the postulated position of the Southeast Asian genus *Kostermanthus* Prance as sister to the rest of the family, followed by the *Parinari-Neocarya* (DC.) Prance clade, which is pantropical. However, limited taxon sampling was a major issue, with the risk that 'rogue' taxa, widely divergent taxa with a poorly supported placement, reduced the overall phylogenetic signal. To minimize this risk, we built an improved dataset, with a threefold increase in species sampling, including all currently recognized genera, and sequencing of the full plastid genome for each of the sampled species.

Here, we ask when and how Chrysobalanaceae arose to become an important component of Amazonian tree communities. To this end, we use a novel approach to date the major events of diversification, including the crown age of the family. We then build a revised biogeographic scenario for the family both pantropically and in the Neotropics. Finally, based on an analysis of changes in diversification rates across the phylogenetic tree, we discuss shifts in diversification rates in Chrysobalanaceae.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

MATERIAL AND METHODS

De novo plastid genome sequencing

We sequenced full plastid genomes for 163 specimens of Chrysobalanaceae, representing 156 species and all 27 currently recognized genera (Supporting Information, Table S1). Two accessions of the following species were sampled: *Bafodeya benna* (Scott-Elliot) Prance, *Chrysobalanus cuspidatus* Griseb. ex Duss, *Couepia bracteosa* Benth., *Hymenopus heteromorphus* (Benth.) Sothers & Prance, *Leptobalanus octandrus* (Hoffmanns. ex Roem. & Schult.) Sothers & Prance, *Maranthes robusta* (Oliv.) Prance and *Parinari excelsa* Sabine. Tissue was taken from herbarium collections or leaf samples dried in silica. Total DNA was extracted using standard methods. A separate Illumina library was prepared for each sample, and the libraries were then multiplexed in groups of 24 or 48. The pools were sequenced on HiSeq 2000–2500 high-throughput sequencers, yielding 101-nucleotide pair-ended DNA reads, or the more recent HiSeq 3000 sequencer, yielding 150 pair-ended reads. Each run produced *c.* 700 gigabases in total. Runs generated for Bardon *et al.* (2016) were treated as new accessions and reassembled *de novo* using the bioinformatic pipeline described below.

Plastome assembly was performed on a local cluster running Linux CentOS v.6.7. We used the NOVOplasty organelle genome assembler v.2.7.2 (Dierckxsens, Mardulyn & Smits, 2016). We assumed a genome size range of 140–180 kbp, K-mer size of 39 and disabled the 'variant detection' option. We obtained unique circularized plastid genomes for 54 specimens. For another 58 specimens, we set a reference sequence to guide the assembly (using *Licania canescens* Benoist available in NCBI NC300566). In some cases, NOVOplasty returns two or more optional genomes, corresponding to large inversions, and we manually selected the option matching gene order. Points of reference of the circularized plastid genomes were aligned using the CSA software (Fernandes, Pereira & Freitas, 2009). For the remaining 51 specimens (30% of our final dataset), we mapped the reads directly against a reference plastid genome using the 'map to reference' option in Geneious v.9.0.5.

Plastomes were then rotated to a common starting point, and aligned using MAFFT v.7.222 (Katoh, Rozewicki & Yamada, 2017). The first raw alignment of 163 plastomes was initially of 207 248 nucleotides. It was manually edited to remove ambiguous characters produced at the assembly stage. We annotated the alignment for Chrysobalanaceae by using that previously published for *Hirtella physophora* Mart. & Zucc. by Malé *et al.* (2014; accession NC 024066 on NCBI). The alignment was realigned with this plastome, and annotations were transferred to all sequences in Geneious using the 'transfer annotations' option. Within the alignment, we also carefully checked the reading frames of the coding DNA sequences (CDS).

Reconstruction and dating of the crown age of Chrysobalanaceae

To reconstruct a dated phylogenetic tree of the family, we used a two-step approach. We estimated divergence times for the crown node of Chrysobalanaceae and the deeper divergences in the family. For this first analysis, we selected high-quality plastid genomes spanning Malpighiales, including representatives of the main clades of the family. We included only CDSs, which resulted in a high-quality global alignment, with 20 taxa, 19 members of Malpighiales (including five Chrysobalanaceae) and one outgroup from Oxalidales [*Sloanea latifolia* (Rich.) K.Schum., Elaeocarpaceae]. We used this first analysis to infer the crown age of Chrysobalanaceae. We conducted a second analysis using the full 163-plastome dataset, in which the crown age inferred from the first analysis was set as a constraint.

The 20-taxon sequence matrix was used to reconstruct phylogenetic relationships using maximum likelihood in RAxML v.8.2.10 (Stamatakis, 2014). The best model was found to be the general time reversible model for site substitution, with a gamma site model (GTR+Γ) for all partitions, as identified using the JmodelTest2 software (Darriba *et al.*, 2012). We considered two scenarios: all CDSs as a single partition; and two partitions, first/second codons and third codons. All analyses gave the same results, so we subsequently analysed only the single partition. We ran the analysis on the CIPRES supercomputing portal (Miller, Pfeiffer & Schwartz, 2010) using rapid bootstrapping with an automatic halting option and searching for the best-scoring tree.

To generate a time-calibrated tree, we relied on Bayesian relaxed molecular clock models as implemented in BEAST v.2.5.1 (Bouckaert *et al.*, 2019) using the DNA matrix together with fossil constraints (also available on CIPRES). The input file was generated using the BEAUTi software. We imported the same alignment as used in the RAxML analysis. We also used the GTR+ Γ model for each partition and assumed a relaxed clock log-normal model (Drummond *et al.*, 2006). We expected a Cenozoic age for the Chrysobalanaceae clade, and because our tree dating only used few internal priors, we set the branching process prior to be the Yule process (or pure birth process; Condamine *et al.*, 2015).

For the age priors, the split between Malpighiales and Oxalidales was constrained by imposing a uniform prior between 103 and 112 Myr (constraint 1; Xi *et al.*, 2012).

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

The dating of the flowering plant tree has recently been revisited using an extensive plastome dataset (Li *et al.*, 2019), and although Celastrales were proposed as a new sister group to Malpighiales, the split between Malpighiales+Celastrales and Oxalidales was inferred at 106.2 Mya (range: 93.2–125.0 Mya), close to the date proposed by Magallón *et al.* (2015), and our prior is consistent with both analyses.

We also used internal dating constraints, each modelled as uniform priors in the BEAST analysis as recommended by Condamine *et al.* (2015): the minimal age was that of the fossil and the maximum age was set at 110 Myr, the stem age of Malpighiales. Fossils based on reproductive structures were selected over other fossils and obtained from carefully documented studies (Parham *et al.*, 2011). The stem age of Clusiaceae was constrained with *Paleoclusia chevalieri* Crepet & Nixon as the oldest fossil assigned to this order, dated at 89 Mya (Crepet & Nixon, 1998; constraint 2). We constrained the stem age of *Parinari* to be > 19 Myr based on endocarp fossils recently found in Panama and reliably assigned to this genus (Jud, Nelson & Herrera, 2016; constraint 3). We also constrained the crown age of the Neotropical Chrysobalanaceae clade based on the fossil flower found in Dominican amber and reliably assigned to *Licania* section *Hymenopus* Benth. (Chambers & Poinar, 2010; constraint 4), which is likely to be > 16 Myr (see below). The stem age of Euphorbiaceae was constrained at > 61 Myr (*Acalypha* pollen type found in China, reported by Xi *et al.*, 2012; constraint 5). The stem age of Salicaceae was constrained to be > 48 Mya (fossil flower of *Pseudosalix handleyi* Boucher, Manchester & Judd; Boucher *et al.*, 2003, constraint 6), the stem age of Caryocaraceae > 55 Myr (pollen of *Retisyncolporites angularis* González-Guzmán; Germeraad, Hopping & Muller, 1968, constraint 7) and the stem age of Humiriacae > 37 Myr (fossil endocarp of *Lacunofructus cuatrecasana* Herrera, Manchester & Jaramillo; Herrera *et al.*, 2012, constraint 8).

The default starting tree for BEAST is incompatible with the imposed age constraints, so we started the MCMC search using an ultrametric starting tree modified from the RAxML output tree. We used the chronos() function of the *ape* package (Paradis, Claude & Strimmer, 2004) in the R software to generate an dated initial tree within age constraints consistent with the constraints described above. In the BEAST input file, we fixed the topology, i.e. we assumed that the input tree had the correct topology, and optimized only the parameters for evolutionary rates and branch lengths. This was achieved by manually removing the 'narrow exchange', 'wide exchange', 'Wilson Balding' and 'subtree slide' operators from the xml input file. MCMC was run for 200 million generations, sampling parameters and trees every 10 000 generations. Convergence was evaluated using Tracer v.1.7.1 (Rambaut *et al.*, 2018). The effective sampling sizes of each parameter were checked at the end of each analysis and considered to be of good quality when > 200. Divergence times were computed using TreeAnnotator v.2.5.1 after the removal of the 25% burn-in part of the MCMC (Bouckaert *et al.*, 2019).

Dated phylogenetic reconstruction for Chrysobalanaceae

To construct a dated phylogenetic hypothesis for 163 specimens of Chrysobalanaceae, we followed the same strategy described above for reconstruction of the dated phylogenetic tree of Chrysobalanaceae. We used a two-partition model. All CDS were treated as a first partition. All other regions, including intergenic regions, intronic regions, rRNA and tRNA, were a second partition. Prior to partitioning the aligned matrix, we removed the second copy of the inverted repeat, to avoid double weighting the phylogenetic signal of the inverted repeat.

Considerable age uncertainty has surrounded the biogeographical history of Chrysobalanaceae, which in part traces back to the attribution of Eocene leaf and pollen fossils from North America to Chrysobalanaceae (Berry, 1916; Wodehouse, 1932), constraining the crown age of Chrysobalanaceae to being > 50 Myr (Davis *et al.*, 2005; Bardon *et al.*, 2012; Xi *et al.*, 2012). However, Jud *et al.* (2016) found no solid evidence for fossils of Chrysobalanaceae prior to the Miocene. In the Neotropics, the first undisputed evidence of fossils of Chrysobalanaceae is demonstrated with flowers and fruits, possibly of *Licania* section *Hymenopus* (Chambers & Poinar, 2010), preserved in amber from the northern mountain range of the Dominican Republic, dated to 15–20 Mya (Iturralde-Vinent & McPhee, 1996; Iturralde-Vinent, 2001), here assumed to be > 16 Mya. We also used three independent fossils of *Parinari*, including wood and fruits from 21 Mya in Panama (Jud *et al.*, 2016), fruits from 19 Mya in Ethiopia (Tiffney, Fleagle & Brown, 1994) and wood from the mid-Miocene in India (Srivastava & Awasthi, 1996).

The dated tree was produced using BEAST v.2.5.1 using the same DNA matrix as for the RAxML analysis, linking the relaxed clock, log-normal models between the two partitions. The crown of Chrysobalanaceae was constrained with a Gaussian prior of mean of 36 Myr with variance of ± 2 Myr, consistent with the first analysis. We also used constraints 3 and 4 of the first analysis: the stem age of *Parinari* was set to be > 19 Myr, and the Neotropical Chrysobalanaceae clade was

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

set to be older than 16 Myr. MCMC was run for 100 million generations, sampling parameters and trees every 10 000 generations. Divergence times were computed using TreeAnnotator after the first 25% of the trees was discarded.

Diversification and biogeographic analyses

We first tested whether the dated phylogenetic tree was consistent with phases of accelerated or decelerated diversification. One prediction is that the major orogenic changes during the Miocene spurred plant diversification (Lagomarsino *et al.*, 2016; Pérez-Escobar *et al.*, 2017). We estimated rates of diversification and shifts in these rates using the Bayesian analysis of macroevolutionary mixture (BAMM, Rabosky *et al.*, 2013; Rabosky, 2014). BAMM tests the hypothesis that diversification has occurred homogeneously across the phylogenetic tree, the alternative being that shifts in diversification rate have occurred on specific branches of the tree. We used the initial control file with priors on rate parameters inferred by the function setBAMMpriors() of the BAMMtools package in R (Rabosky *et al.*, 2014). BAMM also provides an analytical correction for incompletely sampled trees, and here we assigned a sampling weight to each genus. The species sampling rates within the genera ranged from 13% (*Moquilea* Aubl.) to 57% (*Couepia*), for a 29% mean across the family. For *Hymenopus* (Benth.) Sothers & Prance, which appears in two separate clades in our analysis, we assumed an equal sampling of both clades (set at 36%). BAMM can also be used to compute diversification rates within subclades of the tree. We ran the reverse-jump MCMC simulation for ten million iterations to ensure convergence, which was assessed with the EffectiveSize() function of BAMMtools. For each of the four well-sampled Neotropical genera (*Couepia*, *Hirtella*, *Licania*, *Moquilea*), plus *Parinari*, we inferred the speciation rates from the BAMM run. This was done by selecting the subclade using the ape R package (Paradis *et al.*, 2004), and analysing the inferred diversification parameters on it.

We also reconstructed the biogeographic history of Chrysobalanaceae by inferring the most likely ancestral area(s). To this end, we attributed each species to a region. Because of the Neotropical focus of this study, we defined two broad regions outside the Neotropics: Africa and Southeast Asia (including Oceania). In the Neotropics, we defined five regions: Caribbean and Central America (including the Chocó region of Colombia, and southeast USA), savannas and seasonally dry tropical forests (including the Llanos of Colombia and Venezuela, cerrado in Brazil and dry forests such as the Chiquitania in Bolivia and caatinga in Brazil), Atlantic rainforest, the Andes and Amazonia. To reconstruct the ancestral area(s), we used an unconstrained dispersal-extinction-cladogenesis (DEC) model (Ree & Smith, 2008) as implemented in the R package BioGeoBEARS v.1.1.2 (Matzke, 2012, 2014). We did not perform a model comparison approach including the founder-event speciation option of BioGeoBEARS because taxon sampling in our dataset remains too incomplete and because this approach has limits (Ree & Sanmartín, 2018).

RESULTS

The 163 plastomes had a mean length of 162 204 ± 1195 bp and were fully assembled, except for *Licania densiflora* Kleinhoonte (151 268 bp), *Parinari curatellifolia* Planch. ex Benth. (157 972 bp) and *Licania micrantha* Miq. (158 098 bp). Average sequencing depth was 369 ± 233 (range: 30–1110). Manual editing of the DNA sequences focused on just 42 bp or less than 0.03% of the plastome, except for four species: *Kostermanthus heteropetalus* (Scort. ex King) Prance (2479 edits, close to *K. robustus* Prance, which had only 43 edits), *Bafodeya benna*2 (Scott-Elliot) Prance (1726 edits, close to *Bafodeya benna*1 with 286 edits), *Parinariopsis licaniiflora* (Sagot) Sothers & Prance (1075 edits) and *Cordillera platycalyx* (Cuatrec.) Sothers & Prance (1040 edits). The full DNA alignment and dated tree is available on Dryad (https://doi.org/10.5061/dryad.ghx3ffbkp).

The 20-taxon reconstruction was based on an alignment of CDS of 61 953 sites and 7292 patterns (Supporting Information, Fig. S1). The crown age of Chrysobalanaceae was inferred in the late-Eocene, at *c.* 38.9 Mya (95% highest posterior density, 95% HPD: 34.2–43.9 Mya). *Kostermanthus* Prance and *Bafodeya* Prance ex. F.White were sister to the rest of Chrysobalanaceae.

The phylogenetic reconstruction of Chrysobalanaceae was based on 67 317 CDS sites with 3335 patterns and 82 566 non-CDS sites with 13 935 patterns (Figs 1–3). Of the 162 internal nodes, 114 (70%) had bootstrap percentages > 90 and 133 (82%) had bootstrap percentages > 70. Most poorly resolved nodes correspond to within-genus splits and/or recent events. The tree inferred from RAxML with branch lengths showed no heterogeneity in substitution rates across the family (Supporting Information, Fig. S2).

Excluding *Kostermanthus-Bafodeya*, three clades are mainly African (A–C, Fig. 1). Pantropical *Parinari* was inferred to have diversified 7.6 Mya (95% HPD: 5.7–9.3 Mya), although its stem age was 28 Myr (95% HPD: 23–33.2 Myr).

Clade D is predominantly Neotropical and displays a secondary dispersal from the Neotropics

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

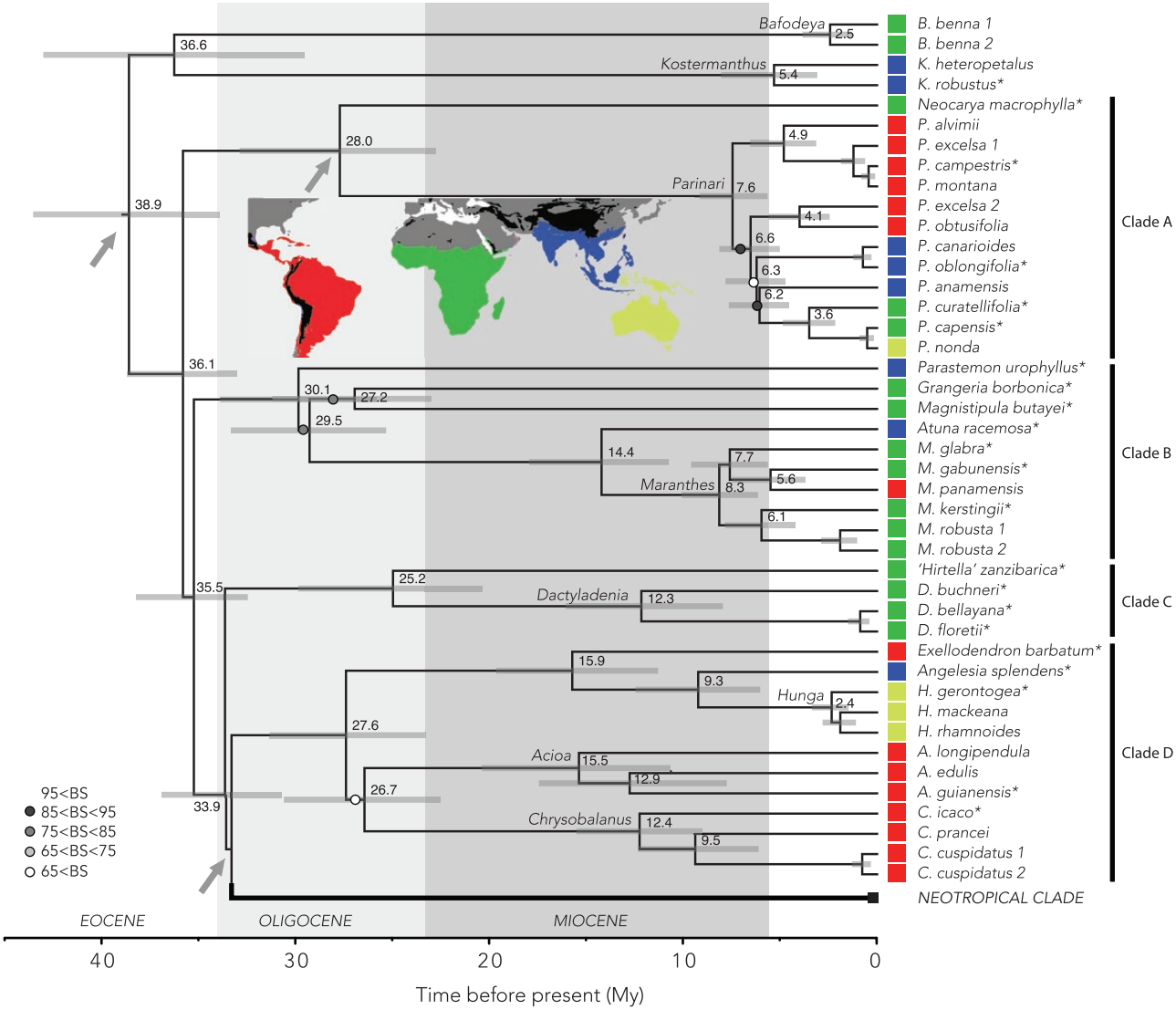

Figure 1. Phylogenetic tree for the early-diverging clades of Chrysobalanaceae, obtained from the software RAxML v.8.2.10, with dating from the software BEAST v.2.5.1. The arrows illustrate the internal calibration points used for tree dating ($N = 3$). Clade support < 95% was colour-coded with circles; all other clades had support percentages > 95%. The numbers next to each node are the inferred ages in millions of years. Grey bars represent the 95% confidence intervals. Asterisks indicate plastid genomes published in Bardon *et al.* (2016).

into Oceania and Southeast Asia (*Hunga* Pancher ex Prance and *Angelesia*), which diverged from Neotropical *Exellodendron* Prance 15.9 Mya (95% HPD: 11.4–19.8 Mya). The Neotropical clade and clade D diversified in the early Oligocene, 33.6 Mya (95% HPD: 30.6–36.9 Mya).

The crown age of the Neotropical clade was inferred at 29.1 Myr (95% HPD: 25.5–32.6 Myr, Fig. 2). *Moquilea*, *Couepia*, *Leptobalanus* (Benth.) Sothers & Prance and *Licania sensu* Sothers *et al.* (2016) were monophyletic. Species-rich genera of the Neotropical clade diversified in the mid- to late Miocene: *Moquilea* 15.1 Mya (95% HPD: 11.9–18.5 Mya), *Couepia* 10.3 Mya (8.4–12.3 Mya), *Leptobalanus* 10.3 Mya (7.6–13.0 Mya) and *Licania* 16.8 Mya (13.6–20 Mya).

Gaulettia (Sothers *et al.*, 2014) and Neotropical *Hirtella* are monophyletic (Fig. 3). The myrmecophilous species of *Hirtella* (*Hirtella* section *Myrmecophila* Prance; seven species), did not form a clade. The sister of *Gaulettia* (23.4 Mya) included *Hirtella* plus a complex of seven groups with low bootstrap support: *Hymenopus* cf. *occultans* (Prance) Sothers & Prance, *Microdesmia* (Benth.) Sothers & Prance, *Hymenopus*1, *Afrolicania* Mildbr., *Cordillera* Sothers & Prance, *Parinariopsis* (Huber) Sothers & Prance and *Hymenopus*2. Except for *Hymenopus*, each genus was

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

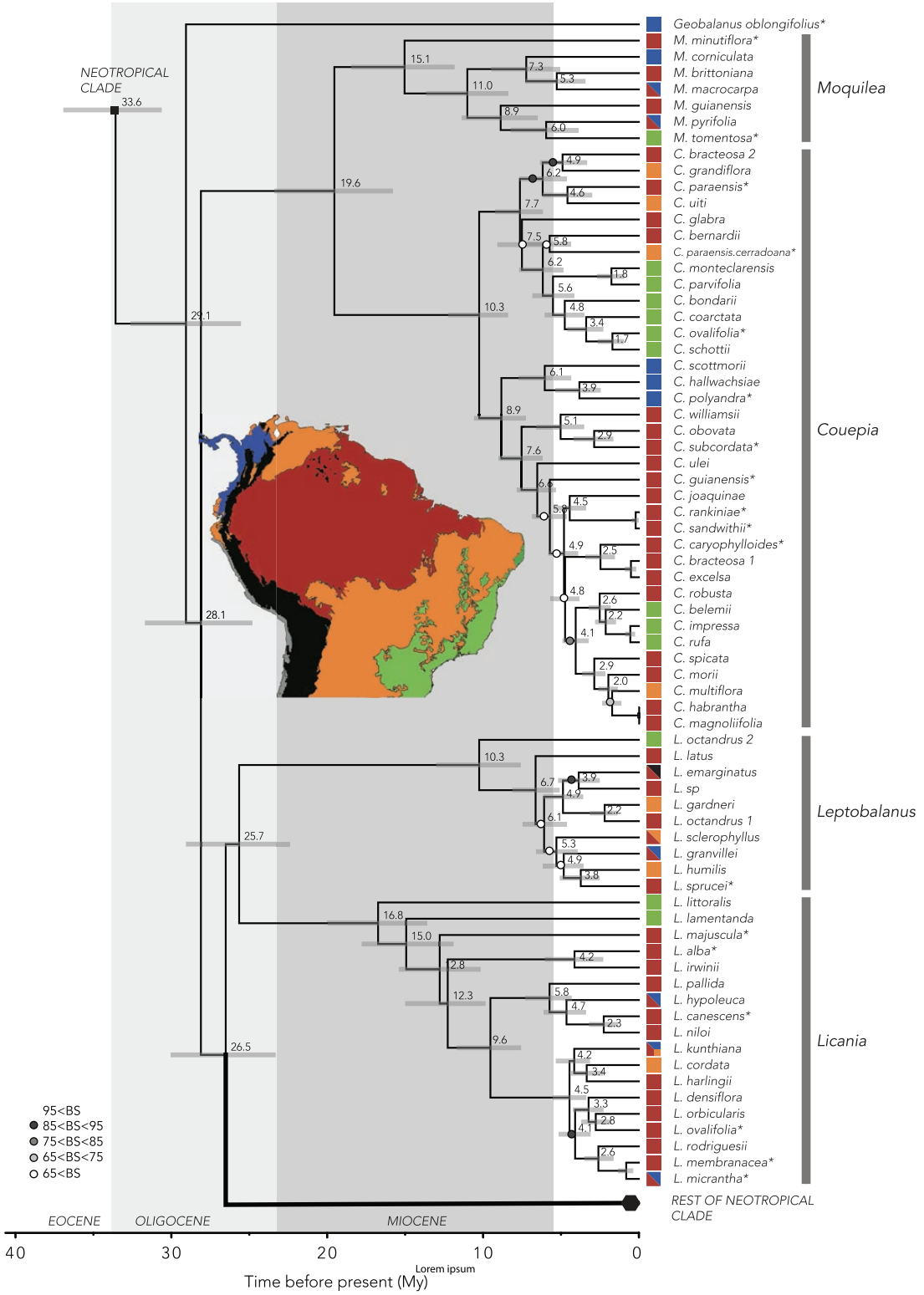

Figure 2. Phylogenetic tree of Chrysobalanaceae, continuation of Figure 1. The biome map is modified from Olson *et al.* (2001). Dark red: Amazonia; orange: savannas and seasonally dry tropical forests; green: Atlantic tropical forests; blue: forests of Central America, the Caribbean and Chocó; black: high-elevation ecosystems (> 1000 m a.s.l.).

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

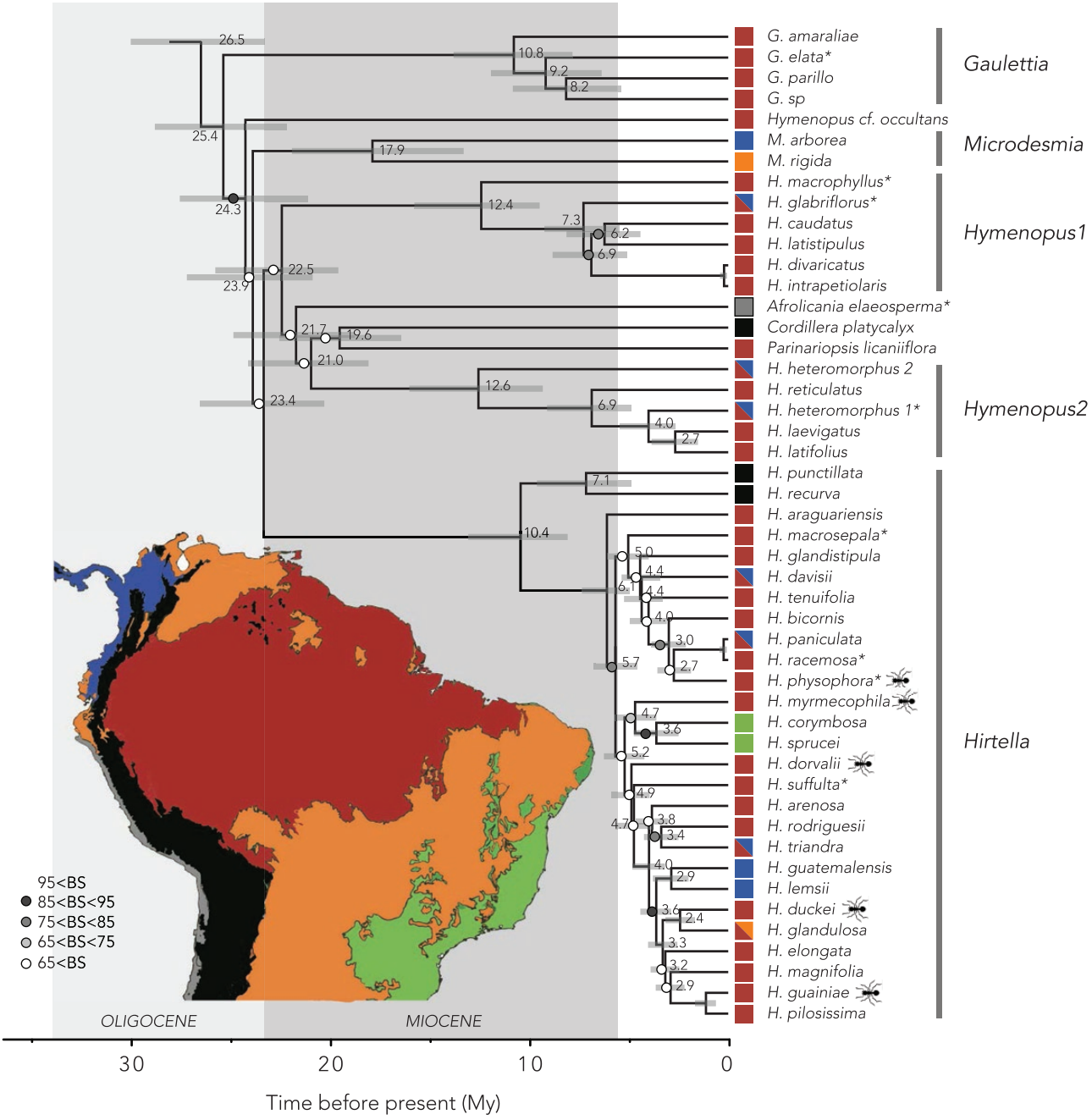

Figure 3. Phylogenetic tree of Chrysobalanaceae, continuation of Figure 2. In the *Hirtella* clade, the ant symbols indicate myrmecophilous species.

monophyletic. However, based on bootstrap support in this analysis, we cannot exclude the possibility that the clades of *Hymenopus* (Fig. 3) form a single group. The Neotropical genera in Figure 3 also diversified in the mid- to late Miocene: *Gaulettia* 10.8 Mya (7.8–13.8 Mya), *Hirtella* 10.4 Mya (8.1–13.1 Mya), *Hymenopus*1 12.6 Mya (9.3–16 Mya) and *Hymenopus*2 12.4 Mya (9.5–15.8 Mya). The position of *Afrolicania*, the only non-Neotropical species of the core Neotropical clade, suggests a single dispersal event from the Neotropics to Africa 24 Mya (95% HPD: 22.4–25.6 Mya).

The BioGeoBEARS analysis detected that the combined Neotropical clade and clade D were unambiguously assigned to Amazonia, with secondary dispersal events into Central America, the Atlantic forest and savannas/dry tropical forests (Fig. 4). Migration events to the cerrado were mainly in the Pliocene, confirming Simon *et al.* (2009).

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

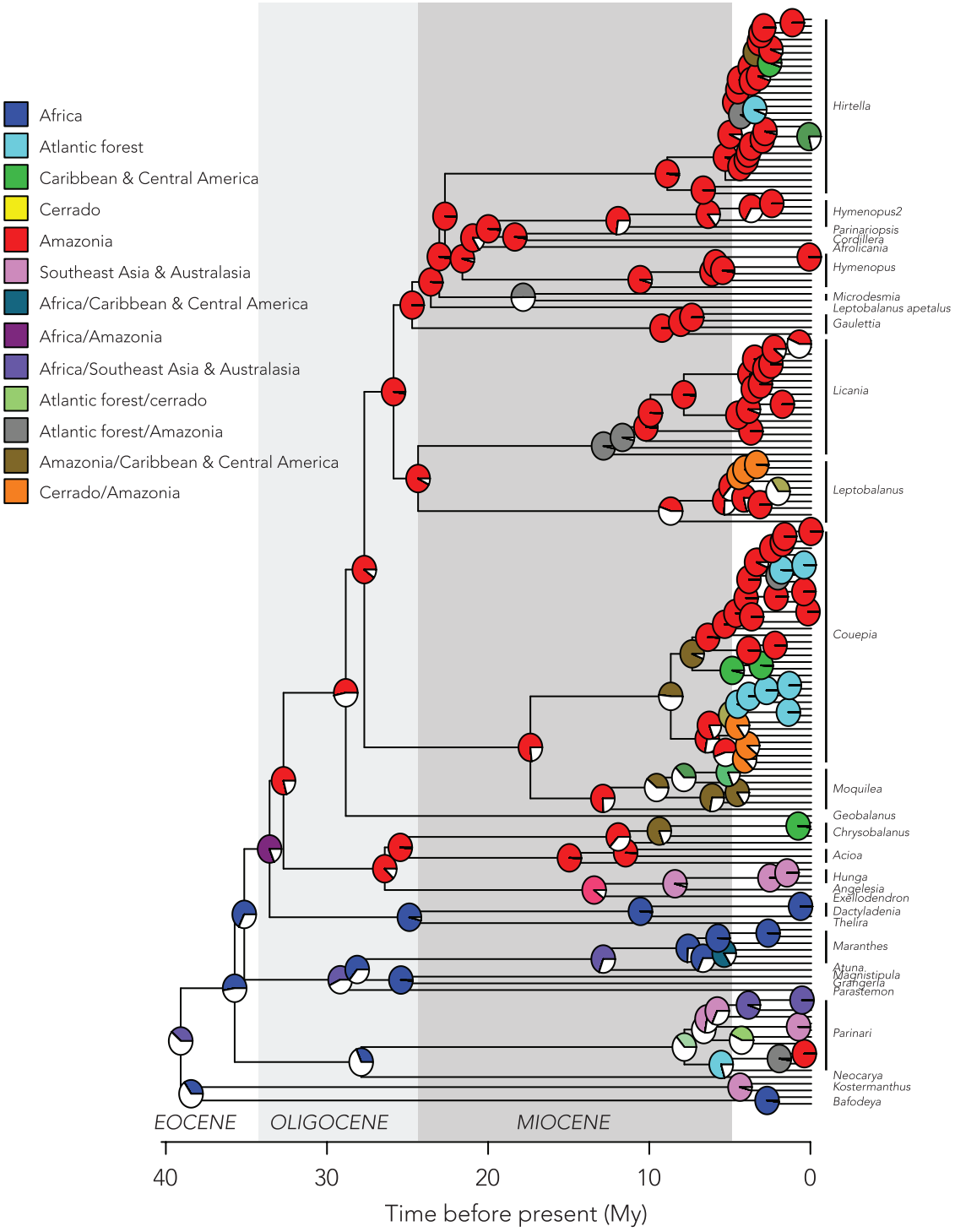

Figure 4. Ancestral area reconstruction of Chrysobalanaceae using the dispersal, extinction and cladogenesis (DEC) algorithm as implemented in the BioGeoBears software. Pie charts indicate relative support for the dominant ancestral area; all other ancestral area probabilities are lumped and represented in white.

The BAMM analysis converged (effect size for number of shifts was > 1000 with a log-likelihood > 400). It identified four shifts in diversification rates as the most likely (Fig. 5). A rate-through time analysis for Chrysobalanaceae demonstrated a clear increase in speciation rates after 10 Mya (Fig. 5). The four outlying

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

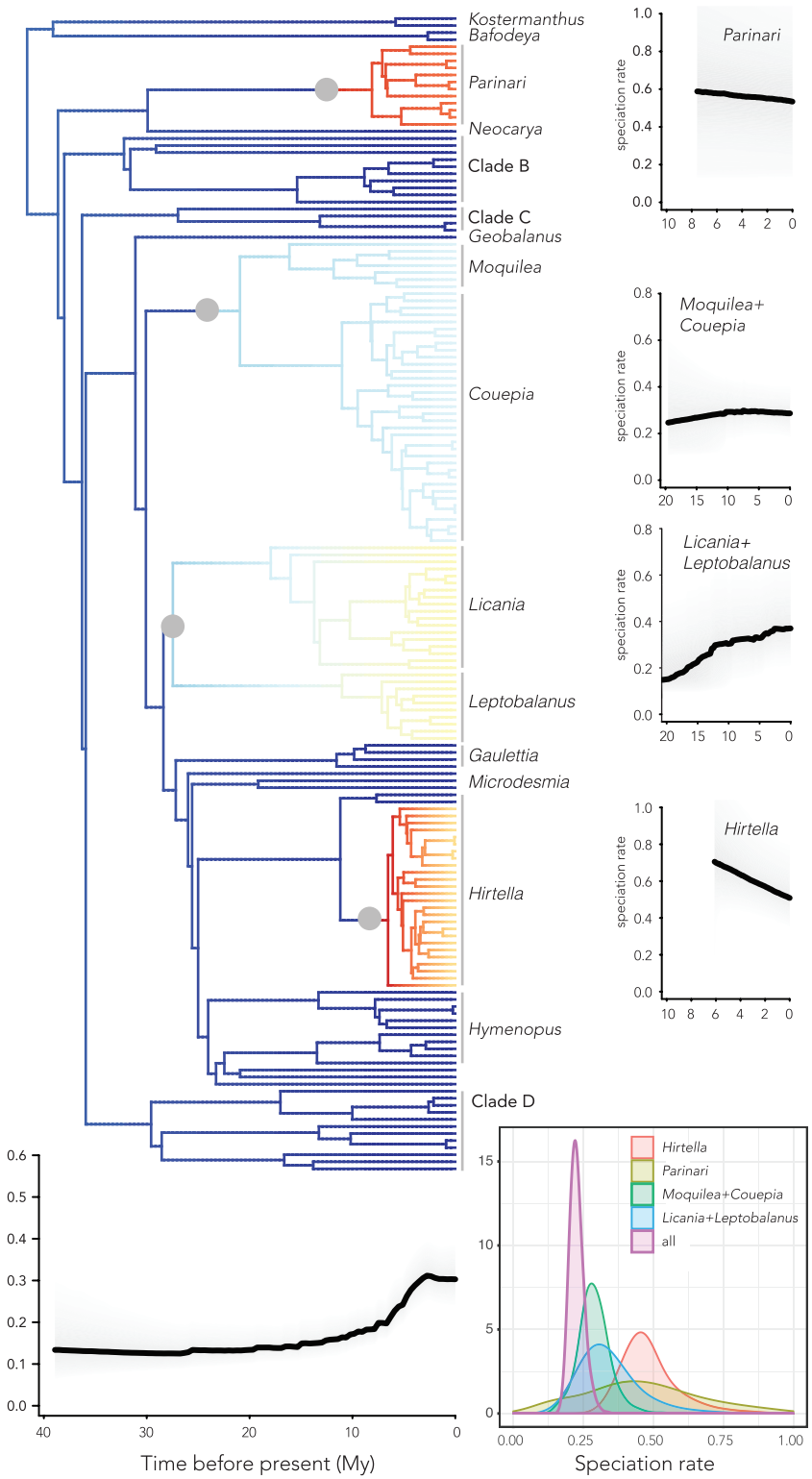

Figure 5. Plot with branches coloured by speciation rate (lineages/Myr), representing a summary of BAMM analysis. Grey circles indicate the positions of regime shifts in the best configuration. Side plots represent the speciation rate-through-time (RTT) plots for each of the four clades with regime shifts, whereas the bottom plot represents the RTT plot for the entire tree. Finally, the bottom-right histograms indicate the distribution of speciation rates for the four regime-shift clades and the entire family.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

clades correspond to *Parinari*, *Moquilea+Couepia*, *Licania+Leptobalanus* and *Hirtella* [excluding *H. recurva* (Spruce ex Prance) Sothers & Prance and *H. punctillata* Ducke]. The most rapid diversification was in *Hirtella* minus *H. punctillata* and *H. recurva*, for which BAMM inferred a speciation rate of λ = 0.60 (in lineages per Myr, 90% confidence interval: 0.42–0.85), compared with a family mean-speciation rate of λ = 0.23 (0.19–0.27), and family mean-speciation rate excluding *Hirtella* of λ = 0.20 (0.17–0.25). A rate-through time analysis for *Hirtella* demonstrates a decline in rate through time (Fig. 5, right panels). *Parinari* was the other clade that exhibited a significantly higher speciation rate compared to the average but with much greater confidence intervals, λ = 0.49 (0.14–0.96).

DISCUSSION

We provide a comprehensive analysis of the evolutionary history of Chrysobalanaceae based on the analysis of 163 fully sequenced plastid genomes, including about a third of the species currently recognized in the family. As discussed in the following, our results provide new support for the Miocene origin of the Neotropical flora, and to our knowledge it is the first to be built on a fully sampled matrix of plastid genome data for a Neotropical tree family. Our study was based on an aligned length of 162 204 bp, far more than recently published studies on Neotropical plant diversification.

In contrast, virtually all existing evolutionary papers of Neotropical plant families have been based on selected plastid markers combined with sequences from the ribosomal cluster (internal transcribed spacer). Important recent studies on the Miocene diversification of Amazonian clades have focused on Annonaceae based on an aligned length of 7960 bp (Pirie *et al.*, 2018), Meliaceae with 5207 bp (Koenen *et al.*, 2015) and Detarioideae (Fabaceae) with 2463 bp (Schley *et al.*, 2018). In our study, the most poorly sampled species had a 93% plastome coverage. Generally, coverage is > 99%, and the matrix was almost complete. We also included 156 species, 131 of which are Neotropical, with much increased taxon and character sampling compared to previous efforts. The results include important new biogeographic and systematic results about Chrysobalanaceae and confirm results for the Neotropical flora, but with much greater confidence than the other studies due to the greater amounts of data included.

Early diversification of Chrysobalanaceae

Bafodeya plus *Kostermanthus* were found to be sister to the rest of the family. Both results were unexpected based on floral morphology and previous molecular results. *Bafodeya* was placed in the *Parinari-Neocarya* clade, whereas *Kostermanthus* was proposed to have a relationship to *Atuna* Raf. or Neotropical *Acioa* Aubl. and African *Dactyladenia* Welw., although none of these relationships was well supported (Yakandawala, Morton & Prance, 2010). *Kostermanthus* occurs in Southeast Asian rainforests in mixed dipterocarp and heath associations. Monospecific *Bafodeya* is endemic to mid-elevation sandstone plateaus of West Africa. *Euphronia* Mart & Zucc., the sole genus of Euphroniaceae (sister family to Chrysobalanaceae; Xi *et al.*, 2012), is endemic to the Guiana Shield and restricted to white sand or rocky areas. Thus, *Kostermanthus* and *Bafodeya* share ecological affinities with *Euphronia*. If the phylogenetic structure proposed here is confirmed, this suggests that the ancestral habitat of Chrysobalanaceae was nutrient-poor and sandier than modern tropical rainforests. It would be important to reassess the position of both genera, and this is a good example where nuclear gene data would be helpful.

The crown age of the *Parinari* clade was inferred at 9 Myr (95% HPD: 8.2–9.8 Myr). *Parinari* possesses the most reliable fossil record of the family, due to the diagnostic features of its endocarp (Jud *et al.*, 2016). The early Miocene *Parinari* fossils in Africa (Tiffney *et al.*, 1994) and South America (Jud *et al.*, 2016) pre-date by *c.* 10 Myr the crown age of *Parinari*, consistent with their high rates of speciation and extinction. Bardon *et al.* (2016) supported an African origin for *Parinari* due to the native African distribution of *Neocarya*, but the increased sampling of our study leads to a less clear-cut result. The two earliest-diverging clades in *Parinari* contain all five Neotropical accessions. One clade with a crown age 7.4 Myr contains all Palaeotropical accessions. Noteworthy is the position of *P. nonda* F.Muell. ex Benth., from tropical Australia and Papua New Guinea, close to African *P. capensis* Harv. (inferred age of 0.37 Myr for the *P. capensis*/*P. nonda* split), which suggests a recent long-distance dispersal event from Africa to Australasia. Overall, if more research on *Parinari* confirms the crown age of < 10 Myr, it would be a striking case of a pantropically distributed genus of long-lived tropical trees with a trans-oceanic dispersal (Renner, 2004). On the whole, our species sampling is currently insufficient to confidently resolve the biogeographical history of *Parinari* because our results are based on sampling of only 11 species of the 39 currently described *Parinari* spp.

Clade B has a strong African component. Increased sampling for *Maranthes* (five of the 12 species now included) produced a date of 14.4 Mya (95% HPD: 10.9–18.1 Mya). *Maranthes* is present on all three continents, including the Neotropical species: *M. panamensis* (Standl.) Prance & F.White, which

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

our analysis places as sister to *M. gabunensis* (Engl.) Prance. This suggests that *M. panamensis* is the product of a recent dispersal from Africa. We also emphasize that *Magnistipula* Engl. remains an unresolved puzzle in Chrysobalanaceae and further research should include more comprehensive taxon coverage including all three subgenera. Clade C also contains only African species: *Dactyladenia* plus the African *Hirtella* spp. (Prance & White, 1988: 149), which need to be revised and segregated from *Hirtella* based on our results.

Clade D contains 28 species (14 Neotropical) and five genera. *Hunga* (endemic to New Caledonia) and *Angelesia* (more broadly Australasian) are sister to Neotropical *Exellodendron* (one species out of five sampled here). Thus, the *Angelesia-Hunga* clade probably results from a long-distance dispersal event from the Neotropics to Australasia between 9.3 and 15.9 Mya. Cases of dispersal from South America to Australasia via Antarctica probably occurred before the cooling event of the mid-Oligocene, *c.* 30 Mya (Siegert, 2008), but our dating seems to reject this scenario, pointing instead to a much more recent dispersal.

In clade D, *Chrysobalanus icaco* L. also has a well-documented amphi-Atlantic distribution. We include for the first time all three species of *Chrysobalanus*, *C. icaco*, *C. cuspidatus* Griseb. and *C. prancei* I.M.Turner (formerly *C. venezuelanus* Prance), which were divergent based on plastid genome information, with an early divergence in the mid-Miocene at *c.* 12.4 Mya (95% HPD: 9.1–15.7 Mya). It would be important to further explore the divergence of the African populations of *C. icaco*, which includes two subspecies. The placement of *Acioa* sister to *Chrysobalanus* differs from that in Bardon *et al.* (2016), and increased taxon sampling proved important to further resolve this clade.

We emphasize that our dated phylogenetic tree is based on limited fossil material and discovery of new fossils could alter these dates. The crown of Chrysobalanaceae was dated at *c.* 38.9 Mya (95% HPD: 34.2–43.9 Mya), slightly older than a previous estimate of 33 Mya (Bardon *et al.*, 2016). This discrepancy is easily explained because this analysis is based on a better sampling of the early-diverging clade in Chrysobalanaceae, and we used flat priors rather than log-normal ones (Condamine *et al.*, 2015).

Structure of the core Neotropical clade

The main focus of this study was to better resolve the evolutionary history of Chrysobalanaceae in the Neotropics. Prior to 2014, the 'core Neotropical clade' (> 99% Neotropical), included only three genera, but no fewer than 395 species (Prance, 1972; Prance & Sothers, 2003). After redefinition of *Couepia* (Sothers *et al.*, 2014) and *Licania* (Sothers *et al.*, 2016), the structure of the core Neotropical clade has been considerably clarified.

Here we recognize 12 genera of Chrysobalanaceae as members of the core Neotropical clade: Neotropical *Hirtella* (105 species), *Geobalanus* Small (three species), *Microdesmia* (two species), *Cordillera* (one species), *Parinariopsis* (one species), *Moquilea* (54 species), *Couepia* (62 species), *Leptobalanus* (31 species), *Licania* (100 species), *Gaulettia* (nine species) and *Hymenopus* (28 species) and *Afrolicania* (one species), the only non-Neotropical genus in this clade. In addition, *Exellodendron* (five species) and *Acioa* (six species) are exclusively Neotropical but outside the core clade.

The stem age of the core Neotropical clade is estimated in the mid-Eocene, and the crown age is in the early Oligocene *c.* 29.1 Mya (25.5–32.6 Mya). According to the biogeographical analysis, this core Neotropical clade diversified in Amazonia. The various non-Amazonian clades (notably *Couepia* in the Atlantic rainforest and beyond the Andes into the Chocó and Central America) are interpreted as dispersal events. This scenario confirms the more general analysis of Antonelli *et al.* (2018), but also shows that it is difficult to attribute extant diversity to a single biome: in the case of Chrysobalanaceae, diversification in Amazonia was preceded by a long extra-Neotropical evolutionary history, even potentially outside the forest biome.

We now turn to the sister clade of *Gaulettia*, which contains some of the unresolved taxonomic issues in Chrysobalanaceae. This clade, dated at 26.3 Myr (95% HPD: 24.7–27.9 Myr), includes *Hirtella* plus seven clades with low bootstrap support for their inter-relationships: *Microdesmia*, *Hymenopus*1, *Afrolicania*, *Cordillera*, *Parinariopsis*, *Hymenopus*2 and *Hymenopus* cf. *occultans*. Our analysis suggests that *Hymenopus* may include two genera (designated as 1 and 2), but further research is needed to confirm this proposal, especially in the light of the weak support for their separation. One hypothesis for the lack of support in this clade is that it may have resulted from a single diversification event giving rise to a variety of new forms associated with the end of the Oligocene. Global warming at this time was associated with a loss of palynofloral diversity in the foothills of the Andes (Jaramillo, Rueda & Mora, 2006), which could be due to the rapid Andean uplift around this time (Hoorn *et al.*, 2010). The current distribution of these genera sheds little light on a possible allopatric diversification scenario, in part because the extant distribution due to cultivation of *Microdesmia* species may not reflect their historical distribution (Sothers *et al.*, 2016), and also because extant species of both *Hymenopus* and *Hirtella* have large modern distributions. Bardon *et al.* (2016)

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

was published before the generic realignments illustrated here. This tree also differs from that in Sothers *et al.* (2016) based on plastid and nuclear (*Xdh*, ITS) DNA sequences.

Evolutionary history of Neotropical genera of Chrysobalanaceae

According to our results, onset of diversification for seven genera spanned the mid-Miocene: *Moquilea* (15.1 Mya), *Couepia* (10.3 Mya), *Leptobalanus* (10.3 Mya), *Licania* (16.8 Mya), *Hirtella ss.* (10.4 Mya), *Gaulettia* (10.8 Mya), *Hymenopus*1 (12.4 Mya) and *Hymenopus*2 (12.6 Mya). This precedes the initiation of the modern Amazon River and demise of the Pebas wetland (Figueiredo *et al.*, 2009).

Against the backdrop of this geological and climatic setting, the evolutionary history of *Couepia* is informative. The genus clearly split into two groups, one with an affinity for dry forests (*paraensis* clade) with a recent (early Pliocene) unique dispersal to the Atlantic rainforest. In the other (*guianensis* clade) there was an early-diverging clade of Central American/Chocó species sister to the rest, which are predominantly Amazonian. From the ancestral area reconstruction analysis, we were unable to assign this clade to a specific region, although the most likely area is Amazonia+Central America/Chocó. Finally, *Couepia* spp. currently found in the Atlantic rainforest of Brazil seem to have resulted from two independent dispersals, both post-Miocene. *Moquilea*, *Licania*, and *Hymenopus* have species in both Amazonia and Central America. However, for these, it would be important to better sample the populations on both sides of the Andes to ensure that these are not divergent 'cryptic' species. If our result is confirmed, cross-Andean dispersals have occurred frequently since the Pleistocene.

Sister to the rest of *Hirtella* is a clade of two species including *H. recurva* (Spruce ex Prance) Sothers & Prance (Sothers *et al.*, 2014), found at > 2000 m in the Ecuadorian Andes, and *H. punctillata* Ducke, collected at > 1000 m in the Serra do Aracá tepui (Prance & Johnson, 1992). The fact that these two species cluster together, although distant and morphologically distinct, is unexpected. Aside from these two species, the remaining large clade of *Hirtella* did not diversify before the end of the Miocene at *c.* 6 Mya. Thus, *Hirtella* is an example of explosive diversification, with a speciation rate inferred *c.* 0.60 lineages per Myr (90% confidence interval: 0.42–0.85). Diversification in *Hirtella* thus has a comparable magnitude to that in *Inga* Miller, which is thought to have diversified *c.* 10 Mya, producing *c.* 300 species (Dexter *et al.*, 2017), and two genera of Meliaceae, *Trichilia* P.Browne and *Guarea* F.Allemão, as reported by Koenen *et al.* (2015). All four genera have their centre of diversity in Amazonia, and it is thus tempting to speculate that the timing of these events is consistent with a westward expansion of Amazonian forests after drainage of the Pebas wetland (Figueiredo *et al.*, 2009). However, it is also possible that ecological attributes of these groups may have played a role: these genera are predominantly understory plants, and their seeds are dispersed by animals (Baker *et al.*, 2014). Like *Inga*, *Hirtella* has a well-documented association with ants, and this could be a major factor in their diversification (Kursar *et al.*, 2009).

We failed to find support for a single myrmecophilous group in *Hirtella*, meaning that the myrmecophilous association has been repeatedly derived in the genus. However, an alternative interpretation is that incomplete lineage sorting is prevalent in this recent clade, and that plastid genomes are unable to uncover such shallow phylogenetic relationships. Greater species and regional sampling would be necessary to confirm relationships in this intriguing group.

This analysis included only a few species with multiple accessions, and some of these revealed surprises. Two accessions of *Parinari excelsa* Sabine (Parque Estadual Cristalino, Brazil, and Saint Laurent du Maroni, French Guiana) fell into separate clades, and so did the accessions of *Couepia bracteosa* (Sinnamary, French Guiana and Manaus, Brazil), *Leptobalanus octandrus* (Manaus and São Paulo, Brazil) and *Hymenopus heteromorphus* (Benth.) Sothers & Prance (Régina, French Guiana and Manaus). In all these cases, the accessions were from distant localities. One explanation may be that there are actually several cryptic, or previously unreported, species within the currently large ranges of these species. With more comprehensive sequencing of targeted species across their distribution, it will be possible to assess the prevalence of such entities in Chrysobalanaceae. This situation probably holds more generally across Amazonian plant families (Misiewicz & Fine, 2014; Loiseau *et al.*, 2019).

On the use of plastid genomes to infer the evolutionary history of tropical flowering plants

High-throughput technologies have greatly facilitated the sequencing of plastid genomes and these have been used in plant phylogenomics for well over a decade (Jansen *et al.*, 2007; Moore *et al.*, 2007; Straub *et al.*, 2012). About 500 complete flowering plant plastomes had been sequenced by 2014 (Wicke & Schneeweiss, 2015), and there were close to 5000 fully sequenced plastomes representing > 1300 genera available on the NCBI website just five years later (July 2019).

Plastomes have been used to infer phylogenetic relationships in Poales (Givnish *et al.*, 2010),

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

Malpighiales (Xi *et al.*, 2012), Zingiberales (Barrett *et al.*, 2013), all angiosperms (Ruhfel *et al.*, 2014), Apocynaceae (Straub *et al.*, 2014), Rosaceae (Zhang *et al.*, 2017) and Caryophyllales (Yao *et al.*, 2019). In such analyses, proper curation of data (Philippe *et al.*, 2011) and appropriate phylogenetic reconstruction methods (Gonçalves *et al.*, 2019) have been crucial to ensure reliable results. Heterogeneity in evolutionary rates should be carefully considered, as it provides insights into modes of evolution (Ruhfel *et al.*, 2014).

Several mechanisms are known to impact the rate of evolution of plastomes. Groups with known symbiotic associations, such as mycoheterotrophy, show different evolutionary rates, due to gene silencing and loss (Wilke *et al.*, 2011). Also, plastid genomes turn out to be biparentally inherited in at least 20% of land plants (Zhang, 2010), suggesting the potential for recombination and therefore a more complex picture than often assumed for evolution of this compartment.

It is not known how often cytonuclear incongruence occurs in the tree of flowering plants, and previously found contradictions between plastid genome data and morphology may be solved using large nuclear gene datasets. For example, because of incomplete lineage sorting, recent and rapidly diversifying clades can be resolved only based on nuclear gene data, such as Andean *Espeletia* Mutis ex Humb. & Bompl. (Asteraceae; Pouchon *et al.*, 2018) and Australian *Nicotiana* L. (Solanaceae; Dodsworth *et al.*, 2020).

Targeted capture of hundreds of nuclear genes could bring even further insight into the question of plant diversification, as has been shown for the Neotropical palm clade Geonomateae (> 3 million bp; Loiseau *et al.*, 2019), Fabaceae (*c.* 1 million bp; Koenen *et al.*, 2020) or land plants (One Thousand Plant Transcriptomes Initiative, 2019). However, nuclear gene information did not radically transform the phylogenetic tree of Geonomateae (Roncal *et al.*, 2012), and plastid genome data were found to be consistent with nuclear gene data in Fabaceae except at the root node, the latter probably caused by incomplete lineage sorting (Koenen *et al.*, 2020). Also, assembling such large nuclear gene datasets represents specific challenges, and phylogenetic reconstruction methods using these data are still in development (Zhang *et al.*, 2018). Although nuclear genes are necessary to resolve parts of the plant tree of life where plastid genomes are insufficiently informative, many Amazonian plant families have not been included in such phylogenetic work, and plastome analyses are a natural step to document systematic relationships and study Amazonian plant diversification.

CONCLUSIONS

Chrysobalanaceae have long been promoted as a model for the study of Neotropical diversification, but unravelling their systematics has represented a major challenge (Prance, 1972; Prance & White, 1988; Yakandawala *et al.*, 2010). Previously, we have proposed a phylogenetic analysis of the family based on 51 species with fully sequenced plastid genomes and an additional 88 species sequenced for only a few markers (Bardon *et al.*, 2016). With a total of 163 sequenced plastomes in Chrysobalanaceae, and 156 species, the present study is a major update of this previous work and demonstrates that a more comprehensive strategy helps gain greater confidence on the monophyly of several genera, even if a few issues remain. In the future, it would be important to: explore whether the *Hymenopus* complex can be clarified, determine if nuclear DNA confirms the position of *Bafodeya* and *Kostermanthus* as sister to the rest of the family, examine the evolutionary history of Neotropical *Hirtella* and of *Parinari* with better species coverage, and add more *Magnistipula* spp. to include all three subgenera. For six lowland Amazonian genera of Chrysobalanaceae, we document accelerated diversification in the wake of the Andean uplift. This study thus provides support for the view that much of the extant Neotropical plant diversity has arisen within the past 15 Myr, Amazonian diversification has played a key role in this diversification process and the majority of diversification events have taken place *in situ*, rather than being the product of intercontinental dispersal.

ACKNOWLEDGEMENTS

We thank Fabien Condamine and Thomas Couvreur for commenting on an earlier manuscript. We gratefully acknowledge funding from 'Programme Investissement d'Avenir' managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; TULIP, ref. ANR-10-LABX-0041).

REFERENCES

Antonelli A, **Sanmartín I. 2011.** Why are there so many plant species in the Neotropics? *Taxon* **60:** 403–414.

Antonelli A, **Zizka A**, **Carvalho FA**, **Scharn R**, **Bacon CD**, **Silvestro D**, **Condamine FL. 2018.** Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences of the United States of America* **115:** 6034–6039.

Bacon CD, **Velásquez-Puentes FJ**, **Hoorn C**, **Antonelli A. 2018.** Iriarteeae palms tracked the uplift of Andean Cordilleras. *Journal of Biogeography* **45:** 1653–1663.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

Baker TR, **Pennington RT**, **Magallón S**, **Gloor E**, **Laurance WF**, **Alexiades M**, **Alvarez E**, **Araujo A**, **Arets EJMM**, **Aymard G**, **de Oliveira AA**, **Amaral I**, **Arroyo L**, **Bonal B**, **Brienen RJW**, **Chave J**, **Dexter KG**, **Di Fiore A**, **Eler E**, **Feldpausch TR**, **Ferreira L**, **Lopez-Gonzalez G**, **van der Heijden G**, **Higuchi N**, **Honorio E**, **Huamantupa I**, **Killeen TJ**, **Laurance S**, **Leaño C**, **Lewis SL**, **Malhi Y**, **Schwantes Marimon B**, **Marimon Junior BH**, **Monteagudo Mendoza A**, **Neill D**, **Peñuela-Mora MC**, **Pitman N**, **Prieto A**, **Quesada CA**, **Ramírez F**, **Ramírez Angulo H**, **Rudas A**, **Ruschel AR**, **Salomão RP**, **de Andrade AS**, **Silva JNM**, **Silveira M**, **Simon MF**, **Spironello W**, **ter Steege H**, **Terborgh J**, **Marisol Toledo M**, **Armando Torres-Lezama A**, **Rodolfo Vasquez R**, **Vieira ICG**, **Vilanova E**, **Vos VA**, **Phillips OL. 2014.** Fast demographic traits promote high diversification rates of Amazonian trees. *Ecology Letters* **17:** 527–536.

Baker WJ, **Couvreur TLP. 2013.** Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. II. Diversification history and origin of regional assemblages. *Journal of Biogeography* **40:** 286–298.

Bardon L, **Chamagne J**, **Dexter KG**, **Sothers CA**, **Prance GT**, **Chave J. 2012.** Origin and evolution of Chrysobalanaceae: insights into the evolution of plants in the Neotropics. *Botanical Journal of the Linnean Society* **171:** 19–37.

Bardon L, **Sothers C**, **Prance GT**, **Malé PJG**, **Xi Z**, **Davis CC**, **Murienne J**, **García-Villacorta R**, **Coissac E**, **Lavergne S**, **Chave J. 2016.** Unraveling the biogeographical history of Chrysobalanaceae from plastid genomes. *American Journal of Botany* **103:** 1089–1102.

Barrett CF, **Specht CD**, **Leebens-Mack J**, **Stevenson DW**, **Zomlefer WB**, **Davis JI. 2013.** Resolving ancient radiations: can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? *Annals of Botany* **113:** 119–133.

Berry EW. 1916. *The lower Eocene floras of southeastern North America*, *Vol. 91*. Washington: US Government Printing Office.

Boucher LD, **Manchester SR**, **Judd WS. 2003.** An extinct genus of Salicaceae based on twigs with attached flowers, fruits, and foliage from the Eocene Green River Formation of Utah and Colorado, USA. *American Journal of Botany* **90:** 1389–1399.

Bouckaert R, **Vaughan TG**, **Barido-Sottani J**, **Duchêne S**, **Fourment M**, **Gavryushkina, A**, **Heled J**, **Jones G**, **Kühnert D**, **De Maio N**, **Matschiner M**, **Mendes FK**, **Müller NF**, **Ogilvie HA**, **du Plessis L**, **Popinga A**, **Rambaut A**, **Rasmussen D**, **Siveroni I**, **Suchard MA**, **Wu C-H**, **Xie D**, **Zhang C**, **Stadler T**, **Drummond AJ. 2019.** BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **15:** e1006650.

Chambers KL, **Poinar GO Jr. 2010.** The Dominican amber fossil *Lasiambix* (Fabaceae: Caesalpiniodeae?) is a *Licania* (Chrysobalanaceae). *Journal of the Botanical Research Institute of Texas* **4:** 217–218.

Condamine FL, **Nagalingum NS**, **Marshall CR**, **Morlon H. 2015.** Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evolutionary Biology* **15:** 65.

Coyne JA, **Orr HA. 2004.** *Speciation*. Sunderland: Sinauer.

Crepet WL, **Nixon KC. 1998.** Fossil Clusiaceae from the Late Cretaceous (Turonian) of New Jersey and implications regarding the history of bee pollination. *American Journal of Botany* **85:** 1122–1133.

Darriba D, **Taboada GL**, **Doallo R**, **Posada D. 2012.** jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Davis CC, **Webb CO**, **Wurdack KJ**, **Jaramillo CA**, **Donoghue MJ. 2005.** Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. *The American Naturalist* **165:** E36–E65.

Dexter KG, **Lavin M**, **Torke BM**, **Twyford AD**, **Kursar TA**, **Coley PD**, **Drake C**, **Hollands R**, **Pennington RT. 2017.** Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of the National Academy of Sciences of the United States of America* **114:** 2645–2650.

Dierckxsens N, **Mardulyn P**, **Smits G. 2016.** NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* **45:** e18.

Dodsworth S, **Christenhusz MJM**, **Conran JG**, **Guignard MS**, **Knapp S**, **Struebig M**, **Leitch AR**, **Chase MW. 2020.** Extensive plastid-nuclear discordance in a recent radiation of *Nicotiana* section *Suaveolentes* (Solanaceae). *Botanical Journal of the Linnean Society* **XX:** xx–yy. https://doi.org/10.1093/botlinnean/boaa024.

Donoghue MJ, **Edwards EJ. 2014.** Biome shifts and niche evolution in plants. *Annual Review of Ecology, Evolution, and Systematics* **45:** 547–572.

Donoghue MJ, **Sanderson MJ. 2015.** Confluence, synnovation, and depauperons in plant diversification. *New Phytologist* **207:** 260–274.

Drummond AJ, **Ho SY**, **Phillips MJ**, **Rambaut A. 2006.** Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4:** e88.

Eiserhardt WL, **Couvreur TL**, **Baker WJ. 2017.** Plant phylogeny as a window on the evolution of hyperdiversity in the tropical rainforest biome. *New Phytologist* **214:** 1408–1422.

Fernandes F, **Pereira L**, **Freitas AT. 2009.** CSA: an efficient algorithm to improve circular DNA multiple alignment. *BMC Bioinformatics* **10:** 230.

Figueiredo J, **Hoorn C**, **van der Ven P**, **Soares E. 2009.** Late Miocene onset of the Amazon River and the Amazon deep-sea fan: evidence from the Foz do Amazonas Basin. *Geology* **37:** 619–622.

Fine PVA, **García-Villacorta R**, **Pitman NC**, **Mesones I**, **Kembel SW. 2010.** A floristic study of the white-sand forests of Peru. *Annals of the Missouri Botanical Garden* **97:** 283–305.

Fine PVA, **Zapata F**, **Daly DC. 2014.** Investigating processes of Neotropical rain forest tree diversification by examining the evolution and historical biogeography of the Protieae (Burseraceae). *Evolution* **68:** 1988–2004.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

Gentry AH. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69:** 557–593.

Germeraad JH, **Hopping CA**, **Muller J. 1968.** Palynology of Tertiary sediments from tropical areas. *Review of Palaeobotany and Palynology* **6:** 189–348.

Givnish TJ, **Ames M**, **McNeal JR**, **McKain MR**, **Steele PR**, **de Pamphilis CW**, **Graham SW**, **Pires JC**, **Stevenson DW**, **Zomlefer WB**, **Briggs BG**, **Devall MR**, **Moore MJ**, **Heaney JM**, **Soltis DE**, **Soltis PS**, **Thiele K**, **Leebens-Mack JH. 2010.** Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales. *Annals of the Missouri Botanical Garden* **97:** 584–617.

Gonçalves DJ, **Simpson BB**, **Ortiz EM**, **Shimizu GH**, **Jansen RK. 2019.** Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Molecular Phylogenetics and Evolution* **138:** 219–232.

Gregory-Wodzicki KM. 2000. Uplift history of the central and northern Andes: a review. *Geological Society of America Bulletin* **112:** 1091–1105.

Haffer J. 2008. Hypotheses to explain the origin of species in Amazonia. *Brazilian Journal of Biology* **68:** 917–947.

Herrera F, **Manchester SR**, **Jaramillo C. 2012.** Permineralized fruits from the late Eocene of Panama give clues of the composition of forests established early in the uplift of Central America. *Review of Palaeobotany and Palynology* **175:** 10–24.

Hoorn C, **Bogotá-A GR**, **Romero-Baez M**, **Lammertsma EI**, **Flantua SG**, **Dantas EL**, **Dino R**, **do Carmo DA**, **Chemale F Jr**. **2017.** The Amazon at sea: onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. *Global and Planetary Change* **153:** 51–65.

Hoorn C, **Wesselingh FP**, **ter Steege H**, **Bermudez MA**, **Mora A**, **Sevink J**, **Sanmartín I**, **Sanchez-Meseguer A**, **Anderson CL**, **Figueiredo JP**, **Jaramillo C**, **Riff D**, **Negri FR**, **Hooghiemstra H**, **Lundberg J**, **Stadler T**, **Särkinen T**, **Antonelli A. 2010.** Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330:** 927–931.

Hughes CE, **Pennington RT**, **Antonelli A. 2012.** Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society* **171:** 1–18.

Iturralde-Vinent MA. 2001. Geology of the amber-bearing deposits of the Greater Antilles. *Caribbean Journal of Science* **37:** 141–167.

Iturralde-Vinent MA, **MacPhee RDE. 1996.** Age and paleogeographical origin of Dominican amber. *Science* **273:** 1850–1852.

Jansen RK, **Cai Z**, **Raubeson LA**, **Daniell H**, **dePamphilis CW**, **Leebens-Mack J**, **Müller KF**, **Guisinger-Bellian M**, **Haberle RC**, **Hansen AK**, **Chumley TW**, **Lee SB**, **Peery R**, **McNeal JR**, **Kuehl JV**, **Boore JL. 2007.** Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 19369–19374.

Jaramillo C, **Romero I**, **D'Apolito C**, **Bayona G**, **Duarte E**, **Louwye S**, **Louwye S**, **Escobar J**, **Luque J**, **Carrillo-Briceño JD**, **Zapata V**, **More A**, **Schouten S**, **Zavada M**, **Harrington G**, **Ortiz J**, **Wesselingh FP. 2017.** Miocene flooding events of western Amazonia. *Science Advances* **3:** e1601693.

Jaramillo C, **Rueda MJ**, **Mora G. 2006.** Cenozoic plant diversity in the Neotropics. *Science* **311:** 1893–1896.

Jud NA, **Nelson CW**, **Herrera F. 2016.** Fruits and wood of *Parinari* from the early Miocene of Panama and the fossil record of Chrysobalanaceae. *American Journal of Botany* **103:** 277–289.

Katoh K, **Rozewicki J**, **Yamada KD. 2017.** MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* **20:** 1160–1166.

Koenen EJ, **Clarkson JJ**, **Pennington TD**, **Chatrou LW. 2015.** Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. *New Phytologist* **207:** 327–339.

Koenen EJ, **Ojeda DI**, **Steeves R**, **Migliore J**, **Bakker FT**, **Wieringa JJ**, **Kidner C**, **Hardy OJ**, **Pennington RT**, **Hughes CE. 2020.** Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies. *New Phytologist* **225:** 1355–1369.

Kursar TA, **Dexter KG**, **Lokvam J**, **Pennington RT**, **Richardson JE**, **Weber MG**, **Murakami ET**, **Drake C**, **McGregor R**, **Coley PD. 2009.** The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 18073–18078.

Lagomarsino LP, **Condamine FL**, **Antonelli A**, **Mulch A**, **Davis CC. 2016.** The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist* **210:** 1430–1442.

Li HT, **Yi TS**, **Gao LM**, **Ma PF**, **Zhang T**, **Yang JB**, **Gitzendanner MA**, **Fritsch PW**, **Cai J**, **Luo Y**, **Wang H**, **van der Bank M**, **Zhang SD**, **Wang QF**, **Wang J**, **Zhang ZR**, **Fu CN**, **Yang J**, **Hollingsworth PM**, **Chase MW**, **Soltis DE**, **Soltis PS**, **Li DZ. 2019.** Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* **5:** 461.

Loiseau O, **Olivares I**, **Paris M**, **de La Harpe M**, **Weigand A**, **Koubinová D**, **Rolland J**, **Bacon CD**, **Balslev H**, **Borchenius F**, **Cano A**, **Couvreur TLP**, **Delnatte C**, **Fardin F**, **Gayot M**, **Mejía F**, **Mota-Machado T**, **Perret M**, **Roncal J**, **Sanin MJ**, **Stauffer F**, **Lexer, C**, **Kessler M**, **Salamin N. 2019.** Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. *Frontiers in Plant Science* **10:** 864.

Magallón S, **Gómez-Acevedo S**, **Sánchez-Reyes LL**, **Hernández-Hernández T. 2015.** A metacalibrated

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* **207:** 437–453.

Malé PJG, **Bardon L**, **Besnard G**, **Coissac E**, **Delsuc F**, **Engel J**, **Lhuillier E**, **Scotti-Saintagne C**, **Tinaut A**, **Chave J. 2014.** Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. *Molecular Ecology Resources* **14:** 966–975.

Matzke NJ. 2012. Founder-event speciation in BioGeoBEARS package dramatically improves likelihoods and alters parameter inference in dispersal-extinction-cladogenesis (DEC) analyses. *Frontiers of Biogeography* **4:** 210.

Matzke NJ. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* **63:** 951–970.

Miller MA, **Pfeiffer W**, **Schwartz T. 2010.** Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments Workshop (GCE), IEEE*, 1–8.

Misiewicz TM, **Fine PV. 2014.** Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). *Molecular Ecology* **23:** 2543–2558.

Moore MJ, **Bell CD**, **Soltis PS**, **Soltis DE. 2007.** Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 19363–19368.

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

Olson DM, **Dinerstein E**, **Wikramanayake ED**, **Burgess ND**, **Powell GV**, **Underwood EC**, **d'Amico J**, **Itoua I**, **Strand HE**, **Morrison JC**, **Loucks CJ**, **Allnutt TF**, **Ricketts TH**, **Kura Y**, **Lamoreux JF**, **Wettengel WW**, **Hedao P**, **Kassem KR. 2001.** Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience* **51:** 933–938.

Paradis E, **Claude J**, **Strimmer K. 2004.** APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Parham JF, **Donoghue PC**, **Bell CJ**, **Calway TD**, **Head JJ**, **Holroyd PA**, **Inoue JG**, **Irmis RB**, **Joyce WG**, **Ksepka DT**, **Patané JSL**, **Smith ND**, **Tarver JE**, **van Tuinen M**, **Yang Z**, **Angielczyk KD**, **Greenwood JM**, **Hipsley CA**, **Jacobs L**, **Makovicky PJ**, **Müller J**, **Smith KT**, **Teodor JM**, **Warnock RCM**, **Benton MJ. 2011.** Best practices for justifying fossil calibrations. *Systematic Biology* **61:** 346–359.

Pennington RT, **Dick CW. 2004.** The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions of the Royal Society of London, B* **359:** 1611–1622.

Pérez-Escobar OA, **Chomicki G**, **Condamine FL**, **Karremans AP**, **Bogarín D**, **Matzke NJ**, **Silvestro D**, **Antonelli A. 2017.** Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist* **215:** 891–905.

Philippe H, **Brinkmann H**, **Lavrov DV**, **Littlewood DTJ**, **Manuel M**, **Wörheide G**, **Baurain D. 2011.** Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biology* **9:** e1000602.

Pirie MD, **Maas PJ**, **Wilschut RA**, **Melchers-Sharrott H**, **Chatrou LW. 2018.** Parallel diversifications of *Cremastosperma* and *Mosannona* (Annonaceae), tropical rainforest trees tracking Neogene upheaval of South America. *Royal Society Open Science* **5:** 171561.

Pouchon C, **Fernández A**, **Nassar JM**, **Boyer F**, **Aubert S**, **Lavergne S**, **Mavárez J. 2018.** Phylogenomic analysis of the explosive adaptive radiation of the *Espeletia* complex (Asteraceae) in the tropical Andes. *Systematic Biology* **67:** 1041–1060.

Prance GT. 1972. Chrysobalanaceae. *Flora Neotropica* **9:** 1–409.

Prance GT. 1982. A review of the phytogeographic evidences for Pleistocene climate changes in the Neotropics. *Annals of the Missouri Botanical Garden* **69:** 594–624.

Prance GT, **Johnson DM. 1992.** Plant collections from the plateau of Serra do Aracá (Amazonas, Brazil) and their phytogeographic affinities. *Kew Bulletin* **47:** 1–24.

Prance GT, **Sothers CA. 2003.** *Part 10, Chrysobalanaceae. Species plantarum flora of the World*. Canberra: Australian Biological Resources Study.

Prance GT, **White F. 1988.** The genera of Chrysobalanaceae: a study in practical and theoretical taxonomy and its relevance to evolutionary biology. *Philosophical Transactions of the Royal Society of London, B* **320:** 1–184.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity dependence on phylogenetic trees. *PLoS One* **9:** e89543.

Rabosky DL, **Grundler M**, **Anderson C**, **Title P**, **Shi JJ**, **Brown JW**, **Huang H**, **Larson JG. 2014.** BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* **5:** 701–707.

Rabosky DL, **Santini F**, **Eastman JM**, **Smith SA**, **Sidlauskas B**, **Chang J**, **Alfaro ME. 2013.** Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* **4:** 1958.

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

Raven PH, **Axelrod DI. 1974.** Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* **61:** 539–673.

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. *Journal of Biogeography* **45:** 741–749.

Ree RH, **Smith SA. 2008.** Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57:** 4–14.

Renner S. 2004. Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* **165:** S23–S33.

Roncal J, **Henderson A**, **Borchsenius F**, **Cardoso SRS**, **Balslev H. 2012.** Can phylogenetic signal, character

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

displacement, or random phenotypic drift explain the morphological variation in the genus *Geonoma* (Arecaceae)? *Biological Journal of the Linnean Society* **106:** 528–539.

Roncal J, **Kahn F**, **Millan B**, **Couvreur TL**, **Pintaud JC. 2013.** Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). *Botanical Journal of the Linnean Society* **171:** 120–139.

Ruhfel BR, **Gitzendanner MA**, **Soltis PS**, **Soltis DE**, **Burleigh JG. 2014.** From algae to angiosperms – inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evolutionary Biology* **14:** 23.

Sacek V. 2014. Drainage reversal of the Amazon River due to the coupling of surface and lithospheric processes. *Earth and Planetary Science Letters* **401:** 301–312.

Schley RJ, **de la Estrella M**, **Pérez-Escobar OA**, **Bruneau A**, **Barraclough T**, **Forest F**, **Klitgård B. 2018.** Is Amazonia a 'museum' for Neotropical trees? The evolution of the *Brownea* clade (Detarioideae, Leguminosae). *Molecular Phylogenetics and Evolution* **126:** 279–292.

Siegert MJ. 2008. Antarctic subglacial topography and ice-sheet evolution. *Earth Surface Processes and Landforms* **33:** 646–660.

Simon MF, **Grether R**, **de Queiroz LP**, **Skema C**, **Pennington RT**, **Hughes CE. 2009.** Recent assembly of the cerrado, a Neotropical plant diversity hotspot, by *in situ* evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 20359–20364.

Smith BT, **McCormack JE**, **Cuervo AM**, **Hickerson MJ**, **Aleixo A**, **Cadena CD**, **Pérez-Emán J**, **Burney CW**, **Xie X**, **Harvey MG**, **Faircloth BC**, **Glenn TC**, **Derryberry EP**, **Prejean J**, **Fields S**, **Brumfield RT. 2014.** The drivers of tropical speciation. *Nature* **515:** 406–409.

Sothers C, **Prance GT**, **Buerki S**, **De Kok R**, **Chase MW. 2014.** Taxonomic novelties in Neotropical Chrysobalanaceae: towards a monophyletic *Couepia*. *Phytotaxa* **172:** 176–200.

Sothers CA, **Prance GT. 2014.** Resurrection of *Angelesia*, a Southeast Asian genus of Chrysobalanaceae. *Blumea* **59:** 103–105.

Sothers CA, **Prance GT**, **Chase MW. 2016.** Towards a monophyletic *Licania*: a new generic classification of the polyphyletic Neotropical genus *Licania* (Chrysobalanaceae). *Kew Bulletin* **71:** 58.

Srivastava R, **Awasthi N. 1996.** Fossil woods from the Neogene of Warkalli Coast and their palaeoecological significance. *Geophytology* **26:** 88–98.

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

ter Steege H, **Pitman NCA**, **Sabatier D**, **Baraloto C**, **Salomao RP**, **Guevara JE**, **Phillips OL**, **Castilho CV**, **Magnusson WE**, **Molino J-F**, **Monteagudo A**, **Nunez Vargas P**, **Montero JC**, **Feldpausch TR**, **Coronado ENH**, **Killeen TJ**, **Mostacedo B**, **Vasquez R**, **Assis RL**, **Terborgh J**, **Wittmann F**, **Andrade A**, **Laurance WF**, **Laurance SGW**, **Marimon BS**, **Marimon B-H**, **Guimaraes Vieira IC**, **Amaral IL**, **Brienen R**, **Castellanos H**, **Cardenas Lopez D**, **Duivenvoorden JF**, **Mogollon HF**, **Matos FDdA**, **Davila N**, **Garcia-Villacorta R**, **Stevenson Diaz PR**, **Costa F**, **Emilio T**, **Levis C**, **Schietti J**, **Souza P**, **Alonso A**, **Dallmeie F**, **Montoya AJD**, **Fernandez Piedade MT**, **Araujo-Murakami A**, **Arroyo L**, **Gribel R**, **Fine PVA**, **Peres CA**, **Toledo M**, **Aymard CGA**, **Baker TR**, **Ceron C**, **Engel J**, **Henkel TW**, **Maas P**, **Petronelli P**, **Stropp J**, **Zartman CE**, **Daly D**, **Neill D**, **Silveira M**, **Paredes MR**, **Chave J**, **Lima Filho DdA**, **Jorgensen PM**, **Fuentes A**, **Schongart J**, **Cornejo Valverde F**, **Di Fiore A**, **Jimenez EM**, **Penuela Mora MC**, **Phillips JF**, **Rivas G**, **van Andel TR**, **von Hildebrand P**, **Hoffman B**, **Zent EL**, **Malhi Y**, **Prieto A**, **Rudas A**, **Ruschell AR**, **Silva N**, **Vos V**, **Zent S**, **Oliveira AA**, **Schutz AC**, **Gonzales T**, **Trindade Nascimento M**, **Ramirez-Angulo H**, **Sierra R**, **Tirado M**, **Umana Medina MN**, **van der Heijden G**, **Vela CIA**, **Vilanova Torre E**, **Vriesendorp C**, **Wang O**, **Young KR**, **Baider C**, **Balslev H**, **Ferreira C**, **Mesones I**, **Torres-Lezama A**, **Urrego Giraldo LE**, **Zagt R**, **Alexiades MN**, **Hernandez L**, **Huamantupa-Chuquimaco I**, **Milliken W**, **Palacios Cuenca W**, **Pauletto D**, **Valderrama Sandoval E**, **Valenzuela Gamarra L**, **Dexter KG**, **Feeley K**, **Lopez-Gonzalez G**, **Silman MR. 2013.** Hyperdominance in the Amazonian tree flora. *Science* **342:** 1243092–1243092. doi:10.1126/science.1243092

Straub SC, **Moore MJ**, **Soltis PS**, **Soltis DE**, **Liston A**, **Livshultz T. 2014.** Phylogenetic signal detection from an ancient rapid radiation: effects of noise reduction, long-branch attraction, and model selection in crown clade Apocynaceae. *Molecular Phylogenetics and Evolution* **80:** 169–185.

Straub SC, **Parks M**, **Weitemier K**, **Fishbein M**, **Cronn RC**, **Liston A. 2012.** Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* **99:** 349–364.

Tiffney BH, **Fleagle JG**, **Brown TM. 1994.** Early to Middle Miocene angiosperm fruits and seeds from Fejej, Ethiopia. *Tertiary Research* **15:** 25–42.

Wallace AR. 1878. *Tropical nature, and other essays*. London: Macmillan.

Wang X, **Edwards RL**, **Auler AS**, **Cheng H**, **Kong X**, **Wang Y**, **Cruz FW**, **Dorale JA**, **Chiang HW. 2017.** Hydroclimate changes across the Amazon lowlands over the past 45 000 years. *Nature* **541:** 204.

Wicke S, **Schneeweiss GM. 2015.** Next-generation organellar genomics: potentials and pitfalls of high-throughput technologies for molecular evolutionary studies and plant systematics. In: Horandl E, Appelhans MS, eds. *Next-generation sequencing in plant systematics*. Königstein: Koeltz.

Wodehouse RP. 1932. Tertiary pollen. I. Pollen of the living representatives of the Green River flora. *Bulletin of the Torrey Botanical Club* **59:** 313–340.

Xi Z, **Ruhfel BR**, **Schaefer H**, **Amorim AM**, **Sugumaran M**, **Wurdack KJ**, **Endress PK**, **Matthews ML**, **Stevens PF**, **Davis CC. 2012.** Phylogenomics and *a posteriori* data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 17519–17524.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020

Yakandawala D, **Morton CM**, **Prance GT. 2010.** Phylogenetic relationships of the Chrysobalanaceae inferred from chloroplast, nuclear, and morphological data. *Annals of the Missouri Botanical Garden* **97:** 259–281.

Yao G, **Jin JJ**, **Li HT**, **Yang JB**, **Mandala VS**, **Croley M**, **Mostow R**, **Douglas NA**, **Chase MW**, **Christenhusz MJM**, **Soltis DE**, **Soltis PS**, **Smith SA**, **Brockington SF**, **Moore MJ**, **Yi TS**, **Li DZ. 2019.** Plastid phylogenomic insights into the evolution of Caryophyllales. *Molecular Phylogenetics and Evolution* **134:** 74–86.

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

Zhang Q. 2010. Why does biparental plastid inheritance revive in angiosperms? *Journal of Plant Research* **123:** 201–206.

Zhang SD, **Jin JJ**, **Chen SY**, **Chase MW**, **Soltis DE**, **Li HT**, **Yang JB**, **Li DZ**, **Yi TS. 2017.** Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytologist* **214:** 1355–1367.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Phylogenetic tree for the 20-taxon dataset, including Chrysobalanaceae, obtained from the software BEAST2. The arrow points to the crown age of Chrysobalanaceae, inferred at *c*. 38.9 Mya. Squares indicates fossil constraints.

Figure S2. Phylogenetic tree for Chrysobalanaceae, obtained from the software RAxML v.8.2.10. Branch support was reported along the branches.

Table S1. Description of the accessions included in this study, with plastid genome length and area coding.

Downloaded from https://academic.oup.com/botlinnean/advance-article/doi/10.1093/botlinnean/boaa052/5890102 by jerome.chave@univ-tlse3.fr on 09 August 2020