



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Phylogeny and biogeography of *Ceiba Mill.* (Malvaceae, Bombacoideae)

### Citation for published version:

Pezzini, FF, Dexter, KG, De Carvalho-sobrinho, JG, Kidner, CA, Nicholls, JA, De Queiroz, LP & Pennington, RT 2021, 'Phylogeny and biogeography of *Ceiba Mill.* (Malvaceae, Bombacoideae)', *Frontiers of Biogeography*, vol. 0, no. 0. <https://doi.org/10.21425/F5FBG49226>

### Digital Object Identifier (DOI):

[10.21425/F5FBG49226](https://doi.org/10.21425/F5FBG49226)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Published In:

Frontiers of Biogeography

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# UC Merced

## Frontiers of Biogeography

### Title

Phylogeny and biogeography of Ceiba Mill. (Malvaceae, Bombacoideae)

### Permalink

<https://escholarship.org/uc/item/2c8639qj>

### Journal

Frontiers of Biogeography, 0(0)

### Authors

Pezzini, Flávia Fonseca  
Dexter, Kyle G.  
de Carvalho-Sobrinho, Jefferson G.  
et al.

### Publication Date

2020

### DOI

10.21425/F5FBG49226

### Supplemental Material

<https://escholarship.org/uc/item/2c8639qj#supplemental>

### License

<https://creativecommons.org/licenses/by/4.0/> 4.0

Peer reviewed

Journal  
Frontiers of Biogeography  
Publication year: 2021  
Issue: 13.2  
Article doi: doi:10.21425/F5FBG49226  
Article number: e49226  
Article format: Long (author names in front page + abstract + 6-10 keywords)  
Article type: Research Article  
Article series: Not applicable to this article  
Title: Phylogeny and biogeography of *Ceiba* Mill. (Malvaceae, Bombacoideae)  
Subtitle: n.a.  
Author list:  
Flávia Fonseca Pezzini Orcid ID: [0000-0001-5988-7361](https://orcid.org/0000-0001-5988-7361)  
Kyle G. Dexter Orcid ID: [0000-0001-9232-5221](https://orcid.org/0000-0001-9232-5221)  
Jefferson G. de Carvalho-Sobrinho Orcid ID: [0000-0003-3605-0707](https://orcid.org/0000-0003-3605-0707)  
Catherine A. Kidner <https://orcid.org/0000-0001-6426-3000>  
James A. Nicholls <https://orcid.org/0000-0002-9325-563X>  
Luciano P. de Queiroz <https://orcid.org/0000-0001-7436-0939>  
R. Toby Pennington <https://orcid.org/0000-0002-8196-288X>

Strapline author: Pezzini et al.

Strapline running title: Evolutionary history of tropical dry forest

Supplementary Materials: Yes (1 table, 2 Figures)

## **Phylogeny and biogeography of *Ceiba* Mill. (Malvaceae, Bombacoideae)**

Flávia Fonseca Pezzini<sup>1,2</sup>, Kyle G. Dexter<sup>3</sup>, Jefferson G. de Carvalho-Sobrinho<sup>4,5</sup>, Catherine A. Kidner<sup>1,2</sup>, James A. Nicholls<sup>6</sup>, Luciano P. de Queiroz<sup>7</sup>, R. Toby Pennington<sup>1,8</sup>

<sup>1</sup> Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom

<sup>2</sup> School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom

<sup>3</sup> School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom

<sup>4</sup> Colegiado de Ciências Biológicas, Universidade Federal do Vale do São Francisco, Petrolina, Brazil

<sup>5</sup> Departamento de Ciências Florestais, Universidade Federal Rural de Pernambuco, Recife, Brazil

<sup>6</sup> Australian National Insect Collection, CSIRO, Acton, Australia

<sup>7</sup> Herbário, Universidade Estadual de Feira de Santana, Feira de Santana, Brazil

<sup>8</sup> Geography, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Corresponding author: Flávia Fonseca Pezzini, [flaviapezzini@gmail.com](mailto:flaviapezzini@gmail.com), 20a Inverleith Row  
Edinburgh, EH3 5LR, UK

## Abstract

The Neotropics is the most species-rich area in the world, and the mechanisms that generated and maintain its biodiversity are still debated. This paper contributes to the debate by investigating the evolutionary and biogeographic history of the genus *Ceiba* Mill. (Malvaceae, Bombacoideae). *Ceiba* comprises 18 mostly Neotropical species, largely endemic to two major biomes, seasonally dry tropical forests (SDTFs) and rain forests. Its species are among the most characteristic elements of Neotropical SDTF, one of the most threatened biomes in the tropics. Phylogenetic analyses of DNA sequence data (from the nuclear ribosomal internal transcribed spacers [nrITS] for 30 accessions representing 14 species of *Ceiba*) recovered the genus as monophyletic. The phylogeny showed geographic and ecological structure in three main clades: (i) a rain forest lineage of nine accessions of *C. pentandra* sister to the remaining species; (ii) a highly supported clade composed of *C. schottii* and *C. aesculifolia* from Central American and Mexican SDTF, plus two accessions of *C. samauma* from semi-humid, inter Andean valleys in Peru; and (iii) a highly supported South American SDTF clade including 10 species showing little sequence variation. Within this South American SDTF clade, no species represented by multiple accessions were resolved as monophyletic. We demonstrate that the patterns of species age, monophyly, and geographic structure previously reported for SDTF species within the Leguminosae family are not shared by *Ceiba*, suggesting that further phylogenetic studies of unrelated groups are required to understand general patterns.

## Highlights

- This paper provides a well-sampled phylogeny of the iconic genus *Ceiba*, one of the most characteristic tree genera of Neotropical seasonally dry tropical forest (SDTF).
- There is a clear phylogenetic signal for biome preference and geographic structure in *Ceiba*.
- We estimate a mid-Miocene origin for *Ceiba*, with the stem node age of the genus estimated at 21.1 (14.7-27.1 [95% HPD]) Ma and the crown node age at 12.7 (8.2-17.6 [95% HPD]) Ma.
- *Ceiba* species have stem ages younger than expected in the SDTF clade and stem ages older than expected for the rain forest species clade.
- Patterns of species age, monophyly, ecological and geographic structure previously reported for SDTF species are only partially shared by *Ceiba*.

**Keywords:** Caatinga, multiple accessions, Neotropics, radiation, rain forests, seasonally dry tropical forests, species monophyly

## Introduction

The Neotropics is the most species-rich region in the world (Antonelli and Sanmartín 2011) and the mechanisms that generated and maintain its biodiversity are constantly debated (Moreau and Bell 2013, Silveira et al. 2016, Rangel et al. 2018, Schley et al. 2018, Sosa et al. 2018, Raven et

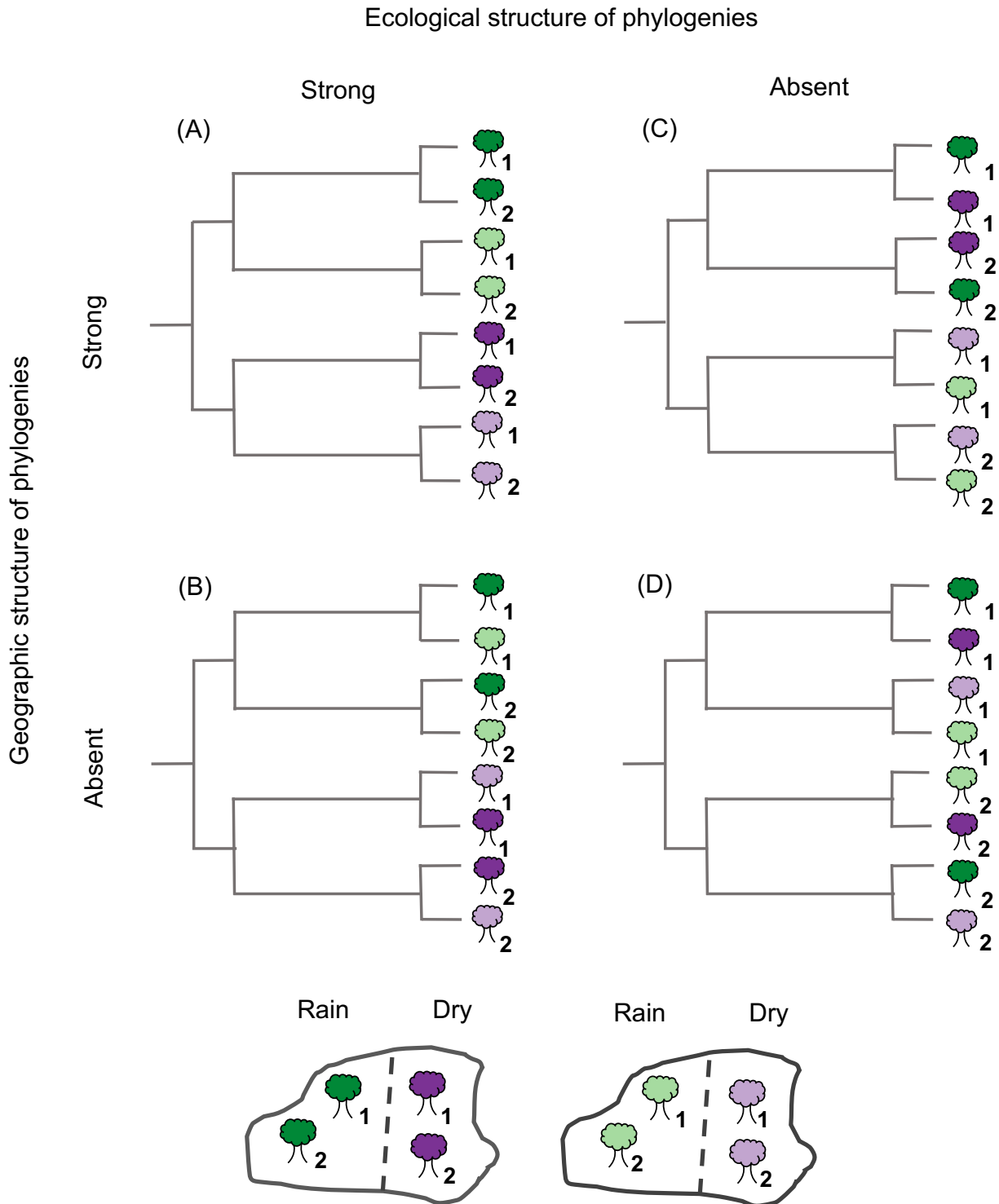
al. 2020). Through evolutionary time, the Neotropics have been climatically and geologically dynamic, resulting in a diversity of biomes, from deserts to tropical rain forests (Hughes et al. 2013, Rangel et al. 2018). To understand the history and dynamics of those biomes, molecular phylogenetic and phylogeographic approaches have been used to infer the historical relationships amongst biomes and areas (Pennington et al. 2006). In recent years the dichotomy regarding the ‘cradle’ vs. ‘museum’ debate (Stebbins 1974) explaining Neotropical diversity has given way to a more nuanced approach, considering that plant diversification patterns may be recent, old, slow or rapid, even within individual clades (McKenna and Farrell 2006, Couvreur et al. 2011, Hughes et al. 2013, Koenen et al. 2015, Eiserhardt et al. 2017). As suggested in the literature more than 10 years ago (Wiens and Donoghue 2004, Pennington et al. 2006), this heterogeneity in diversification timing and rate within and among clades may be related not only to climatic and geological events, but also to the age and ecological differences of the biomes. For example, geologically old biomes (e.g., tropical rain forest) are likely to have provided lineages that colonised newer biomes (e.g., savannas), and the relative difficulty of evolving adaptations such as drought tolerance or the ability to survive fire might determine whether a lineage could adapt to such a new biome (niche evolution; Simon et al. 2009, Pennington & Lavin 2016) or remain confined to the same biome (niche conservatism) over evolutionary timescales (Crisp et al. 2009, Segovia et al. 2020).

Clades endemic to two of the major Neotropical biomes, seasonally dry tropical forests (SDTFs) and tropical rain forests (hereafter rain forests), give good examples of different and distinctive phylogenetic and biogeographic patterns, suggesting an interaction of ecology and phylogeny over evolutionary timescales (Pennington et al. 2011, Pennington and Lavin 2016, Dexter et al. 2017). SDTFs occur on fertile soils and are characterized by the absence of fire adaptation in the flora and a predominantly continuous tree canopy, which becomes more open in the drier sites, with plants shedding up to 90–95% of their leaves during the five to six month long dry season (Murphy and Lugo 1986, Pennington et al. 2009). This biome has been one of the least studied, but is one of the most threatened in the tropics (Miles et al. 2006, DRYFLOR 2016). It occurs in disjunct areas throughout the Neotropics (Figs 2, 4) and has high beta-diversity and plant species endemism (Pennington et al. 2009, DRYFLOR 2016). The Bombacoideae clade (Malvaceae) is often common, distinctive, and includes some of the most dominant and frequent species in SDTF, including members of the genera *Cavanillesia*, *Ceiba*

and *Pseudobombax*. Other important members of the SDTF include species within the Caesalpinia group (Leguminosae) and *Handroanthus* (Bignoniaceae) (Silva and Scariot 2004a, Silva and Scariot 2004b, Linares-Palomino and Ponce Alvarez 2005, Ballesteros et al. 2009, Pennington et al. 2009, Carvalho and Felfili 2011, Gagnon et al. 2019). By contrast, the Amazon is the largest rain forest in the world (Sioli 1984), dating back to the Paleocene (Burnham and Johnson 2004). The vast continuous areas of this biome harbours ca. 16,000 tree species, although only 227 ‘hyperdominant’ species account for half of the trees in the Amazon (ter Steege et al. 2013). The contrasting difference in the ecology of tropical dry and rain forests underpins the hypothesis that major Neotropical biomes are distinct ‘evolutionary theatres’, in which the course of diversification and biogeography may have been different (Lavin et al. 2004, Pennington et al. 2009, Pennington and Lavin 2016) due to distinct ecological processes and attributes specific to organisms acting over evolutionary time scales (eg., Baker et al. 2014, Smith et al. 2014, Honorio Coronado et al. 2019).

SDTF-confined clades contain species that often resolve as monophyletic and with old stem ages in DNA-sequence-based phylogenies (Pennington et al. 2010, Pennington et al. 2011, Särkinen et al. 2012, Lavin et al. 2018). de Queiroz and Lavin (2011) showed that multiple accessions of *Coursetia caatingicola* and *C. rostrata* (Leguminosae) from the Brazilian Caatinga, the largest continuous area of SDTF, coalesce, i.e., are recovered as monophyletic. This pattern might be explained by the fact that the coalescence of conspecific accessions depends on the effective population size ( $N_e$ ), and the time to coalesce in small populations, such as the ones in the small isolated areas of SDTF across the Neotropics, will be much faster when compared to larger populations (Naciri and Linder 2015, Pennington and Lavin 2016). In addition, closely related lineages inhabiting SDTF tend to occur geographically closer when compared to more distantly related lineages. Pennington et al. (2010) used nrITS to show that this geographic structure can be detected even at the population level for the legume shrub *Cyathostegia*, of which the northern and southern populations in the inter-Andean valleys from Ecuador and Peru are reciprocally monophyletic sister to a clade containing populations from southern Peru. The geographically structured phylogenetic pattern characteristic of SDTF clades suggests dispersal-limitation caused by the stable ecological conditions of the biome maintained over long evolutionary timescales (Pennington et al. 2010, Hughes et al. 2013). By contrast, tree clades confined to the Amazon rain forest are suggested to contain more non-monophyletic

species, more species with young stem ages, and clades that lack geographic phylogenetic structure (Dexter et al. 2017). These rain forest patterns might be explained by frequent dispersal and subsequent successful colonization over evolutionary timescales (Pennington and Lavin 2016 Fig. 1).



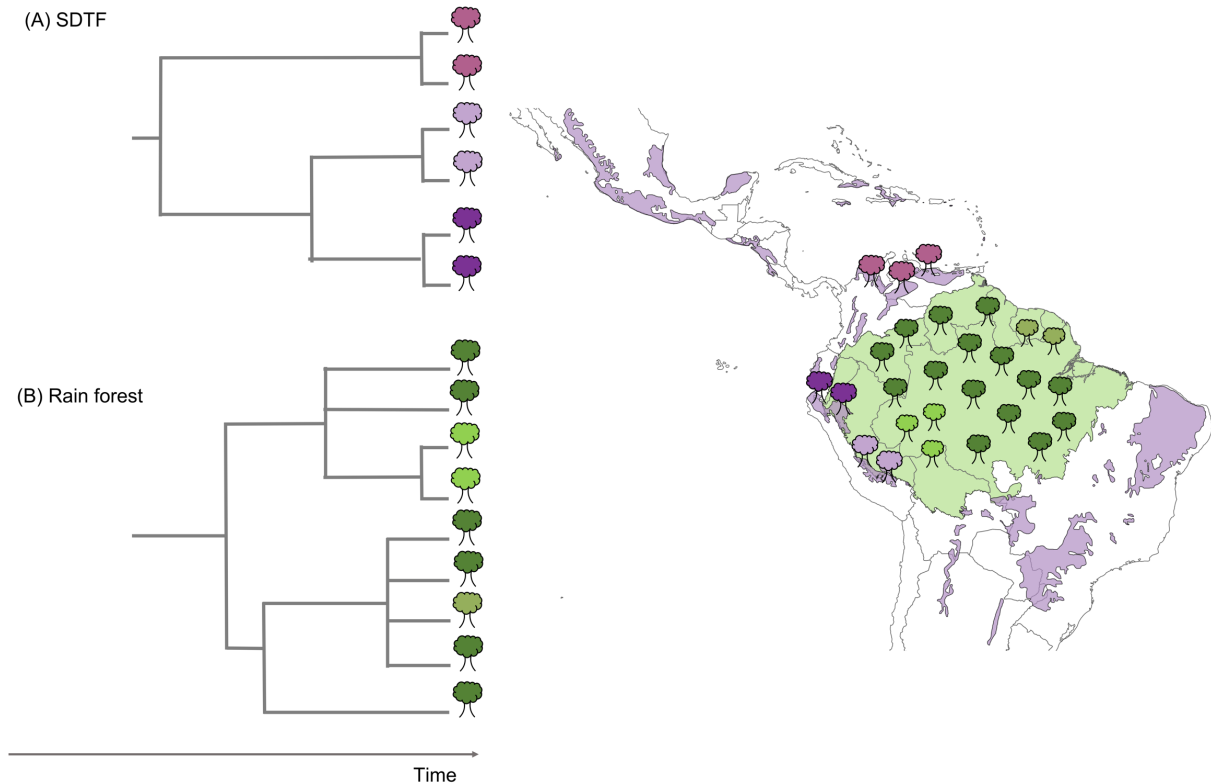
**Figure 1.** Two hypothetical islands, each with an area of seasonally dry tropical forest and rain forest. In total, eight different species represented by different colours (dark green species 1, dark green species 2, dark purple 3, etc.) occur in the two biomes and the different islands. Hypothetical phylogenies showing patterns of presence or absence of geographic and ecological structures (design of this figure after Graham and Fine 2008).

Species non-monophyly can only be investigated by including multiple accessions per species in phylogenetic trees. Examples of such densely sampled phylogenies revealed discordance between morphological and genetic species delimitation for animals (Wiens and Penkrot 2002) and plants (Dainou et al. 2014, Prata et al. 2018), thus reviving the discussion about whether species should be monophyletic (Rieseberg and Brouillet 1994). Different modes of speciation such as allopatric, parapatric, or sympatric result in different phylogenetic patterns (Rieseberg and Brouillet 1994, Knowles and Carstens 2007). Rieseberg and Brouillet (1994) predicted that allopatric speciation is likely to produce monophyletic daughter species in contrast to parapatric and sympatric speciation that often result in a paraphyletic progenitor and monophyletic derivative species. In this sense, speciation is a process producing a paraphyletic entity (the ancestral species) that only becomes monophyletic over time. A key aspect is that the time to achieve monophyly of a paraphyletic ancestral species will vary depending on factors such as effective population size, level of gene flow, and type of genetic data (Rieseberg and Brouillet 1994). For example, nuclear genes take longer to coalesce when compared to chloroplast or mitochondrial genes (Rieseberg and Brouillet 1994, Hudson and Coyne 2002). Likewise, large census population sizes are more likely to cause longer coalescence times (Hudson and Coyne 2002) because they reflect large effective population size, which helps the persistence of ancestral polymorphism (Knowles and Carstens 2007). Therefore, large population sizes mean a longer time to achieve monophyly (Rieseberg and Brouillet 1994, Hudson and Coyne 2002).

Our aim here is to assess whether different biomes, and the different ecological processes within them, underlie distinct phylogenetic patterns within *Ceiba*. Specifically, Pennington and Lavin (2016) hypothesised that rain forest species will often be recovered as paraphyletic with narrowly-distributed species nested within broadly distributed species and that SDTF species will be recovered as monophyletic (Fig. 2). Skeels and Cardillo (2019) used simulations



combining geographic ranges of species and phylogenetic trees to infer speciation modes for different taxa, indicating that the signature of distinct past speciation events is detectable in current species distributions and phylogenetic patterns.



**Figure 2.** Contrasting phylogenetic and distribution patterns of six hypothetical species: (A) three from seasonally dry tropical forest (SDTF) showing species restricted to different SDTF fragments and recovered as monophyletic; (B) three from tropical rain forest, one broadly distributed species recovered as paraphyletic due to two narrowly distributed species nested within it. SDTF areas are represented in purple in the map, and tropical rain forest in green.

The Neotropical genus *Ceiba* Mill. (Malvaceae: Bombacoideae) comprises 18 tree species grouped into taxonomic sections *Ceiba* and *Campylanthera* (Schott & Endl.) K. Schum. based on morphological characters of pollen and the presence and vascularisation of staminal appendages. It is one of the most characteristic elements of many Neotropical SDTFs (Fig. 3). However, it also contains species confined to the Amazon rain forest and is therefore a good case

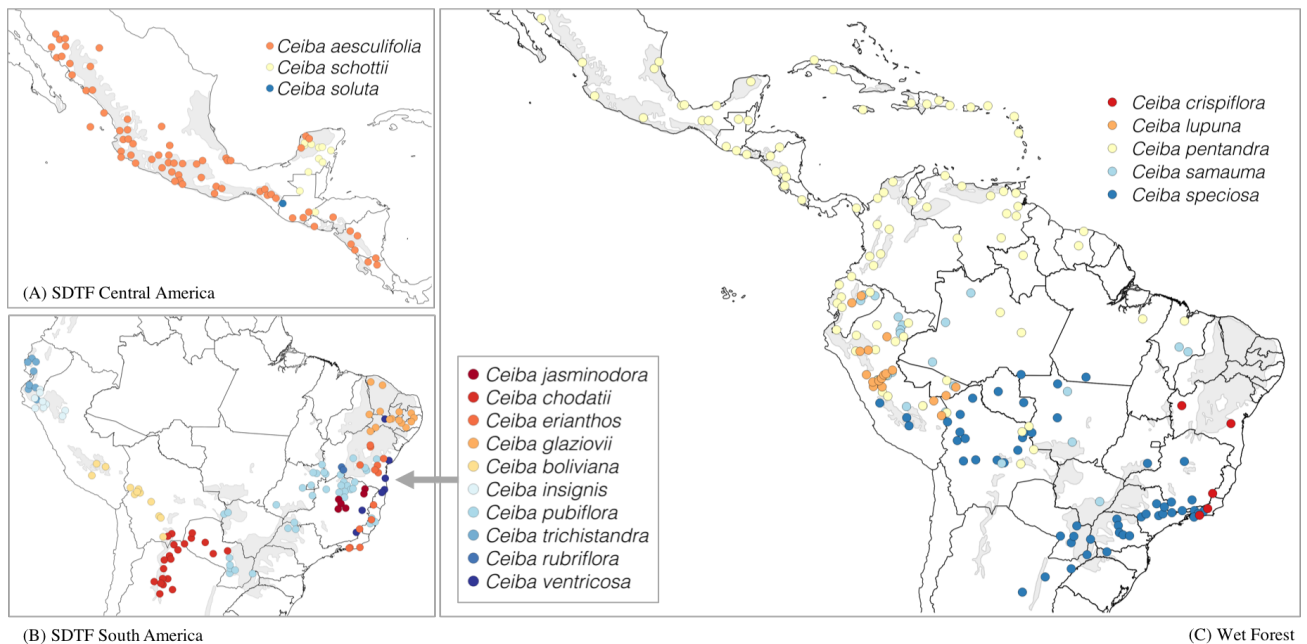
study to investigate biome-specific differences in the nature of species and their diversification trajectories.



**Figure 3.** Different species of *Ceiba* from Brazil. *Ceiba pubiflora* from Minas Gerais (a, h) and Goiás (i); *Ceiba jasminodora* from Minas Gerais (b); *Ceiba rubriflora* from Minas Gerais (c); *Ceiba glaziovii* from Bahia (d, e); *Ceiba speciosa* from São Paulo (f); *Ceiba samauma* from Brazil (g). Photos: F. Pezzini.

*Ceiba* species have digitately compound leaves, aculeate trunk and branches and can vary from 50 m canopy emergents in seasonally flooded várzea forests in the Amazon (*C. pentandra*) to 2 m treelets on rocky outcrops (campos rupestres) in Minas Gerais, Brazil (*C. jasminodora*) (Gibbs and Semir 2003). In some species (*C. chodatii*, *C. glaziovii*, *C. pubiflora*, *C. speciosa*) the trunk can be ventricose (swollen), explaining its vernacular names barriguda (“swollen belly”; Brazil) and palo borracho (“drunken tree”; Peru, Bolivia, and Argentina; Fig. 3). Most species are deciduous and flower when leafless. Fruits are woody loculicidal capsules with endocarp modified in trichomes (“kapok”). These trichomes have been historically associated with the wind dispersal of small, numerous, and lightweight seeds, and may also play a role in both water dispersal and germination of rain forests species (Kubitzki and Ziburski 1994, Fischer 1997).

*Ceiba* species occur mostly in SDTF, with the exception of the widespread *C. samauma*, *C. speciosa* and *C. pentandra*, all of which occur also in more humid environments, and *C. lupuna*, which is the only species restricted to rain forests (Fig. 4). *Ceiba jasminodora* is restricted to rocky outcrops of the Brazilian Espinhaço mountain range, thus a species from dry habitats grouped with the SDTF species throughout this article (Table S1). On average, each of the thirteen SDTF species has a narrower geographic distribution when compared to the five rain-forest-inhabiting species (Fig. 4).



**Figure 4.** Distribution of 18 *Ceiba* species in three maps: (A) three species from SDTFs in Central and North America, (B) ten species from SDTFs in South America, and (C) five species from rain forests in Latin America. Grey areas represent the distribution of SDTF following DRYFLOR (2016). Occurrence records adapted from Gibbs and Semir (2003).

Previous Bayesian analyses of sequence data from the nuclear ribosomal internal and external transcribed spacers (nrITS and nrETS) and plastid markers (*matK*, *trnL-F*, *trnS-trnG*) for 13 species recovered *Ceiba* as monophyletic and sister to *Neobuchia paulinae* Urb. (Duarte et al. 2011, Carvalho-Sobrinho et al. 2016). Together with *Spirotheca* Ulbr., *Pochota fendleri* (Seem.) W.S. Alverson & Duarte, and *Pseudobombax* Dugand, these taxa form the well-supported “striated bark” clade (Carvalho-Sobrinho et al. 2016). However, relationships within

*Ceiba* were poorly resolved and only one individual per species was included in the phylogenetic tree. *Ceiba* has a historically complex taxonomy with species boundaries still unclear, especially within the *Ceiba insignis* species complex (Gibbs and Semir 2003). This is aggravated by the fact that herbarium specimens are often incomplete because individuals produce flowers and fruits when leafless. Therefore, a well sampled phylogenetic tree with multiple accessions per species would be a useful tool to explore the nature of species in *Ceiba*, that is, assess whether the DNA sequences of different individuals representing a single taxonomically recognised species coalesce, i.e., are resolved as monophyletic.

This paper investigates the evolutionary history of *Ceiba* species. We aim to assess whether the *Ceiba* phylogenetic tree is geographically or ecologically structured, and if species confined to SDTFs are resolved differently in the phylogenetic tree as compared with rain forest species (i.e., monophyletic on long stem lineages).

## Methods

### *Taxon sampling*

We present the best sampled phylogeny of the genus *Ceiba* to date, covering 30 accessions representing 14 of the 18 species accepted for the genus (Gibbs and Semir 2003, Carvalho-Sobrinho and de Queiroz 2008). Critically, this study is the first to sample multiple individuals per species for six species (Table S1). As outgroups, we included 10 accessions representing species of the closest sister clades among Bombacoideae (Carvalho-Sobrinho et al. 2016): *Eriotheca*, *Pseudobombax*, *Pochota fendleri*, and *Spirotheca*. The full data set represents a combination of new sequence data from field surveys as well as from herbarium specimens, doubling the number of accessions of *Ceiba* in relation to the previous study by Carvalho-Sobrinho et al. (2016).

### *DNA sequence data*

We used the nuclear ribosome internal transcribed spacer region (i.e., the 5.8S and flanking ITS1 and ITS2 spacers, hereafter nrITS) region to investigate species relationships in *Ceiba*. In Bombacoideae, this region has been widely explored to help elucidate relationships among

genera and species (Baum et al. 1998, Duarte et al. 2011, Carvalho-Sobrinho et al. 2016), and to investigate genetic structure among populations (Dick et al. 2007). Previous systematic studies in Bombacoideae used a combination of nuclear (nrITS, nrETS) and plastid (*matK*, *trnS-trnG* and *trnL-trnF*) markers, and nrITS had the highest number of informative characters (Duarte et al. 2011, Carvalho-Sobrinho et al. 2016). In spite of having drawbacks related to potential paralogous copies (Álvarez and Wendel 2003, Buckler et al. 1997), nrITS can still play an important role in the investigation of species relationships if analysed carefully, for example identifying pseudogenes and assessing orthology in the case of intra-individual polymorphism (Bailey 2003, Feliner and Rosselló 2007).

Genomic DNA extraction was performed for 36 herbarium and silica-gel dried leaf samples (Table S1) using Qiagen DNeasy Plant Mini Kits and following the manufacturer's protocol, with the following changes: twice the volume of buffer AP1 in addition to a small amount of PVPP (polyvinyl polypyrrolidone) was added at the lyse step followed by an incubation of 30 minutes; addition of 1uL of Riboshredder in the lysate solution followed by incubation at 37°C for 20 minutes; addition of twice the volume of buffer P3; and final elution in 46 uL of EB buffer run through the column twice to increase yield. Each 20 uL PCR amplification reaction contained 0.5 uL of template, 2 uL of dNTPs (2 mM), 2 uL of 10x reaction buffer, 1 uL of MgCl<sub>2</sub> (50 mM), 0.65 uL of both forward primer and reverse primer solutions (10 uM), 0.1 uL of Taq polymerase, 4 uL of CES buffer and 9.1 uL of ddH<sub>2</sub>O. Amplification followed the same procedure described in Carvalho-Sobrinho et al. (2016) using 17SE and 26 SE as amplification primers (Sun et al. 1994) and ITS4 (5'- TCC TCC GCT TAT TGA TAT GC-3') and ITS92 (5-'AAG GTT TCC GTA GGT GAA C-3') as sequencing primers. Samples were submitted to the Edinburgh Genomics laboratory at the University of Edinburgh for Sanger sequencing. For low quality sequences, we tested variations of the protocol (e.g., diminishing the amount of template in the PCR reaction or varying the sequencing primer). High quality sequences were recovered for 13 out of the 36 samples from which DNA was extracted for the nrITS region. The remaining samples were discarded due to absence of DNA or short fragment length (less than 300bp).

All the inter-accession polymorphisms detected were validated visually by checking the electropherograms. Sequences were edited with Sequencher 5.4.1 (Gene Codes Corp., Ann

Arbor, Michigan) and alignments were performed manually in Mesquite (Maddison and Maddison 2015). We investigated the potential presence of nrITS pseudogenes by conducting phylogenetic analysis in separate partitions representing the 5.8S (conserved region) and the ITS 1 and ITS 2 regions (fast evolving regions). In pseudogenes, the conserved and the fast evolving regions are expected to show similar rate of evolution, whereas in functional genes the conserved region is expected to have a lower rate of evolution compared to the fast evolving regions (Bailey et al. 2003). We assigned partitions by comparing to the annotated accession of *Pseudobombax tomentosum* (Mart.) A. Robyns (GenBank KM453206) and checked for differences in rates of evolution in the partitioning scheme proposed with PartitionFinder2 version 2.1.1 (Lanfear et al. 2017) using PhyML version 3.0 (Guindon et al. 2010) and the greedy search algorithm (Lanfear et al. 2012).

### *Phylogenetic analysis and molecular dating*

We implemented maximum likelihood (ML) and Bayesian Inference (BI) analysis. To determine the best fitting model of sequence evolution for each of the three partitions, we used PartitionFinder2 version 2.1.1 (Lanfear et al. 2017) in the ML analysis and the reversible jump model selection (RB) implemented in BEAST2 version 2.5.1 (Bouckaert et al. 2019) under the Bayesian framework. IQ-TREE version 2.0.3 (Nguyen et al. 2015, Minh et al. 2020) was used to run the ML analysis with 1,000 bootstrap replicates and using the partition model option (-p) (Chernomor et al. 2016), with substitution models specified as follows: GTR+G for ITS1, K80+G for 5.8S, and HKY+G for ITS2, as inferred with PartitionFinder2 (see above). BEAST2 version 2.5.1 (Bouckaert et al. 2019) was used to perform BI analysis and temporally calibrate the phylogeny. Different combinations of relaxed clock models (Uncorrelated Exponential Distribution [UCED] and Uncorrelated Log Normal Distribution [UCLD]; Drummond et al. 2006) and tree priors (Yule and Birth-Death) were compared. Few studies have objectively contrasted the effect of different models in the divergence time estimation, and a poorly inferred time-calibrated phylogeny can have serious consequences for our understanding of diversification history of lineages (Louca and Pennell 2020). For example, different tree priors resulted in impressive differences in age estimation for cycads, with the Yule prior inferring ages three times older than the Birth-Death prior (Condamine et al. 2015). In Bayesian analysis, the most suitable model can be selected by comparing the Bayes Factor (BF). The BF is equal to the

ratio of the Marginal Likelihood Estimate (MLE) of two models ( $BF = MLE1/MLE2$ ) or to the difference of MLEs in log space, that is,  $\log(BF) = \log(MLE1) - \log(MLE2)$ . Positive values of BF would favour MLE1, and different values have different strengths. Values above five indicates that one model is significantly favoured over the other (Kass and Raftery 1995, Condamine et al. 2015), values above 20 indicate strong support and values above 150 overwhelming support. We estimated the Marginal Likelihood using the Nested Sampling algorithm (Skilling 2006) implemented in the NS package version 1.0.4 (Russel et al. 2019) for BEAST2 version 2.5.1 with 60 particles and 10,000 chain length. The NS package also calculates the standard deviation (SD) of the estimated Marginal Likelihood, which allows us to have confidence in the BF values calculated. For each combination of priors, we ran two independent runs of 10 million generations, sampled every 1,000 generations, visually inspected convergence of MCMC, and ensured effective sample size  $> 200$  for all parameters of each run using Tracer v1.7.1 (Rambaut et al. 2018). Resulting trees and log files from each run were combined using LogCombiner with a burn-in of 10% and the Maximum Clade Credibility Tree was summarized in TreeAnnotator with node heights as mean heights. We used r8s (Sanderson 2004) to implement the penalised likelihood method (Sanderson 2002) and calculate substitutions rates. We used the phylogram derived from the ML analysis as input and conducted a cross-validation analysis to find the best smoothing parameter.

We used a fossil flower of *Eriotheca prima* (Duarte 1974) from the middle to late Eocene (de Lima and Salard-Chebouldaëff 1981) of Brazil as a primary calibration for our BEAST2 analysis. The flower was identified as *Eriotheca* based on its small size (*Pachira* have larger flowers) and androecium organisation, which characterizes the extant species of the genus (Robyns 1963, Duarte et al. 2011, Carvalho-Sobrinho et al. 2016). In previous studies, *Eriotheca* was resolved as sister to a clade comprising *Pseudobombax*, *Spirotheca*, *Ceiba*, and *Pochota fendleri* (Duarte et al. 2011, Carvalho-Sobrinho et al. 2016). Because the dating of this fossil is imprecise (middle to late Eocene: 33-56 million years old [Ma]), we assigned the offset age of 33 Ma as a minimum age to the stem node (Renner 2005, Pennington et al. 2006) of *Eriotheca* (the crown node of the clade comprising *Ceiba*, *Eriotheca*, *Pochota fendleri*, *Pseudobombax*, and *Spirotheca*), which is equivalent to the root node of the outgroups and ingroup of this study. We assigned a log-normal distribution with a mean of 1.542 and standard deviation of 1.5. This fossil calibration is conservative, with 95% of the prior distribution comprised between 33 and 47 Ma

and thus the ages estimated here are considered minimum ages estimates. In order to explore the effects of using the medium and maximum ages of the *Eriotheca* fossil on phylogenetic age estimates, we also ran analyses assigning minimum ages of 47 Ma and 56 Ma to the *Eriotheca* stem (Figs S1 and S2). We followed the dates on the Geologic Time Scale v. 5.0 (Gradstein et al. 2012).

### *Phylogenetic signal test*

We tested for strength of phylogenetic signal for the binary traits related to ecology (rain vs dry forests) and geography (Central and North America vs South America) using the *D value* proposed by Fritz and Purvis (2010) and the Caper package (v. 1.0.1; Orme 2013) in R, with 5,000 permutations. Under a null model of Brownian motion evolution of a binary trait, *D* has an expected value of 0. A negative *D* value indicates a strongly clustered phylogenetic pattern for a given binary trait (perhaps due to some process of evolutionary constraint), a value of one indicates a completely random pattern with respect to the phylogeny (i.e., no correlation between phylogeny and the trait at all), and values above one indicate an overdispersed phylogenetic pattern (perhaps due to divergent selection). We assigned species to ecological and geographic categories following Gibbs and Semir (2003; Table S1). Despite occurring mainly in rain forests, *Ceiba speciosa* has also been recorded in dry semi-deciduous forests (Fig. 4). To check for possible bias in the ecological affinity of *C. speciosa*, we also conducted a phylogenetic signal analysis assigning this species to dry forest, observing no difference when comparing to the analysis run assigning the species to rain forest.

## **Results**

The total length of the aligned sequences was 814 nucleotides, of which 283 were variable and 178 (22%) were parsimony-informative characters. The ML and BI trees showed congruent topologies (Figs 5 and 6). *Ceiba* was strongly supported as monophyletic, with posterior probability (pp) = 1 and bootstrap value (bs) = 100 and was recovered as sister to *Pseudobombax* (Figs 5 and 6). The UCLD and UCED clock models and Yule and Birth-Death tree models inferred similar crown and stem ages with overlapping credibility intervals (95% Highest Posterior Density [HDP]), Table 1). For the UCLD clock prior, the Bayes Factor value of 6.15



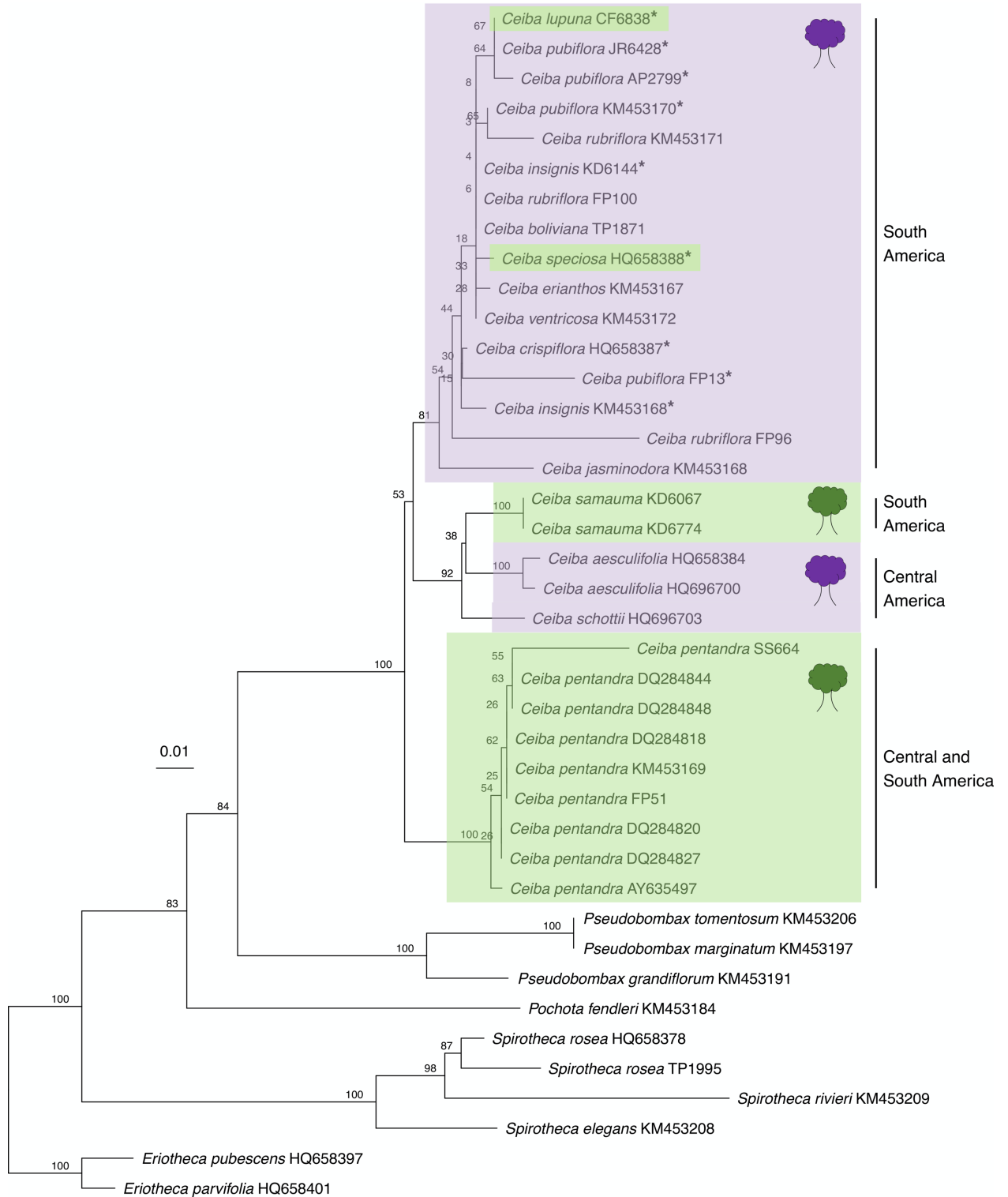
supported the Yule tree prior as the most suitable tree model. For the UCED clock prior, the Bayes Factor of 1.15 indicated that neither tree model was favoured (Table 1). Therefore, results shown onwards for Bayesian analysis are those inferred using the UCLD clock model and the Yule tree model. Using the 33 Ma fossil calibration, the stem node age of *Ceiba* is 21.1 (14.7-27.1 [95% HPD]) Ma, the crown node age is 12.7 (8.2-17.6 [95% HPD]) Ma (Table 1, Fig. 6), and substitution rates estimated as  $1.592 \times 10^{-9}$  substitutions per site per year (s/s/y).

**Table 1.** Absolute ages estimate (Ma) for ten nodes under different tree (Yule and Birth-Death) and clock priors (Uncorrelated Log Normal Distribution [UCLD] and Uncorrelated Exponential Distribution [UCED]). Ages are reported in million years as mean ages followed by the 95% Highest Posterior Density (HDP) as a result of the combined independent runs for each tree and clock priors. MLE (SD): marginal likelihood estimated followed by standard deviation; BF: Bayes factor calculated as the difference between the MLE of the Yule and the Birth-Death prior for each clock prior UCLD and UCED. Values above five indicate that one model is significantly favoured over the other.

Clock prior Tree prior	UCLD		UCED	
	Yule	Birth-death	Yule	Birth-death
<b>MLE (SD)</b>	-3844.31 (1.85)	-3838.15 (1.79)	-3838.74 (1.88)	-3837.62 (1.76)
<b>BF</b>	6.15		1.15	
<b><i>Ceiba</i> crown</b>	12.7 (8.2-17.6)	10.9 (6.5-15.7)	13.5 (8.6-18.8)	11.5 (6.1-17.4)
<b><i>Ceiba</i> stem</b>	21.1 (14.7-27.1)	20.3 (13.5-26.9)	21.0 (14.4-27.5)	19.1 (11.3-26.9)
<b>SDTF SA crown</b>	8.6 (5.0-12.5)	7.0 (3.6-10.7)	9.6 (5.6-14.1)	7.8 (3.8-12.5)
<b>SDTF SA stem</b>	11.2 (7.2-15.6)	9.5 (5.6-13.9)	12.0 (7.5-16.7)	10.0 (5.2-15.4)
<b><i>C. aesculifolia</i> crown</b>	1.5 (0.2-3.2)	1.2 (0.2-2.3)	1.7 (0.3-3.9)	1.3 (0.2-2.9)
<b><i>C. aesculifolia</i> stem</b>	6.3 (3.1-9.8)	5.1 (2.1-8.3)	7.6 (3.7-12.0)	4.7 (1.7-8.3)
<b><i>C. samauma</i> crown</b>	0.4 (0.0001-1.4)	0.3 (0.0001-1.0)	0.5 (0.0001-1.7)	0.3 (0.0001-1.2)
<b><i>C. samauma</i> stem</b>	6.3 (3.1-9.8)	5.1 (2.1-8.3)	6.3 (2.6-10.2)	6.1 (2.4-10.3)
<b><i>C. pentandra</i> crown</b>	3.9 (1.6-6.5)	3.0 (1.1-5.2)	5.5 (2.3-9.2)	4.2 (1.5-7.5)
<b><i>C. pentandra</i> stem</b>	12.7 (8.2-17.6)	10.9 (6.5-15.7)	13.5 (8.6-18.9)	11.5 (6.1-17.4)

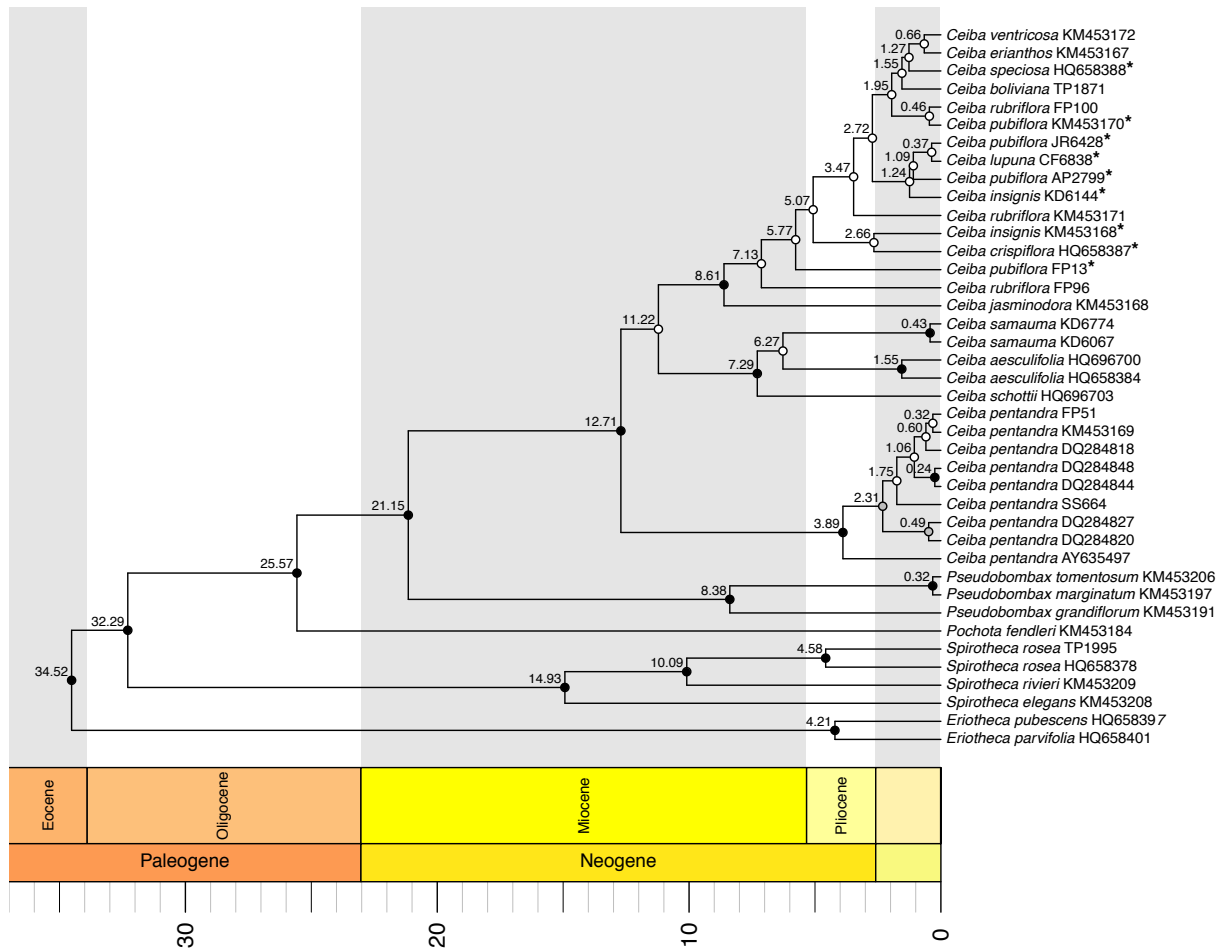
*Ceiba* comprises three main clades: (i) a rain forest lineage of the nine accessions of *C. pentandra*, which are strongly supported as monophyletic (pp = 1 and bs = 100) and sister to the remaining species, with stem node age of 12.7 (8.2-17.6 [95% HPD]) Ma and crown node age of 3.9 (1.6–6.5 [95% HPD]) Ma; (ii) a highly supported clade (pp = 1 and bs = 92) composed of *C. schottii* and *C. aesculifolia* from Central American and Mexican SDTF, plus two accessions of *C. samauma* from inter-Andean valleys in Peru, with stem node age of 11.2 (7.2-15.6 [95% HPD]) Ma and crown age of 7.3 (3.7-11.0 [95% HPD]) Ma; and (iii) a highly supported (pp = 0.99 and bs = 81) South American SDTF clade, including 10 species showing little sequence variation, with stem node age of 11.2 (7.2–15.6 [95% HPD]) Ma and crown node of 8.6 (5.0–12.5 [95% HPD]) Ma. Within this South American clade, *C. rubriflora*, *C. pubiflora* and *C. insignis*, which were represented by multiple accessions, were not resolved as monophyletic. The South American clade represents SDTF species, except for *C. lupuna*, a species with a distribution restricted to rain forest (Figs 5 and 6). *Ceiba aesculifolia* was the only SDTF species recovered as monophyletic, with stem node age of 6.3 (3.1–9.8 [95% HPD]) Ma and crown node of 1.5 (0.2–3.2 [95% HPD]) Ma. The stem node age of *C. jasminodora* was estimated at 8.6 (5.0-12.5[95%HPD]) Ma.

The *D* test showed significant phylogenetic signal for both ecological preference ( $D = 0.1247542$ ,  $P(D=1) = 0.001$ ,  $P(D=0) = 0.3422$ ) and geographic occurrence ( $D = 0.2204991$ ,  $P(D=1) = 0.0176$ ,  $P(D=0) = 0.3152$ ). Both *D* values were statistically indistinguishable from zero, which indicates that closely related species are more likely to show the same ecological preference or geographic occurrence, as expected under a Brownian model of evolution, whereby there would be a constant rate of state switching over time and any given lineage is more likely to stay within the same biome and geographic region per unit time than to switch to the alternative biome or geographic region.



**Figure 5.** Maximum likelihood phylogram derived from analysis of nuclear ribosomal ITS sequence data sets for 14 species of *Ceiba*. Species with asterisk belong to the *Ceiba insignis* species aggregate. Values above branches represent bootstrap values for internal nodes. Tree

symbols in front of accessions represent species occurring in SDTF (purple) and rain forests (green).



**Figure 6.** Maximum clade credibility tree resulting from BEAST2 analysis of nuclear ribosomal ITS sequence data sets for 14 species of *Ceiba*, 33 Ma offset calibration, using Yule tree prior and the Uncorrelated Lognormal Distribution clock model. Values above branches represent nodes' ages reported in million years as mean ages. Circles represent posterior probabilities for internal nodes: black  $\geq 0.95$ ; grey  $< 0.95$  and  $\geq 0.75$ , and white  $< 0.75$ . Species with asterisk belong to the *Ceiba insignis* species aggregate.

## Discussion

We investigated the evolutionary and biogeographic history of *Ceiba*, an iconic genus of Neotropical SDTF, one of the most threatened biomes in the tropics. We showed that *Ceiba*'s phylogenetic tree is geographically and ecologically structured, and that recognised taxonomic species confined to SDTFs are younger and do not resolve as monophyletic whilst rain forest species are older and monophyletic. Such results of younger, non-monophyletic SDTF species and older, monophyletic rain forest species contrast with recent predictions that rain forest species may, on average, have more recent origins than SDTF species and will more often be non-monophyletic. It demonstrates that densely sampled studies of other clades using multiple accessions of each species and a multi-locus approach are needed if we are to understand the nature of species and their boundaries, as well as the diversification process in Neotropical trees.

### *Geographic and ecological structure*

For most *Ceiba* species, diversification took place beginning in the late Miocene and Pliocene, similarly to *Bursera* Jacq. ex L. (De-Nova et al. 2012) in North America. Likewise, Pennington et al. (2004) showed that diversification happened beginning in the late Miocene and Pliocene for five lineages highly endemic to SDTF in South America: *Ruprechtia* C.A.Mey (Polygonaceae), robinoid legumes (Fabaceae), *Chaetocalyx* DC. and *Nissolia* Jacq. (Fabaceae), and *Loxopterygium* Hook.f. (Anacardiaceae). Our substitution rate estimated as  $1.592 \times 10^{-9}$  s/s/y places *Ceiba* around the mean substitution rate observed for nrITS in woody plants of  $1.64 \times 10^{-9}$  s/s/y, with a range from  $0.5$  to  $3.3 \times 10^{-9}$  s/s/y (Dick et al. 2013). *Ceiba* can be contrasted with faster substitution rates among Neotropical tree genera, such as the rain forest genus *Inga* Mill. (Leguminosae) estimated as  $2.34 \times 10^{-9}$  s/s/y (Richardson et al. 2001) and  $7.1-7.9 \times 10^{-9}$  s/s/y (Lavin 2006), and the dry forest genus *Coursetia* DC. (Leguminosae) estimated as  $5.0-8.2 \times 10^{-9}$  s/s/y (Lavin 2006). Conversely, our *Ceiba* substitution rate is in line with the slower rate of  $1 \times 10^{-9}$  s/s/y estimated for *Loxopterygium* (Anacardiaceae), which like *Ceiba*, is a woody and slow growing Neotropical tree genus with long generation time (Pennington et al. 2004).

The D test suggests phylogenetic signal for ecological preference and geographic phylogenetic structure with Central and South American clades in *Ceiba*. However, these tests should be interpreted with some caution. Phylogenetic signal tests are ideally performed on phylogenies that have one tip per species (Freckleton et al. 2002, Blomberg et al. 2003, Fritz and

Purvis 2010), although this does not allow for the fact that species are often not monophyletic, as is the case here. In addition, preliminary simulations showed that the D statistic could not successfully distinguish between evolutionary models (e.g., random phylogenetic structure versus that expected under Brownian motion) in datasets with 14 species, which is the number we have sampled here. Future studies should aim to test for ecological and geographic structure across Bombacoideae, where a greater sample size of taxa can be obtained. In any case, our data suggest multiple shifts from dry to rain forests within *Ceiba* (Fig. 5) because rain forest species are nested within the two dry forest clades. For example, the two accessions representing *C. samauma*, occurring in rain and riverine forest in South America, are nested within the Central American and Mexican clade, and the rain forest species *C. lupuna* and *C. speciosa* are nested within the South American SDTF clade (Fig. 1 c and d).

The geographic structure between the Mexican and Central American clade and the South American clade for *Ceiba* is in line with the pattern reported for other tropical forest lineages (Fig. 1 a and c). Several tropical rain and dry forest lineages show patterns of a clade containing largely South American species recovered as sister to a clade containing largely Central, North American and Caribbean species, for example *Guatteria* Ruiz & Pav. (Annonaceae; Erkens et al. 2007), Neotropical *Manilkara* Adans. (Sapotaceae; Armstrong et al. 2014), Anacardiaceae (Weeks et al. 2014), Caricaceae (Carvalho and Renner 2012), and Lecythidaceae (Vargas and Dick 2020). Moreover, the direction of migration from South America to Central and North America has been well documented for characteristic elements of dry forests such as Malpighiaceae (Willis et al. 2014) and for the rain forest genus *Manilkara* (Armstrong et al. 2014). In fact, Willis et al. (2014) showed a striking increase by six fold in the frequency of migration events of dry forest lineages between South America and Mexico during the mid-Miocene (ca. 23.9 Ma). By contrast, the opposite direction of migration has been reported for the rain forest genus *Guatteria* in the late Miocene (Erkens et al. 2007) and for Caricaceae (Carvalho and Renner 2012). Our data suggest a dispersal event from North and Central America to South America in the late Miocene as *C. samauma* is nested within Mexican and Central American species. Interestingly, in addition to a long-distance dispersal event, our data suggest a biome shift, as *C. samauma* is a rain forest species nested within a dry forest clade. *Ceiba samauma* might represent an interesting exception to the well-established hypothesis by Donoghue (2008) that “it is easier to move than to evolve” (Fig. 1 b). Our results echo previous

studies showing the importance of dispersal in the assembly of tropical forests, which might be more frequent in rain forests (Dexter et al. 2017, Honorio Coronado et al. 2019) than in dry forests (Lavin et al. 2004, Pennington and Lavin 2016, Dick and Pennington 2019). Furthermore, biomes shifts from dry to rain forests have been reported for Bombacoideae, although less common than in the opposite direction (Zizka et al. 2020). Although the geographic and ecological structure in *Ceiba* is in line with other lineages, it is important to note the low support values between (1) the SDTF South American clade and the clade containing the SDTF Central America species plus *C. samauma*, (2) between *C. samauma* and *C. aesculifolia*, and (3) within the SDTF South America clade. The weakly supported backbones prevent us from accurately pinpointing ancestral areas, major lineages dispersal events or biome shifts for the genus and species within *Ceiba*. Nevertheless, the indication that *C. samauma* has moved and evolved combined with the fact that it was recovered as monophyletic makes it a very interesting species for further investigation of Neotropical evolution.

### *Biome-specific differences in the nature of species and their diversification trajectories*

For species in the SDTF clade, our results showed younger crown and stem ages than expected for SDTF lineages. Rain forest species such as *Ceiba pentandra* and *C. samauma* show patterns of long stems with shallow crown groups (Table 1). These patterns contrast to previous studies of individual SDTF species that showed legume species to be older, with stem ages of 5-10 Ma (e.g., Pennington et al. 2010, de Queiroz & Lavin 2011), and run contrary to the prediction of Pennington and Lavin (2016) that rain forest species might, on average, tend to have more recent origins. *Inga* (Leguminosae, c. 300 species) and *Guatteria* (Annonaceae, c. 265 species) are two of the largest genera of Neotropical rain forest lowland trees. Speciation events in *Inga* took place in the last 10 Ma, with many species arising in the past 2 Ma (Richardson et al. 2001). Erkens et al. (2007) estimated the crown age of *Guatteria* at 11.4 Ma and the crown age of the South America clade, containing the highest number of species in the genus, at between 6.1 - 7.4 Ma. Both genera show patterns of lack of coalescence for multiple accessions within species (Fig. 2 b). De-Nova et al. (2012) demonstrated contrast between the average age of dry and rain forest clades within *Bursera*, with those inhabiting SDTF at 7.49 Ma (SD = 5.31) and those inhabiting tropical rain forest at 11.16 Ma (SD = 4.75). However, De-Nova et al. (2012) did not

represent individual species with multiple accessions. Furthermore, using a densely sampled nrITS dataset for the legume genus *Luetzelburgia* Harms, de Queiroz et al. (2017) show that there are cases of non-coalescence of conspecific accessions. *Luetzelburgia* shows strong association with SDTF (Cardoso et al. 2014) and has a crown age estimated at 6.3 Ma (de Queiroz et al. 2017), much younger than the 12.7 Ma estimated here for *Ceiba* (8.2-17.6 [95% HPD]). The pattern of non-monophyly recovered for *Luetzelburgia* species might therefore be due to the relatively recent radiation of this genus (de Queiroz et al. 2017).

The stem age of *C. pentandra* was estimated at 12.7 Ma (8.2-17.6 [95% HPD], see Table 1). The long stem and shallow crown suggest this is an old rain forest species with low levels of extant genetic diversity among populations (see also Dick et al. 2007). Likewise, *C. samauma* was recovered as monophyletic, and has a crown node age estimated at 0.4 Ma (0.0001-1.4 [95% HPD]) and a stem node age of 6.3 Ma (3.1-9.8 [95% HPD]). Both species therefore contrast with the suggested predominant patterns of non-monophyly of conspecific accessions and lack of geographic structure for rain forest species of other Angiosperm plant families (Dexter et al. 2010, Honorio Coronado et al. 2014, Pennington and Lavin 2016). Our result, recovering *C. pentandra* as monophyletic, with low sequence divergence amongst accessions, is consistent with that of Dick et al. (2007), who showed *C. pentandra* to have extremely weak phylogeographic structure based on nrITS and chloroplast *psbB-psbF* for 51 individuals. In addition, the disjunct distribution of this species in Africa was demonstrated to be due to relatively recent long-distance dispersal because of low genetic divergence of the African populations. *Ceiba pentandra* occurs mainly in the vast continuous areas of the Amazon (Gibbs and Semir 2003, Dick et al. 2007) and, although large geographic areas have been correlated with species-rich clades (Donoghue and Sanderson 2015), it was recovered as a depauperate lineage. The Amazon was the stage of major landscape transformations, such as changes in the course of rivers (Hoorn and Wesselingh 2009), and those events may have caused extinctions over evolutionary time. Furthermore, Dick et al. (2007) reported that *Ceiba pentandra* has the weakest phylogeographic structure when compared to other widespread rainforest tree species. The weak phylogeographic structure implies a high level of gene flow that would lower the likelihood of speciation events (Coyne and Orr 2004). Therefore, a low diversification rate



combined with high extinction rate might be the cause of the depauperate lineage in *C. pentandra*, which may represent the only surviving lineage of a once more diverse clade. *Ceiba samauma* is an interesting case. It is widely and discontinuously distributed in humid and riverine forest from Bolivia and Peru to the Brazilian Amazon (Gibbs and Semir 2003). Two samples sequenced in this study came from populations occurring in semi-deciduous forests by the slopes of Andean valleys in Peru (KD6774 and KD6067). *Ceiba samauma* was recovered as a depauperate lineage. It is possible that *C. samauma* has effective pollen and seed flow similar to *C. pentandra*, which would make speciation events less likely. However, the literature on this species is scarce and further studies are necessary.

*Ceiba jasminodora* was recovered with stem age of 8.6 Ma (5.0-12.5 [95% HPD], see Fig. 6). *Ceiba jasminodora* occurs in the granitic campos rupestres (rocky upland vegetation) of the Serra do Espinhaço mountain range in Minas Gerais state, southeastern Brazil (Gibbs and Semir 2003). Campos rupestres are hyperdiverse habitats occurring in isolated patches throughout Brazil (Neves et al. 2018). Although campos rupestres harbour ca. 15% of the Brazilian vascular flora (Neves et al. 2018), little is known about the evolutionary history of these endangered landscapes (Hughes et al. 2013, Silveira et al. 2016). Campos rupestres are thought to be old and stable, containing a combination of old lineages and young species diversified *in situ* (Silveira et al. 2016, Inglis and Cavalcanti 2018), thus contrasting with the Cerrado biome, which is floristically highly related (Neves et al. 2018) but assembled relatively recently, with most lineages diversifying at 4 Ma or less (Simon et al. 2009). The biomes surrounding campos rupestres, such as the Cerrado, Amazon, Atlantic Forest, and SDTF, might be the source of plant lineages that colonised there (Neves et al. 2018). For example, the genus *Calliandra* Benth., common in campos rupestres, seems to have a SDTF origin (de Souza et al. 2013). However, unlike the Cerrado, colonisation of campos rupestres from surrounding biomes may not be evolutionarily recent. The stem nodes of *Calliandra* species occurring in campos rupestres are dated to the Miocene (de Souza et al. 2013). Our phylogenetic tree shows low support values in backbone of the clade containing *C. jasminodora* (Figs 5 and 6), and a better genetic sample combined with more accessions of this species will allow us to make future inferences about biome shifts in this lineage.

Within the two predominantly SDTF clades, there is little evidence of (1) lineages older than rain forest lineages, and (2) lineages representing morphologically recognized species with

long stems and monophyletic crown groups, as predicted by Pennington and Lavin (2016, Fig. 2 a). The crown age of the South American SDTF clade, containing 10 species, is estimated at 8.6 Ma (5.0–12.5 [95% HPD]) and the Mexican SDTF clade, containing 2 species, is estimated at 7.3 (3.7–11.0 [95% HPD]) Ma, with a stem age for both estimated at 11.2 (7.2–15.6 [95% HPD]) Ma. Only one species from SDTF, *Ceiba aesculifolia*, was recovered as monophyletic with a crown age estimated at 1.5 (0.2–3.2 [95% HPD]) Ma and a stem age at 6.3 Ma (3.1–9.8 [95% HPD]). Even when assigning a minimum age of 56 Ma to the *Eriotheca* stem, the same pattern is observed (Fig. S2). The crown age of the South American SDTF clade is estimated at 14.9 (8.1–22.0 [95% HPD]) Ma and the Mexican SDTF clade, containing two species, is estimated at 12.6 (6.5–19.5 [95% HPD]) Ma, with a stem age for both estimated at 19.5 (12.1–22.3 [95% HPD]) Ma. The crown age of *Ceiba aesculifolia* was estimated at 2.7 (0.4–5.6 [95% HPD]) Ma and stem age at 10.79 (5.1–17.0 [95% HPD]) Ma (Fig. S2).

The lack of resolution among the dry forest accessions, with most species being recovered as non-monophyletic, indicates absence of intraspecific coalescence for the nrITS locus. Explanations for this include incomplete lineage sorting after speciation events, paralogous gene copies, inaccurate species delimitation, and/or hybridisation followed by introgression (Naciri and Linder 2015, Pennington and Lavin 2016). We eliminated sequences with possible paralogues by visually inspecting the electropherograms and by comparing substitution rates along branch lengths following Bailey et al. (2003). Some species of *Ceiba* are hypothesised to be interfertile and hybridise (Gibbs and Semir 2003), especially within the *C. insignis* species aggregate. However, it is also suggested that those species diverge in time of anthesis and pollinator type as well, and we have seen no evidence of putative hybrids in the field (Pezzini, pers. obs.). Eight out of the ten species within the South American SDTF clade are from Brazil and of these, four are distributed in the *Caatinga*, the largest area of SDTF in the Neotropics (700,000 km<sup>2</sup>; Silva de Miranda et al. 2018). *Ceiba* species such as *C. pubiflora* are often widespread (Fig. 4) and abundant (Lima et al. 2010). Taken together, this evidence suggests that the non-monophyly of *Ceiba* species found in SDTF such as *C. pubiflora* may be a reflection of large effective population sizes and hence a longer time to coalescence (Naciri and Linder 2015, Pennington and Lavin 2016), rather than due to hybridisation or nrITS paralogy. Our study illustrates that the general patterns of species age, monophyly, and geographic structure reported for species belonging to the Leguminosae family in SDTF (Pennington and

Lavin 2016) are not shared by one of the most characteristic SDTF tree genera and suggests that further phylogenetic studies of unrelated groups are required.

### *Taxonomic implications*

Our data support: (i) the circumscription of *Chorisia* within *Ceiba*, as proposed by Gibbs et al. (1988), Ravenna (1998) and Gibbs and Semir (2003), and confirmed by recent molecular studies (Carvalho-Sobrinho et al. 2016); (ii) the non-monophyly of the *C. insignis* aggregate species as circumscribed by Gibbs & Semir (2003); and (iii) the monophyly of the section *Campylanthera* (Schott & Endl.) K.Schum. that includes the Central American species *C. aesculifolia*, *C. schottii* and the widespread Amazonian *C. samauma*. Our data suggest that *C. boliviana*, *C. erianthos* and *C. rubriflora*, not included by Gibbs and Semir (2003), are also part of the *insignis* complex (Fig. 5). Five of the seven species within this complex are restricted to the SDTF patches of South America, while *C. speciosa* is widespread, and *C. lupuna* occurs in riverine rain forests in the Peruvian and Brazilian Amazon (Fig. 4).

## **Acknowledgments**

This paper is part of the PhD thesis of F.F.P. who acknowledges the support of the Brazilian National Research Council (CNPq) for funding her PhD at the University of Edinburgh and the Royal Botanic Garden Edinburgh under the Science without Borders programme (206954/2014-0 - GDE). Fieldwork was partially supported by Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB - process APP0006/2011). F.F.P. thanks F.C. Diniz for fieldwork collaboration, A. Griffiths for help with the phylogenetic signal analysis, E. Koenen for insightful discussions, and M. Lavin for invaluable suggestions on earlier versions of the manuscript. We are thankful to R. Erkens and an anonymous reviewer for their thorough and constructive comments that have greatly improved the manuscript.

## **Author Contributions**

F.F.P. and R.T.P conceived the original idea; F.F.P. executed fieldwork, data collection, laboratory work and data analysis with input from K.G.D; F.F.P. led the writing of the

manuscript with input, comments and review from K.G.D., J.G.C.-S., C.A.K., J.A.N., L.P.Q., and R.T.P.

## Data Accessibility Statement

The data used in this study are archived at GenBank (Table S1). Scripts for analysis conducted here and for making figures are available in the GitHub page of F.F.P., available at <https://github.com/fpezzini>.

## Supplementary Material

The following materials are available as part of the online article at <https://escholarship.org/uc/fb>:

**Table S1.** Collection details of each accession and ecological preference for the species of *Ceiba*

**Figure S1.** Maximum clade credibility tree resulting from BEAST2 analysis of nuclear ribosomal ITS sequence data sets for 14 species of *Ceiba* for 47 Ma offset calibration, using Yule tree prior and the Uncorrelated Lognormal Distribution clock model

**Figure S2.** Maximum clade credibility tree resulting from BEAST2 analysis of nuclear ribosomal ITS sequence data sets for 14 species of *Ceiba* for 56 Ma offset calibration, using Yule tree prior and the Uncorrelated Lognormal Distribution clock model.

## References

- Álvarez, I. & Wendel, J.F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29, 417–434.
- Antonelli, A. & Sanmartín, I. (2011) Why are there so many plant species in the Neotropics? *Taxon*, 60, 403–414.
- Armstrong, K.E., Stone, G.N., Nicholls, J.A., Valderrama, E., Anderberg, A.A., Smedmark, J., Gautier, L., Naciri, Y., Milne, R. & Richardson, J.E. (2014) Patterns of diversification amongst tropical regions compared: a case study in Sapotaceae. *Frontiers in Genetics*, 5, 131–144.
- Bailey, C. (2003) Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Molecular Phylogenetics and Evolution*, 29, 435–455.
- Baker, T.R., Pennington, R.T., Magallon, S., et al. (2014) Fast demographic traits promote high diversification rates of Amazonian trees. *Ecology Letters*, 17, 527–536.
- Ballesteros, J., Reyes, K. & Racero, J. (2009) Estructura poblacional y etología de *Bradypus*

- variegatus* en fragmento de bosque seco tropical, Córdoba-Colombia. *Revista MVZ Córdoba*, 14, 1812–1819.
- Baum, D.A., Small, R.L. & Wendel, J.F. (1998) Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Systematic Biology*, 47, 181–207.
- Blomberg, S.P., Garland, T. & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57, 717–745.
- Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., et al. (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 15, e1006650.
- Buckler, E.S., Ippolito, A. & Holtsford, T.P. (1997) The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics*, 145, 821–32.
- Burnham, R.J. & Johnson, K.R. (2004) South American palaeobotany and the origins of Neotropical rainforests. *Philosophical Transactions of the Royal Society B*, 359, 1595–1610.
- Cardoso, D.B.O.S., de Queiroz, L.P. & de Lima, H.C. (2014) A taxonomic revision of the South American papilionoid genus *Luetzelburgia* (Fabaceae). *Botanical Journal of the Linnean Society*, 175, 328–375.
- Carvalho-Sobrinho, J.G., Alverson, W.S., Alcantara, S., Queiroz, L.P., Mota, A.C. & Baum, D.A. (2016) Revisiting the phylogeny of Bombacoideae (Malvaceae): novel relationships, morphologically cohesive clades, and a new tribal classification based on multilocus phylogenetic analyses. *Molecular Phylogenetics and Evolution*, 101, 56–74.
- Carvalho-Sobrinho, J.G. & de Queiroz, L.P. (2008) *Ceiba rubriflora* (Malvaceae: Bombacoideae), a new species from Bahia, Brazil. *Kew Bulletin*, 63, 649–653.
- Carvalho, F.A. & Felfili, J.M. (2011) Variações temporais na comunidade arbórea de uma floresta decidual sobre afloramentos calcários no Brasil Central: composição, estrutura e diversidade florística. *Acta Botanica Brasilica*, 25, 203–214.
- Carvalho, F.A. & Renner, S.S. (2012) A dated phylogeny of the papaya family (Caricaceae) reveals the crop's closest relatives and the family's biogeographic history. *Molecular Phylogenetics and Evolution*, 65, 46–53.
- Chernomor, O., von Haeseler, A. & Minh, B.Q. (2016) Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, 65, 997–1008.
- Condamine, F.L., Nagalingum, N.S., Marshall, C.R. & 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, 1–18.
- Couvreur, T.L., Forest, F. & Baker, W.J. (2011) Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. *BMC Biology*, 9, 44.
- Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer & Associates, Sunderland, Massachusetts, USA.
- Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, 458, 754–756.
- Daïnou, K., Mahy, G., Duminil, J., Dick, C.W., Doucet, J.L., Donkpégan, A.S.L., Pluijgers, M., Sinsin, B., Lejeune, P. & Hardy, O.J. (2014) Speciation slowing down in widespread and long-living tree taxa: insights from the tropical timber tree genus *Milicia* (Moraceae). *Heredity*, 113, 74–85.

- De-Nova, J.A., Medina, R., Montero, J.C., Weeks, A., Rosell, J.A., Olson, M.E., Eguiarte, L.E. & Magallón, S. (2012) Insights into the historical construction of species-rich Mesoamerican seasonally dry tropical forests: the diversification of *Bursera* (Burseraceae, Sapindales). *New Phytologist*, 193, 276–287.
- Dexter, K.G., Lavin, M., Torke, B.M., Twyford, A.D., Kursar, T.A., Coley, P.D., Drake, C., Hollands, R. & Pennington, R.T. (2017) Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of the National Academy of Sciences USA*, 114, 2645–2650.
- Dexter, K.G., Pennington, T.D. & Cunningham, C.W. (2010) Using DNA to assess errors in tropical tree identifications: how often are ecologists wrong and when does it matter? *Ecological Monographs*, 80, 267–286.
- Dick, C.W., Bermingham, E., Lemes, M.R. & Gribel, R. (2007) Extreme long-distance dispersal of the lowland tropical rainforest tree *Ceiba pentandra* L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology*, 16, 3039–49.
- Dick, C.W., Lewis, S.L., Maslin, M. & Bermingham, E. (2013) Neogene origins and implied warmth tolerance of Amazon tree species. *Ecology and Evolution*, 3, 162–169.
- Dick, C.W. & Pennington, R.T. (2019) History and geography of Neotropical tree diversity. *Annual Review of Ecology, Evolution, and Systematics*, 50, 279–301.
- Donoghue, M.J. (2008) A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences USA*, 105, 11549–11555.
- Donoghue, M.J. & Sanderson, M.J. (2015) Confluence, synnovation, and depauperons in plant diversification. *New Phytologist*, 207, 260–274.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence D. Penny (ed.). *PLoS Biology*, 4, e88.
- DRYFLOR (2016) Plant diversity patterns in Neotropical dry forests and their conservation implications. *Science*, 353, 1383–1387.
- Duarte, L. (1974) Sobre uma flor de Bombacaceae da Bacia Terciária de Fonseca, MG. *Anais da Academia Brasileira de Ciências*, 46, 407–411.
- Duarte, M.C., Esteves, G.L., Salatino, M.L.F., Walsh, K.C. & Baum, D. A. (2011) Phylogenetic analyses of *Eriotheca* and related genera (Bombacoideae, Malvaceae). *Systematic Botany*, 36, 690–701.
- Eiserhardt, W.L., Couvreur, T.L.P. & Baker, W.J. (2017) Plant phylogeny as a window on the evolution of hyperdiversity in the tropical rainforest biome. *New Phytologist*, 214, 1408–1422.
- Erkens, R.H.J., Chatrou, L.W., Maas, J.W., van der Niet, T. & Savolainen, V. (2007) A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. *Molecular Phylogenetics and Evolution*, 44, 399–411.
- Feliner, G.N. & Rosselló, J.A. (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution*, 44, 911–919.
- Fischer, E.A. (1997) The role of plumes in *Eriotheca pentaphylla* (Bombacaceae) seed survival in south-eastern Brazil. *Journal of Tropical Ecology*, 13, 133–138.
- Freckleton, Harvey & Pagel (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, 160, 712.
- Fritz, S.A. & Purvis, A. (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, 24, 1042–

1051.

- Gagnon, E., Ringelberg, J.J., Bruneau, A., Lewis, G.P. & Hughes, C.E. (2019) Global succulent biome phylogenetic conservatism across the pantropical Caesalpinia Group (Leguminosae). *New Phytologist*, 222, 1994–2008.
- Gibbs, P. & Semir, J. (2003) A taxonomic revision of the genus *Ceiba* Mill. (Bombacaceae). *Anales del Jardín Botánico de Madrid*, 60, 259–300.
- Gibbs, P.E., Semir, J. & Da Cruz, N.D. (1988) A proposal to unite the genera *Chorisia* Knuth with *Ceiba* Miller (Bombacaceae). *Notes from the Royal Botanic Garden Edinburgh*, 45, 125–136.
- Gradstein, F.M., Ogg, J.G., Schmitz, M.D. & Ogg, G.M. (2012) *The geologic time scale*, Elsevier. 1176 pp.
- Graham, C.H. & Fine, P.V.A. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, 11, 1265–1277.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59, 307–321.
- Honorio Coronado, E.N., Dexter, K.G., Hart, M.L., Phillips, O.L. & Pennington, R.T. (2019) Comparative phylogeography of five widespread tree species: insights into the history of western Amazonia. *Ecology and Evolution*, 7333–7345.
- Honorio Coronado, E.N., Dexter, K.G., Poelchau, M.F., Hollingsworth, P.M., Phillips, O.L. & Pennington, R.T. (2014) *Ficus insipida* subsp. *insipida* (Moraceae) reveals the role of ecology in the phylogeography of widespread Neotropical rain forest tree species. *Journal of Biogeography*, 41, 1697–1709.
- Hoorn, Carina & Wesselingh, F. (2010) *Amazonia: landscape and species evolution: a look into the past*. Wiley-Blackwell Publishing Ltd., Oxford, UK. 464 pp.
- Hudson, R.R. & Coyne, J.A. (2002) Mathematical consequences of the genealogical species concept. *Evolution*, 56, 1557–1565.
- Hughes, C.E., Pennington, R.T. & Antonelli, A. (2013) Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society*, 171, 1–18.
- Inglis, P.W. & Cavalcanti, T.B. (2018) A molecular phylogeny of the genus *Diplusodon* (Lythraceae), endemic to the campos rupestres and cerrados of South America. *Taxon*, 67, 66–82.
- Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.
- Knowles, L.L. & Carstens, B.C. (2007) Delimiting species without monophyletic gene trees. *Systematic Biology*, 56, 887–895.
- Koenen, E.J.M., Clarkson, J.J., Pennington, T.D. & Chatrou, L.W. (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.
- Kubitzki, K. & Ziburski, A. (1994) Seed dispersal in flood plain forests of Amazonia. *Biotropica* 26, 30.
- Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29, 1695–1701.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological

- phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773.
- Lavin, M. (2006) Floristic and geographical stability of discontinuous seasonally dry tropical forests explains patterns of plant phylogeny and endemism. In: *Neotropical savannas and seasonally dry forests: plant diversity, biogeography, and conservation* (ed. by R.T. Pennington, G.P. Lewis and J.A. Ratter), pp. 433–447. CRC Press, Boca Raton.
- Lavin, M., Pennington, R.T., Hughes, C.E., Lewis, G.P., Delgado-Salinas, A., Duno de Stefano, R., de Queiroz, L.P., Cardoso, D. & Wojciechowski, M.F. (2018) DNA sequence variation among conspecific accessions of the legume *Coursetia caribaea* reveals geographically localized clades here ranked as species. *Systematic Botany*, 43, 664–675.
- Lavin, M., Schrire, B.P., Lewis, G., Pennington, R.T., Delgado-Salinas, A., Thulin, M., Hughes, C.E., Matos, A.B. & Wojciechowski, M.F. (2004) Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philosophical Transactions of the Royal Society B*, 359, 1509–1522.
- de Lima, M.R. & Salard-Cheboldaeff, M. (1981) Palynologie des bassins de Gandarela et Fonseca (Eocene de l'état de Minas Gerais, Bresil). *Boletim IG*, 12, 33–54.
- Lima, M.S. de, Damasceno-Júnior, G.A. & Tanaka, M.O. (2010) Aspectos estruturais da comunidade arbórea em remanescentes de floresta estacional decidual, em Corumbá, MS, Brasil. *Revista Brasileira de Botânica*, 33, 437–453.
- Linares-Palomino, R. & Ponce Alvarez, S.I. (2005) Tree community patterns in seasonally dry tropical forests in the Cerros de Amotape Cordillera, Tumbes, Peru. *Forest Ecology and Management*, 209, 261–272.
- Louca, S. & Pennell, M.W. (2020) Extant timetrees are consistent with a myriad of diversification histories. *Nature*, 580, 502–505.
- Maddison, W. P. and D.R. Maddison. (2015) Mesquite: a modular system for evolutionary analysis. Version 3.2 <http://www.mesquiteproject.org>
- McKenna, D.D. & Farrell, B.D. (2006) Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. *Proceedings of the National Academy of Sciences USA*, 103, 10947–10951.
- Miles, L., Newton, A.C., DeFries, R.S., Ravilious, C., May, I., Blyth, S., Kapos, V. & Gordon, J.E. (2006) A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, 33, 491–505.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. & Lanfear, R. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534.
- Moreau, C.S. & Bell, C.D. (2013) Testing the museum versus cradle tropical biological diversity hypothesis: phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution*, 67, 2240–2257.
- Murphy, P.G. & Lugo, A.E. (1986) Ecology of tropical dry forest. *Annual Review of Ecology and Systematics*, 17, 67–88.
- Naciri, Y. & Linder, H.P. (2015) Species delimitation and relationships: the dance of the seven veils. *Taxon*, 64, 3–16.
- Neves, D.M., Dexter, K.G., Pennington, R.T., Bueno, M.L., de Miranda, P.L.S. & Oliveira-Filho, A.T. (2018) Lack of floristic identity in campos rupestres — a hyperdiverse mosaic of rocky montane savannas in South America. *Flora*, 238, 24–31.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular*



- Biology and Evolution, 32, 268–274.
- Orme, C.D.L. (2013) The caper package: comparative analyses in phylogenetics and evolution in R. [Http://Caper.R-Forge.R-Project.Org/](http://Caper.R-Forge.R-Project.Org/), 1–36.
- Pennington, R.T., Daza, A., Reynel, C. & Lavin, M. (2011) *Poissonia eriantha* (Leguminosae) From Cuzco, Peru: an overlooked species underscores a pattern of narrow endemism common to seasonally dry Neotropical vegetation. *Systematic Botany*, 36, 59–68.
- Pennington, R.T. & Lavin, M. (2016) The contrasting nature of woody plant species in different Neotropical forest biomes reflects differences in ecological stability. *New Phytologist*, 210, 25–37.
- Pennington, R.T., Lavin, M. & Oliveira-Filho, A. (2009) Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics*, 40, 437–457.
- Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K. & Butterworth, C.A. (2004) Historical climate change and speciation: Neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society B*, 359, 515–538.
- Pennington, R.T., Lavin, M., Sarkinen, T., Lewis, G.P., Klitgaard, B.B. & Hughes, C.E. (2010) Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences USA*, 107, 13783–13787.
- Pennington, R.T., Richardson, J.E. & Lavin, M. (2006) Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist*, 172, 605–616.
- Prata, E.M.B., Sass, C., Rodrigues, D.P., Domingos, F.M.C.B., Specht, C.D., Damasco, G., Ribas, C.C., Fine, P.V.A. & Vicentini, A. (2018) Towards integrative taxonomy in Neotropical botany: disentangling the *Pagamea guianensis* species complex (Rubiaceae). *Botanical Journal of the Linnean Society*, 188, 213–231.
- de Queiroz, L.P., Cardoso, D.B.O.S., Fernandes, M.F. & Moro, M.F. (2017) Diversity and evolution of flowering plants of the Caatinga Domain. In: *Caatinga: the largest tropical dry forest region in South America* (ed. by J.M.C. da Silva, I.R. Leal, & M. Tabarelli), pp. 23–63. Springer International Publishing, Cham.
- de Queiroz, L.P. & Lavin, M. (2011) *Coursetia* (Leguminosae) from eastern Brazil: nuclear ribosomal and chloroplast DNA sequence analysis reveal the monophyly of three Caatinga-inhabiting species. *Systematic Botany*, 36, 69–79.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67, 901–904.
- Rangel, T.F., Edwards, N.R., Holden, P.B., Diniz-Filho, J.A.F., Gosling, W.D., Coelho, M.T.P., Cassemiro, F.A.S.S., Rahbek, C. & Colwell, R.K. (2018) Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science*, 361, eaar5452.
- Raven, P.H., Gereau, R.E., Phillipson, P.B., Chatelain, C., Jenkins, C.N. & Ulloa Ulloa, C. (2020) The distribution of biodiversity richness in the tropics. *Science Advances*, 6, eabc6228.
- Ravenna, P. (1998) On the identity, validity, and actual placement in *Ceiba* of several *Chorisia* species (Bombacaceae), and description of two new south american species. *Onira* 3, 42–51.
- Renner, S.S. (2005) Relaxed molecular clocks for dating historical plant dispersal events. *Trends*

- in *Plant Science*, 10, 550–558.
- Richardson, J.E., Pennington, R.T., Pennington, T.D. & Hollingsworth, P.M. (2001) Rapid diversification of a species-rich genus of Neotropical Rain Forest trees. *Science*, 293, 2242–2245.
- Rieseberg, L.H. & Brouillet, L. (1994) Are many plant species paraphyletic? *Taxon*, 43, 21.
- Robyns, A. (1963) Essai de monographie du genre *Bombax* s.l. (Bombacaceae) (Suite). *Bulletin du Jardin botanique de l'État a Bruxelles*, 33, 145.
- Russel, P.M., Brewer, B.J., Klaere, S. & Bouckaert, R.R. (2019) Model selection and parameter inference in phylogenetics using nested sampling. *Systematic Biology*, 68, 219–233.
- Sanderson, M.J. (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, 19, 101–109.
- Sanderson, M.J. (2004) *r8s User's Manual*. Version 1.70, 1–42.
- Särkinen, T., Pennington, R.T., Lavin, M., Simon, M.F. & Hughes, C.E. (2012) Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. *Journal of Biogeography*, 39, 884–900.
- Schley, R.J., de la Estrella, M., Pérez-Escobar, O.A., 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.
- Segovia, R.A., Pennington, R.T., Baker, T.R., Coelho de Souza, F., Neves, D.M., Davis, C.C., Armesto, J.J., Olivera-Filho, A.T. & Dexter, K.G. (2020) Freezing and water availability structure the evolutionary diversity of trees across the Americas. *Science Advances*, 6, eaaz5373.
- Silva de Miranda, P.L., Oliveira-Filho, A.T., Pennington, R.T., Neves, D.M., Baker, T.R. & Dexter, K.G. (2018) Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America. *Global Ecology and Biogeography*, 27, 899–912.
- Silva, L.Á. da & Scariot, A. (2004a) Composição e estrutura da comunidade arbórea de uma floresta estacional decidual sobre afloramento calcário no Brasil central. *Revista Árvore*, 28, 69–75.
- Silva, L.Á. da & Scariot, A. (2004b) Comunidade arbórea de uma floresta estacional decídua sobre afloramento calcário na Bacia do rio Paran . *Revista  rvore*, 28, 61–67.
- Silveira, F.A.O., Negreiros, D., Barbosa, N.P.U., et al. (2016) Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. *Plant and Soil*, 403, 129–152.
- Simon, M.F., Grether, R., de Queiroz, L.P., Skema, C., Pennington, R.T. & Hughes, C.E. (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 USA*, 106, 20359–20364.
- Sioli, H. ed. (1984) *The Amazon: limnology and landscape ecology of a mighty tropical river and its basin*. Springer, Dordrecht, Netherlands.
- Skeels, A. & Cardillo, M. (2019) Reconstructing the geography of speciation from contemporary biodiversity data. *The American Naturalist*, 193, 240–255.
- Skilling, J. (2006) Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1, 833–859.
- Smith, B.T., McCormack, J.E., Cuervo, A.M., et al. (2014) The drivers of tropical speciation.

Nature, 515, 406–409.

- Sosa, V., De-Nova, J.A. & Vásquez-Cruz, M. (2018) Evolutionary history of the flora of Mexico: dry forests cradles and museums of endemism. *Journal of Systematics and Evolution*, 56, 523–536.
- de Souza, É.R., Lewis, G.P., Forest, F., Schnadelbach, A.S., van den Berg, C. & de Queiroz, L.P. (2013) Phylogeny of *Calliandra* (Leguminosae: Mimosoideae) based on nuclear and plastid molecular markers. *Taxon*, 62, 1200–1219.
- Stebbins, G.L. (1974) Flowering plants: evolution above the species level. The Belknap Press of Harvard University Press, Cambridge.
- ter Steege, H., Pitman, N.C.A., Sabatier, D., et al. (2013) Hyperdominance in the Amazonian tree flora. *Science*, 342, 1243092–1243092.
- Sun, Y., Skinner, D.Z., Liang, G.H. & Hulbert, S.H. (1994) Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics*, 89, 26–32.
- Vargas, O.M. & Dick, C.W. (2020) Diversification history of Neotropical Lecythidaceae, an ecologically dominant tree family of Amazon rain forest. In: *Neotropical diversification: patterns and processes* (ed. by V. Rull and A.C. Carnaval), pp. 791–809. Springer International Publishing, Cham, Switzerland.
- Weeks, A., Zapata, F., Pell, S.K., Daly, D.C., Mitchell, J.D. & Fine, P.V.A. (2014) To move or to evolve: contrasting patterns of intercontinental connectivity and climatic niche evolution in “Terebinthaceae” (Anacardiaceae and Burseraceae). *Frontiers in Genetics*, 5.
- Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, 19, 639–644.
- Wiens, J.J. & Penkrot, T.A. (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Systematic Biology*, 51, 69–91.
- Willis, C.G., Franzone, B.F., Xi, Z., Davis, C.C., Pennington, T. & Botanic, R. (2014) The establishment of Central American migratory corridors and the biogeographic origins of seasonally dry tropical forests in Mexico. *Frontiers in Genetics*, 5, 1–14.
- Zizka, A., Carvalho-Sobrinho, J.G., Pennington, R.T., Queiroz, L.P., Alcantara, S., Baum, D.A., Bacon, C.D. & Antonelli, A. (2020) Transitions between biomes are common and directional in Bombacoideae (Malvaceae). *Journal of Biogeography*, 47, 1310–1321.

Submitted: 23 July 2020

First decision: 24 August 2020

Accepted: 24 November 2020

Edited by Robert J. Whittaker