

1 **Formatted for: *Journal of Biogeography* (Deadline 4th of October 2021)**

2 **Running title: Dispersal and biome shifts in *Pterocarpus***

3

4 **The diversification of *Pterocarpus* (Leguminosae: Papilionoideae)**  
5 **was influenced by biome-switching and infrequent long-distance dispersal**

6 Rowan J. Schley<sup>1\*</sup>, Ming Qin<sup>2</sup>, Mohammad Vatanparast<sup>3,4</sup>, Panagiota Malakasi<sup>1</sup>, Manuel de la  
7 Estrella<sup>5</sup>, Gwilym Lewis<sup>1</sup>, Bente Klitgård<sup>1</sup>

8 <sup>1</sup>*Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK*

9 <sup>2</sup>*South China Botanical Garden, Chinese Academy of Sciences, 1190 Tianyuan Rd, Tianhe, Guangzhou, Guangdong, China*

10 <sup>3</sup>*Department of Geosciences and Natural Resource Management, University of Copenhagen, 1958, Denmark*

11 <sup>4</sup>*Department of Botany, National Museum of Natural History, Smithsonian Institution, MRC 166, Washington, DC 20560, USA*

12 <sup>5</sup>*Departamento de Botánica, Ecología y Fisiología Vegetal, Facultad de Ciencias, Universidad de Córdoba, 14014, Córdoba, Spain*

13 <sup>6</sup>*University of Exeter, North Park Road, Exeter, EX4 4QE, UK*

14

15

16 \*Corresponding author. *E-mail address:* [rowan.schley@gmail.com](mailto:rowan.schley@gmail.com) (R. J. Schley).

17

18

19

20 **Acknowledgements**

21 All the authors would like to thank the PAFTOL program at RBG, Kew for funding the sequencing of  
22 *Pterocarpus* clade accessions. MQ's lab visit was funded by the National Natural Science Foundation  
23 of China (grant no. 31670193) and the Science and Technology Planning Project of Guangdong  
24 Province (grant no. 2019B030316020).

25 **Abstract**

26 **Aim**

27 Phenotypes which evolved for dispersal over ecological timescales may lead to significant  
28 macroevolutionary consequences, such as infrequent long-distance dispersal and diversification in  
29 novel biomes. We aimed to reconstruct the phylogenetic history of *Pterocarpus* (Leguminosae/  
30 Fabaceae) to assess whether seed dispersal phenotypes and biome switching explain the current  
31 biogeographical patterns of this group.

32

33 **Location**

34 Pantropical

35

36 **Taxon**

37 The *Pterocarpus* clade, particularly *Pterocarpus* (Leguminosae/Fabaceae)

38

39 **Methods**

40 We sequenced ~300 nuclear loci captured using *Angiosperms-353*, a genomic ‘bait set’ for flowering  
41 plants, from which we generated a time-calibrated phylogenomic tree. To corroborate this, we also  
42 generated a time-calibrated phylogenetic tree from data-mined Sanger-sequencing data. We then  
43 collated distribution data and fruit dispersal morphology traits to compare trait-dependent and trait-  
44 independent biogeographical models, allowing us to assess whether dispersal traits influenced the  
45 spatio-temporal evolution of *Pterocarpus*. Finally, using the results of these model tests, we estimated  
46 the ancestral ranges and biomes of *Pterocarpus* species to better understand their biogeographical  
47 history, and assessed the degree and direction of biome switching over the course of *Pterocarpus*’  
48 diversification history.

49

50 **Results**

51 We recovered well-supported phylogenetic relationships within *Pterocarpus*, within which there were  
52 two subclades – one Neotropical and the other Palaeotropical. Our divergence date estimates  
53 suggested that *Pterocarpus* largely diversified from around 12 Ma, during the Miocene.  
54 Trait-dependent biogeographical models were rejected for both range and biome evolution within  
55 *Pterocarpus*, but models parameterising dispersal were supported. *Pterocarpus*’ ancestral node shared  
56 a range across the new-world and old-world tropics, followed by divergence into two clades, one  
57 palaeotropical and one neotropical. Biome switching occurred most frequently into rainforest and  
58 grassland.

59

## 60 **Main conclusions**

61 Overall, our analyses suggest that *Pterocarpus* underwent infrequent cross-continental dispersal and  
62 establishment into novel biomes. While this was minimally impacted by fruit dispersal syndromes,  
63 biome switching precipitated by long-distance dispersal and environmental change have played an  
64 important role in diversification within *Pterocarpus* since the Miocene.

65 Keywords: Angiosperms, Biomes, Dispersal, Diversification, Legumes, Phylogenomics, Traits, Tropics

66

## 67 Introduction

68 Dispersal is a major process structuring the geographic ranges of organisms, both at ecological and  
69 evolutionary timescales (Brown et al., 1996, Sanmartín & Ronquist, 2004). Indeed, adaptations which  
70 evolved to promote dispersal locally may end up being significant evolutionary innovations,  
71 influencing the biogeographical fate of lineages by allowing dispersal across continents and  
72 subsequent diversification over millions of years (Klaus & Matzke, 2020).

73 While vicariance has a major role in explaining biogeographical patterns, the role of dispersal was  
74 seen as being somewhat less significant (Sanmartín, 2012). However, there is now ample evidence  
75 that dispersal, and especially trans-oceanic dispersal, plays a major role in the biogeographical  
76 dynamics of a great number of terrestrial lineages (De Queiroz, 2005, Gillespie et al., 2012, Harris et  
77 al., 2018). This is especially evident in plants due to the wide range of traits that they have evolved to  
78 promote dispersal of their fruits, seeds and spores (Cousens et al., 2008). Seeds in particular have  
79 developed a wide range of morphological traits which allow dispersal over long distances, including  
80 wings and other aerodynamic structures (Cain et al., 2000), long dormancy periods (van der Pijl,  
81 1982), thick seed coats and air pockets which render the seeds buoyant, thus facilitating dispersal  
82 across oceans (Bellot et al., 2020, Renner, 2004). Multiple recent studies have highlighted that the  
83 biogeographical distributions of modern plant lineages are largely explained by traits which they  
84 evolved for effective dispersal in their ecological context (e.g. Podocarpaceae (Klaus & Matzke,  
85 2020); Annonaceae (Onstein et al., 2019)) using trait-dependent biogeographical modes. Indeed, it has  
86 been well established that a wide range of phenotypic traits have a significant impact on  
87 macroevolutionary dynamics (e.g. Rabosky et al., 2014).

88 Given their superlative biological diversity, the tropics are one of the most important places to  
89 examine the impact of dispersal on biogeography and diversification. Dispersal is particularly  
90 important for terrestrial organisms in the tropics because they are subject to both geographical and  
91 ecological constraints to their distribution. The tropics are largely separated into three zones  
92 (Neotropical, African and Asian/Australasian) by large areas of ocean, bounded to the north and south  
93 by increasingly unsuitable climates with increasing distance from the equator (Corlett & Primack,  
94 2011, Morley, 2000). There is ample evidence of disjunct ranges across continents within tropical  
95 organisms, many of which likely result from dispersal, particularly in plants (e.g., Arecaceae  
96 (Eiserhardt et al., 2011), Podocarpaceae (Klaus & Matzke, 2020), Leguminosae/Fabaceae (Schrire et  
97 al., 2005, Vatanparast et al., 2013), Vitaceae (Nie et al., 2012)).

98 Following dispersal, plants may subsequently encounter powerful selective pressures in contrasting  
99 habitats to those in which they evolved. This can drive adaptation, and eventually speciation, to these  
100 novel environments (Waser & Campbell, 2004). Biomes which span climatic extremes are found in  
101 relatively close proximity within the tropics, being defined by their ecological community  
102 composition (Pennington & Dick, 2004, Sanmartín, 2014) as well as by the physiognomy of their  
103 dominant plant species (Pennington & Ratter, 2006, Woodward et al., 2004). These biomes are often  
104 seen as independent evolutionary arenas (Hughes et al., 2013, Nürk et al., 2020, e.g. Pennington et al.,  
105 2009, Ringelberg et al., 2020). This biome concept reflects that environmental pressures select for  
106 species with similar functional attributes (Echeverría-Londoño et al., 2018), further implying that  
107 colonisation of contrasting biomes can promote speciation (Pennington & Dick, 2010). For example,  
108 many genera with species typical of seasonally dry tropical forest and savanna also contain species  
109 from rainforest (Pennington & Ratter, 2006, Pennington et al., 2000), suggesting that biome switching  
110 has driven ecological speciation.

111 The relative invasibility of certain biomes by immigrants may be higher than in others. Globally,  
112 lowland tropical ecosystems such as rainforest and savanna tend to be most permeable to colonisation  
113 by taxa from other biomes (Dexter et al., 2015), acting as ‘lineage sinks’ (Donoghue & Edwards,  
114 2014, Pennington & Hughes, 2014) because plant communities in these environments experience high  
115 species turnover through exposure to drought and fire (da Costa et al., 2010, Pennington & Lavin,  
116 2016). Such an imbalance in the direction of biome switching in tropical plant taxa is also likely to be  
117 an outcome of the relative size of the adaptive peak to be overcome by immigrants, coupled with the  
118 pre-existence of adaptive traits (Donoghue & Edwards, 2014). Again, drought and fire are likely to be  
119 modulators of the shape of such adaptive peaks, requiring specialised phenotypes to survive them (De  
120 Micco & Aronne, 2012, Midgley & Bond, 2013) and constraining plant evolution in the tropics  
121 (Olmstead, 2013).

122 Pantropical clades which occur in different biomes are an ideal system to study the biogeographical  
123 processes outlined above. The legume family (Leguminosae/Fabaceae) is replete with such clades,  
124 present in every terrestrial biome in the tropics (Lewis, G. P. et al., 2005) and dominant in many of  
125 them (de la Estrella et al., 2017, Gentry, 1988, ter Steege et al., 2013, ter Steege et al., 2020). The  
126 genus *Pterocarpus* Jacq. belongs to the largest subfamily of legumes (Papilionoideae) and consists of  
127 around 33 species (Klitgård & Lavin, 2005, Lewis, G. P., 1987, Rojo, 1972), many of which are  
128 important timber species or are used in traditional medicine (Saslis-Lagoudakis et al., 2011). This  
129 genus is diagnosed by either winged or corky (coriaceous) fruits adapted for dispersal by wind or  
130 water, respectively, and this is likely one reason why *Pterocarpus* is found pantropically (Klitgård &  
131 Lavin, 2005). *Pterocarpus* forms a part of the broader, pantropical ‘*Pterocarpus* clade’ which contains  
132 22 genera and ~200 species, of which most are trees (Klitgård et al., 2013).

133 Given its pantropical distribution, its presence in many contrasting biomes and its diversity of  
134 dispersal syndromes, *Pterocarpus* is an ideal system for evolutionary studies of biogeography, biome-  
135 switching and dispersal within the tropics (Klitgård et al., 2013). Accordingly, we aim to examine the  
136 impact of dispersal traits on the diversification of *Pterocarpus* to provide insights into how trait  
137 evolution can structure the biogeographical distribution of species, and so influence the evolution of  
138 tropical diversity. We also aim to infer the patterns of biome switching which may result from such

139 dispersal events, to help clarify how shifts between biomes may accumulate or generate diversity. To  
140 do this, we will reconstruct the relationships between species within the *Pterocarpus* clade using novel  
141 next-generation sequencing techniques, and infer a time-calibrated phylogeny for *Pterocarpus* using  
142 multiple approaches. Following this, we will test whether trait-dependent or trait-independent  
143 biogeographical models best explain extant biogeographical patterns, both in terms of continental  
144 realm and biome, and assess whether transitions between certain biomes are more frequent than  
145 others. As such, we predict that:

- 146 1) Extant biogeographical patterns in *Pterocarpus* were significantly influenced by their seed  
147 dispersal traits, and hence that a trait-dependent model will best describe their spatiotemporal  
148 evolution.
- 149 2) Biome shifts will be more frequent into rainforests and savannas, given their ecologically  
150 dynamic nature and resulting high invasibility by taxa dispersing into them.

151

## 152 Materials and Methods

### 153 Taxon sampling

154 Novel phylogenomic data were generated using target capture sequencing for 26/33 accepted  
155 *Pterocarpus* species (75%), which were sampled across the distributional range of the genus. In total,  
156 27 accessions were sampled within *Pterocarpus*, detailed in Supporting Information, Appendix S1,  
157 Table S1.1. A further 22 outgroup taxa from the 21 other genera belonging to the *Pterocarpus* clade  
158 within which *Pterocarpus* is nested were sampled for phylogenetic analysis. In addition, Sanger  
159 sequence data from one nuclear locus (nrITS) and four plastid loci (*matK*, *ndhF-rpl32*, *rbcl*, *trnL-*  
160 *trnF*) were downloaded from NCBI GenBank <https://www.ncbi.nlm.nih.gov/genbank/> (Appendix S1,  
161 Table S1.2) for comparison between phylogenetic dating methods. These data represented 28/33  
162 *Pterocarpus* species, along with 8 outgroup species from the *Pterocarpus* clade.

163 A species list of all accepted *Pterocarpus* taxa was compiled using IPNI (IPNI, 2020) and a generic  
164 monograph of *Pterocarpus* (Rojo, 1972) to ensure that taxonomically accepted species were included  
165 in analyses, and to account for synonyms and infraspecific taxa. Voucher specimens were examined,  
166 and their determination updated as appropriate. Newly described and recently reinstated species from  
167 within the *P. rohrii* complex were also sampled (Klitgård et al., *in prep*). Leaf material for DNA  
168 extraction was acquired from both silica material ( $N= 1$ ) and herbarium specimens ( $N= 48$ ) collected  
169 from the BM, K and MO herbaria.

170

### 171 Library preparation and sequencing

172 DNA was extracted from 20 mg of dried leaf material with the CTAB method (Doyle & Doyle, 1987),  
173 following which DNA concentrations were measured using a Quantus fluorometer (Promega,  
174 Wisconsin, USA). Libraries were prepared using the NEBNext® Ultra™ II DNA Kit (New England  
175 Biolabs, Massachusetts, USA). Samples with high molecular weight DNA (>1000 bp, measured with

176 a TapeStation 4200 (Agilent Technologies, California, USA)) were sheared with a Covaris focussed  
177 ultrasonicator M220 (Covaris, Massachusetts, USA). Following this, end-preparation and Illumina  
178 adaptor ligation were undertaken according to the NEBNext protocol, including a 400bp size-  
179 selection step with Agencourt Ampure XP magnetic beads (Beckman Coulter, California, USA).  
180 Libraries containing a range of insert sizes were then amplified using PCR, following the protocol  
181 outlined in Appendix S1, Table S1.2. Targeted bait capture was performed using the MyBaits protocol  
182 (Arbor Biosciences, Michigan, USA) to target 353 phylogenetically informative nuclear genes with  
183 *Angiosperms-353*, a bait kit designed for all Angiosperms. The final library pools were sequenced  
184 using a paired-end 150bp run on the Illumina HiSeq platform by Macrogen 154 Inc. (Seoul, South  
185 Korea).

186

### 187 Quality filtering, read assembly and alignment

188 DNA sequencing reads were quality-checked with FASTQC v0.11.3 (Andrews, 2010) and were  
189 trimmed using TRIMMOMATIC v.0.3.6 (Bolger et al., 2014) to remove adapter sequences and quality-  
190 filter reads. TRIMMOMATIC settings permitted <4 mismatches, a palindrome clip threshold of 30 and a  
191 simple clip threshold of 6. Bases with a quality score <28 and reads shorter than 36 bases long were  
192 removed from the dataset. Following quality-filtering, loci were assembled using SPADES v3.11.1  
193 (Bankevich et al., 2012) by the HybPiper pipeline v1.2 (Johnson et al., 2016) with a coverage cut-off  
194 of 8x. All loci with <50% recovery among taxa were removed, and potentially paralogous loci were  
195 removed using the Python (Python Software Foundation, 2010) script '*paralog\_investigator.py*',  
196 distributed with the HybPiper pipeline. Recovery of different loci was visualised using the  
197 '*gene\_recovery\_heatmap.R*' script distributed with HybPiper.

198 Following this, targeted loci were aligned by gene region (excluding those with potential paralogs)  
199 using 1,000 iterations in MAFFT (Katoh & Standley, 2013) with the '*—adjustdirectionaccurately*'  
200 option to incorporate reversed sequences. These alignments were then cleaned using the `□automated1`  
201 flag in TRIMAL (Capella-Gutiérrez et al., 2009), following which they were visually inspected for poor  
202 sequence quality using *Geneious* v. 8.1.9 (<https://www.Geneious.com>). This was done in order to  
203 remove taxa with mostly missing data and to prevent the inclusion of poorly recovered loci into the  
204 dataset, resulting in 303 refined alignments.

205 The nrITS and four plastid Sanger sequencing loci downloaded from NCBI GenBank were refined to  
206 include 68 sequences belonging to 36 species from nine genera (GenBank numbers for these  
207 sequences are listed in Appendix S1, Table S1.3). These data were aligned using MAFFT with the  
208 same parameters as above. These sequences were also subject to visual data inspection and trimming  
209 in *Geneious* to minimise missing sequence data where multiple accessions of a species were present.

210

### 211 Phylogenomic inference and divergence dating

212 Three hundred and three gene trees were inferred using RAXML v.8.0.26 (Stamatakis, 2014) with  
 213 1,000 rapid bootstrap replicates and the GTRCAT model of nucleotide substitution. A species tree  
 214 was generated based on the best-scoring RAXML trees using ASTRAL v.5.6.1 under the default  
 215 parameters, with monophyly not enforced ('-a' flag) (Zhang, C. et al., 2018). To compare between  
 216 phylogenetic methods, we then concatenated all 303 target capture locus alignments into a single  
 217 alignment using AMAS (Borowiec, 2016), and inferred a phylogenetic tree using RAXML HPC2  
 218 (Stamatakis, 2014) on the CIPRES web portal ((Miller et al., 2010), <https://www.phylo.org/>).  
 219 Inference was performed using 1,000 rapid bootstrap replicates and the GTRCAT model of nucleotide  
 220 substitution.

221 MCMCTREE, from the PAML package (Yang, 2007), was used to date our phylogenomic tree. This  
 222 program allows Bayesian divergence time estimation from large phylogenomic datasets based on a  
 223 pre-determined phylogenetic tree (Dos Reis & Yang, 2013). As such, the ASTRAL tree was rooted  
 224 with an appropriate outgroup (*Casuarinia astragalina*) and the tree's branch lengths were removed  
 225 with the R (R Development Core Team, 2013) package 'PhyTools' (Revell, 2012). This rooted,  
 226 branchless tree and the concatenated 303-locus target capture alignment were then used as inputs for  
 227 MCMCTREE. Fossil calibrations (listed in Table 1) were added using skew-normal priors as  
 228 uncertainty is distributed asymmetrically across the mean age in fossil calibrations (Ho & Phillips,  
 229 2009) using the R package 'MCMCtreeR' (Puttick, 2019).

230 The MCMCTREE analysis was run by first calculating the gradient (b) and Hessian (h) parameters  
 231 using the HKY substitution model, from which the posteriors of divergence times and rates were  
 232 estimated using two independent runs of 20 million generations to ensure MCMC convergence. A  
 233 third, identical run with only priors was performed and compared to the runs with data to assess  
 234 whether prior settings were adequate. MCMC convergence for posterior estimation was checked using  
 235 TRACER (Rambaut et al., 2015), as well as in the R packages 'ape' and 'bppr' (Angelis & Dos Reis,  
 236 2015, Paradis & Schliep, 2019). The final time-calibrated phylogenetic tree estimated with  
 237 MCMCTREE was then plotted with 'MCMCtreeR'.

238

239 **Table 1:** Fossils used to generate a time-calibrated phylogenomic tree of the Pterocarpus clade in MCMCTREE  
 240 and a dated phylogenetic tree in BEAST v.1.8.0. Skewnormal calibration priors were used for fossil ages in  
 241 MCMCTREE to provide a minimum age estimate, and the corresponding Lognormal priors were used for the  
 242 BEAST analysis. Mean ages are shown for calibrations, with 95% confidence intervals in parentheses. All ages  
 243 are in millions of years (Ma).

| Calibration point                          | Age (Ma)            | Prior distribution (MCMCTREE) | Prior distribution (BEAST) | Fossil  |
|--|---------------------|-------------------------------|----------------------------|---|
| Pterocarpus clade                          | 37.5<br>(34-41)     | Skewnormal                    | Lognormal                  | <i>Luckowcarpa gunnii</i><br>(Martínez, 2018) |
| <i>Pterocarpus-Riedeliella</i> crown group | 21.8<br>(15.9-27.8) | Skewnormal                    | Lognormal                  | <i>Pterocarpus tertarius</i><br>(Buzek, 1992) |



|                               |                  |            |                                |   |
|-------------------------------|------------------|------------|--------------------------------|---|
| <i>Tipuana</i><br>crown group | 8.2<br>(7.9-8.5) | Skewnormal | Not used<br>(no suitable taxa) | <i>Tipuana ecuatoriana</i><br>(Burnham, 1995) |
|-------------------------------|------------------|------------|--------------------------------|---|

244

245 The GenBank dataset of five Sanger-sequenced loci was used to perform corroborative divergence  
246 time estimation in BEAST v.1.8.0 (Drummond & Bouckaert, 2015) on the CIPRES web portal.  
247 Substitution rate models were chosen for each of the five loci using AICc in JMODELTEST (Darriba et  
248 al., 2012), and are outlined in Appendix S1, Table S1.4. Tree model choice was based on path  
249 sampling analysis output from BEAST, using 10 million generations with 30 path steps. Two fossil  
250 calibration points from the MCMCTREE analysis were used to provide date estimates in BEAST  
251 (Table1). BEAST comprised three independent runs of 50 million generations which were combined  
252 using LOGCOMBINER v.1.8.2 with a burn-in level of 10%, using a birth-death tree model and a  
253 lognormal clock model. Convergence of runs was assessed using effective sample size (ESS) of all  
254 parameters in TRACER v. 1.6 with a threshold value of 200. Final ultrametric trees were generated in  
255 TREEANNOTATOR v1.8.2 and visualised using FIGTREE v1.4.4 (Rambaut, 2014). Further information  
256 on BEAST analysis is available in Appendix S1, Supplementary Methods.

257

### 258 Biogeographical analyses

259 We tested trait-dependent biogeographical models in the R package ‘BioGeoBEARS’ (Matzke, 2013)  
260 to understand the relationships between seed dispersal traits, biome switching and the spatio-temporal  
261 evolution of *Pterocarpus*. Fruit morphology traits were collected from the first monograph of  
262 *Pterocarpus* (Rojo, 1972), a taxonomic revision circumscribing new species within the *P. rohrii*  
263 complex (Klitgård *et al. in prep*), and taxonomic treatments of various *Pterocarpus* species (de  
264 Candolle, 1825, Klitgård *et al.*, 2000, Zamora, 2000). Since no fruiting specimen of *P. rohrii* var.  
265 *rubiginosus* exists, it was inferred to be winged since it belongs to the *P. rohrii* species complex. Fruit  
266 dispersal traits were then plotted on the *Pterocarpus* ASTRAL tree with biome and realm for each  
267 species using the *dotplot()* function in the R package ‘PhyTools’.

268 Geographical distribution data for non-cultivated *Pterocarpus* specimens were collected from GBIF  
269 ([www.GBIF.org](http://www.GBIF.org) (06 November 2019), GBIF Occurrence Download  
270 <https://doi.org/10.15468/dl.tr3h2y>) based on the curated taxonomic name list described above. This  
271 dataset was then further refined using the R package ‘CoordinateCleaner’ ((Zizka *et al.*, 2019);  
272 <https://github.com/ropensci/CoordinateCleaner>) and the input for our *BioGeoBears* analysis was  
273 produced from these data with Alex Zizka’s ‘biogeography in R’ scripts  
274 ([https://github.com/azizka/Using\\_biodiversity\\_data\\_for\\_biogeography](https://github.com/azizka/Using_biodiversity_data_for_biogeography)). These records and the  
275 specimens underpinning them were then examined in detail by taxonomic experts on the *Pterocarpus*  
276 clade (B. B. Klitgaard and G. P. Lewis) to further remove records which were mis-identified.

277 Based on these data, we assigned species to biogeographical regions using the *wwfLoad()* function in  
278 ‘speciesgeocodeR’ (Töpel *et al.*, 2017), based on the WWF ecoregions of Olson *et al.* (2001). These  
279 ecoregions represented five tropical terrestrial ‘realms’ (Australasia, Afrotropics, Indo-Malaya,  
280 Neotropics and Oceania) and six tropical terrestrial biomes (tropical moist forest (i.e., ‘rainforest’),



281 montane grassland, desert, mangrove, tropical dry forest and tropical grassland (i.e., ‘savanna’),  
282 which are displayed in Appendix S1, Fig. S1.1a. The assignments for each species were cross-  
283 referenced with monographs and previous studies of *Pterocarpus* (Rojo, 1972, Saslis-Lagoudakis et  
284 al., 2011, Klitgård et al., in prep), and with the Plants Of the World Online database  
285 (<http://powo.science.kew.org/>). All traits (fruit morphology, realm assignment and biome assignment)  
286 are then visualised onto the *Pterocarpus* phylogenetic tree in Appendix S1, Fig. S1.1b.

287 We then compared the fit of trait-independent and trait-dependent biogeographical models outlined by  
288 Klaus and Matzke (2020) for biogeographical realms and biomes in BioGeoBEARS. For realms, we  
289 tested between three default models in BioGeoBEARS, all of which incorporate dispersal, extinction  
290 and range switching. These models were DEC (Ree & Smith, 2008), which additionally accounts for  
291 vicariance and ‘subset’ sympatric speciation, DIVALIKE (Ronquist, 1997), which also accounts for  
292 vicariance, and BAYAREALIKE (Landis et al., 2013), which additionally accounts for widespread  
293 sympatric speciation but not vicariance. For biomes, we assessed the fit of Markov-*k* (Mk) models  
294 (Lewis, P. O., 2001) (i.e., BAYAREALIKE a+, d=e=0 in BioGeoBEARS) to accommodate equal  
295 rates of character evolution, after Kriebel et al. (2019). Markov-*k* more accurately models the  
296 inheritance of ecological preferences in daughter species and avoids bias towards large ancestral  
297 ranges compared to models such as DEC. We also included four trait-dependent models in our  
298 comparisons. The first two models only parameterised trait switching (‘*Trait\_1Rate*’ and  
299 ‘*Trait\_2Rates*’). Then, for realm, a trait-dependent extension of the DEC model was used to include  
300 trait switching and multipliers on dispersal rate (*DEC\_t12\_t21\_m2*), and for biome a trait-dependent  
301 implementation of the Markov-*k* model was used, which included the same trait switching and  
302 dispersal parameters (‘*Markov-k\_t12\_t21\_m2*’).

303 The best-fit model was chosen using AICc and Akaike weights, from which ancestral ranges were  
304 estimated across *Pterocarpus*. Ancestral ranges were plotted onto the time-calibrated tree from  
305 MCMCTREE using the function ‘*plot\_BioGeoBears\_results()*’ for realm and biome independently.  
306 Finally, for biome, we counted the number of transitions between the most likely ancestral biomes at  
307 each node, and produced a matrix of biome shifts. This matrix was visualised as a transition plot using  
308 the *plotmat* function in the R package ‘*diagram*’ (Soetaert, 2012). Further information regarding  
309 biogeographical data collection and analysis is available in Appendix S1, Supplementary Methods.

310

## 311 Results

### 312 *Phylogeny and divergence dating*

313 A mean of 17.7% of reads were successfully assembled onto the *Angiosperms-353* bait sequences,  
314 with a range from 6.9% to 30.5%. Gene recovery is shown as a heatmap in Appendix S1, Fig. S1.2.  
315 Of the 353 loci targeted by *Angiosperms-353*, we discarded 35 loci which were potentially paralogous  
316 and 15 which were assembled poorly or had over 50% missing data across taxa.

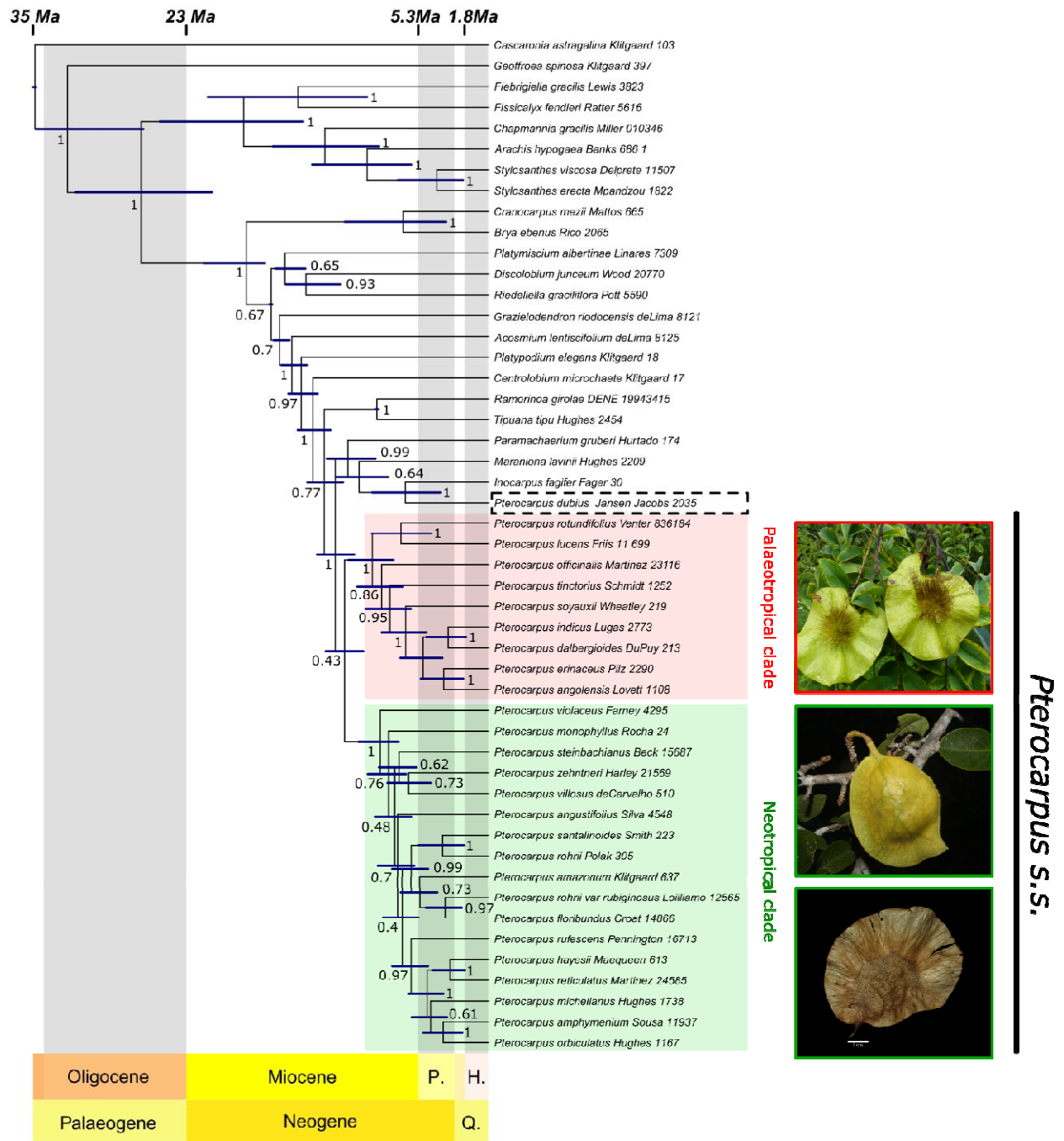
317 Our ASTRAL analysis indicated that most bipartitions were well supported with local posterior  
318 probabilities (LPP) >0.8, with 1 being full support (Fig. 1) and a quartet score of 0.68. Within this  
319 tree, most generic splits are well supported (LPP >0.9) and largely support the monophyly of

320 *Pterocarpus* with the exception of one species, *Pterocarpus dubius* (formerly *Etaballia dubia*).  
321 *Pterocarpus dubius* was resolved in a clade containing *Inocarpus*, *Maraniona* and *Paramachaerium*,  
322 which was in a sister position to *Pterocarpus*. Within *Pterocarpus*, there were two major subclades,  
323 one containing species from the Neotropics (the neotropical clade) and one containing species from  
324 the Palaeotropics (the palaeotropical subclade). However, this split received a relatively low support  
325 (LPP=0.43, quartet score = 34.48%). The concatenated RAXML analysis of the same 303 loci  
326 (Appendix S1, Fig. S1.3) recovered a nearly identical topology to the ASTRAL species tree (Fig. 1),  
327 and was much better resolved, with nearly all nodes having bootstrap values >95. Moreover, the split  
328 between neotropical and palaeotropical *Pterocarpus* species received full support (BS=100) in the  
329 inference based on the concatenated dataset.

330 Divergence time estimation carried out in MCMCTREE provided a robust, time-calibrated tree of the  
331 *Pterocarpus* clade. The two MCMC runs were convergent, with every node showing a suitable  
332 effective sample size (> 200), and posterior estimates fit the priors well as indicated in Appendix S1,  
333 Fig. S1.4a, b and c. The time-calibrated phylogenomic tree in Fig. 1 suggests that the *Pterocarpus*  
334 clade arose ~35 Ma, during the late Eocene, and diversification of present-day genera occurred during  
335 the Miocene, from 23-5.3 Ma. Of these, *Pterocarpus (sensu stricto)* diverged from its sister genera  
336 ~12 Ma, and diversified mostly in the late Miocene. It is worth noting that divergence date estimates  
337 are described by larger HPD intervals in the earlier divergence events within the *Pterocarpus* clade,  
338 and this increased uncertainty is mirrored in Appendix S1, Fig. S1.5. This figure indicates that  
339 uncertainty is greater in the earlier divergence events inferred, as well as being much reduced around  
340 fossil calibration points, as would be expected.

341

342 **Figure 1:** Fossil-calibrated MCMCTREE of the *Pterocarpus* clade. Node bars show 95% highest posterior  
343 density (HPD) for node ages. Local posterior probabilities (LPP), a measure of topological support, are shown  
344 for each node, and were taken from the ASTRAL analysis which provided the tree topology for the MCMCTREE  
345 analysis. Shading on the background of the phylogram represents geological epochs, which are labelled beneath  
346 the tree. The red and green shading indicates species belonging to the palaeotropical and neotropical subclades  
347 of *Pterocarpus*, respectively, and the extent of the genus *sensu stricto* is shown by the black bar. The dashed  
348 box around *Pterocarpus dubius* highlights its paraphyly with the rest of the genus. Photographs show winged  
349 fruits of *Pterocarpus erinaceus* (© Gwilym Lewis), coriaceous fruits of *Pterocarpus monophyllus* (© Domingos  
350 Cardoso) and winged fruits of *Pterocarpus violaceus* (© Reinaldo Aguilar, CC BY-NC-SA 2.0).



351

352

353 The topology of the BEAST trees (Appendix S1, Fig. S1.6) inferred using Sanger-sequencing data  
 354 from GenBank recovered a broadly similar topology to the MCMCTREE analysis (Fig. 1). Moreover,  
 355 the divergence time estimates of this phylogeny were largely congruent with those from Fig. 1,  
 356 including an origin of the Pterocarpus clade within the late Eocene (~38 Ma), followed by most  
 357 divergence between genera, and diversification of *Pterocarpus*, within the Miocene (~23-5.3 Ma).  
 358 However, the age estimate for the initial divergence of *Pterocarpus* is later than that inferred with  
 359 MCMCTREE, at around 25 Ma, again during the Miocene. In addition, the BEAST analysis displayed

360 a much broader degree of uncertainty on age estimates, as indicated by the breadth of the 95% HPDs,  
361 and recovered *Pterocarpus dubius* as nested within *Pterocarpus*, whereas in the MCMCTREE analysis  
362 *P. dubius* is in sister clade position to the rest of *Pterocarpus*.

363

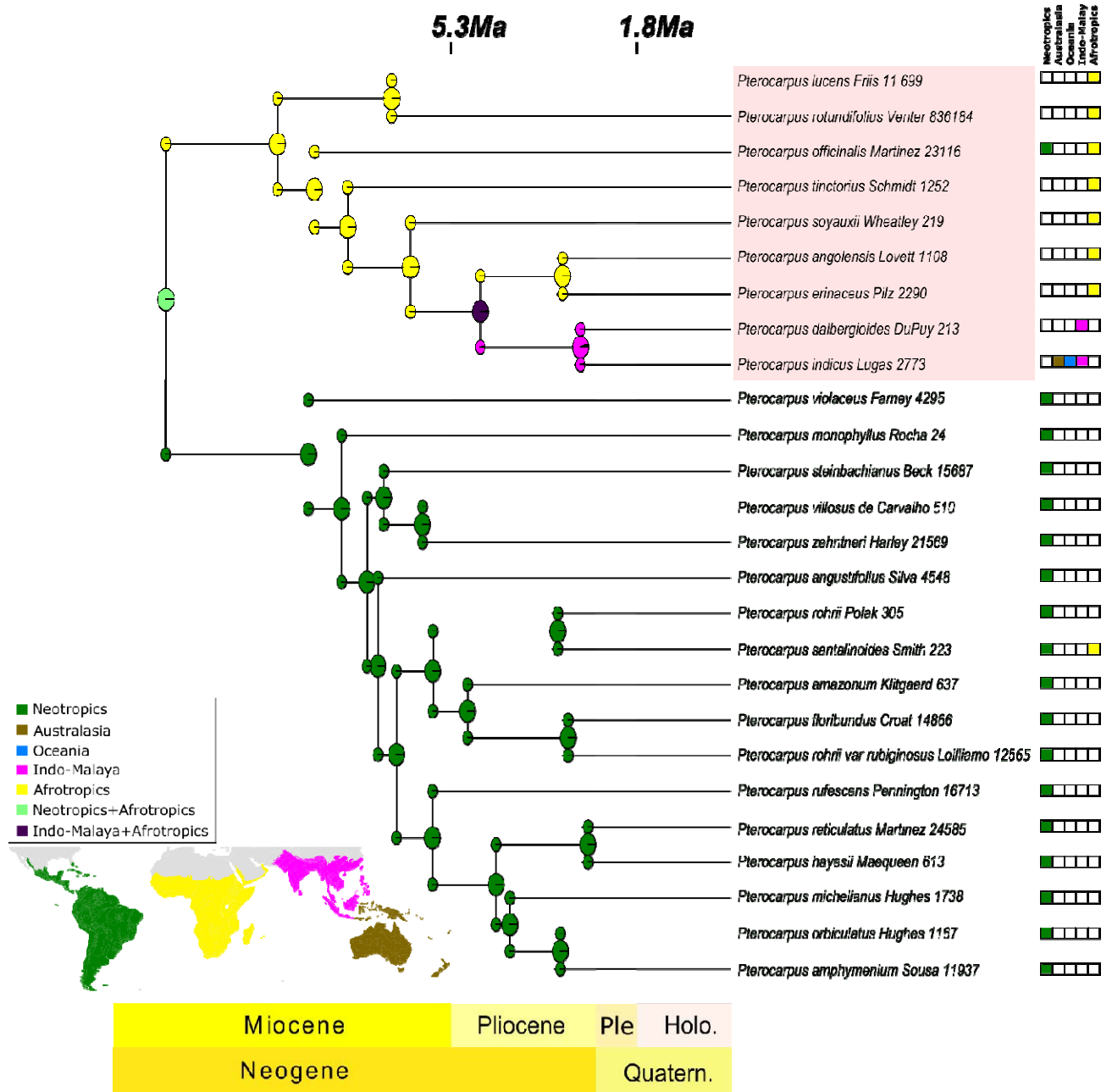
#### 364 Biogeography

365 AICc comparison indicated that the best-fitting BioGeoBEARS model was DIVALIKE (AICc=  
366 52.785) (Appendix S1, Table S1.5). Using DIVALIKE, we estimated a shared ancestral range of both  
367 the Neotropics and Afrotropics for the common ancestor of *Pterocarpus*, with subsequent splitting  
368 into neotropical and palaeotropical clades (Fig. 2a). The neotropical clade diversified only within the  
369 Americas and gave rise to the most species. Within the palaeotropical clade the common ancestor of  
370 *P. indicus* and *P. dalbergioides* appears to have dispersed into Asia, and subsequently into Australasia  
371 and Oceania, having a shared Afro-Asian range. It is interesting to note that within both neotropical  
372 and palaeotropical clades, there are species which share ranges, being found in both eastern South  
373 America and west Africa (*P. santalinoides* and *P. officinalis*). A shared characteristic of these two  
374 amphi-Atlantic species are thick-walled, wingless, buoyant fruits which allow water-borne dispersal.  
375 The current distribution of *Pterocarpus* species (both in terms of biome and realm), as well as the  
376 distribution of fruit dispersal traits across the ASTRAL tree, are visualised in Appendix S1, Fig.  
377 S1.1b.

378 For biome evolution, the best-fit model was the trait-independent implementation of Markov-*k*  
379 (AICc= 328.218, Appendix S1, Table S1.5). Moist forest, dry forest and grassland were inferred as  
380 being nearly equally likely as the ancestral biome for the MRCA of *Pterocarpus* (Fig. 2b). Within the  
381 neotropical clade of *Pterocarpus*, there appears to be multiple polymorphic biome preferences, with  
382 species being found in dry forest as well as one other habitat type (desert, grassland or moist forest).  
383 However, for the majority of stem nodes within the neotropical *Pterocarpus* species, the most likely  
384 ancestral biome appears to have been dry forest, with only the clade containing *P. rohrii*, *P.*  
385 *santalinoides*, *P. amazonum* and *P. floribundus* containing species only found in moist forest. Within  
386 the palaeotropical *Pterocarpus* clade, the most probable state for the majority of ancestral nodes was  
387 shared between moist forest and grassland, with most species being found in grassland in addition to  
388 other biomes, mirroring the high polymorphism in neotropical *Pterocarpus* species. It is also of  
389 interest that one of the two *Pterocarpus* species which inhabit both South America and Africa (*P.*  
390 *officinalis*) occurs both in tropical rainforest as well as mangroves.

391 We inferred the most shifts between biome states across the *Pterocarpus* tree into moist forests (i.e.,  
392 rainforest) and grassland (i.e., savanna). Specifically, we recovered the most shifts from dry forest to  
393 moist forest (n=4, Fig. 3), followed by shifts from rainforest to grassland (n=3). Transitions also  
394 occurred between drier habitats (e.g. inferred shifts from dry forest to desert and grassland (n=2), or  
395 grassland to dry forest and desert (n=1)), between similar habitats (e.g., grassland into montane  
396 grassland (n=2), from moist forest to mangrove (n=1)) and in the contrasting direction to the most  
397 frequent shifts (e.g. moist forest to dry forest (n=2) and grassland to moist forest (n=1)).

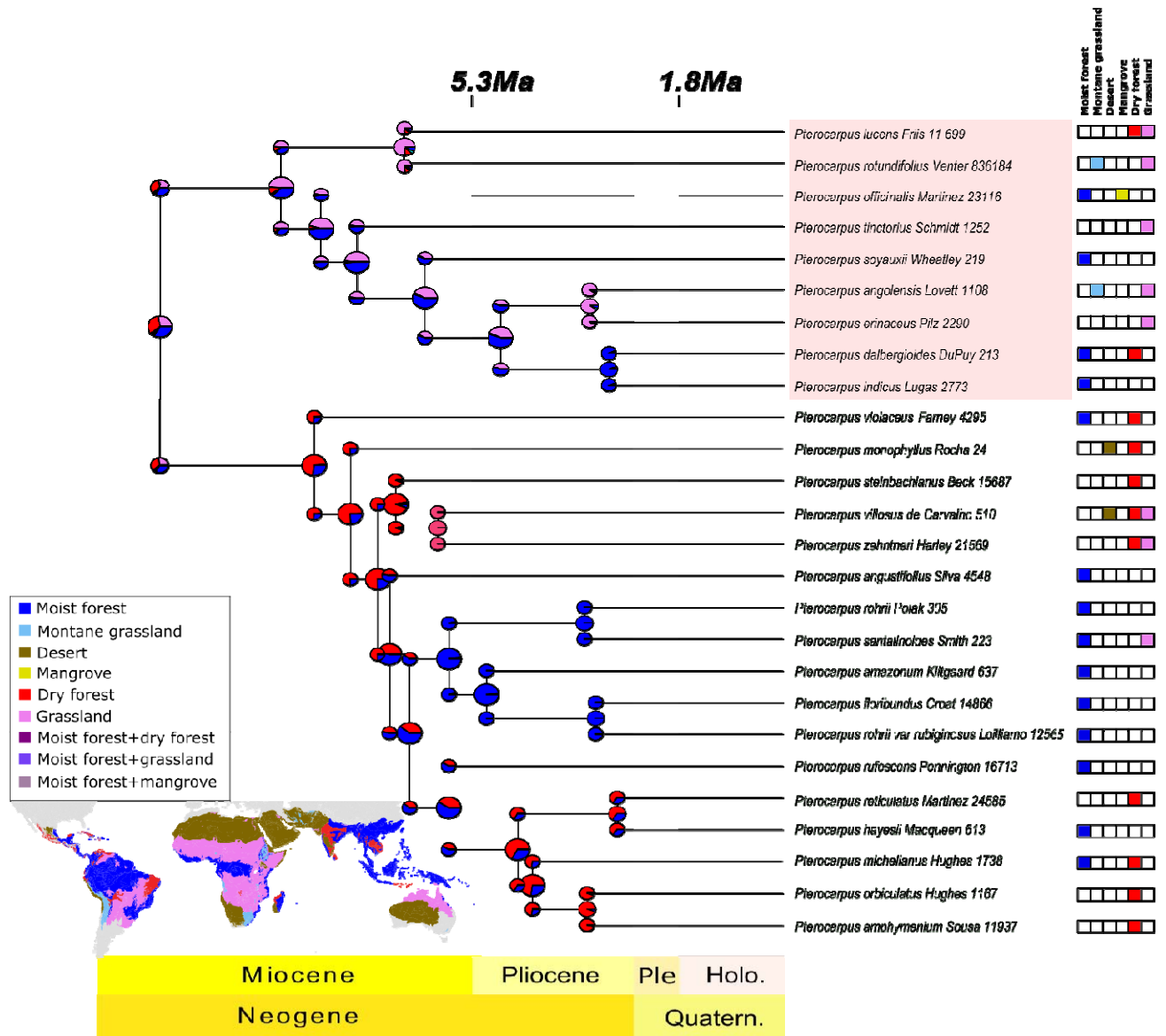
398



399

400 **Figure 2a)** Biogeographical range estimation of *Pterocarpus*. Pie charts represent likelihoods of ranges derived  
 401 from ancestral range estimations run on the MCMCTREE topology, using the R package 'BioGeoBEARS'. The  
 402 legend represents five biogeographical realms defined by Olson *et al.* (2001) and shown in Appendix S1, Fig.  
 403 S1.1a. Shading on the background of the phylogenetic tree represents geological epochs, and within the bottom  
 404 bar 'Ple' represents 'Pleistocene', 'Holo.' represents 'Holocene' and 'Quatern.' represents 'Quaternary'. Inset is  
 405 a map displaying biogeographical realms as defined by Olson *et al.* (2001).

406



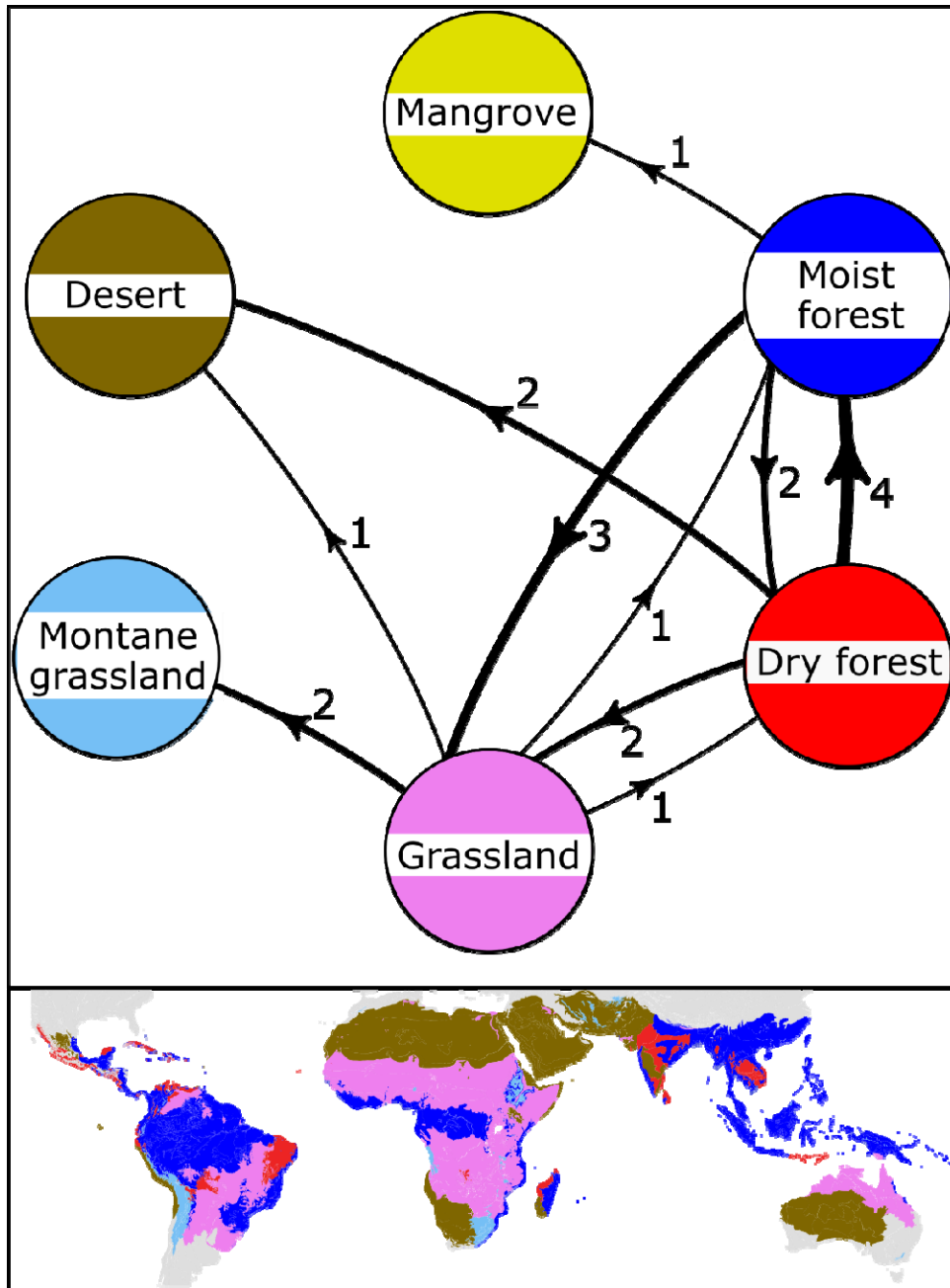
407

408 **Figure 2b)** Historical biome estimation of *Pterocarpus*. Pie charts represent likelihoods of ranges derived from  
409 ancestral biome estimations run on the MCMCTREE topology, using the R package 'BioGeoBEARS'. The  
410 legend represents six biomes defined by Olson *et al.* (2001) and shown in Appendix S1, Fig. S1.1a. Shading on  
411 the background of the phylogenetic tree represents geological epochs, and within the bottom bar 'Ple' represents  
412 'Pleistocene', 'Holo.' represents 'Holocene' and 'Quatern.' represents 'Quaternary'. Inset is a map displaying  
413 biomes as defined by Olson *et al.* (2001).

414

415





416

417 **Figure 3)** Biome transitions between species within *Pterocarpus*, based on the ancestral biome reconstruction  
418 with the Mk model shown in Fig. 2b. Arrows on lines show the direction of the biome transition, while numbers  
419 and thickness of the lines represent the number of transitions between each biome. Inset is a map displaying  
420 biomes as defined by Olson *et al.* (2001).

421

422

423

424

## 425 Discussion

426 Divergence dating indicated that the *Pterocarpus* clade began to diversify in the late Eocene (~35 Ma),  
427 followed by extensive diversification during the Miocene (23-5.3 Ma). During the Miocene, the  
428 common ancestor of *Pterocarpus* diverged into two clades. One clade was palaeotropical and  
429 diversified largely in Africa with subsequent dispersal into Indo-Malaya, Australasia and Oceania.  
430 The second clade was neotropical, which diversified only in the tropical Americas and was more  
431 speciose than the palaeotropical clade. The diversification of *Pterocarpus* appears to have involved  
432 multiple biome shifts, largely into lowland habitats (namely, savanna and tropical moist forest), and  
433 while dispersal across large geographical areas played a role in range expansion, the fruit dispersal  
434 traits we examined were not significantly associated with macroevolutionary dynamics according to  
435 our model comparisons.

436 Overall, it appears that traits for hydrochoric dispersal evolved multiple times independently in  
437 *Pterocarpus*, likely in concert with adaptation to waterlogged habitats such as mangroves. These traits  
438 which evolved to promote dispersal in the ecological theatre may have then led to rare long-distance  
439 dispersal across oceans (as we see in extant species such as *P. santalinoides* and *P. officinalis*). A  
440 similar pattern has been shown in the Bombacoideae (Malvaceae) (Zizka et al., 2020), as well as in  
441 *Centrolobium* and *Platymiscium*, which are close relatives of *Pterocarpus* (Klitgård, 2005, Pirie et al.,  
442 2009, Saslis-Lagoudakis et al., 2008).

443

## 444 Phylogenetic analysis of the *Pterocarpus* clade

445 Most *Pterocarpus* clade lineages, including *Pterocarpus* itself, arose during the Miocene, between 23-  
446 5.3 Ma (Fig. 1; Appendix S1, Fig. S1.6). This was congruent across methods (MCMCTREE, BEAST)  
447 and across data types (target capture data, Sanger-sequencing data). While our divergence dating  
448 analyses were broadly congruent with each other across methods and data type, they differed in their  
449 date estimates for the origin of *Pterocarpus*. MCMCTREE analysis inferred an origin date of 12Ma  
450 (Fig. 1), whereas the BEAST analysis suggested that *Pterocarpus* arose around 25Ma (Appendix S1,  
451 Fig. S1.6). The most likely reason for this discrepancy is the size of the dataset, since MCMCTREE  
452 was run on a nuclear target capture dataset comprising 303 loci, while BEAST was run on pre-  
453 existing Sanger sequence data for 5 loci, resulting in a much less informative sequence matrix.

454 It is also likely that the inclusion of plastid loci within the BEAST analysis extended the date estimate  
455 for the origin of *Pterocarpus*, due to deeper divergences within the plastid genome (Wolfe et al.,  
456 1987). This conflicting signal would have also led to the much higher degree of uncertainty on age  
457 estimates in the BEAST analysis, as indicated by the breadth of the 95% HPDs. In addition, the nested  
458 position of *Pterocarpus dubius* within *Pterocarpus*, as shown in the BEAST analysis (Appendix S1,  
459 Fig. S1.6) contrasts with the sister clade position of *P. dubius* to *Pterocarpus* in the MCMCTREE  
460 analysis (Fig. 1). This species was previously circumscribed as *Etaballia dubia* and is highly  
461 divergent in terms of floral morphology, however previous work with Sanger-sequencing loci  
462 resolved this species within *Pterocarpus* (Klitgård et al., 2013). It is possible that the inclusion of  
463 chloroplast loci in the BEAST tree caused this incongruence, due to the lower mutation rate of

464 chloroplast regions (Wolfe et al., 1987), or through plastid capture resulting from hybridisation  
465 (Naciri & Linder, 2015). While there is undoubtedly useful signal within these plastid loci, further  
466 work using plastomes would help to accurately disentangle these discordances.

467

#### 468 *Spatio-temporal evolution of Pterocarpus*

469 The Miocene-era diversification of the Pterocarpus clade, and particularly *Pterocarpus*, reflects  
470 patterns shown in other tropical tree genera (e.g., the mahoganies (Meliaceae) (Koenen et al., 2015),  
471 the Brownea clade (Leguminosae) (Schley et al., 2018) and the Daniellia clade (Leguminosae) (Choo  
472 et al., 2020)). This diversification was likely concurrent with major climatic, geological and  
473 ecological changes which occurred in the tropics during the Miocene (Keeley & Rundel, 2005,  
474 Morley, 2000) such as was the Miocene global cooling event (Zachos et al., 2001, 2008). This cooling  
475 was precipitated by the closure of the Tethys seaway and a subsequent change in ocean currents  
476 (Rommerskirchen et al., 2011, Zhang, Z. et al., 2014), leading to a decrease in global temperature,  
477 desiccation of the continents and a retreat of tropical vegetation, especially on the African continent  
478 (Couvreur et al., 2021, Morley & Richards, 1993, Senut et al., 2009). Such cooling likely also  
479 explains the long-distance dispersal events and biome-switching we observed in our ancestral range  
480 reconstructions for *Pterocarpus* (Fig. 2a & 2b), particularly within the palaeotropical clade, as  
481 discussed below.

482

#### 483 i. *The palaeotropical clade*

484 The palaeotropical clade appears to have diversified largely within Africa over the last ~10Ma, in the  
485 late Miocene and into the early Pliocene (Fig. 2a), with multiple biome shifts from moist forests into  
486 grasslands (n=3, Fig. 3). These events were likely concurrent with the cooling and desiccation of the  
487 African continent which lead to the expansion of grasslands during this period (Couvreur et al., 2021).  
488 This palaeotropical clade was also less speciose than its neotropical counterpart, which may be  
489 explained by higher levels of extinction in rainforest lineages, again precipitated by the expansion of  
490 grassland in Africa's past. This has led to the well-documented 'odd man out' pattern of rainforest  
491 tree diversity, with the Afrotropics being significantly less diverse than other tropical regions  
492 (Couvreur, 2015, Richards, 1973). Moreover, most of the extant *Pterocarpus* species within the  
493 palaeotropical clade are now found in grassland and are wind-dispersed (Fig. 2b; Appendix S1, Fig.  
494 S1.1b), suggesting historical adaptation to novel grassland environments.

495 Within the Palaeotropical clade, the two Asian sister species (*P. dalbergioides* and *P. indicus*) only  
496 dispersed to the Asian tropics and beyond into Australasia and Oceania around 4Ma, during the  
497 Pliocene (Fig. 2a). It is likely that the divergence of these species was caused by infrequent eastward  
498 seed dispersal, promoted by the east African coastal current (EACC) between Africa and Asia which  
499 formed with the closing of the Tethys seaway in the late Miocene (Scotese, 2004). It is interesting to  
500 note that these species are found in rainforests, as rainforest tree species tend to be excellent  
501 dispersers (Dexter et al., 2017, Pennington & Lavin, 2016), and are often water dispersed. This  
502 mirrors the highly dispersible nature of the other rainforest-dwelling member of the Palaeotropical

503 clade (*P. officinalis*) which is distributed in both Africa and South America and is also found in  
504 mangroves. Moreover, *P. officinalis* is derived from Miocene African ancestors (Fig. 2a), is present in  
505 the Neotropical fossil record (Graham, 1995) and includes a West African subspecies (subsp. *gilletii*),  
506 further suggesting that the corky exocarp of the fruit, which it evolved for dispersal at an ecological  
507 scale (Appendix S1, Fig. S1.1b), is responsible for rare dispersal across the Atlantic and the  
508 foundation of a Neotropical population (Rivera-Ocasio et al., 2002).

509

510 ii. *The neotropical subclade*

511 Within neotropical *Pterocarpus* we inferred multiple events of biome-switching from dry forest into  
512 rainforest (Fig. 2b; n=4, Fig. 3)), with the most likely ancestral biome for this clade being dry forest.  
513 These events mainly occurred between the late Miocene and the Pliocene (~8 - 2 Ma). In particular,  
514 we see the switch from dry forest in stem nodes into clades which are found nearly exclusively in wet  
515 forest towards the tips, mirroring findings in closely related genera (e.g. *Platymiscium*  
516 (Saslis-Lagoudakis et al., 2008)). Indeed, biome switching between dry forest and moist forest,  
517 which are often interdigitated, is well documented in the Neotropics (Dexter et al., 2018, Ireland et al.,  
518 2010, Klitgård, 2005, Pennington & Dick, 2010, Pezzini et al., 2020).

519 The majority of Neotropical dry forests formed during the Miocene (Pennington & Ratter, 2006) as  
520 shown by fossil evidence (e.g. Burnham, 1995, Burnham & Carranco, 2004). This was followed by  
521 periods of intense climatic flux and the contraction of dry forests resulting from increasingly wet  
522 Pleistocene interglacial periods, leading to the ‘refugia’ of dry forest which we observe today within  
523 inter-Andean valleys, Central America and the ‘dry arc’ south of the Amazon in Brazil and Bolivia  
524 (Pennington et al., 2004). The biome switching we inferred within the Neotropics may be explained  
525 by these oscillations, as the expansion of rainforests would have resulted in a much higher likelihood  
526 of immigration into rainforest by plant species dispersing from dry forest refugia.

527

528 iii. *Asymmetrical biome transitions in Pterocarpus*

529 Biome transitions across *Pterocarpus* occurred mainly into moist forest (i.e., rainforest) and grassland  
530 (i.e., savanna) (Fig. 3). The higher relative frequency of immigration into both moist forests and  
531 grasslands may be explained by the fact that these communities experience persistent disturbance  
532 (e.g., fire in tropical savannas) or because they contain species which are vulnerable to climatic  
533 perturbation (e.g., drought in rainforests) (da Costa et al., 2010, Pennington & Lavin, 2016). In  
534 rainforests, intermediate disturbance increases the likelihood of seedling establishment across entire  
535 communities due to intense competition for light (Hubbell et al., 1999), and within savannas, seedling  
536 establishment of local species is greatly reduced by frequent fire (Hoffmann, 1996, Setterfield, 2002,  
537 Wakeling et al., 2011). This may promote the competitive success of tree seeds dispersing into such  
538 environments from adjacent habitats. Thus, the highly dispersible nature of *Pterocarpus* fruits could  
539 have promoted immigration into these novel biomes, and the moderate disturbance which occurs to  
540 plant communities within these biomes could have facilitated establishment, adaptation and  
541 speciation.

542 The frequency of biome shifts into savanna and moist forests also implies that their wide extent  
543 renders them more likely to exchange seeds with adjacent biomes (Donoghue & Edwards 2014). This  
544 is particularly true for the ‘donation’ of immigrants from restricted biomes which interdigitate with  
545 savannas or rainforest, such as dry forest. Overall, our results highlight that biome switching may play  
546 an important role in diversification of tropical trees and is more common than previously thought, as  
547 recently highlighted across neotropical taxa (Antonelli et al., 2018).

548

#### 549 Caveats

550 Despite the utility of biogeographical models for understanding spatio-temporal evolution, they are  
551 not without their drawbacks. In particular, biogeographical models may be prone to over-estimating  
552 long-distance dispersal when model assumptions are violated (e.g., *DEC+J* in BioGeoBEARS (Ree &  
553 Sanmartín, 2018)), highlighting the importance of appropriate model choice (Matzke, 2014). As far as  
554 is possible, we have attempted to account for these biases by, for example, avoiding the use of the *+J*  
555 parameter in biogeographical analyses.

556 It is also possible that small sample sizes may be problematic, since with larger datasets parameter  
557 estimates are likely to be more accurate (e.g., in species distribution modelling (Wisz et al., 2008)).  
558 However, as in our study, this can be avoided by having a balanced sampling strategy which accounts  
559 for the nearly all species in studies to <50 taxa (e.g., in Araliaceae (Valcárcel & Wen, 2019), penguins  
560 (Vianna et al., 2020) and cat-eyed snakes (Weinell et al., 2021)). Finally, our specimens were  
561 identified by taxonomic experts on *Pterocarpus* and underwent multiple corroborative phylogenetic  
562 analyses, thereby minimising error as a result of misidentification or phylogenetic uncertainty.  
563 However, even while accounting for such biases, it is important to interpret biogeographical model  
564 outputs cautiously, and to realise that they are approximate estimates of biogeographic history.

565

#### 566 Conclusion

567 Overall, our analyses suggest that while a dispersal trait-dependent model of range macroevolution  
568 was not supported, the biogeographical history of *Pterocarpus* was likely mediated by cross-  
569 continental dispersal, associated range change and diversification into adjacent biomes. Indeed, biome  
570 switching between moist and arid environments, likely precipitated by infrequent long-distance  
571 dispersal events and environmental change, appear to have promoted ecological differentiation and  
572 diversification within *Pterocarpus* since the early Miocene.

573

#### 574 Data availability statement

575 The target capture sequencing data generated for this project are publicly available in the NCBI  
576 sequence read archive under the BioProject number PRJNA728569, with the accession numbers  
577 SAMN19092119 - SAMN19092167.

578

579

580

## References

- 581 Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Available  
582 online at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 583 Angelis, K. & Dos Reis, M. (2015) The impact of ancestral population size and incomplete lineage  
584 sorting on Bayesian estimation of species divergence times. *Current Zoology* **61**(5) 874-885.
- 585 Antonelli, A., Zizka, A., Carvalho, F.A., Scharn, R., Bacon, C.D., Silvestro, D. & Condamine, F.L.  
586 (2018) Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National*  
587 *Academy of Sciences of the United States of America* **115**(23) 6034-6039.
- 588 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S. ... Prjibelski, A.D.  
589 (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.  
590 *Journal of Computational Biology* **19**(5) 455-477.
- 591 Bellot, S., Bayton, R.P., Couvreur, T.L., Dodsworth, S., Eiserhardt, W.L., Guignard, M.S. ... Baker,  
592 W.J. (2020) On the origin of giant seeds: the macroevolution of the double coconut (*Lodoicea*  
593 *maldivica*) and its relatives (Borasseae, Arecaceae). *The New Phytologist* **228**(3) 1134.
- 594 Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence  
595 data. *Bioinformatics* **30**(15) 2114-2120.
- 596 Borowiec, M.L. (2016) AMAS: a fast tool for alignment manipulation and computing of summary  
597 statistics. *PeerJ* **4**e1660.
- 598 Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996) The geographic range: size, shape, boundaries,  
599 and internal structure. *Annual Review of Ecology and Systematics* **27**(1) 597-623.
- 600 Burnham, R.J. (1995) A new species of winged fruit from the Miocene of Ecuador: *Tipuana*  
601 *ecuatoriana* (Leguminosae). *American Journal of Botany* **82**(12) 1599-1607.
- 602 Burnham, R.J. & Carranco, N.L. (2004) Miocene winged fruits of *Loxopterygium* (Anacardiaceae)  
603 from the Ecuadorian Andes. *American Journal of Botany* **91**(11) 1767-1773.
- 604 Buzek, C. (1992) Fruits of *Pterocarpus tertiaris* Weyland from the North-Bohemian Basin,  
605 Czechoslovakia. *Advances in legume systematics: Part 4* (ed. by P.S. Herendeen & D.L. Dilcher), pp.  
606 19-31.
- 607 Cain, M.L., Milligan, B.G. & Strand, A.E. (2000) Long-distance seed dispersal in plant populations.  
608 *American Journal of Botany* **87**(9) 1217-1227.
- 609 Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. (2009) trimAl: a tool for automated  
610 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**(15) 1972-1973.
- 611 Choo, L.M., Forest, F., Wieringa, J.J., Bruneau, A. & de la Estrella, M. (2020) Phylogeny and  
612 biogeography of the *Daniellia* clade (Leguminosae: Detarioideae), a tropical tree lineage largely  
613 threatened in Africa and Madagascar. *Molecular phylogenetics and evolution* **146**106752.



- 614 Corlett, R.T. & Primack, R.B. (2011) *Tropical rain forests: an ecological and biogeographical*  
615 *comparison* John Wiley & Sons.
- 616 Cousens, R., Dytham, C. & Law, R. (2008) *Dispersal in plants: a population perspective* Oxford  
617 University Press.
- 618 Couvreur, T.L. (2015) Odd man out: why are there fewer plant species in African rain forests? *Plant*  
619 *Systematics and Evolution* **301**(5) 1299-1313.
- 620 Couvreur, T.L., Dauby, G., Blach  $\square$  Overgaard, A., Deblauwe, V., Dessein, S., Droissart, V. ... Ley,  
621 A.C. (2021) Tectonics, climate and the diversification of the tropical African terrestrial flora and  
622 fauna. *Biological Reviews* **96**(1) 16-51.
- 623 da Costa, A.C., Galbraith, D., Almeida, S., Portela, B.T.T., da Costa, M., Junior, João de Athaydes  
624 Silva ... Fisher, R. (2010) Effect of 7 yr of experimental drought on vegetation dynamics and biomass  
625 storage of an eastern Amazonian rainforest. *New Phytologist* **187**(3) 579-591.
- 626 Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new  
627 heuristics and parallel computing. *Nature Methods* **9**(8) 772.
- 628 de Candolle, A.P. (1825) *Prodromus Systematis Naturalis Regni Vegetabilis 2. Parisii: Sumptibus*  
629 *Sociorum Treuttel et Wu  $\square$  rtz.*
- 630 de la Estrella, M., Forest, F., Wieringa, J.J., Fougere-Danezan, M. & Bruneau, A. (2017) Insights on  
631 the evolutionary origin of Detarioideae, a clade of ecologically dominant tropical African trees. *New*  
632 *Phytologist* **214**(4) 1722-1735.
- 633 De Micco, V. & Aronne, G. (2012) Morpho-anatomical traits for plant adaptation to drought. *Plant*  
634 *responses to drought stress*(ed. by Anonymous ), pp. 37-61. Springer.
- 635 De Queiroz, A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in*  
636 *Ecology & Evolution* **20**(2) 68-73.
- 637 Dexter, K.G., Smart, B., Baldauf, C., Baker, T.R., Balinga, M.P., Brienen, R. ... Muledi, J.I. (2015)  
638 Floristics and biogeography of vegetation in seasonally dry tropical regions. *International Forestry*  
639 *Review* **17**(2) 10-32.
- 640 Dexter, K.G., Pennington, R.T., Oliveira-Filho, A.T., Bueno, M.L., Silva de Miranda, Pedro L &  
641 Neves, D.M. (2018) Inserting tropical dry forests into the discussion on biome transitions in the  
642 tropics. *Frontiers in Ecology and Evolution* **6**104.
- 643 Dexter, K.G., Lavin, M., Torke, B.M., Twyford, A.D., Kursar, T.A., Coley, P.D. ... Pennington, R.T.  
644 (2017) Dispersal assembly of rain forest tree communities across the Amazon basin. *Proceedings of*  
645 *the National Academy of Sciences of the United States of America* **114**(10) 2645-2650.
- 646 Donoghue, M.J. & Edwards, E.J. (2014) Biome shifts and niche evolution in plants. *Annual Review of*  
647 *Ecology, Evolution, and Systematics* **45**547-572.
- 648 Dos Reis, M. & Yang, Z. (2013) MCMCTree  
649 tutorials. <http://abacus.gene.ucl.ac.uk/software/MCMCTree.Tutorials.pdf>.
- 650 Doyle, J.J. & Doyle, J.L. (1987) Genomic plant DNA preparation from fresh tissue- CTAB method.  
651 *Phytochemical Bulletin* **19**(11) 11-15.

- 652 Drummond, A.J. & Bouckaert, R.R. (2015) *Bayesian evolutionary analysis with BEAST* Cambridge  
653 University Press.
- 654 Echeverría-Londoño, S., Enquist, B.J., Neves, D.M., Violle, C., Boyle, B., Kraft, N.J. ... Sandel, B.  
655 (2018) Plant functional diversity and the biogeography of biomes in North and South America.  
656 *Frontiers in Ecology and Evolution* **6**219.
- 657 Eiserhardt, W.L., Svenning, J., Kissling, W.D. & Balslev, H. (2011) Geographical ecology of the  
658 palms (Arecaceae): determinants of diversity and distributions across spatial scales. *Annals of Botany*  
659 **108**(8) 1391-1416.
- 660 Garamszegi, L.Z. & Møller, A.P. (2011) Nonrandom variation in within-species sample size and  
661 missing data in phylogenetic comparative studies. *Systematic Biology* **60**(6) 876-880.
- 662 Gentry, A.H. (1988) Tree species richness of upper Amazonian forests. *Proceedings of the National*  
663 *Academy of Sciences* **85**(1) 156-159.
- 664 Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R. & Roderick, G.K. (2012) Long-  
665 distance dispersal: a framework for hypothesis testing. *Trends in ecology & evolution* **27**(1) 47-56.
- 666 Graham, A. (1995) Diversification of Gulf/Caribbean mangrove communities through Cenozoic time.  
667 *Biotropica* 20-27.
- 668 Harris, A.J., Ickert-Bond, S. & Rodríguez, A. (2018) Long distance dispersal in the assembly of  
669 floras: A review of progress and prospects in North America. *Journal of Systematics and Evolution*  
670 **56**(5) 430-448.
- 671 Ho, S.Y. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of  
672 evolutionary divergence times. *Systematic Biology* **58**(3)367-380.
- 673 Hoffmann, W.A. (1996) The effects of fire and cover on seedling establishment in a neotropical  
674 savanna. *Journal of Ecology* 383-393.
- 675 Hubbell, S.P., Foster, R.B., O'Brien, S.T., Harms, K.E., Condit, R., Wechsler, B., Wright, S.J. & De  
676 Lao, S.L. (1999) Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical  
677 forest. *Science* **283**(5401) 554-557.
- 678 Hughes, C., Pennington, R.T. & Antonelli, A. (2013) Neotropical plant evolution: assembling the big  
679 picture. *Botanical Journal of the Linnean Society* **171**(1) 1-18.
- 680 IPNI. (2020) The international plant names index. **2019**(10 December) .
- 681 Ireland, H.E., Kite, G.C., Veitch, N.C., Chase, M.W., Schrire, B., Lavin, M., Linares, J. &  
682 Pennington, R.T. (2010) Biogeographical, ecological and morphological structure in a phylogenetic  
683 analysis of *Ateleia* (Swartziaeeae, Fabaceae) derived from combined molecular, morphological and  
684 chemical data. *Botanical Journal of the Linnean Society* **162**(1) 39-53.
- 685 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J. & Wickett,  
686 N.J. (2016) HybPiper: Extracting coding sequence and introns for phylogenetics from  
687 high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* **4**(7)  
688 1600016.

- 689 Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7:  
690 improvements in performance and usability  
691 . *Molecular Biology and Evolution* **30**(4) 772-780.
- 692 Keeley, J.E. & Rundel, P.W. (2005) Fire and the Miocene expansion of C4 grasslands. *Ecology*  
693 *Letters* **8**(7) 683-690.
- 694 Klaus, K.V. & Matzke, N.J. (2020) Statistical Comparison of Trait-Dependent Biogeographical  
695 Models Indicates That Podocarpaceae Dispersal Is Influenced by Both Seed Cone Traits and  
696 Geographical Distance. *Systematic Biology* **69**(1) 61-75.
- 697 Klitgård, B.B. & Lavin, M. (2005) Tribe Dalbergieae sens. lat. *Legumes of the world*(ed. by G.P.  
698 Lewis, B. Schrire, B. Mackinder & M. Lock), pp. 307-335. Kew Publishing, Richmond, Surrey.
- 699 Klitgård, B.B., de Queiroz, L.P. & Lewis, G.P. (2000) A remarkable new species of Pterocarpus  
700 (Leguminosae: Papilionoideae: Dalbergieae) from Bahia, Brazil. *Kew Bulletin* 989-992.
- 701 Klitgård, B.B. (2005) Platymiscium (Leguminosae: Dalbergieae): Biogeography Systematics,  
702 Morphology, Taxonomy and Uses. *Kew Bulletin* 321-400.
- 703 Klitgård, B.B., Forest, F., Booth, T.J. & Saslis-Lagoudakis, C.H. (2013) A detailed investigation of  
704 the Pterocarpus clade (Leguminosae: Dalbergieae): Etaballia with radially symmetrical flowers is  
705 nested within the papilionoid-flowered Pterocarpus. *South African Journal of Botany* **89**128-142.
- 706 Koenen, E.J., Clarkson, J.J., Pennington, T.D. & Chatrou, L.W. (2015) Recently evolved diversity and  
707 convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest  
708 hyperdiversity. *New Phytologist* **207**(2) 327-339.
- 709 Kriebel, R., Drew, B.T., Drummond, C.P., González-Gallegos, J.G., Celep, F., Mahdjoub, M.M. ...  
710 Walker, J.B. (2019) Tracking temporal shifts in area, biomes, and pollinators in the radiation of Salvia  
711 (sages) across continents: leveraging anchored hybrid enrichment and targeted sequence data.  
712 *American Journal of Botany* **106**(4) 573-597.
- 713 Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013) Bayesian analysis of  
714 biogeography when the number of areas is large. *Systematic Biology* **62**(6) 789-804.
- 715 Lewis, G.P. (1987) *Legumes of Bahia*Kew Publishing, Richmond, Surrey.
- 716 Lewis, G.P., Schrire, B., Mackinder, B. & Lock, M. (2005) *Legumes of the World*Kew Publishing,  
717 Richmond, Surrey.
- 718 Lewis, P.O. (2001) A likelihood approach to estimating phylogeny from discrete morphological  
719 character data. *Systematic Biology* **50**(6) 913-925.
- 720 Martínez, C. (2018) Dalbergieae (Fabaceae) Samara Fruits from the Late Eocene of Colombia.  
721 *International journal of plant sciences* **179**(7) 541-553.
- 722 Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation  
723 is a crucial process in island clades. *Systematic Biology* **63**(6) 951-970.
- 724 Matzke, N.J. (2013) BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary  
725 Analysis in R Scripts. <https://rdr.io/cran/BioGeoBEARS/>.

- 726 Midgley, J.J. & Bond, W.J. (2013) Plant adaptations to fire: an evolutionary perspective. *Fire*  
727 *Phenomena and the Earth System: An Interdisciplinary Guide to Fire Science*(ed. by C. M. Becher),  
728 pp. 125-134. Wiley Online Library.
- 729 Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference  
730 of large phylogenetic trees. *Proceedings of the Gateway Environments Workshop (GCE)* 1-8.
- 731 Moore, B.R., Hohna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016) Critically evaluating the  
732 theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the*  
733 *National Academy of Sciences of the United States of America* **113**(34) 9569-9574.
- 734 Morley, R. (2000) *Origin and evolution of tropical rain forests*. John Wiley & Sons.
- 735 Morley, R. & Richards, K. (1993) Gramineae cuticle: a key indicator of Late Cenozoic climatic  
736 change in the Niger Delta. *Review of palaeobotany and palynology* **77**(1-2) 119-127.
- 737 Naciri, Y. & Linder, H.P. (2015) Species delimitation and relationships: the dance of the seven veils.  
738 *Taxon* **64**(1) 3-16.
- 739 Nie, Z., Sun, H., Manchester, S.R., Meng, Y., Luke, Q. & Wen, J. (2012) Evolution of the  
740 intercontinental disjunctions in six continents in the Ampelopsis clade of the grape family (Vitaceae).  
741 *BMC evolutionary biology* **12**(1) 1-13.
- 742 Nürk, N.M., Linder, H.P., Onstein, R.E., Larcombe, M.J., Hughes, C.E., Piñeiro Fernández, L. ...  
743 Cutts, V. (2020) Diversification in evolutionary arenas—Assessment and synthesis. *Ecology and*  
744 *Evolution* **10**(12) 6163-6182.
- 745 Olmstead, R.G. (2013) Phylogeny and biogeography in Solanaceae, Verbenaceae and Bignoniaceae: a  
746 comparison of continental and intercontinental diversification patterns. *Botanical Journal of the*  
747 *Linnean Society* **171**(1) 80-102.
- 748 Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C. ...  
749 Morrison, J.C. (2001) Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global  
750 map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience*  
751 **51**(11) 933-938.
- 752 Onstein, R.E., Kissling, W.D., Chatrou, L.W., Couvreur, T.L.P., Morlon, H. & Sauquet, H. (2019)  
753 Which frugivory-related traits facilitated historical long-distance dispersal in the custard apple family  
754 (Annonaceae)? *Journal of Biogeography* **46**(8) 1874-1888.
- 755 Paradis, E. & Schliep, K. (2019) ape 5.0: an environment for modern phylogenetics and evolutionary  
756 analyses in R. *Bioinformatics* **35**(3) 526-528.
- 757 Pennington, R.T. & Lavin, M. (2016) The contrasting nature of woody plant species in different  
758 neotropical forest biomes reflects differences in ecological stability. *New Phytologist* **210**(1) 25-37.
- 759 Pennington, R.T. & Hughes, C.E. (2014) The remarkable congruence of New and Old World savanna  
760 origins. *New Phytologist* **204**(1) 4-6.
- 761 Pennington, R.T. & Dick, C.W. (2010) Diversification of the Amazonian flora and its relation to key  
762 geological and environmental events: a molecular perspective. *Amazonia, Landscape and Species*  
763 *evolution*(ed. by C. Hoorn & F.P. Wesselingh)Blackwell Publishing, Oxford, UK.

- 764 Pennington, R.T. & Ratter, J.A. (2006) *Neotropical savannas and seasonally dry forests: plant*  
765 *diversity, biogeography, and conservation*. CRC press, Boca Raton, Florida, United States.
- 766 Pennington, R.T. & Dick, C.W. (2004) The role of immigrants in the assembly of the South American  
767 rainforest tree flora. *Philosophical Transactions of the Royal Society of London Series B: Biological*  
768 *sciences* **359**(1450) 1611-1622.
- 769 Pennington, R.T., Lavin, M. & Oliveira-Filho, A. (2009) Woody plant diversity, evolution, and  
770 ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology,*  
771 *Evolution, and Systematics* **40**437-457.
- 772 Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K. & Butterworth, C.A. (2004)  
773 Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of  
774 both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society of*  
775 *London. Series B: Biological Sciences* **359**(1443) 515-538.
- 776 Pennington, R.T., Prado, D.E. & Pendry, C.A. (2000) Neotropical seasonally dry forests and  
777 Quaternary vegetation changes. *Journal of Biogeography* **27**261-273.
- 778 Pezzini, F.F., Dexter, K.G., de Carvalho-Sobrinho, J.G., Kidner, C.A., Nicholls, J.A., De Queiroz,  
779 L.P. & Pennington, R.T. (2020) Phylogeny and biogeography of *Ceiba* Mill.(Malvaceae,  
780 Bombacoideae). *Frontiers of Biogeography* .
- 781 Pirie, M.D., Klitgaard, B.B. & Pennington, R.T. (2009) Revision and Biogeography of *Centrolobium*  
782 (Leguminosae-Papilionoideae). *Systematic Botany* **34**(2) 345-359.
- 783 Puttick, M.N. (2019) MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior  
784 ages on trees. *Bioinformatics* **35**(24) 5321-5322.
- 785 Python Software Foundation. (2010) Python Language Reference, version 2.7. <http://www.python.org>  
786 .
- 787 R Development Core Team. (2013) R: A language and environment for statistical computing.  
788 Available at: <http://www.R-project.org/>.
- 789 Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017) Is BAMM flawed? Theoretical and practical  
790 concerns in the analysis of multi-rate diversification models. *Systematic Biology* .
- 791 Rabosky, D.L., Donnellan, S.C., Grundler, M. & Lovette, I.J. (2014) Analysis and visualization of  
792 complex macroevolutionary dynamics: an example from Australian scincid lizards. *Systematic*  
793 *Biology* **63**(4) 610-627.
- 794 Rambaut, A. (2014) FigTree, a graphical viewer of phylogenetic trees ([tree.bio.ed.ac.](http://tree.bio.ed.ac.uk/software/figtree)  
795 [uk/software/figtree](http://tree.bio.ed.ac.uk/software/figtree)). **1.4.2**.
- 796 Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A.J. (2015) Tracer.  
797 <http://tree.bio.ed.ac.uk/software/tracer/> **1.6**.
- 798 Ree, R.H. & Sanmartín, I. (2018) Conceptual and statistical problems with the DEC J model of  
799 founder event speciation and its comparison with DEC via model selection. *Journal of Biogeography*  
800 **45**(4) 741-749.

- 801 Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by  
802 dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**(1) 4-14.
- 803 Renner, S. (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International*  
804 *Journal of Plant Sciences* **165**(S4) S23-S33.
- 805 Revell, L.J. (2012) phytools: an R package for phylogenetic comparative biology (and other things).  
806 *Methods in Ecology and Evolution* **3**(2) 217-223.
- 807 Richards, P.W. (1973) Africa, the 'odd man out'. *Tropical forest ecosystems in Africa and South*  
808 *America: a comparative review*.(ed. by B.J. Meggers, E.S. Ayensu, and W.D. Duckworth), pp. 21-26.  
809 Smithsonian Institution Press, Washington, DC, USA.
- 810 Ringelberg, J.J., Zimmermann, N.E., Weeks, A., Lavin, M. & Hughes, C.E. (2020) Biomes as  
811 evolutionary arenas: Convergence and conservatism in the trans-continental succulent biome. *Global*  
812 *Ecology and Biogeography* **29**(7) 1100-1113.
- 813 Rivera-Ocasio, E., Aide, T.M. & McMillan, W.O. (2002) Patterns of genetic diversity and  
814 biogeographical history of the tropical wetland tree, *Pterocarpus officinalis* (Jacq.), in the Caribbean  
815 basin. *Molecular ecology* **11**(4) 675-683.
- 816 Rojo, J.P. (1972) *Pterocarpus (Leguminosae - Papilionaceae): Revised for the world*.Cramer, Lehre,  
817 Germany.
- 818 Rommerskirchen, F., Condon, T., Mollenhauer, G., Dupont, L. & Schefuss, E. (2011) Miocene to  
819 Pliocene development of surface and subsurface temperatures in the Benguela Current system.  
820 *Paleoceanography* **26**(3) .
- 821 Ronquist, F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical  
822 biogeography. *Systematic Biology* **46**(1) 195-203.
- 823 Sanmartín, I. (2014) Biogeography. *The Tree of Life*.(ed. by P. Vargas and R. Zardoya), pp. 555-576.  
824 Sinauer Associates, Inc., Sunderland, MA, USA.
- 825 Sanmartín, I. (2012) Historical Biogeography: Evolution in Time and Space. *Evolution education &*  
826 *outreach* **5**(4) 555-568.
- 827 Sanmartín, I. & Ronquist, F. (2004) Southern hemisphere biogeography inferred by event-based  
828 models: plant versus animal patterns. *Systematic Biology* 216-243.
- 829 Saslis-Lagoudakis, C., Chase, M.W., Robinson, D.N., Russell, S.J. & Klitgaard, B.B. (2008)  
830 Phylogenetics of neotropical Platymiscium (Leguminosae: Dalbergieae): systematics, divergence  
831 times, and biogeography inferred from nuclear ribosomal and plastid DNA sequence data. *American*  
832 *Journal of Botany* **95**(10) 1270-1286.
- 833 Saslis-Lagoudakis, C.H., Klitgaard, B.B., Forest, F., Francis, L., Savolainen, V., Williamson, E.M. &  
834 Hawkins, J.A. (2011) The use of phylogeny to interpret cross-cultural patterns in plant use and guide  
835 medicinal plant discovery: an example from *Pterocarpus* (Leguminosae). *PloS one* **6**(7) e22275.
- 836 Schley, R.J., de la Estrella, M., Pérez-Escobar, O.A., Bruneau, A., Barraclough, T., Forest, F. &  
837 Klitgård, B. (2018) Is Amazonia a 'museum' for Neotropical trees? The evolution of the *Brownea*  
838 clade (Detarioideae, Leguminosae). *Molecular Phylogenetics and Evolution* **126**279-292.



- 839 Schrire, B.D., Lavin, M. & Lewis, G.P. (2005) Global distribution patterns of the Leguminosae:  
840 insights from recent phylogenies. 375-422.
- 841 Scotese, C.R. (2004) A continental drift flipbook. *The Journal of geology* **112**(6) 729-741.
- 842 Senut, B., Pickford, M. & Ségalen, L. (2009) Neogene desertification of Africa. *Comptes Rendus*  
843 *Geoscience* **341**(8-9) 591-602.
- 844 Setterfield, S.A. (2002) Seedling establishment in an Australian tropical savanna: effects of seed  
845 supply, soil disturbance and fire. *Journal of Applied Ecology* **39**(6) 949-959.
- 846 Karline Soetaert(2012) *R Package shape: functions for plotting graphical shapes, colors..., version 1.*  
847 Karline Soetaert.
- 848 Stamatakis, A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of  
849 Large Phylogenies. *Bioinformatics* **Open access**.Published online: 10.1093/bioinformatics/btu033.
- 850 ter Steege, H., Prado, P.I., de Lima, R.A., Pos, E., de Souza Coelho, L., de Andrade Lima Filho,  
851 Diogenes ... Castilho, C.V. (2020) Biased-corrected richness estimates for the Amazonian tree flora.  
852 *Scientific reports* **10**(1) 1-13.
- 853 ter Steege, H., Pitman, N.C., Sabatier, D., Baraloto, C., Salomao, R.P., Guevara, J.E. ... Silman, M.R.  
854 (2013) Hyperdominance in the Amazonian tree flora. *Science* **342**(6156) 1243092.
- 855 Töpel, M., Zizka, A., Calió, M.F., Scharn, R., Silvestro, D. & Antonelli, A. (2017) SpeciesGeoCoder:  
856 fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and  
857 evolution. *Systematic Biology* **66**(2) 145-151.
- 858 Valcárcel, V. & Wen, J. (2019) Chloroplast phylogenomic data support Eocene amphipacific early  
859 radiation for the Asian Palmate core Araliaceae. *Journal of Systematics and Evolution* **57**(6) 547-560.
- 860 van der Pijl, L. (1982) Ecological Developments in Leguminous Fruits. *Principles of Dispersal in*  
861 *Higher Plants*(ed. by Anonymous ), pp. 150-159. Springer.
- 862 Vatanparast, M., Klitgård, B.B., Adema, F.A., Pennington, R.T., Yahara, T. & Kajita, T. (2013) First  
863 molecular phylogeny of the pantropical genus *Dalbergia*: implications for infrageneric  
864 circumscription and biogeography. *South African Journal of Botany* **89**143-149.
- 865 Vianna, J.A., Fernandes, F.A., Frugone, M.J., Figueiró, H.V., Pertierra, L.R., Noll, D. ... Parker, P.  
866 (2020) Genome-wide analyses reveal drivers of penguin diversification. *Proceedings of the National*  
867 *Academy of Sciences* **117**(36) 22303-22310.
- 868 Wakeling, J.L., Staver, A.C. & Bond, W.J. (2011) Simply the best: the transition of savanna saplings  
869 to trees. *Oikos* **120**(10) 1448-1451.
- 870 Waser, N.M. & Campbell, D.R. (2004) Ecological speciation in flowering plants. *Adaptive*  
871 *speciation*(ed. by Anonymous ), pp. 264-277. Cambridge University Press Cambridge.
- 872 Weinell, J.L., Barley, A.J., Siler, C.D., Orlov, N.L., Ananjeva, N.B., Oaks, J.R., Burbrink, F.T. &  
873 Brown, R.M. (2021) Phylogenetic relationships and biogeographic range evolution in cat-eyed snakes,  
874 *Boiga* (Serpentes: Colubridae). *Zoological Journal of the Linnean Society* **192**(1) 169-184.

- 875 Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A. & NCEAS Predicting  
876 Species Distributions Working Group. (2008) Effects of sample size on the performance of species  
877 distribution models. *Diversity and Distributions* **14**(5) 763-773.
- 878 Wolfe, K.H., Li, W.H. & Sharp, P.M. (1987) Rates of nucleotide substitution vary greatly among  
879 plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of  
880 Sciences of the United States of America* **84**(24) 9054-9058.
- 881 Woodward, F.I., Lomas, M.R. & Kelly, C.K. (2004) Global climate and the distribution of plant  
882 biomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*  
883 **359**(1450) 1465-1476.
- 884 Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and  
885 evolution* **24**(8) 1586-1591.
- 886 Zachos, J., Dickens, G.R. & Zeebe, R.E. (2008) An early Cenozoic perspective on greenhouse  
887 warming and carbon-cycle dynamics. *Nature* **451**(7176) 279-283.
- 888 Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001) Trends, rhythms, and aberrations  
889 in global climate 65 Ma to present. *Science* **292**(5517) 686-693.
- 890 Zamora, N. (2000) Nuevas especies y combinaciones en Leguminosas de Mesoamérica. *Novon* 175-  
891 180.
- 892 Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018) ASTRAL-III: polynomial time species tree  
893 reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**(6) 153.
- 894 Zhang, Z., Ramstein, G., Schuster, M., Li, C., Contoux, C. & Yan, Q. (2014) Aridification of the  
895 Sahara desert caused by Tethys Sea shrinkage during the Late Miocene. *Nature* **513**(7518) 401-404.
- 896 Zizka, A., Carvalho □ Sobrinho, J.G., Pennington, R.T., Queiroz, L.P., Alcantara, S., Baum, D.A.,  
897 Bacon, C.D. & Antonelli, A. (2020) Transitions between biomes are common and directional in  
898 Bombacoideae (Malvaceae). *Journal of Biogeography* **47**(6) 1310-1321.
- 899 Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D. ... Scharn, R. (2019)  
900 CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases.  
901 *Methods in Ecology and Evolution* **10**(5) 744-751.

902

### 903 Significance statement

904 Dispersal involves organisms spreading from one location to another, and is one of the major  
905 processes defining the distribution of organisms. Sometimes, structures which evolved for local  
906 dispersal may lead to rare, long-distance dispersal, and this may allow species to colonise new  
907 landmasses and diversify in novel biomes. As such, here we build an evolutionary tree to understand  
908 relationships between a genus of rosewood species (*Pterocarpus*), and test whether dispersal traits and  
909 biome switching have impacted their evolution. We found that dispersal and biome switching, likely  
910 precipitated by climate change in the Miocene, impacted the modern distribution of *Pterocarpus*, but  
911 that this was not significantly explained by dispersal structures.

912

913 Biosketch

914 Rowan Schley is a postdoctoral researcher interested in the evolution, biogeography and conservation  
915 of tropical biodiversity. His research is broadly centred around speciation with a particular focus on  
916 the effects of introgression, using phylogenomic and population genomic approaches to address these  
917 questions in tropical trees.

918 R.J.S. conceived the hypotheses, led data analysis and wrote the manuscript. M.Q. and P.M. led the  
919 phylogenomic data generation. M.V., M.d.l.E. and G.L. provided many comments and background  
920 information on the Legume family for the manuscript. B.K. helped conceive the study and led sample  
921 collection, and provided many comments on the manuscript. All authors contributed to editing the  
922 manuscript.