

1 **Phylogenomic data reveal how a climatic inversion and glacial refugia shape**
2 **patterns of diversity in an African rain forest tree species**

3

4 Andrew J. Helmstetter^{1*}, Biowa E. N. Amoussou^{1,2}, Kevin Bethune¹, Narcisse G.
5 Kandem³, Romain Glèlè Kakai⁴, Bonaventure Sonké³, Thomas L. P. Couvreur^{1,3,5}

6

7 ¹ IRD, UMR DIADE, Univ. Montpellier, Montpellier, France

8 ² Faculté des Sciences Agronomiques, Université d'Abomey-Calavi, 04 BP 1525,
9 Cotonou, Benin

10 ³ Université de Yaoundé I, Ecole Normale Supérieure, Département des Sciences
11 Biologiques, Laboratoire de Botanique systématique et d'Ecologie, B.P. 047,
12 Yaoundé, Cameroon

13 ⁴ Laboratoire de Biomathématiques et d'Estimations Forestières, Université
14 d'Abomey-Calavi, 03 BP 2819, Cotonou, Bénin

15 ⁵ Facultad de Ciencias Exactas y Naturales, Pontificia Universidad Católica del
16 Ecuador, Av. 12 de Octubre 1076 y Roca, Quito, Ecuador

17

18 * Corresponding Author

19 Email: andrew.j.helmstetter@gmail.com

20 Address:

21 UMR DIADE, équipe DYNADIV

22 Institut de recherche pour le développement (IRD)

23 911 avenue d'Agropolis, BP 64501

24 34394 Montpellier Cedex 5

25 France

26 **ABSTRACT**

27 The world's second largest expanse of tropical rain forest is in Central Africa and
28 contains incredible species diversity. Population genetic studies have consistently
29 revealed significant structure across central African rain forest plants, in particular a
30 North-South genetic discontinuity close to the equator at the level of a climatic
31 inversion. Here, we take a phylogeographic approach using 351 nuclear markers in
32 112 individuals across the distribution of the African rain forest tree species *Annickia*
33 *affinis* (Annonaceae). We show for the first time that the North-South divide is the
34 result of a single major colonisation event across the climatic inversion from an
35 ancestral population located in Gabon. We suggest that environmental differences
36 across the inversion and associated trait divergence may have contributed to this
37 phylogenetic discontinuity. We find evidence for inland dispersal, predominantly in
38 northern areas, and variable demographic histories among genetic clusters, indicating
39 that populations responded differently to past climate change. We show how newly-
40 developed genomic tools can provide invaluable insights into our understanding of
41 tropical rain forest evolutionary dynamics.

42

43 Keywords: Phylogenomics, phylogeography, rain forest, sequence capture, Africa,
44 dispersal

45

46 **Abbreviations**

47 TRF = Tropical rain forest

48 CAR = central African rain forest

49 LGM = last glacial maximum

50 SFS = site frequency spectrum

51

52 **1 INTRODUCTION**

53 Tropical rain forests (TRFs) possess an incredibly diverse flora and fauna making up
54 half of the world's biodiversity. Understanding how this diversity is generated is
55 critical if we are to protect it (Mace et al. 2003). Central Africa hosts the world's
56 second largest continuous extent of TRF (Linder 2001). Climatic fluctuations during
57 the Pleistocene and associated glacial forest refugia are suggested to have influenced
58 how genetic diversity is distributed in Central African rain forests (CAR) (Hardy et al.
59 2013). However, the nature (Anhuf et al. 2006; Diamond and Hamilton 1980; Maley
60 1996; Bonnefille 2007) and importance of forest refugia continue to be intensely
61 debated (Cowling et al. 2008; Lezine et al. 2019). Population genetic studies within
62 CAR plant species document differing levels of response to past climatic fluctuations
63 (reviewed in Hardy et al. 2013). Conversely, one major phylogeographic pattern
64 common to many CAR plant species studied is the existence of a phylogeographic
65 barrier along a North-South axis around 0-3 degrees north (see Fig. 1A; Hardy et al.
66 2013; Faye et al. 2016; Heuertz et al. 2014). There appears to be no visible geographic
67 barrier to explain this break as continuous rain forest exists across the entire area. This
68 North-South phylogeographic barrier corresponds, however, to the central African
69 climatic hinge, an inversion zone between the northern and southern rainy seasons
70 (Hardy et al; 2013). Interestingly, this barrier is rarely recovered in phylogeographic

71 studies of animals and thus seems to affect a greater effect on plants groups (e.g.
72 Fuchs and Bowie 2015; Bohoussou et al. 2015; Bell et al. 2017; Blatrix et al. 2017).
73
74 Three main hypotheses have been suggested to explain how this North-South genetic
75 discontinuity originated (Hardy et al. 2013). First, TRF might have disappeared
76 (repeatedly) along the climatic hinge during past climatic fluctuations, isolating
77 allopatric north/south populations, which subsequently recolonised the area. Second,
78 because seasons are inverted across the hinge, flowering times might be displaced
79 between northern and southern populations preventing interbreeding and gene flow.
80 Third, successful colonisation across the climatic hinge may be limited by factors
81 such as environmental differences (e.g. changing levels of water stress). At present,
82 little is known about the relative importance of these three scenarios in generating the
83 observed genetic discontinuity.

84

85 The phylogeographic approach (Avice et al. 1987) can unravel the history of
86 populations, and ultimately uncover which processes have shaped current patterns of
87 diversity. It is therefore an ideal framework to study how the climatic inversion, and
88 other factors, have shaped patterns of intraspecific diversity in CAR plants. High-
89 throughput sequencing allows the generation of phylogeographic datasets consisting
90 of large numbers of independently segregating nuclear loci that can account for
91 coalescent stochasticity (Edwards et al. 2016) where studies with small numbers of
92 markers fall short. Phylogenomic data can also be used to reconstruct the spatial
93 evolutionary history of species by inferring phylogenetic trees among populations and
94 dispersal dynamics. These approaches can help us to understand relationships among

95 populations on either side of the climatic inversion and how often lineages traversed
96 this barrier.

97

98 If glacial refugia have played an important role in CAR plant dynamics we would
99 expect to find evidence of dispersal inland because most putative CAR refugia are
100 located in the Atlantic Guineo-Congolian region (Fig. 1A; Maley 1996; Anhuf et al.
101 2006). For example, within the palm species *Podococcus barteri*, modelling past
102 ranges and genetic data supported the hypothesis of one large coastal refugia in
103 western Gabon and southwestern Cameroon (Faye et al. 2016). Furthermore, we
104 would expect population size to increase towards the present as populations spread
105 out from climatically stable areas. 

106

107 To develop upon our current understanding of the phylogeographic patterns
108 introduced above, we present, for the first time using nuclear phylogenomic
109 approaches, the evolutionary dynamics of a central African tree species, *Annickia*
110 *affinis*. This species belongs to the pantropical plant family Annonaceae (Chatrou et
111 al. 2012) growing up to 30 metres tall and typically inhabits primary, secondary and
112 degraded rain forests (Versteegh and Sosef 2007). The species is widespread and
113 common across Lower Guinea, from southern Nigeria to the western tip of
114 Democratic Republic of Congo and is therefore ideal for studying CAR
115 phylogeography and the nature of the climatic inversion as a phylogeographic barrier.

116

117 Here, we used a newly developed baiting kit (Couvreur et al. 2019) to sequence
118 hundreds of nuclear markers in 112 individuals covering most of the distribution of *A.*
119 *affinis*. First, we identified the major genetic clusters within *A. affinis* and their

120 distribution to test if *A. affinis* shows a North-South genetic structuring along the
121 climatic inversion. Second, we built a phylogenetic hypothesis of relationships among
122 genetic clusters and conducted spatiotemporal diffusion analyses to test if dispersal
123 has been frequent across the climatic inversion, or if it has been restricted over time.
124 We also test for an inland (west to east) dispersal pattern, congruent with expansion
125 out of climatically stable areas. Finally, we reconstruct effective population size
126 through time to infer the past demography of each identified genetic cluster to test for
127 recent population expansion, and if demographic histories are congruent among
128 clusters.

129

130 **2 MATERIAL AND METHODS**

131 **2.1 Sample collection**

132 A total of 112 individuals of *Annickia affinis* were sampled across most of the species
133 distribution range in Central Africa (Table S1). In addition, two individuals were
134 sampled from the sister species *Annickia polycarpa* as outgroups (Couvreur et al.
135 2019).

136

137 **2.2 DNA extraction, gene capture and sequencing**

138 DNA was extracted from silica gel dried leaves using the MATAB (Sigma-Aldrich,
139 Saint Louis, Missouri, USA) and chloroform separation methods following Couvreur
140 et al. (2019). Nuclear markers were captured using the Annonaceae bait kit (Couvreur
141 et al. 2019) made of 11,583 baits 120 bp long targeting 469 exonic regions. Barcoded
142 Illumina libraries were constructed based on a modified protocol of Rohland and
143 Reich (2012). See supplementary methods for details.

144

145 **2.3 Read filtering, contig assembly and multi-sequence alignment**

146 Reads were cleaned and filtered following the protocol in Couvreur et al. (2019) and
147 HybPiper (Johnson et al. 2016) was used to prepare sequence data for phylogenetic
148 inference. Alignments were conducted using MAFFT (Katoh and Standley 2013) and
149 cleaned with GBLOCKS (Castresana 2000). Putative paralogs for *A. affinis* that were
150 flagged by HybPiper were verified and removed during this processes. Further
151 information can be found in the supplementary methods.

152

153 **2.4 Phylogenetic inference**

154 We filtered our dataset by choosing only those exons that had 75% of their length
155 reconstructed in 75% of *A. affinis* individuals. We then used the corresponding
156 supercontigs (i.e. targeted regions and surrounding off-target sequences) for
157 phylogenetic inference. We added empty sequences when individuals were missing
158 from locus alignments and we concatenated loci with the pxcat function from phyx
159 (Brown et al. 2017). We assigned a different GTR+GAMMA model to each locus to
160 account for differences in substitution rates. We then ran RAxML (v8.2.9)
161 (Stamatakis 2014) using the ‘-f a’ option with 100 replicates. The tree was rooted
162 using *A. polycarpa* as outgroup. For comparison we also conducted a coalescent-
163 based phylogenetic analysis using ASTRAL-III (Zhang et al. 2017), which uses
164 individual gene trees to infer a species tree. Finally, we constructed a phylogenetic
165 network using splitstree (v4.14.6; Huson and Bryant 2006) and the full SNP dataset
166 (see below) using the neighbour-net algorithm.

167

168 **2.5 Phylogeographic Diffusion in Continuous Space**

169 We reconstructed the spatiotemporal dynamics of *A. affinis* using BEAST v1.8.4
170 (Drummond and Rambaut 2007) and spread3 v0.9.6 (Bielejec et al. 2011). As this
171 analysis is computationally intensive we used a subset of our dataset by selecting only
172 the five most informative loci based on number of phylogenetic informative sites. We
173 added a partition consisting of longitude and latitude coordinates as continuous trait
174 data. We used a HKY+G substitution model and a strict clock model for each genetic
175 locus and an exponential growth coalescent tree prior. We ran the analysis for 100
176 million generations and assessed effective sample sizes (ESS) using Tracer v1.7
177 (Rambaut et al. 2018). We then used spread3 to visualize the output at several time
178 points during the history of *A. affinis*. We repeated this analysis with the next five
179 most informative loci to ensure similar patterns were recovered across datasets.

180

181 **2.6 SNP calling**

182 We used SeCaPr (v1.1.4; Andermann et al. 2018) to call SNPs as it generates a
183 pseudoreference made up of the consensus sequences for each target locus (paralogs
184 removed) that is tailored to the given dataset, which is more efficient than the bait kit
185 reference made from distantly related Annonaceae species. We then mapped our
186 cleaned, paired reads to this pseudoreference using BWA (v0.7.12; Li and Durbin
187 2009). Duplicates were removed and SNPs were called using HaplotypeCaller in
188 GATK (v4.0; McKenna et al. 2010). We used bcftools (v1.8; Li 2011) to apply
189 thresholds of mapping quality (>40%) depth (>25), quality by depth (>2) to filter
190 SNPs. We also removed those SNPs with a minor allele frequency < 0.01, kept only
191 biallelic SNPs and excluded monomorphic sites .

192

193 **2.7 Population genetic clustering and statistics**

194 We examined the genetic structure of *A. affinis* using three approaches. First, we
195 undertook a Discriminant Analysis of Principal Components (DAPC) (Jombart et al.
196 2010). We used the function *find.clusters* in the R package ‘adegenet’ (Jombart 2008)
197 to identify clusters using successive K-means with 100,000 iterations per value of k
198 and a maximum k value of 20. We identified the most appropriate number of clusters
199 by examining the change in Bayesian Information Criterion (BIC) with increasing
200 values of k (number of clusters). We then used the function *dapc* in order to define the
201 diversity between the groups identified using *find.clusters*. We performed cross-
202 validation of our DAPC analysis to ensure our chosen number of PCs was reliable.
203 We used cluster membership inferred using DAPC to define populations for
204 calculating summary statistics (see supplementary methods) and downstream
205 analyses.

206

207 Second, we used TESS3 - an approach that takes into account geographic location
208 information when inferring population clusters (Caye et al. 2016). TESS3 was
209 implemented using the R package ‘tess3r’ (Caye et al. 2016). We used the projected
210 least squares algorithm and a maximum k of 20. We examined the cross-validation
211 score for each value of K to identify the appropriate number of clusters.

212

213 Third, we used fastSTRUCTURE (Raj et al., 2014) on a reduced set of unlinked SNPs
214 (one per locus). We ran fastSTRUCTURE using the default settings and the simple
215 prior. The script ‘chooseK.py’ was used to identify the number of clusters that
216 maximized the marginal likelihood and explained the structure in the data.

217

218 **2.8 Demographic history**

219 We used stairway plot (v2; Liu and Fu 2015), a model-flexible approach that uses site
220 frequency spectra (SFS) to infer changes in effective population size (N_e) through
221 time. We generated filtered VCF files representing each cluster as detailed above but
222 did not apply a minor allele frequency filter. We then calculated folded SFS for each
223 cluster. Stairway plot uses SNP counts to estimate the timing of events and changes in
224 N_e so the removal of SNPs with missing data may skew counts. To overcome this we
225 modified each SFS by first calculating the minor allele frequency at each SNP and
226 then multiplying this by the mean number of sequences (haploid samples) at each site.
227 This results in a new SFS that makes use of all observed site frequencies and
228 minimizes the number of SNPs removed. The total of samples is slightly reduced
229 based on the amount of missing data. The number of random breakpoints were
230 calculated as recommended in the manual. We used 67% of sites for training and
231 performed 200 bootstrap replicates. The number of observed sites was calculated as
232 the total length of the pseudoreference. We used an angiosperm wide mutation rate of
233 5.35×10^{-9} sites/year (De la Torre et al. 2017) and a generation time of 15 years
234 based on the generation time of the Annonaceae species *Annona crassiflora*
235 (Collevatti et al. 2014). In addition, sequencing error can skew the SFS by inflating
236 the number of singletons. We ran analyses using the entire SFS, and then reran with
237 singletons removed to ensure similar histories were reached.

238

239 **2.9 Modelling of current and past ranges**

240 Current and Last Glacial Maximum (21k years ago; LGM) potential distributions
241 were modelled using MaxEnt (v3.3.3; Phillips et al. 2006) as implemented in
242 ‘biomod2’ (v3.3-7.1; Thuiller et al. 2009). Current and LGM (MIROC global
243 circulation model) climatic data were downloaded from WordClim ver. 1.4 (Hijmans

244 et al. 2005) at a resolution of 10*10 arc-minute. The LGM period represents the latest
245 unfavourable climate for tropical species and is therefore a good period to model the
246 impact of past climate change on potential range. A total of 346 presence data points
247 (Table S1) covering the known distribution of *A. affinis* were spatially filtered to one
248 point per cell to avoid overfitting due to sampling bias. Model overfitting was
249 constrained by using the β regularization parameter in Maxent, which limits model
250 complexity (Radosavljevic and Anderson 2014), and was set to 2.00 and 4.00, rather
251 than the default MaxEnt value of 1.00. Modelling with all 19 bioclim variables
252 produced unrealistic results and failed to properly model the current species range
253 independent of the regularization parameter (results not shown). Using just eight
254 bioclim variables (four precipitation and four temperature, see supplementary
255 methods) greatly improved the accuracy of the models to the known distribution.
256 Model performance was evaluated using a cross-validation procedure (Ponder et al.
257 2001, Muscarella et al. 2014, see supplementary methods). Model fit was assessed
258 using area under curve (AUC; Elith et al. 2006) and the true skill statistics (TSS,
259 Allouche et al. 2006). The best fitting model was then projected into the LGM.

260

261 **3 RESULTS**

262 **3.1 Sequencing**

263 A total of 124.7 million reads were generated for 112 *A. affinis* individuals at an
264 average coverage depth of 77.5x across all targeted loci. Using HybPiper we
265 identified 366 loci where 75% of the exon length was recovered in at least 75% of
266 individuals. A total of 15 loci showed signs of paralogy and were removed, leading to
267 a final dataset of 351 supercontigs totalling 756 kb of sequence data. After cleaning

268 and filtering our SNP calling approach yielded 6,787 high-quality SNPs from 280
269 different loci.

270

271 **3.2 How are populations structured across the range of *A. affinis*?**

272 After cross-validation, we chose to keep 40 PCA axes as this was shown to be
273 appropriate for accurately inferring clusters (Fig. S1A). Changes in BIC greatly
274 decreased after $k = 4$ (Fig. S1B) suggesting that four clusters best fit our data (Fig.
275 1A, Fig. S2). Two major clusters contained 35 and 63 individuals that were located
276 primarily in Western Gabon (cluster WG) and Cameroon (cluster CA) respectively.
277 Two smaller clusters of seven individuals each were located in eastern Gabon (cluster
278 EG) and Gabon / Republic of Congo (cluster GC). There is a clear discontinuity in
279 genetic structure across the equator, separating cluster CA from the rest, except for a
280 pair of individuals belonging to cluster EG.

281

282 The TESS3 analysis also found that four clusters best defined our data (Fig. 1B, S3)
283 with geographic discontinuities generally congruent between analyses (Fig. 1; Fig.
284 S4). Individual admixture proportions revealed limited mixed ancestry within samples
285 (Fig. 1B), except at a single location. The two northern most individuals belonging to
286 cluster EG, found in at Meyo Centre in Cameroon (labelled in Fig. 1A), had a
287 considerable proportion of their ancestry from cluster CA (Fig. 1B). The inverse was
288 true of the two individuals from cluster CA that were from the same location. The
289 fastSTRUCTURE analysis supported aforementioned analyses, even with a reduced
290 SNP dataset. We found that k was between 3 and 5 and results closely mirrored
291 DAPC clusters (Fig. S5). However, there was little evidence for admixture when

292 using this approach. Our inferred phylogenetic network (Fig. 1C) also revealed four
293 major clusters, and that clusters WG & CA and EG & GC were grouped together.

294

295 **3.3 How did populations of *A. affinis* disperse across central Africa?**

296 The RAxML phylogenetic tree (Fig. 1D) was highly supported at deeper nodes,
297 giving a reliable evolutionary history between major clades. The tree topology
298 reflected our clustering inferences, lending further support to our four identified
299 clusters and robust evidence for phylogeographic structuring. The topology of our
300 ASTRAL tree was very similar to the RAxML tree (Fig. S6).

301

302 We assessed the geographic locations in these clusters at a finer scale by mapping
303 each tip of the RAxML phylogenetic tree to its collection site (Fig. S7). We
304 subdivided clusters WG and CA into major clades (I-IV) as points of reference (see
305 Fig. 1). In cluster WG (Fig. S7C) the earliest diverging individuals are found at Mt.
306 Cristal and in coastal rain forests in northwest Gabon. The remaining individuals in
307 cluster WG formed a monophyletic group (clade I) and are found to the South and
308 East, as far as the southern tip of the Republic of Congo. We then examined the
309 geographic locations in cluster CA and identified three clades with distinct geographic
310 distributions (Fig. S7D) going up the Cameroon's Atlantic coast. The middle and
311 largest cluster extended inland. Given this structure we repeated DAPC clustering
312 analyses using only cluster CA individuals and revealed fine-scale genetic structure
313 that supported these three clades (Fig. S8).

314

315 Our diffusion analysis was based on 47.3kb of sequence data across five partitions
316 and converged with the ESS > 200 for all parameters after 100 million generations.

317 The root was inferred to be around central Gabon (Fig. 2A). We estimated a single
318 lineage crossed the climatic inversion, from South to North, establishing cluster CA
319 (Fig. 2). Late in the evolutionary history of *A. affinis* another dispersal event crossed
320 the barrier at Meyo Centre (see Fig. 1A). Our repetition of the diffusion analysis with
321 different loci matched these patterns (Fig. S9) indicating our results are reliable and
322 unlikely to have been biased by particular gene histories.

323

324 **3.4 Do different populations share similar demographic histories?**

325 We estimated the demographic history of the four DAPC clusters (Fig. 3; Fig S10-13
326 for full plots). Over the last 100 thousand years (Ka) three clusters (GC, WG and CA)
327 experienced similar demographic histories with population decline around 60-80 Ka
328 followed by an increase in N_e towards the present. We found this increase began at
329 different times across these three clusters though all show rapid increases in
330 population size close to the end of the LGM, around 20 Ka. Cluster CA showed
331 evidence of a rapid growth very close to the present, in the last 4 Ka. Cluster EG had a
332 much different history, exhibiting a relatively constant population size throughout its
333 history with a gradual decline in the last 10 Ka. Results were very similar when
334 singletons were removed indicating that sequencing errors were not affecting our
335 analyses (results not shown).

336

337 **3.5 Which areas have remained climatically stable over time?**

338 A total of 113 data points were retained after filtration. The best predictors were
339 Precipitation of Wettest Month (Bio13) and Precipitation Seasonality (Bio15) (Table
340 S2). A regularization multiplier of 2 generated a better model fit than with 4, showing
341 a good visual match with the known distribution of the species at present (Fig. 4A).

342 The mean value of the AUC for the training and test data were respectively 0.77 and
343 0.76. The mean value of TSS was 0.454, indicating that the model is better than a
344 random model. During the LGM, the highest presence probabilities were all located
345 along the Atlantic coast in Cameroon, Equatorial Guinea and Gabon (Fig. 4B).

346

347 **4 DISCUSSION**

348 **4.1 Limited dispersal across the climatic hinge**

349 Intraspecific diversity, based on phylogenomic nuclear sequence data, within the
350 widespread tree species *Annickia affinis* is highly structured with a clear North-South
351 divide between identified genetic clusters (Fig. 1). This is the first time this has been
352 observed in plants using genomic data and adds to the growing evidence of an
353 important phylogeographic barrier around a climatic hinge between 0 and 3°N in
354 numerous CAR distributed plants (Hardy et al. 2013; Heuertz et al. 2014; Faye et al.
355 2016; Ley et al. 2017; Pineiro et al. 2017). This North-South discontinuity is,
356 however, generally not recovered in CAR distributed animals except in rare cases
357 (e.g. Portik et al. 2017). This suggests that the processes taking place in relation to
358 this barrier affect the flora of CARs more than the fauna. Indeed, Blatrix et al. (2017)
359 showed that this barrier was more abrupt in the studied tree species (*Barteria*
360 *fistulosa*) than within the associated symbiotic ants. However, the reasons for this
361 genetic break in a seemingly continuous rain forest region remain little understood
362 (Hardy et al. 2013).

363

364 Here, we show that, throughout the evolutionary history of *A. affinis*, a single major
365 northward cross-hinge colonisation event occurred leading to the successful
366 establishment of the Cameroon population (Fig. 2C). This result lends support to the

367 **third hypothesis** of Hardy et al. (2013), that possible environmental differences have
368 prevented multiple establishments of populations crossing the hinge. Indeed, the small
369 red to black fleshy fruits of *Annickia affinis* (Versteegh & Sosef, 2007) are frugivore-
370 dispersed (Poulsen et al. 2001; Holbrook and Smith 2000) and can potentially travel
371 long distances (> 500 m) for example by hornbills. Thus, the genetic structure of *A.*
372 *affinis* in general, and the North-South divide in particular, is not linked to seed
373 dispersal limitation *per se*.

374

375 We detected that one genetic cluster (cluster EG) extends across the climatic hinge
376 into south Cameroon (Meyo Centre site, Fig. 1A) leading to a more recent,
377 northwards migration event into the climatic hinge area (Fig. 2D). This indicates that
378 the barrier is not entirely impassable, agreeing with other studies (Hardy et al. 2013;
379 Duminil et al. 2015; Pineiro et al. 2017). The Meyo Centre site lies within the
380 inversion zone and several individuals with mixed ancestry are found here (Fig. 1A,
381 B). A similar result was found in *B. fistulosa*, with 20% of individuals sampled near
382 1°N being hybrids (Blatrix et al. 2017). In addition, the general lack of evidence for
383 admixture found between genetic clusters on either side of the inversion (Fig. 1B; Fig.
384 S5) suggest that gene flow is nevertheless rare. The existence of hybrids in the
385 absence of gene flow between clusters could be the result of intrinsic (developmental;
386 environment independent) or extrinsic (environment dependent) post-zygotic
387 isolation due to lower fitness of hybrids (Turelli et al. 2001; Blatrix et al. 2017). The
388 example of Meyo Centre provides some evidence that even if dispersal across the
389 hinge is possible, it doesn't result in the successful establishment of new populations,
390 reinforcing the phylogenetic break over time. Hardy et al.'s (2013) hypotheses were
391 not mutually exclusive and given that there is no clear barrier to dispersal of pollen or

392 seeds, there may be a role for divergence in traits such as flowering time causing
393 reproductive isolation. This may in turn be linked to why successful establishment
394 across the inversion is rare. However, more fine scale sampling and information on
395 ecological differences among populations across the hinge will be needed to test these
396 ideas. Interestingly, similar north-south phylogeographic breaks are also known from
397 the Atlantic rain forests of Brazil, due to differing climatic regimes and floral
398 compositions (Carnaval et al. 2014; Leite et al. 2016). This suggests that similar
399 processes, though not necessarily driven by exactly the same factors, might be driving
400 patterns of intraspecific diversity in different TRF regions.

401

402 **4.2 Out-of-refugia migration in northern forests**

403 The inferred evolutionary dynamics of *A. affinis* support a role for Pleistocene forest
404 oscillations in shaping intraspecific genetic diversity patterns. The four retrieved
405 clusters are found in allopatry or parapatry (Fig. 1A). This supports the hypothesis of
406 incomplete mixing after post-glacial expansion and is similar to patterns found in
407 other CAR species (Hardy et al. 2013) and within species from other TRF regions
408 (Carnaval et al. 2009; Leite and Rogers 2013). Evidence was found for recent
409 demographic expansion in three clusters (CA, GC and WG, Fig. 3), as would be
410 expected if *A. affinis* expanded out of refugia. These expansions were estimated to
411 have taken place 15-25 Ka but we note that further work is needed to determine a
412 more accurate mutation rate and generation time for *A. affinis* to verify the timing of
413 these events. Therefore we avoid interpreting the exact timing of demographic events
414 and instead focussed on the population size trends. Sampling sizes were also small (n
415 =7) for clusters GC and EG meaning we are less confident in the patterns
416 reconstructed for these clusters. Similar patterns of recent expansion were detected in

417 populations of central African plants (Pineiro et al. 2017) and animals (Bell et al.
418 2017) as well as in studies on neotropical flora (Vitorino et al. 2016) and fauna
419 (Batalha-Filho et al. 2012). The refuge hypothesis has received support from
420 population genetic studies of CAR plants, showing concordance between putative
421 refugia and regions of high or unique allele/haplotype diversity (Lowe et al. 2010;
422 Dauby et al. 2014; Heuertz et al. 2014; Faye et al. 2016). However, cluster EG
423 showed constant population size with a slight decline towards the present, perhaps
424 indicating that refugia have not played an important role in its demographic history.
425 Similar demographic patterns were found in populations of two central African
426 *Erythrophleum* species (Duminil et al. 2015) though these exhibited a more
427 pronounced decline in the last 50 Ka. Overall, our results indicate that demographic
428 responses to past climate change have been different among populations of *A. affinis*
429 across central Africa. Similar patterns of recent expansion were detected in
430 populations of central African plants (Pineiro et al. 2017) and animals (Bell et al.
431 2017) as well as in studies on neotropical flora (Vitorino et al. 2016) and fauna
432 (Batalha-Filho et al. 2012). The refuge hypothesis has received support from
433 population genetic studies of CAR plants, showing concordance between putative
434 refugia and regions of high or unique allele/haplotype diversity (Lowe et al. 2010;
435 Dauby et al. 2014; Heuertz et al. 2014; Faye et al. 2016).

436
437 While our results are mixed, we do find evidence to support the scenario presented by
438 Anhof et al. (2006) who proposed that coastal rain forests in central Africa acted as
439 refugia during the LGM. The modelled LGM distribution of *A. affinis* indicates that
440 suitable habitat was concentrated continuously along the coast, from Cameroon to
441 Gabon (Fig. 4B), like in the understory palm species *Podococcus barteri* (Faye et al.

442 2016). In addition, we uncovered fine-scale genetic structure and evidence for
443 dispersal eastwards in Cameroon (Fig. 2), demonstrating a possible out-of-refugia
444 pattern in this area.

445

446 In contrast, an inland pattern of migration was not found in Gabon where dispersal
447 was both towards the east and west from a central area. This may be because there
448 was a large amount of highly-suitable area (>0.8) during the LGM that extended
449 further from the coast in Gabon than in Cameroon (Fig. 4), meaning that populations
450 could persist and expand out of this area. In addition, we inferred more pronounced
451 East-West clustering (Fig. 1) in Gabon than in Cameroon, which has been observed in
452 at least four other CAR tree species (Hardy et al. 2013). Bringing our results together,
453 it appears that refugia may have played a different role for populations in different
454 areas, and that each has responded to past climate in change in its own way.

455

456 **5 CONCLUSIONS**

457 This study uncovered the evolutionary dynamics and demographic history of the CAR
458 tree species *Annickia affinis*. Our approach is the first to use genome-wide data from
459 hundreds of nuclear loci to infer population-level phylogeographic patterns in CAR
460 plants. We found high levels of genetic structure consistent with a pattern of North-
461 South genetic discontinuity often recovered in this region. We highlighted how a
462 climatic inversion limits colonisation and shapes patterns of population structure
463 across a continuous rain forest region. We also show that the current distribution of
464 extant populations is the result of different demographic histories and, in northern
465 regions, migration inland from putative refugia in coastal rain forests. This study
466 provides a proof-of-concept for future work taking advantage of recent genomic

467 resources, such as the sequence capture kit used here, to improve our understanding of
468 TRF evolution, at the population level and above.

469

470 **FUNDING**

471 This study was supported by the Agence Nationale de la Recherche (grant number
472 ANR-15- CE02-0002-01 to TLPC).

473

474 **ACKNOWLEDGEMENTS**

475 We thank Prof. Moutsambote, Raoul Niangadouma, Théophile Ayole for help in the
476 field. The authors acknowledge the IRD itrop HPC (South Green Platform) at IRD
477 Montpellier for providing HPC resources. We are grateful to the Centre National de la
478 Recherche Scientifique et Technique (CENAREST), the Agence National des Parques
479 Nationaux (ANPN) and Prof. Bourobou Bourobou for research permits (AR0020/16;
480 AR0036/15 (CENAREST) and AE16014 (ANPN). Fieldwork in Cameroon was
481 undertaken under the “accord cadre de cooperation” between the IRD and Ministère
482 de la Recherche Scientifique et Technique (MINRESI). Prof. Bouka Biona of the
483 Institut National de Recherche en Sciences Excates et Naturelles (IRSEN) of the
484 Republic of Congo is thanked for research permits.

485

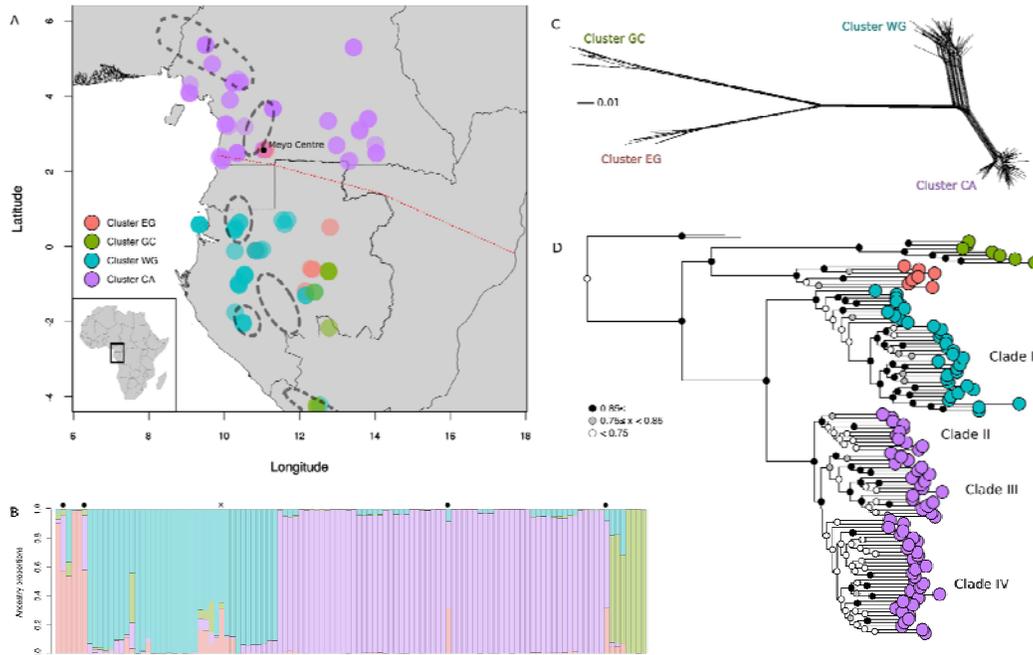
486 The authors have no competing interests.

487

488 **FIGURES**

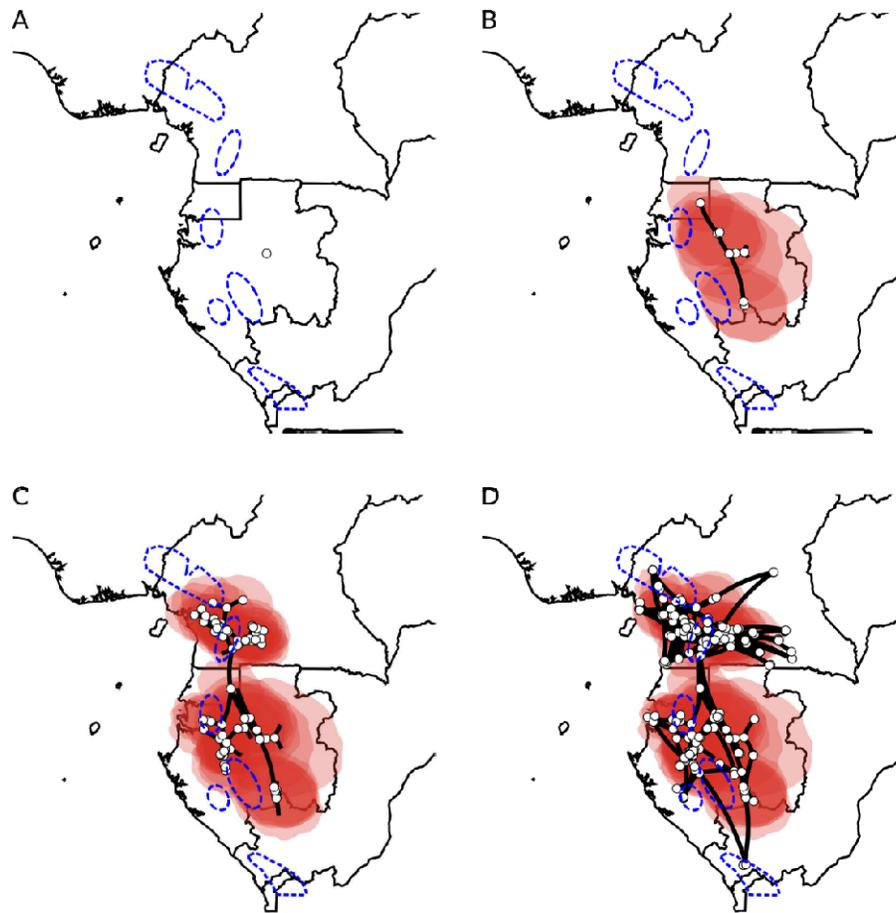
489 Colour should be used for all figures

490



491 **Fig. 1.** (A) Map of the study region showing genetic clusters inferred using
492 Discriminant analysis of principal components (DAPC, $k=4$). Individuals are colour
493 coded by cluster membership. Superimposed upon the map are the locations of
494 putative glacial refugia (adapted from Faye et al. 2016). The climatic hinge is shown
495 by a dashed red line. Inset is a map of the African continent showing highlighting the
496 study area. (B) Barplot of ancestry proportions inferred using TESS ($k=4$). Colours
497 were made to correspond to those in Fig. 1A as clustering was almost identical
498 between approaches. A single individual, “A_affinis_Ndjole_5” (marked with an x),
499 was inferred as cluster EG (red) in DAPC but TESS suggests the majority of its
500 ancestry is instead from cluster WG (blue). Individuals from the “Meyo Centre”
501 collection site (marked with a black circle in panel A) show evidence of admixture
502 between clusters across the North-South climatic inversion. (C) Phylogenetic network

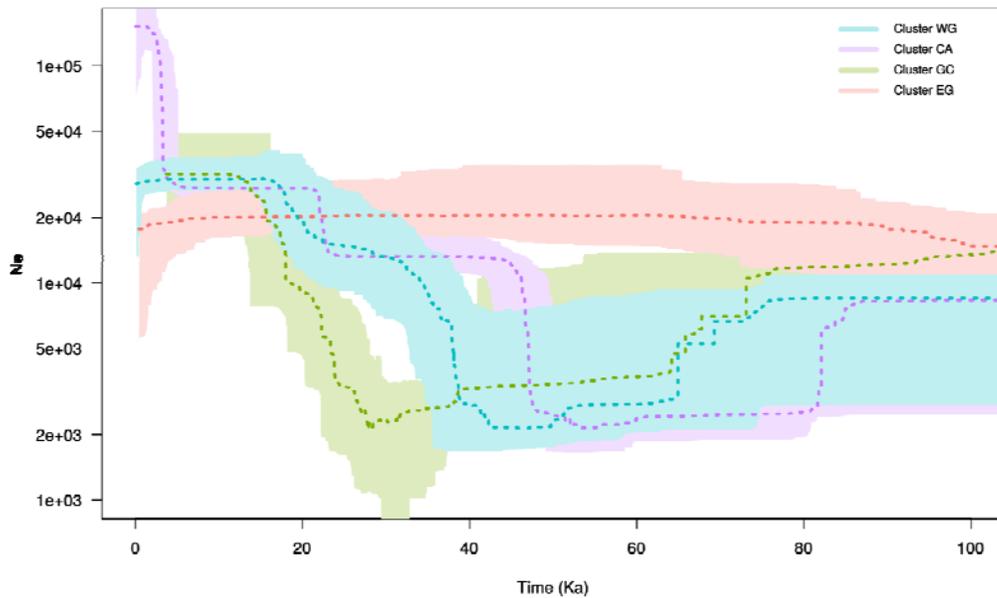
503 among *A. affinis* individuals constructed in splitstree using NeighbourNet algorithm
504 based on 6787 SNPs. (D) RAxML tree representing relationships among *Annickia*
505 *affinis* samples, rooted on two *A. polycarpa* samples. Support values are shown and
506 tips are coloured based on genetic clustering results.
507



508

509 **Fig. 2.** Phylogeographic diffusion analysis split into four time slices. Images were
510 rendered using spread3 and move forward through time starting from the
511 (uncalibrated) time of the most recent common ancestor (A) to the present day (D).
512 White circles represent ancestrally estimated geographic locations for nodes in the
513 inferred phylogenetic tree, as well as current, real locations at tips. Polygons around
514 points represent uncertainty of estimated ancestral locations at 80% highest posterior
515 density (HPD). Putative refugia following Maley (1996) are shown in dashed blue
516 lines.

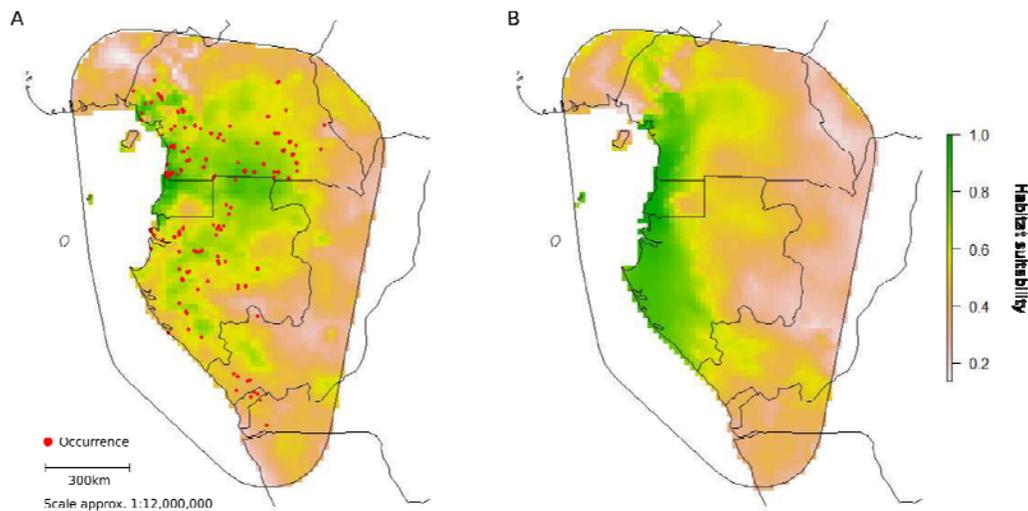
517



518

519 **Fig. 3.** Plots of effective population size through time for each of the four clusters
520 inferred using stairway plot. The present is located on the left side of each graph. The
521 dotted line represents the median population size and the shaded polygon represents
522 the 80% central posterior density intervals. Colours correspond to the colours used in
523 figure 1. Full plots of each species can be found in figures S10-13.

524



525

526 **Fig. 4.** Species distribution models (SDMs) for the present (A) and projected into the
527 past, during the last glacial maximum (B). SDMs were constructed using MaxEnt and
528 bioclimatic variables. The colour scale represents habitat suitability for each cell
529 where green indicates more suitable cells. Red circles in (A) indicate sites where *A.*
530 *affinis* individuals were collected and used in building the model.

531

532

533 **REFERENCES**

- 534 Allouche O, Tsoar A, Kadmon R. 2006 Assessing the Accuracy of Species
535 Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS).
536 *Journal of Applied Ecology* 43(6):1223–32. [https://doi.org/10.1111/j.1365-](https://doi.org/10.1111/j.1365-2664.2006.01214.x)
537 [2664.2006.01214.x](https://doi.org/10.1111/j.1365-2664.2006.01214.x).
- 538 Andermann, T, Cano Á, Zizka A, Bacon C, and Antonelli A. 2018 SECAPR—a
539 Bioinformatics Pipeline for the Rapid and User-Friendly Processing of
540 Targeted Enriched Illumina Sequences, from Raw Reads to Alignments.
541 *PeerJ* 6:e5175. <https://doi.org/10.7717/peerj.5175>.
- 542 Anhuf D, Ledru M-P, Behling H, Da Cruz FW, Cordeiro RC, Van der Hammen
543 T, Karmann I et al. 2006 Paleo-Environmental Change in Amazonian and
544 African Rainforest during the LGM. *Palaeogeography, Palaeoclimatology,*
545 *Palaeoecology* 239(3–4):510–27.
546 <https://doi.org/10.1016/J.PALAEO.2006.01.017>.
- 547 Avisé, JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA,
548 and Saunders NC. 1987. Intraspecific Phylogeography: The Mitochondrial
549 DNA Bridge Between Population Genetics and Systematics. *Annual Review*
550 *of Ecology and Systematics* 18(1):489–522.
551 <https://doi.org/10.1146/annurev.es.18.110187.002421>.
- 552 Batalha-Filho H, Cabanne GS, Miyaki CY. 2012 Phylogeography of an Atlantic
553 Forest Passerine Reveals Demographic Stability through the Last Glacial
554 Maximum. *Molecular Phylogenetics and Evolution* 65(3):892–902.
555 <https://doi.org/10.1016/j.ympev.2012.08.010>.
- 556 Bell, RC, MacKenzie JB, Hickerson MJ, Chavarría KL, Cunningham M,
557 Williams S, and Moritz C. 2011 Comparative Multi-Locus Phylogeography

- 558 Confirms Multiple Vicariance Events in Co-Distributed Rainforest Frogs.
559 *Proceedings of the Royal Society B: Biological Sciences* 279(1730):991–99.
560 <https://doi.org/10.1098/rspb.2011.1229>.
- 561 Bielejec F, Rambaut A, Suchard MA and Lemey P. 2011 SPREAD: spatial
562 phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics*
563 27(20):2910-2912.
- 564 Blatrix R, Peccoud J, Born C, Piatscheck F, Benoit L, Sauve M, Djiéto-Lordon C
565 et al. 2017 Comparative Analysis of Spatial Genetic Structure in an Ant–
566 Plant Symbiosis Reveals a Tension Zone and Highlights Speciation
567 Processes in Tropical Africa. *Journal of Biogeography* 44(8):1856–68.
568 <https://doi.org/10.1111/jbi.12972>.
- 569 Bohoussou KH, Cornette R, Akpatou B, Colyn M, Peterhans JK, Kennis J,
570 Šumbera R et al. 2015 The Phylogeography of the Rodent Genus
571 *Malacomys* Suggests Multiple Afrotropical Pleistocene Lowland Forest
572 Refugia. *Journal of Biogeography* 42(11):2049–61.
573 <https://doi.org/10.1111/jbi.12570>.
- 574 Bonnefille R. 2007. Rainforest responses to past climate changes in tropical
575 Africa. In: Tropical rainforest responses to climate change (eds Bush MB,
576 Flenley JR), pp.117–170, Praxis Publishing, Chichester
- 577 Brown, JW, Walker JF, Smith SA. 2017 Phyx: Phylogenetic Tools for Unix.
578 *Bioinformatics* 33(12):1886–88.
579 <https://doi.org/10.1093/bioinformatics/btx063>.
- 580 Carnaval, AC, Hickerson MJ, Haddad CFB, Rodrigues MT, and Moritz C. 2009
581 Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot.
582 *Science* 323 (5915):785–89. <https://doi.org/10.1126/science.1166955>.

- 583 Carnaval AC, Waltari E, Rodrigues MT, Rosauer D, VanDerWal J, Damasceno
584 R, Prates I et al. 2014 Prediction of Phylogeographic Endemism in an
585 Environmentally Complex Biome. *Proceedings of the Royal Society B:*
586 *Biological Sciences* 281(1792):20141461.
587 <https://doi.org/10.1098/rspb.2014.1461>.
- 588 Castresana J. 2000. Selection of Conserved Blocks from Multiple Alignments for
589 Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* 17
590 (4): 540–52. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>.
- 591 Caye K, Deist TM, Martins H, Michel O, François O. 2016 TESS3: Fast
592 Inference of Spatial Population Structure and Genome Scans for Selection.
593 *Molecular Ecology Resources* 16(2):540–48. [https://doi.org/10.1111/1755-](https://doi.org/10.1111/1755-0998.12471)
594 [0998.12471](https://doi.org/10.1111/1755-0998.12471).
- 595 Chatrou LW, Pirie MD, Erkens RHJ, Couvreur TLP, Neubig KM, Abbott JR,
596 Mols JB, Maas JW, Saunders RMK, Chase MW. 2012 A New Subfamilial
597 and Tribal Classification of the Pantropical Flowering Plant Family
598 Annonaceae Informed by Molecular Phylogenetics. *Botanical Journal of*
599 *the Linnean Society* 169(1):5–40. [https://doi.org/10.1111/j.1095-](https://doi.org/10.1111/j.1095-8339.2012.01235.x)
600 [8339.2012.01235.x](https://doi.org/10.1111/j.1095-8339.2012.01235.x).
- 601 Collevatti RG, Telles MP, Lima JS, Gouveia FO, Soares TN. 2014 Contrasting
602 spatial genetic structure in *Annona crassiflora* populations from fragmented
603 and pristine savannas. *Plant systematics and evolution* 300(7):1719-1727.
- 604 Couvreur TLP, Helmstetter AJ, Koenen EJM, Bethune K, Brandão RD, Little
605 SA, Sauquet H, Erkens RHJ. 2019 Phylogenomics of the Major Tropical
606 Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes.

- 607 *Frontiers in Plant Science* 9 (January). Frontiers:1941.
608 <https://doi.org/10.3389/fpls.2018.01941>.
- 609 Cowling SA, Cox PM, Jones CD, Maslin MA, Peros M, Spall SA. 2008
610 Simulated glacial and interglacial vegetation across Africa: implications for
611 species phylogenies and trans-African migration of plants and animals.
612 *Global Change Biology* 14:827–840. [https://doi.org/10.1111/j.1365-](https://doi.org/10.1111/j.1365-2486.2007.01524.x)
613 [2486.2007.01524.x](https://doi.org/10.1111/j.1365-2486.2007.01524.x)
- 614 Dauby G, Duminil J, Heuertz M, Koffi GK, Stévant T, Hardy OJ. 2014
615 Congruent Phylogeographical Patterns of Eight Tree Species in Atlantic
616 Central Africa Provide Insights into the Past Dynamics of Forest Cover.
617 *Molecular Ecology* 23(9):2299–2312. <https://doi.org/10.1111/mec.12724>.
- 618 De La Torre AR, Li Z, Van De Peer Y, Ingvarsson PK. 2017 Contrasting rates of
619 molecular evolution and patterns of selection among gymnosperms and
620 flowering plants. *Molecular Biology and Evolution* 34(6):1363–1377.
- 621 Diamond AW, Hamilton AC. 1980 The Distribution of Forest Passerine Birds
622 and Quaternary Climatic Change in Tropical Africa. *Journal of Zoology* 191
623 (3): 379–402. <https://doi.org/10.1111/j.1469-7998.1980.tb01465.x>.
- 624 Drummond AJ, Rambaut A. 2007 BEAST: Bayesian Evolutionary Analysis by
625 Sampling Trees. *BMC Evolutionary Biology* 7(1):214.
626 <https://doi.org/10.1186/1471-2148-7-214>.
- 627 Duminil J, Mona S, Mardulyn P, Doumenge C, Walmacq F, Doucet JL, Hardy
628 OJ. 2015 Late Pleistocene Molecular Dating of Past Population
629 Fragmentation and Demographic Changes in African Rain Forest Tree
630 Species Supports the Forest Refuge Hypothesis. *Journal of Biogeography*
631 42(8):1443–54. <https://doi.org/10.1111/jbi.12510>.

- 632 Edwards SV, Potter S, Schmitt CJ, Bragg JG, Moritz C. 2016 Reticulation,
633 Divergence, and the Phylogeography–Phylogenetics Continuum.
634 *Proceedings of the National Academy of Sciences* 113(29):8025–32.
635 <https://doi.org/10.1073/pnas.1601066113>.
- 636 Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ et
637 al. 2006 Novel Methods Improve Prediction of Species’ Distributions from
638 Occurrence Data. *Ecography* 29(2):129–51.
639 <https://doi.org/10.1111/j.2006.0906-7590.04596.x>.
- 640 Faye A, Deblauwe V, Mariac C, Richard D, Sonké B, Vigouroux Y, Couvreur
641 TLP. 2016 Phylogeography of the Genus *Podococcus* (Palmae/Arecaceae)
642 in Central African Rain Forests: Climate Stability Predicts Unique Genetic
643 Diversity. *Molecular Phylogenetics and Evolution* 105:126–38.
644 <https://doi.org/10.1016/j.ympev.2016.08.005>.
- 645 Fuchs J, Bowie, RC. 2015 Concordant genetic structure in two species of
646 woodpecker distributed across the primary West African biogeographic
647 barriers. *Molecular phylogenetics and evolution* 88:64–74.
- 648 Hardy OJ, Born C, Budde K, Daïnou K, Dauby G, Duminil J, Ewédjé EBK et al.
649 2013 Comparative Phylogeography of African Rain Forest Trees: A Review
650 of Genetic Signatures of Vegetation History in the Guineo-Congolian
651 Region. *Comptes Rendus - Geoscience* 345(7–8):284–96.
652 <https://doi.org/10.1016/j.crte.2013.05.001>.
- 653 Heuertz M, Duminil J, Dauby G, Savolainen V, Hardy OJ. 2014 Comparative
654 Phylogeography in Rainforest Trees from Lower Guinea, Africa. *PLoS*
655 *ONE* 9(1):e84307. <https://doi.org/10.1371/journal.pone.0084307>.

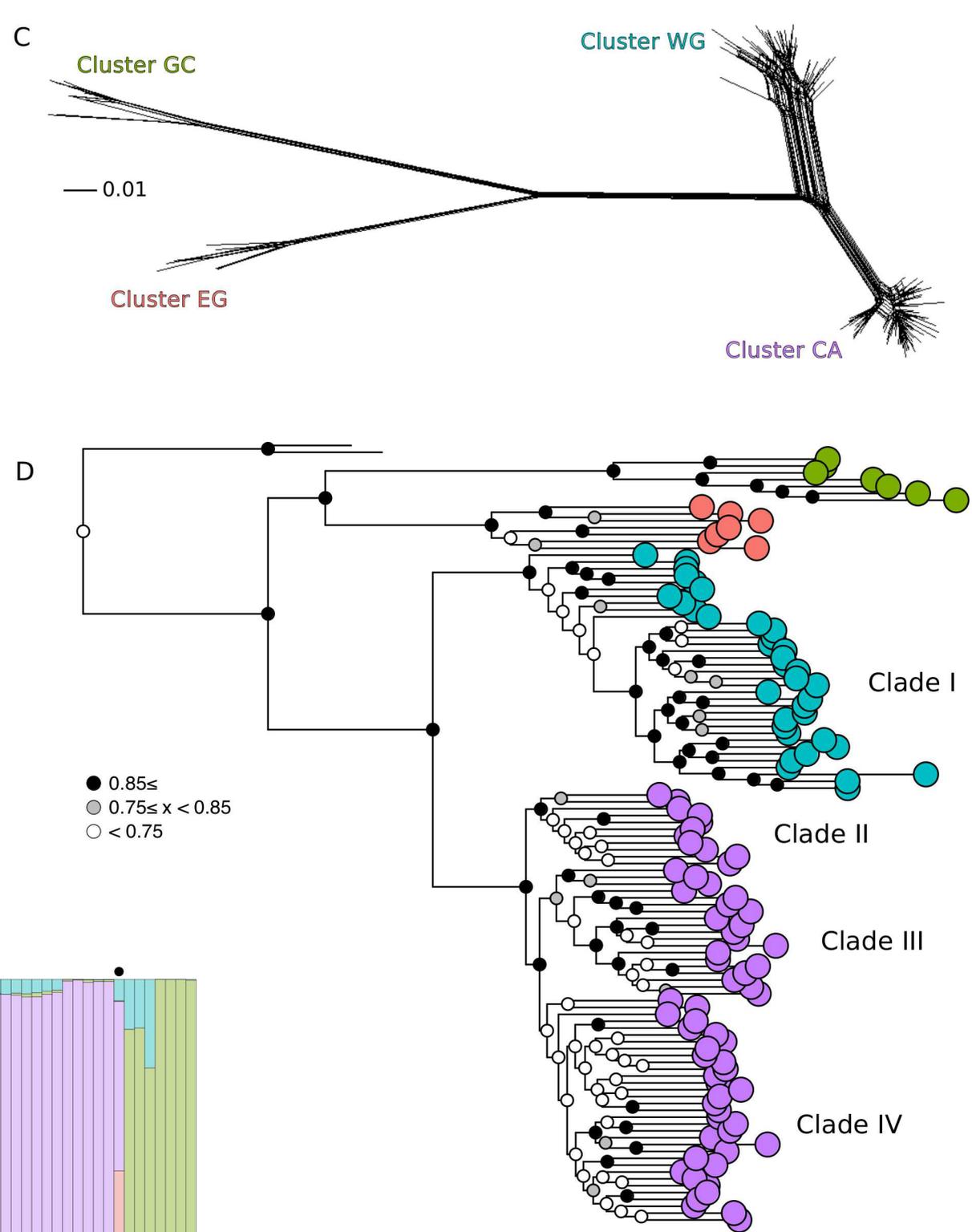
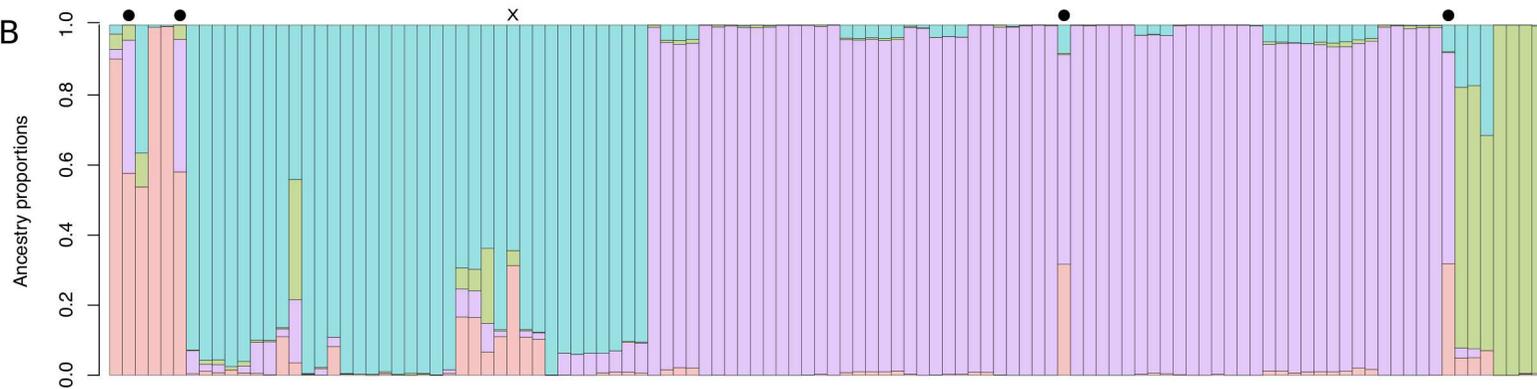
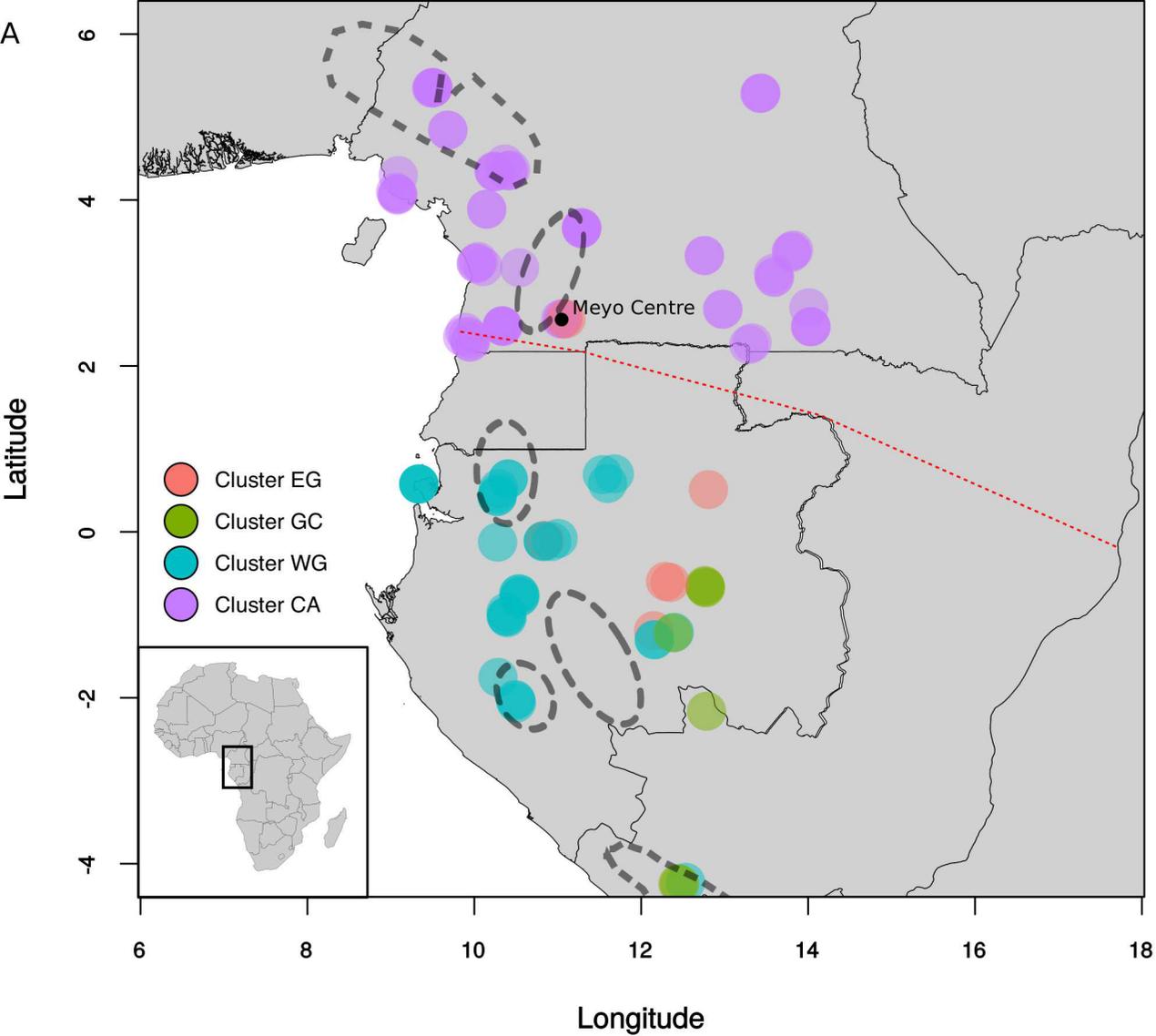
- 656 Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Very High
657 Resolution Interpolated Climate Surfaces for Global Land Areas.
658 *International Journal of Climatology* 25(15):1965–78.
659 <https://doi.org/10.1002/joc.1276>.
- 660 Holbrook KM, Smith TB. 2000 Seed dispersal and movement patterns in two
661 species of *Ceratogymna* hornbills in a West African tropical lowland forest.
662 *Oecologia* 125:249–257. <https://doi.org/10.1007/s004420000445>
- 663 Huson DH, Bryant D. 2006 Application of Phylogenetic Networks in
664 Evolutionary Studies. *Molecular Biology and Evolution*. 23(2):254-267.
665 <https://doi.org/10.1093/molbev/msj030>.
- 666 Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC,
667 Wickett NJ. 2016 HybPiper: Extracting Coding Sequence and Introns for
668 Phylogenetics from High-Throughput Sequencing Reads Using Target
669 Enrichment. *Applications in Plant Sciences* 4(7):1600016.
670 <https://doi.org/10.3732/apps.1600016>.
- 671 Jombart T. 2008 Adegenet: A R Package for the Multivariate Analysis of
672 Genetic Markers. *Bioinformatics* 24(11):1403–5.
673 <https://doi.org/10.1093/bioinformatics/btn129>.
- 674 Jombart T, Devillard S, Balloux F. 2010 Discriminant Analysis of Principal
675 Components: A New Method for the Analysis of Genetically Structured
676 Populations. *BMC Genetics* 11(1):94. [https://doi.org/10.1186/1471-2156-](https://doi.org/10.1186/1471-2156-11-94)
677 [11-94](https://doi.org/10.1186/1471-2156-11-94).
- 678 Katoh K, Standley DM. 2013 MAFFT Multiple Sequence Alignment Software
679 Version 7: Improvements in Performance and Usability. *Molecular Biology*
680 *and Evolution* 30 (4): 772–80. <https://doi.org/10.1093/molbev/mst010>.

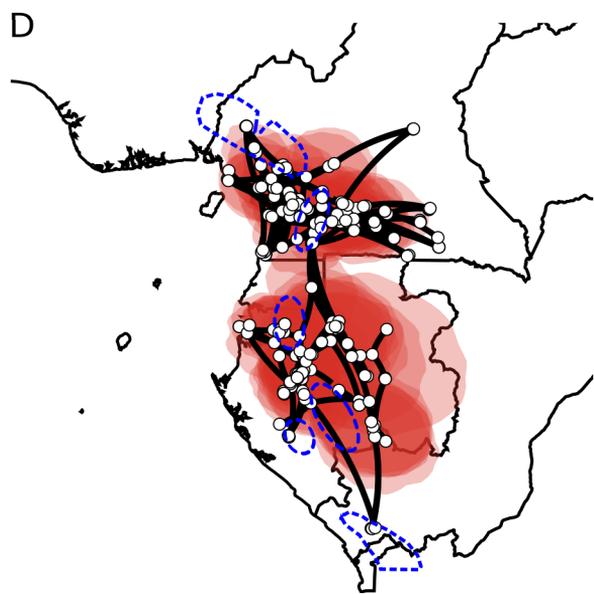
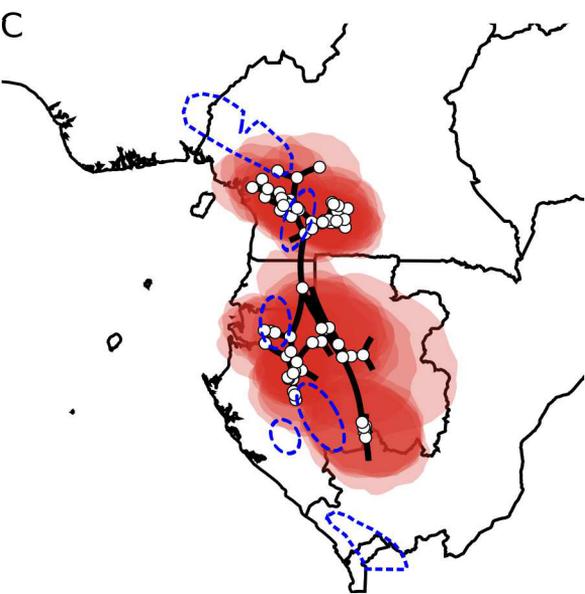
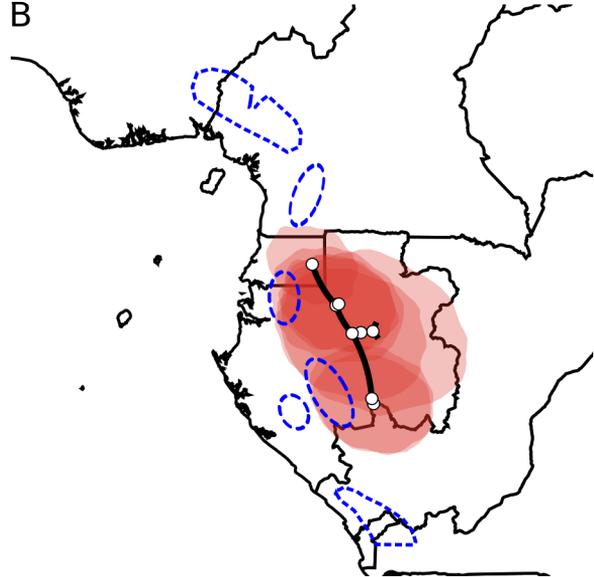
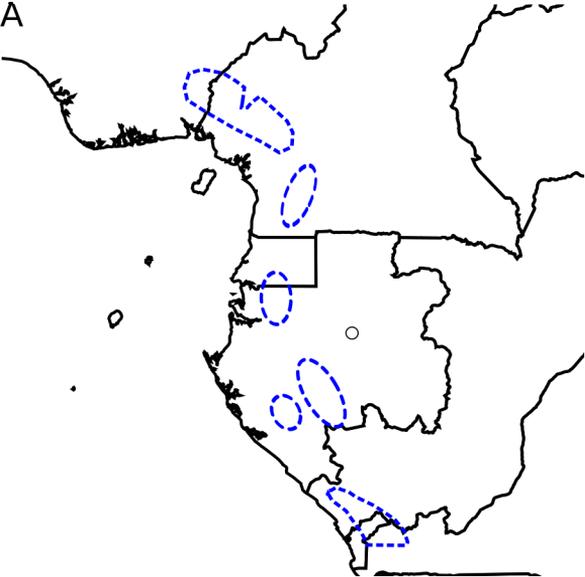
- 681 Leite RN, Rogers DS. 2013 Revisiting Amazonian Phylogeography: Insights into
682 Diversification Hypotheses and Novel Perspectives. *Organisms Diversity &*
683 *Evolution* 13(4):639–64. <https://doi.org/10.1007/s13127-013-0140-8>.
- 684 Leite, YLR, Costa LP, Loss AC, Rocha RG, Batalha-Filho H, Bastos AC,
685 Quaresma VS et al. 2016 Neotropical Forest Expansion during the Last
686 Glacial Period Challenges Refuge Hypothesis. *Proceedings of the National*
687 *Academy of Sciences* 113(4):1008–13.
688 <https://doi.org/10.1073/pnas.1513062113>.
- 689 Ley AC, Heuertz M, Hardy OJ. 2017 The Evolutionary History of Central
690 African Rain Forest Plants: Phylogeographical Insights from Sister Species
691 in the Climber Genus *Haumania* (Marantaceae). *Journal of Biogeography*
692 44(2):308–21. <https://doi.org/10.1111/jbi.12902>.
- 693 Lézine AM, Izumi K, Kageyama M, Achoundong G. 2019 A 90,000-year record
694 of Afromontane forest responses to climate change. *Science* 363(6423):177-
695 181.
- 696 Li H. 2011 A statistical framework for SNP calling, mutation discovery,
697 association mapping and population genetical parameter estimation from
698 sequencing data. *Bioinformatics* 27(21), 2987-2993.
- 699 Li H, Durbin R. 2009 Fast and Accurate Short Read Alignment with Burrows-
700 Wheeler Transform. *Bioinformatics* 25(14):1754–60.
701 <https://doi.org/10.1093/bioinformatics/btp324>.
- 702 Liu X, Fu YX. 2015 Exploring population size changes using SNP frequency
703 spectra. *Nature Genetics* 47, 555-559.

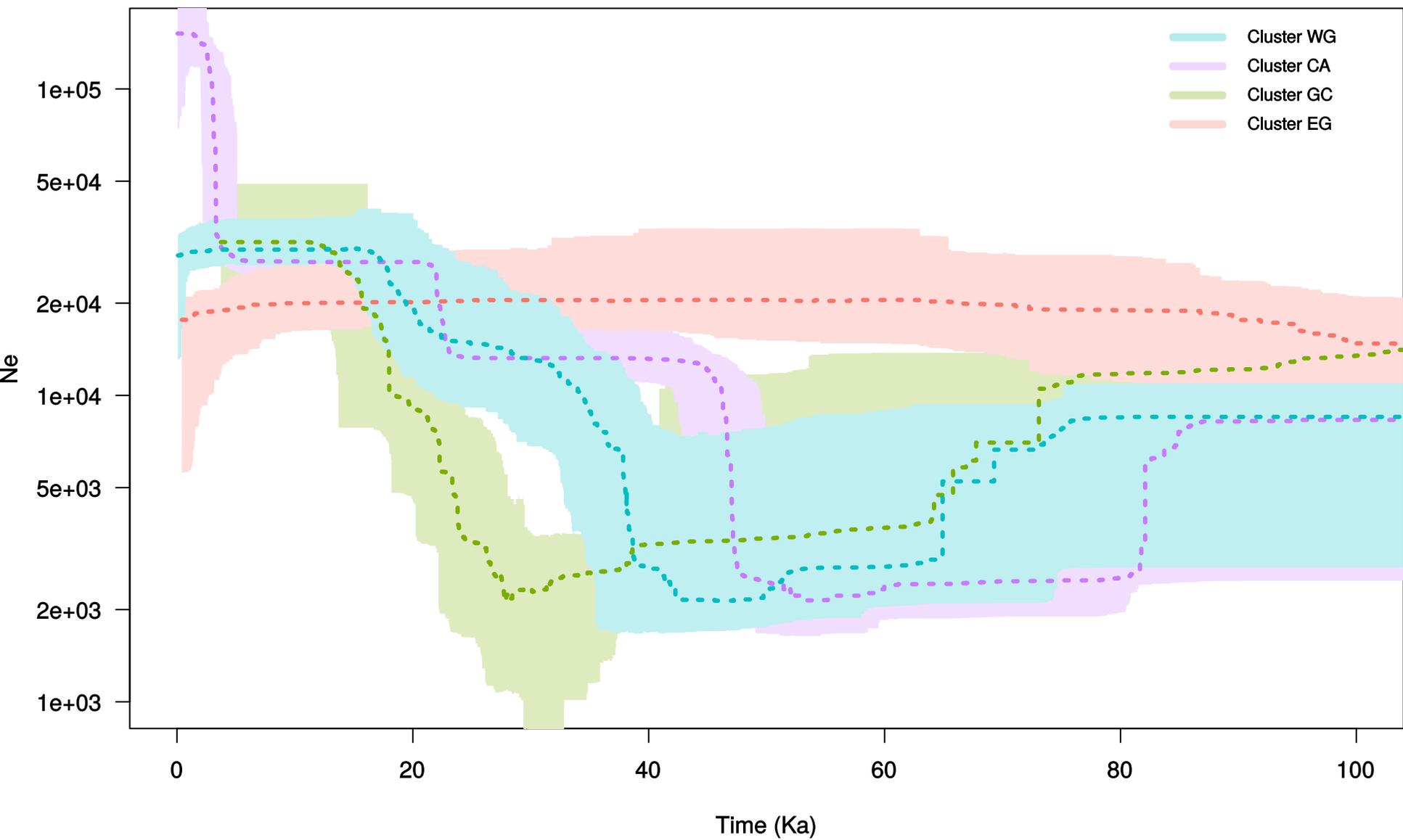
- 704 Linder HP. 2001 Plant Diversity and Endemism in Sub-Saharan Tropical Africa.
705 In *Journal of Biogeography*, 28:169–82. <https://doi.org/10.1046/j.1365->
706 2699.2001.00527.x.
- 707 Lowe AJ, Harris D, Dormontt E, Dawson IK. 2010 Testing Putative African
708 Tropical Forest Refugia Using Chloroplast and Nuclear DNA
709 Phylogeography. *Tropical Plant Biology* 3(1):50–58.
710 <https://doi.org/10.1007/s12042-010-9045-2>.
- 711 Mace GM, Gittleman JL, Purvis A. 2003 Preserving the Tree of Life. *Science*
712 300 (5626):1707–9. <https://doi.org/10.1126/science.1085510>.
- 713 Maley J. 1996 The African Rain Forest - Main Characteristics of Changes in
714 Vegetation and Climate from the Upper Cretaceous to the Quaternary.
715 *Proceedings of the Royal Society of Edinburgh Section B: Biological*
716 *Sciences* 104:31–73. <https://doi.org/10.1017/S0269727000006114>.
- 717 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K et al. 2010 The
718 Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-
719 Generation DNA Sequencing Data. *Genome Research* 20(9):254–60.
720 <https://doi.org/10.1101/gr.107524.110.20>.
- 721 Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M,
722 Anderson RP. 2014 ENMeval: An R Package for Conducting Spatially
723 Independent Evaluations and Estimating Optimal Model Complexity for
724 Maxent Ecological Niche Models. *Methods in Ecology and Evolution*
725 5(11):1198–1205. <https://doi.org/10.1111/2041-210x.12261>.
- 726 Phillips SJ, Anderson RP, Schapire, RE. 2006 Maximum Entropy Modeling of
727 Species Geographic Distributions. *Ecological Modelling* 190(3–4):231–59.
728 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.

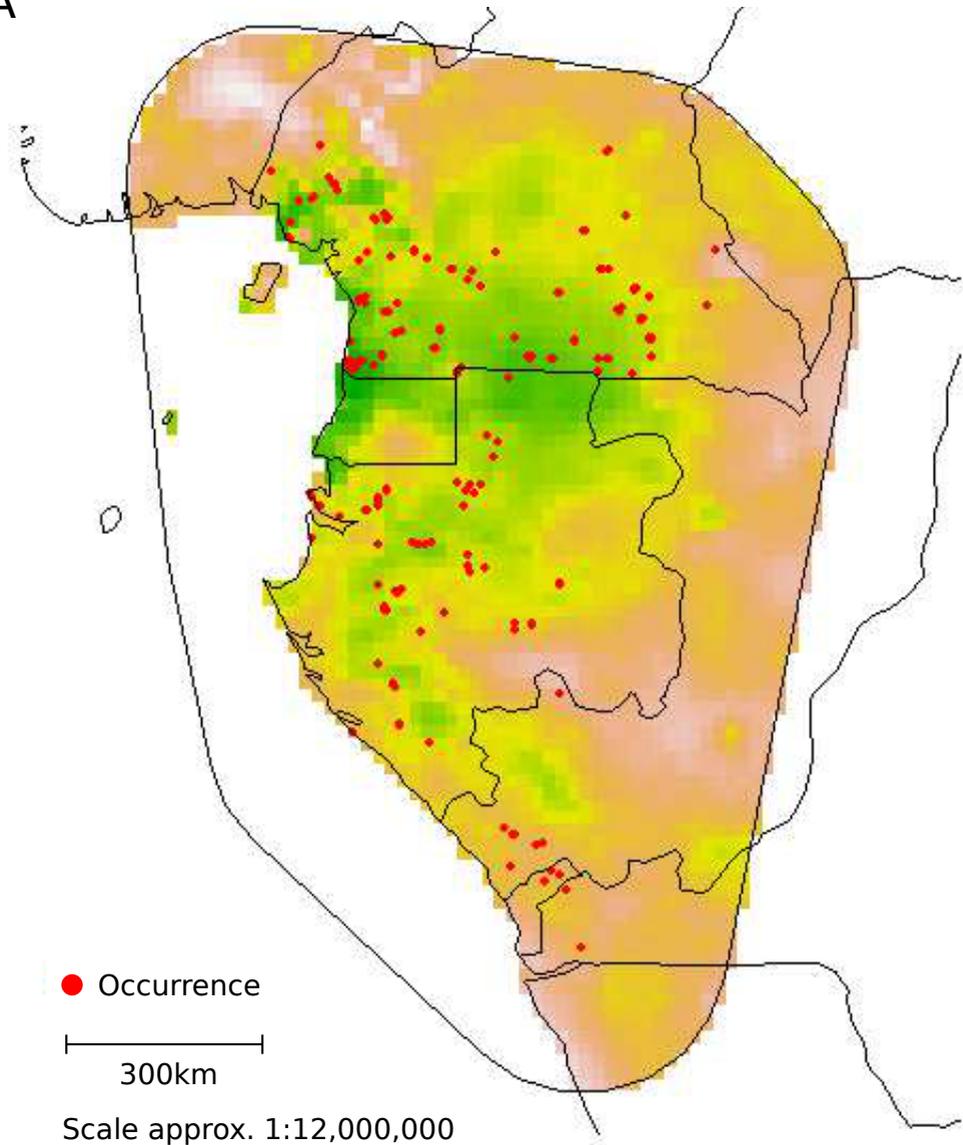
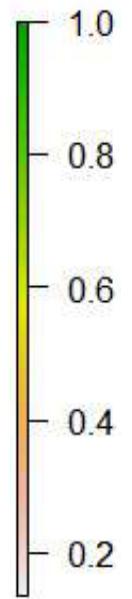
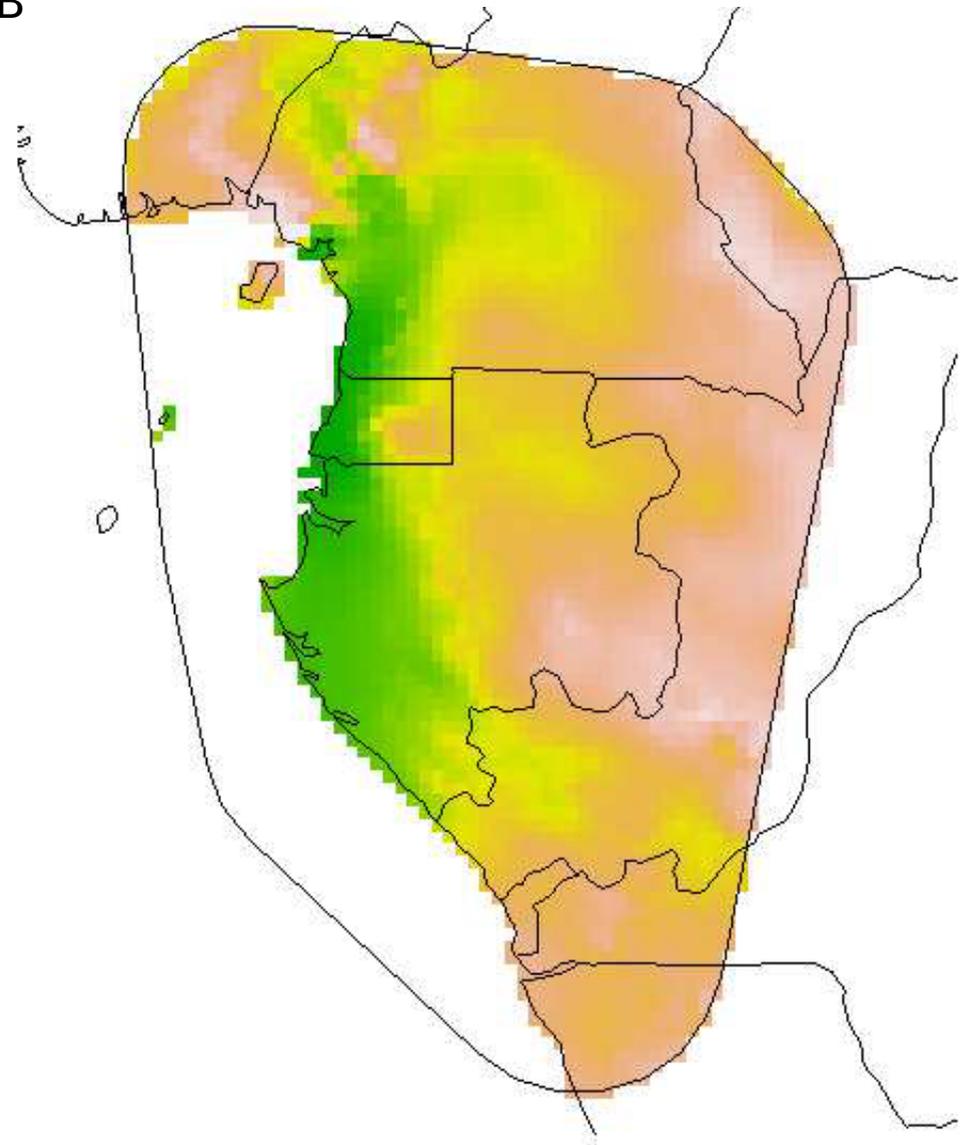
- 729 Piñeiro R, Dauby G, Kaymak E, Hardy OJ. 2017 Pleistocene Population
730 Expansions of Shade-Tolerant Trees Indicate Fragmentation of the African
731 Rainforest during the Ice Ages. *Proceedings of the Royal Society B:*
732 *Biological Sciences* 284(1866):20171800.
733 <https://doi.org/10.1098/rspb.2017.1800>.
- 734 Portik DM, Leaché AD, Rivera D, Barej MF, Burger M et al. 2017 Evaluating
735 mechanisms of diversification in a Guineo-Congolian tropical forest frog
736 using demographic model selection. *Molecular Ecology* 26(19):5245-5263.
- 737 Ponder WF, Carter GA, Flemons P, Chapman RR. 2001 Evaluation of Museum
738 Collection Data for Use in Biodiversity Assessment. *Conservation Biology*
739 15 (3): 648–57. <https://doi.org/10.1046/j.1523-1739.2001.015003648.x>.
- 740 Poulsen JR, Clark CJ, Smith TB. 2001 Seed Dispersal by a Diurnal Primate
741 Community in the Dja Reserve, Cameroon. *Journal of Tropical Ecology*
742 17:787–808.
- 743 Radosavljevic A, Anderson RP. 2014 Making better Maxent models of species
744 distributions: complexity, overfitting and evaluation. *Journal of*
745 *Biogeography* 41:629–643. <https://doi.org/10.1111/jbi.12227>
- 746 Raj A, Stephens M, Pritchard JK. 2014 FastSTRUCTURE: Variational Inference
747 of Population Structure in Large SNP Data Sets. *Genetics* 197(2):573–89.
748 <https://doi.org/10.1534/genetics.114.164350>.
- 749 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018 Posterior
750 summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic*
751 *Biology*. syy032. doi:10.1093/sysbio/syy032

- 752 Rohland N, Reich D. 2012. Cost-Effective, High-Throughput DNA Sequencing
753 Libraries for Multiplexed Target Capture. *Genome Research* 22(5):939–46.
754 <https://doi.org/10.1101/gr.128124.111>.
- 755 Stamatakis A. 2014 RAxML Version 8: A Tool for Phylogenetic Analysis and
756 Post-Analysis of Large Phylogenies. *Bioinformatics* 30(9):1312–13.
757 <https://doi.org/10.1093/bioinformatics/btu033>.
- 758 Thuiller W, Lafourcade B, Engler R, Araújo MB. 2009 BIOMOD – a platform
759 for ensemble forecasting of species distributions. *Ecography* 32:369–373.
760 <https://doi.org/10.1111/j.1600-0587.2008.05742.x>
- 761 Turelli M, Barton NH, Coyne JA. 2001 Theory and speciation. *Trends in*
762 *Ecology & Evolution* 16:330–343. [https://doi.org/10.1016/S0169-](https://doi.org/10.1016/S0169-5347(01)02177-2)
763 [5347\(01\)02177-2](https://doi.org/10.1016/S0169-5347(01)02177-2)
- 764 Versteegh CP, Sosef M. 2007 Revision of the African Genus *Annickia*.
765 *Systematics and Geography of Plants* 77(1):91–118.
766 <https://doi.org/10.1600/036364411X553108>.
- 767 Vitorino LC, Lima-Ribeiro MS, Terribile LC, Collevatti RG. 2016
768 Demographical History and Palaeodistribution Modelling Show Range Shift
769 towards Amazon Basin for a Neotropical Tree Species in the LGM. *BMC*
770 *Evolutionary Biology* 16(1):1–15. [https://doi.org/10.1186/s12862-016-](https://doi.org/10.1186/s12862-016-0779-9)
771 [0779-9](https://doi.org/10.1186/s12862-016-0779-9).
- 772 Zhang C, Sayyari E, Mirarab S. 2017 ASTRAL-III: Increased Scalability and
773 Impacts of Contracting Low Support Branches. In *Lecture Notes in*
774 *Computer Science (Including Subseries Lecture Notes in Artificial*
775 *Intelligence and Lecture Notes in Bioinformatics)* 10562 LNBI:53–75.
776 <https://doi.org/10.1007/978-3-319-67979->







A**B**

● Occurrence

300km

Scale approx. 1:12,000,000