Running Title: Population structure in Goodman's mouse lemur

Population genomic structure in Goodman's mouse lemur reveals long-standing separation of Madagascar's Central Highlands and eastern rainforests

George P. Tiley[1,†], Marina B. Blanco[1,2], José M. Ralison[3], Rodin M. Rasoloarison[3,4], Amanda R. Stahlke[5], Paul A. Hohenlohe[5], Anne D. Yoder[1]

[1]Department of Biology, Duke University, Durham, NC 27708, USA

[2]Duke Lemur Center, Duke University, Durham, NC 27705, USA

[3]Département de Biologie Animale, Université d'Antananarivo, BP 906, Antananarivo 101, Madagascar

[4]Behavioral Ecology and Sociobiology Unit, German Primate Centre, 37077 Göttingen, Germany

[5]Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA

[†]Author for correspondence: george.tiley@duke.edu

Keywords: Central Highland Savanna, Coalescent, Historical Demography, Fragmentation, RAD sequencing, Mouse Lemur

Abstract

The Central Highland Plateau of Madagascar is largely composed of grassland savanna, interspersed with patches of closed-canopy forest. Conventional wisdom has it that these grasslands are anthropogenic in nature, having been created very recently via human agricultural practices. Yet, the ancient origins of the endemic grasses suggest that the extensive savannas are natural biomes, similar to others found around the globe. We use a phylogeographic approach to compare these two competing scenarios. By sampling multiple populations of Goodman's mouse lemur (*Microcebus lehilahytsara*), a small-bodied nocturnal primate, we reconstruct the phylogeographic and demographic history of these "environmental metronomes" to estimate the time at which their populations diverged, and thus proximally, when their habitats would have become fragmented. We applied coalescent methods to RADseq data to infer phylogenetic relationships, population structure, and migration corridors among sampling sites. These analyses indicate that forest fragmentation occurred rapidly during a period of decreased precipitation near the last glacial maximum and would have affected both the Central Highlands and eastern forests. Though there is clear genomic structure separating the populations of the Central Highland from those of the eastern rainforests, there is also evidence of historical migration between them. Findings support the hypothesis that the Central Highland savanna predates human arrival, indicating that it is a natural landscape that has long impacted the population dynamics of Goodman's mouse lemur, and by extension, other forest-dwelling organisms in Madagascar.

Introduction

Madagascar is an iconic biodiversity hotspot with high levels of species diversity and endemism (Myers et al., 2000) with the majority of its biodiversity at extreme risk of extinction due to deforestation and forest fragmentation (Greene & Sussman, 1990; Harper, Steininger, Tucker, Juhn, & Hawkins, 2007; Vieilledent et al., 2018). It is widely accepted that the rapid rate of biodiversity loss is due largely to human activity and agricultural practices (Farris et al., 2017; Styger, Rakotondramasy, Pfeffer, Fernandes, & Bates, 2007). Even so, the timing of and degree to which humans have transformed and fragmented Madagascar's closed-canopy forests has been a point of contention. Much of central Madagascar is now covered with savanna; traditionally thought to be strictly the product of fires from slash-and-burn agriculture and cattle farming (Perrier de la Bâthie, 1921; Humbert, 1927; Gade, 1996). The best available archeological evidence places humans in Madagascar between two thousand years ago (KYA) and ten KYA (Dewar & Wright, 1993; Dewar et al., 2013; Hansford et al., 2018; Mitchell, 2019). This timeline coincides with evidence of increased frequency of fire based on isotope records from stalagmites (Railsback et al., 2020) and charcoal records from sediment cores covering the late to mid Holocene (Burney, 1987a; Gasse & Van Campo, 1998; Virah-Sawmy, Willis, & Gillson, 2010). Pollen abundances from these same sediment cores also imply that the Central Highland vegetation transitioned from woody taxa to grasses during this timeframe, thus supporting the hypothesis that the Central Highland Savanna grasslands are the consequence of human activities. Confusingly, however, it appears that the transition to savanna likely began following Pleistocene glaciations (Burney, 1987b; Gass & Van Campo, 1998), long before the arrival of humans, and that similar shifts in plant composition were not limited to the Central Highlands (Burney, 1993; Virah-Sawmy, Willis, & Gillson 2009). This therefore implies that the Central Highland Savanna (CHS) is not strictly the product of human agricultural practices, and instead, may be the consequence of global climate fluctuations, perhaps facilitated by the natural processes of fire ecology (Bowman & Franklin, 2005).

In support of this second scenario, ecological and phylogenetic studies of Madagascar's grasses suggest the that the CHS originated well before the Holocene, with levels of endemism comparable to the putatively natural grasslands of Africa (Bond, Silander Jr, Ranaivonasy, & Ratirarson, 2008; Vorontsova et al., 2016). And though few C_4 grass lineages such as *Aristida* and *Loudetia* dominate the CHS (Koechlin, 1993; Kull, 2003; Solofondranohatra et al., 2018), diversity among endemic C_4 grass clades would be more consistent with the global diversification of C_4 grasses during the Miocene (e.g. Edwards & Smith, 2010) rather than as a consequence of a large number of independent migration events (Hackel et al., 2018). We are therefore left with a complex model of Pleistocene climate change, perhaps driving the origins of the CHS as a savanna, but overlaid with deforestation and fragmentation, likely associated with early human activity. Consequently, we hypothesize that present-day distributions of forest-dwelling species, as well as the genetic diversity and connectivity among their populations, should reflect *both* ancient climactic changes as well as recent human-meditated processes. It is our goal to examine phylogenomic patterns among populations of Goodman's mouse lemur (*Microcebus lehilahytsara*) to disentangle the impacts of these two environmental scenarios.

The geographic distribution of *M. lehilahytsara* is uniquely suited for this purpose. Though it is found in eastern montane forests, it has largely been considered to be a highland specialist (Radespiel et al., 2012; Blanco et al., 2017). Indeed, it is the only mouse lemur species known to inhabit the isolated forest fragments in the CHS (Blanco et al., in review). The extended range of *M. lehilahytsara* may be due to its ability to enter a state of prolonged torpor (Blanco et al., 2017) thus reducing metabolic demands during the period of extended low temperatures in the CHS during the winter season. While prolonged torpor or hibernation is obligatory in *Cheirogaleus*, another member of the family Cheirogaleidae (e.g. Dausmann, Glos, Ganzhorn, & Heldmaier, 2005; Blanco et al., 2018), mouse lemurs vary in their expression of daily torpor or

hibernation, with some species reported to display daily torpor (less than 24 hours) exclusively and unable to deposit fat stores to sustain prolonged torpor (e.g. Schmid, 2001; Blanco et al., 2018). Although losses in genetic diversity of endangered species putatively due to human action are observable (Olivieri, Sousa, Chikhi, & Radespiel, 2008; Andriaholinirina et al., 2014a; Andriaholinirina et al., 2014b; Baden et al., 2014), a better reconstruction of pre-human conditions can reveal if some populations were already vulnerable, especially given that some populations have been robust to recent habitat loss and fragmentation (Sgarlata et al., 2018).

A recent phylogeographic study of mouse lemurs that included a very limited representation of *M. lehilahytsara* from the CHS suggested that it diverged from its sister species *M. mittermeieri* in the eastern rainforests near 50 KYA (Yoder et al., 2016). Subsequently, the species status of *M. mittermeieri* has been called into question with compelling genetic evidence indicating that it should be synonymized with *M. lehilahytsara* (Poelstra et al., 2020). This suggests the intriguing possibility that rather than be a highland specialist, as was originally hypothesized, *M. lehilahytsara* may instead be a widely distributed generalist, capable of existing in a broad range of habitat types occurring at varying elevations. Thus, demographic studies of these mouse lemurs may illuminate pre-human conditions in Madagascar given that they can arguably be described as "environmental metronomes," expanding and contracting their populations to match climatically driven pulses in forest distributions.

Here, we sampled populations of putative *M. lehilahytsara* from multiple sites in the CHS as well as from eastern forests. Using phylogeographic approaches and coalescent methods to analyze a large RADseq dataset, we explored patterns of genetic connectivity among populations. We also estimated divergence times and effective population size (N_e) among and between them. Findings show that population structure formed well-before any human influence, with isolation of populations in the CHS and eastern forests occurring before the LGM. However, post-

divergence gene flow among the populations suggests that migration was possible between the CHS and eastern forests, posing no more of a geographic barrier than major rivers. Moreover, population sizes were largely robust to past climate change, and have actually increased through the Holocene, even within a 1 Km^2 forest patch.

Methods

Study sites and sampling

Microcebus lehilahytsara individuals were captured at five field sites. Three of the locations, from north to south, Riamalandy (S16° 17’ 06.0” E48° 48’ 54.0” 850 m), Ambatovy (S18° 49’ 30.15” E48° 19’ 06.48” 1100 m) and Tsinjoarivo (S19° 40’ 48.5” E47° 46’ 14.5” 1609 m) are mid-elevation rainforests, situated along the eastern escarpment. Forests are fragmented along the corridor; although, the distance between patches has been exacerbated by deforestation since the 1950s (Vieilledent et al. 2018). Annual rainfall is 1400 mm and daily temperatures vary between 14 and 23 °C at Riamalandy (Goodman, Raherilalao, & Wohlauser, 2018), while Tsinjoarivo and Ambatovy, receive 1700-2000 mm of rainfall with temperatures between 15 and 23 °C (Goodman et al. 2018). At their lowest, temperature can approach freezing conditions at Tsinjoarivo (Dausmann & Blanco, 2016). The other two locations, Ambohitantely and Ankafobe (S18° 06’ 23.1” E47° 11’ 13.2” 1472 m) are located within the Central Highlands. Their vegetational components are characteristic of moist evergreen forests, although more open and drier (Goodman et al., 2018). Annual rainfall is around 1460 mm, and daily temperature between 12.5 and 23 °C, though in winter temperatures can drop substantially, below 6 °C (Goodman et al., 2018). In addition to generally drier conditions at Ambohitantely and Ankafobe compared to Ambatovy and Tsinjoarivo, these sites experience greater seasonality of precipitation too (Supplementary Figure S1). Ambohitantely is a 5600 ha protected area 130 Km northwest of Antananarivo that incorporates natural forest, savanna, and exotic tree plantations, but is subject to annual fires and threated by forest loss. Samples were collected in 2004

146 (Riamalandy), 2006 (Ambohitantely), 2007 (Tsinjoarivo), between 2009 and 2015 (Ambatovy),
147 and between 2015 and 2016 (Ankafobe). Mouse lemurs were captured using live Sherman
148 traps and released back to the forest later the same day. For description of capture methods
149 see Blanco et al. (2017).
150
151 *RADseq library preparation and sequencing*
152 *Microcebus lehilahytsara* samples were gathered from existing collections and a combination of
153 genomic DNA from available tissue and previous whole-genome amplifications were used in
154 order to maximize sampling of these rare and vulnerable species. Library preparations for RAD
155 sequencing followed Ali et al. (2016), using a single-digest protocol with the restriction enzyme
156 SbfI. A median insert size of 380 bp was selected for paired-end sequencing at the Duke
157 Sequencing Core on an Illumina Hi-Seq 4000 with an anticipated 15x coverage per individual.
158
159 *Assembly and orthology evaluation*
160 Data were assembled using STACKS v2.0b (Catchen, Amores, Hohenlohe, Cresko, &
161 Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) using *de novo*
162 methods (Supplementary Methods). Final phased loci (we use locus or loci to refer to a
163 contiguous RAD fragment while sites refer to individual base pairs throughout) output by stacks
164 were filtered such that at least three of the five sample geographic locations were represented
165 with a maximum of 50% missing data (Supplementary Methods). These filtering options were
166 applied in order to maintain computational feasibility of downstream model-based hypothesis
167 testing procedures in addition to not biasing our analyses due to uneven sampling from
168 localities. Additionally, most RADseq-based studies rely solely on heuristic measures such as
169 numbers of individuals and some level of heterozygosity for identifying orthologous clusters; we
170 use a different approach for evaluating orthology that we argue is less sensitive to any
171 clustering parameters. We used syntenic evidence from the *Microcebus murinus* reference

172 genome, with BLASTN (Altschul et al.,1990; Camacho et al., 2009) search of assembled RAD
173 loci, to identify clusters susceptible to paralogy as well as merge orthologous clusters that were
174 over-split by STACKS (Supplementary Methods).
175
176 An important note regarding our sequence data and library preparation method is that we
177 expect even depth of coverage for the forward read with the restriction enzyme cut site, but
178 lowered coverage for the reverse read with the staggered cut. Therefore, while the longer
179 contigs assembled with the paired-end data are useful for circumscribing orthologous loci, we
180 are skeptical of our ability to consistently call heterozygous loci with high confidence beyond the
181 first 151 bp of each locus; many of our downstream hypothesis tests assume we can accurately
182 call these heterozygous sites. Therefore, for each sequence for each orthologous cluster, we
183 extracted 145 bp, the first 151 bp minus the 6 bp sequenced overhang from the SbfI restriction
184 enzyme, for all future analyses. These 145 bp fragments with heterozygous base calls
185 underwent a final alignment with MUSCLE (Edgar, 2004) v3.8.1551 using default parameters.
186
187 *Phylogenetic analyses*
188 We estimated a phylogenies of all *M. lehilahytsara* individuals or of sampling sites using four
189 approaches: 1) concatenated ML analysis with RAxML (Stamatakis, 2014) v8.2.11, 2)
190 topological invariants of quartets approximated with SVDquartets (Chifman & Kubatko, 2014)
191 implemented through PAUP* (Swofford, 2003) v4.0a build 165, 3) a Bayesian implementation of
192 the multispecies coalescent (MSC) model for SNP data (Bryant et al., 2012) through BEAST2
193 (Bouckaert et al., 2014) v2.4.8, 4) full-likelihood implementation of the MSC with joint gene tree
194 estimation using BPP v4.0 (Flouri, Jiao, Rannala, & Yang, 2018). Applying these four different
195 methods allows us to explore uncertainty in phylogeographic history due to model assumptions
196 and the underlying data used for each method.
197

198 RAxML analyses were unpartitioned and used the GTR+Γ model of molecular evolution
199 (Tavaré, 1986; Yang, 1994). We conducted 20 unconstrained ML searches and performed 100
200 bootstrap replicates. We also used the ML framework to perform topological hypothesis testing.
201 Considering that samples from Ambohitantely and Ankafobe are well-supported clade across all
202 analyses, our phylogenetic question reduces to a single quartet. We then tested three
203 competing hypotheses regarding the evolutionary history of Riamalandy, Ambatovy, and
204 Tsinjoarivo. Constrained optimizations of topologies and site likelihoods were performed with
205 RAxML and we evaluated support among our three competing topologies with an approximately
206 unbiased (AU; Shimodaira, 2004) test implemented in CONSEL (Shimodaira & Hasegawa,
207 2001) v0.20.
208
209 RADseq data were evaluated using multiple approaches for topological estimation from
210 invariants. First, we used full-length RAD loci with SVDquartets, which is robust to violations of
211 independence of loci (Chifman & Kubatko, 2014). Second, we analyzed only variable sites that
212 represented at least four individuals. To further examine sensitivity of topology estimation with
213 SVDquartets to our underlying data, we analyzed 10 datasets that only used biallelic SNPS,
214 randomly sampling one SNP per locus. Analyses sampled all quartets and performed 100
215 bootstrap replicates. Distributions of bootstrap trees from SVDqaurtets and ML analyses were
216 compared using unweighted Robinson-Foulds (RF) distances (Robinson & Foulds, 1981)
217 computed with RAxML.
218
219 For SNAPP, we analyzed a single dataset using one biallelic SNP per RAD locus. To reduce the
220 amount of computing time needed for each analysis, we selected only two individuals per
221 sampling location. Two replicates were run to evaluate convergence. Mixing and effective
222 sample sizes were analyzed with TRACER v1.7 (Rambaut et al., 2018), while potential scale
223 reduction factors were calculated with the R package CODA (Plummer, Best, Cowles, & Vines,

224 2006). Bayes factor analyses were used to explore competing topological hypotheses,
225 complementary to our AU tests in the ML framework. We used stepping-stone integration (Xie et
226 al., 2011) to obtain marginal likelihoods for 20, 40, and 60 steps. It is difficult to predict how
227 many steps are necessary for marginal likelihood estimation, so we used a range to observe
228 little change in differences of marginal likelihoods with increasing steps. Because differences in
229 marginal likelihoods were large, we presented natural log Bayes factors (Kass & Raferty, 1995)
230 as opposed to computing model probabilities. Prior settings and details regarding MCMC
231 options are in the Supplementary Methods (Supplementary Material).
232
233 Species tree analyses with BPP used fixed species delimitations. We used full-length RAD loci,
234 considering at least two individuals for sampling location. Eight independent chains collected
235 2500 posterior samples, which were combined for a posterior of 20000 samples total.
236 Robustness of species tree estimation to prior choice on common ancestor ages and population
237 sizes was explored with intentionally inappropriate priors (Supplementary Material).
238
239 *Delimiting population structure and testing migration hypotheses*
240 To elucidate migration patterns between *M. lehilahytsara* populations in the CHS and eastern
241 forests, we developed competing migration hypotheses and tested models with Bayes factors.
242 First, we analyzed biallelic SNPs from RAD loci with STRUCTURE (Pritchard, Stephens, &
243 Donnelly, 2000) v. 2.3.4, using the admixture model and default priors while treating each site
244 unlinked and uncorrelated. We created 20 datasets by sampling 10000 unlinked SNPS per
245 replicate. For each replicate, we ran 20 independent MCMC chains for one to eight clusters
246 assumed *a priori* (k) in order to calculate Δk for estimating the most likely number of Hardy-
247 Weinberg groups (Evanno, Regnaut, & Goudet, 2005). Each chain ran for one million
248 generations after a 100000 generation burn-in.
249

250 Population structure was also evaluated with fixation index (F_{ST}) statistics. Pairwise F_{ST} was
251 calculated between populations as described by Weir & Cockerham (1984) for SNPS with less
252 than 15% missing data in ARLEQUIN (Excoffier & Lischer, 2010) v. 3.5.2.2. Pairwise F_{ST} were
253 used to perform an AMOVA analysis (Excoffier, Smouse, & Quattro, 1992) for Central Highlands
254 (Ambohitantely and Ankafobe) versus eastern (Riamalandy, Ambatovy, and Tsinjoarivo) *M.*
255 *lehilahytsara*. A Mantel test was used to test the presence of isolation-by-distance using 9999
256 permutations. Physical distances between populations were calculated with the Geographic
257 Distance Matrix calculator v 1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg;
258 Last Accessed 2 January 2020).
259
260 While STRUCTURE analyses can estimate the posterior probabilities of admixture among
261 Hardy-Weinberg groups, this is not a hypothesis test. As an independent evaluation of gene
262 flow between *M. lehilahytsara* populations, we computed Bayes factors for isolation-with-
263 migration models using MIGRATE v4.2.14 (Beerli & Palczewski, 2010). Notably, we did not use
264 the default uniform prior for nucleotide diversity (θ) estimates. Instead, we used independent
265 effective population size (N_e) estimates from three mouse lemur species prior to forest
266 fragmentation (Olivieri et al., 2008) to develop a prior distribution for θ (Supplementary Material).
267 Two sets of migration hypothesis testing were conducted on full-length RAD loci. First, we used
268 all individuals binned into either a CHS or eastern population, despite recognizable population
269 substructure. We were limited in the number of loci computationally feasible for obtaining
270 marginal likelihood estimates, and thus analyzed 100 data sets of 100 random RAD loci;
271 missing data was allowed as long as at least two individuals per sampling site were present.
272 The second analysis attempts to differentiate between east-to-west migration and west-to-east
273 migration among Ambatovy, Ambohitantely, and Ankafobe, and does not allow for missing data.
274 The three-population analyses used four diploid individuals per population and compared
275 isolation-with-migration models to those that also included divergence. Posterior estimates of θ

276 was rescaled N_e to using a mutation rate of 1.64 x 10^{-8}, which is based on a direct germline
277 estimate for mouse lemurs (Campbell et al., 2020).
278
279 *Estimating divergence time of eastern and Central Highland mouse lemurs*
280 We used the multispecies coalescent (MSC) model for divergence time estimation among *M.*
281 *lehilahytsara* populations with BPP (Rannala & Yang, 2003; Flouri et al., 2018) v4.0. However,
282 estimating divergence times and population sizes efficiently requires a fixed and rooted
283 topology. We used the topology recovered in the majority of analyses here, and used several
284 lines of evidence to support rooting with Riamalandy (Supplementary Material), including a
285 phylogeny of *M. lehilahytsara* individuals and outgroups using previously sequenced
286 mitochondrial data (Table S2). A partitioned ML analysis and bootstrapping was conducted with
287 PAUP* v4.0a build 163 (Swofford, 2003). Phased haploid sequences were used for BPP and at
288 least two individuals per population were required for each locus. We ran eight chains of 2500
289 samples and evaluated convergence. Final MSC parameter estimates were based on combined
290 chains. Coalescent branch lengths were rescaled to absolute time using a direct germline
291 mutation rate estimate in mouse lemurs (1.64×10^{-8}; Campbell et al., 2020) and generation time
292 estimate (3.75 years/generation) based on wild (Radespiel et al., 2019) and captive (Zohdy et
293 al., 2014) populations. Robustness of divergence times to prior choice was also evaluated
294 (Supplementary Material).
295
296 *Characterizing population size changes during Pleistocene glaciations*
297 We estimated change in population size over time with extended Bayesian skyline plots (Heled
298 & Drummond 2008) with BEAST2 (Bouckaert et al., 2014). Similar analyses have also been
299 successful in discovering king penguin population increases that coincide with Holocene
300 warming (Trucchi et al., 2014). For all sampling locations with four or more individuals, we
301 sampled 100 random RAD loci eight times, with no missing data. Sequences were treated as

302 phased and haploid. Each sampled dataset was run twice to check convergence. Large
303 posteriors were required for agreement on the number of change points and population size
304 estimates between replicates; each chain ran for 500 million generations after a 100 million
305 generation burnin while sampling every 10000 generations. Operators for increasing the
306 sampling of change points followed recommendations by Trucchi et al. (2014). We used linear
307 piecewise reconstructions of population size changes.
308
309 **Results**
310 *Circumscribing orthologous RAD loci*
311 All individuals included in our analyses had an average depth of coverage greater than 10x per
312 locus assembled with STACKS. In total, 92,889 loci were recovered across all individuals with
313 an average of 49,201 loci per individual (Supplementary Table S1). On average, 31% of loci
314 were shared between pairs of individual (Figure 1). The number of loci shared between
315 individuals are likely not due to phylogenetic signal, as there was only a slight negative
316 correlation between evolutionary distance and the proportion of shared loci (Pearson's r = -0.19;
317 p = 0.001). Rather, this pattern was driven, at least in part, by the number of reads, given a
318 strong positive correlation between the number of loci and the number of reads per individual
319 (Pearson's r = 0.86; p = 3.54e-08). Although 92,889 RAD loci were recovered, only 54,018 loci
320 contained at least one biallelic SNP; other loci were completely invariable or consisted of only
321 singletons.
322
323 *Support for Central Highland and eastern Microcebus lehilahytsara*
324 All phylogenetic analyses recovered a well-supported split between *M. lehilahytsara* from
325 Central Highland (Ambohitantely and Ankafobe) and eastern forest (Riamalandy, Ambatovy,
326 and Tsinjoarivo) localities (Figure 2; Supplementary Figures S2-S8). But the relationships
327 among eastern sampling locations was contentious across analyses. The ML, SNAPP, and BPP

328 analyses supported Ambatovy as sister to Tsinjoarivo (H_{TA}; Supplementary Figures S2- S5),
329 although posterior probabilities from SNAPP revealed uncertainty in the bipartition in question.
330 SVDquartets strongly conflicted the predominant H_{TA} topology, favoring Riamalandy sister to
331 Tsinjoarivo (H_{RT}) in all bootstrap replicates when analyzing all variable sites (Supplementary
332 Figure S6) or complete RAD loci (Supplementary Figure S7). However, the H_{RT} topology, even
333 with perfect bootstrap support, is sensitive to the underlying data. Analyses of biallelic SNPs
334 alone with SVDquartets recovered a third topology (H_{RA}); also with high bootstrap support at
335 bipartitions of sampling locations (Supplementary Figure S8). In one out of ten resampled
336 datasets, when treating all individuals as independent lineages, Ambatovy was not recovered as
337 a monophyletic clade. Topological hypothesis testing differentiated among the three recovered
338 phylogenetic hypotheses: H_{TA}, H_{RT}, and H_{RA}.
339
340 AU tests rejected H_{RT}, which came from SVDquartets with all variant sites or full loci. But the AU
341 test could not reject H_{RA}, despite H_{TA} being present in the majority of ML bootstrap trees and
342 post burn-in trees (Figure 3). Bayes factors from SNAPP provided insights into the prevalence
343 of both hypotheses across analyses. SNAPP supported lumping Riamalandy and Ambatovy
344 together even with increasing step sizes and arguably more precise marginal likelihood
345 estimates as opposed to lumping Ambatovy and Tsinjoarivo (Table 1). This result is surprising,
346 considering H_{TA} was the maximum clade credibility tree from the posterior sample in the
347 unconstrained analyses (Supplementary Figures S3 and S4). All phylogenetic evidence
348 suggests a shared common ancestor for Ambatovy and Tsinjoarivo, but Tsinjoarivo may be
349 more isolated due to barriers to gene flow and therefore more evolutionarily distant than
350 expected. Phylogenetic analyses showed strong support for both CHS and eastern clades, and
351 at least no very recent gene flow in the last few generations as evident by monophyly of our five
352 sampling locations. However, such analyses do not rule out gene flow historical gene flow
353 between the CHS and eastern forests.

354

355 *Migration between eastern and Central Highland forests*

356 Population genetic analyses revealed admixture between eastern and Central Highland *M.*
357 *lehilahytsara* as well as population structure within the eastern rainforest (Tsinjoarivo versus
358 Ambatovy and Riamalandy) that mirrored findings of SNAPP Bayes factors (Table 1).
359 STRUCTURE analyses recovered an optimal k = 3 in 19 out of 20 sampled datasets
360 (Supplementary Figures S9-S11). The remaining replicate recovered an optimal k = 2
361 (Supplementary Figure S11). All analyses supported a clustering of Central Highland and
362 eastern *M. lehilahytsara*, with some degree of admixture between the two. The Ambohitantely
363 site, while phylogenetically sister to Ankafobe, shares about half of its genotypes with Ankafobe
364 and the other half with Ambatovy/Riamalandy. Tsinjoarivo often forms its own unique cluster,
365 except in the one replicate where k = 2 and Tsinjoarivo clusters with other eastern forest
366 populations.

367

368 Admixture between Central Highland and eastern *M. lehilahytsara* was also supported by
369 analyses of genetic variation with F_{ST} statistics. Similar to STRUCTURE analyses, admixture
370 between the CHS and eastern forests is largely limited to Ambohitantely and Ambatovy,
371 respectively. Pairwise F_{ST} revealed similar levels of genetic divergence between Ambatovy and
372 Ambohitantely to that of Ambatovy and Tsinjoarivo (Supplementary Table S3). The geographic
373 structure between Central Highland and eastern *M. lehilahytsara* explains a small fraction of
374 genetic variation, with most variation explained within populations (Supplementary Table S4). A
375 Mantel test revealed little evidence for isolation-by-distance (pearson's r = 0.09; p = 0.39),
376 suggesting barriers to gene flow aside distance alone (Supplementary Figure S12).

377

378 Bayes factors were computed for different isolation-with-migration models to further investigate
379 the patterns of admixture observed between the Ambatovy, Ambohitantely, and Ankafobe

genetic clusters. The *n-islands* model was unambiguously found to be the best model for both the analyses that included all individuals modeled as two populations with a subset of data (Supplementary Figure 13; Supplementary Table 5) or three populations with all of the data (Figure 4). There were no large discrepancies in the estimated migration rates; however, the selected model suggested lower effective population sizes in Ankafobe and Ambohitantely compared to Ambatovy. Assuming a mutation rate of 1.64×10^{-8}, MIGRATE estimated contemporary population sizes to be 12,652 for Ankafobe and Ambohitantely, and 17,226 for Ambatovy. Their population size estimates, although influenced by the prior, are not driven by the prior alone (Supplementary Figures S14-S16). Additionally, a higher population size in Ambatovy is supported across all models regardless of the migration and divergence scenarios (Figure 4). Because of Ambatovy's higher effective population size, the number of migrants leaving towards Ambohitantely or Ankafobe per generation is approximately 1.2, compared to 0.8 effective migrants per generation leaving Ambohitantely or Ankafobe towards the east.

Dating Population Divergences

Using the H_{TA} topology supported by Bayes factor analyses, divergence times were estimated using the MSC model. The topology was rooted with Riamalandy based on the analysis of mitochondrial data from *Microcebus lehilahytsara* and outgroup taxa (Supplementary Figure S17) and inference of clock-like trees from BPP (Supplementary Figure 5). Most MSC parameters converged across eight chains with potential scale reduction factors approaching 1 (Supplementary Table S6). A few parameters still retained high variation between chains for both θ and τ, but this discrepancy among chains was largely due to one or two chains with inefficient mixing for different prior combinations (Supplementary Figure S18-S25). Divergence times were not strongly influenced by the priors (Figure 5) and imply that divergence between the CHS and eastern forest, as well as the forest patches within them, occurred almost

simultaneously. Absolute estimates inferred that divergence among our sampled lineages occurred near 25 KYA (Table 2).

Population declines in the CHS

Population sizes experienced some slight contractions until 25 KYA (Figure 6) that coincide with the divergence of sampling locations (Table 2). Population sizes then increased through the Holocene and until the present except in Ankafobe. Ambatovy had consistently higher N_e estimates across replicates. Ankafobe had the lowest N_e estimates. Intermediate N_e estimates are observed for individuals at Ambohitantely and Tsinjoarivo. There is high uncertainty in N_e based on comparisons across replicates (Supplementary Figure S26), but individual analyses tended to converge (Supplementary Figures S27-S30).

Discussion

Population structure of M. lehilahytsara predates human colonization of Madagascar

Our results demonstrated that *M. lehilahytsara* populations have been present in the CHS at least before the LGM. Phylogenetic analyses strongly supported a split between CHS and eastern forest sampling locations (Figures 2 and 3; Supplementary Figures S2-S8), although analyses of genetic structure (Figure 2; Supplementary Figure S11; Supplementary Figure 12; Supplementary Table S3) and migration rates (Figure 4) implied gene flow after the initial divergence. If, on the other hand, the divergence of *M. lehilahytsara* in the CHS and eastern forest had been recent and human mediated, we would expect at least some phylogenetic uncertainty due to low levels of genetic divergence, similar to Tsinjoarivo, Ambatovy, and Riamalandy. Three competing topological hypotheses were uncovered concerning the relationships among the eastern rainforest sampling locations, but the best possible species tree estimation and hypothesis testing implied the northern-most sampling location, Riamalandy,

431 is sister to more southern sampling locations. All phylogenetic evidence suggested that *M.*
432 *lehilahytsara* expanded as a ring around the higher elevated regions of the CHS and that the
433 CHS acted to some extent as a barrier to gene flow during a period of vegetational change and
434 aridification at the LGM (Burney et al., 2004; Samonds et al., 2019). Although, the northern end
435 of this barrier may have remained permeable even through climate change, watersheds may
436 have buffered the region from the aridification that is more evident in the southern extent of the
437 CHS (Wilmé, Goodman, & Ganzhorn, 2006; Muldoon et al., 2012; Yoder et al., 2016; Samonds
438 et al., 2019).
439
440 Rivers are known to function as barriers to gene flow in mouse lemurs (Olivieri et al., 2008), and
441 our data suggest they are more significant in impeding mouse lemur migration than open
442 grasslands. Tsinjoarivo likely represents a population isolated from Ambatovy and Riamalandy
443 by biogeographic barriers. *M. lehilahytsara* from Tsinjoarivo was previously recognized as
444 genetically distinct based on a study of few mitochondrial and nuclear loci, although Bayes
445 factors did not indicate Tsinjoarivo to be an independent lineage under the MSC (Hotaling et al.,
446 2016). The findings of Hotaling et al. (2016) were recapitulated here with our RADseq data, with
447 little to no allele sharing between Tsinjoarivo and Ambatovy or Riamalandy (Figure 2;
448 Supplementary Figure S11). The isolation of Tsinjoarivo from other eastern sampling locations
449 is plausible, because in addition to geographic distance, our sampling locations are separated
450 by the Mangoro River (Figure. 2). BPP analyses estimate the time of separation between
451 Tsinjoarivo and Ambatovy to be near 27 KYA, which again coincides with a time of cool and dry
452 conditions during the LGM (Scholz et al,. 2011; Clark et al., 2009).
453
454 *The mosaic of grassland and wooded savanna was permeable by mouse lemurs*
455 Three lines of evidence support the hypothesis that *M. lehilahytsara* had a large ancestral
456 distribution that covered both the CHS and eastern rainforest escarpment. Within this large

457 geographic area, population substructure would have been formed due to paleoclimatic
458 oscillations but with the maintenance of some level of post-divergence gene flow. First, based
459 on our coalescent estimates of divergence times, the initial divergence between the CHS and
460 eastern forests would have formed during a period coincident with the LGM (Figure 5 and Table
461 2). Mean ambient temperatures would have been 4°C cooler and precipitation much lower at
462 this time (Burney et al., 2004), which may have caused some individuals to retreat towards
463 watershed refugia (Wilmé et al., 2006) while others remained in the CHS. The divergences
464 between *M. lehilahytsara* from Ankafobe and Ambohitantely, Ambatovy and Tsinjoarivo, and
465 their common ancestors occurred near-simultaneously, and this result is not an artifact of prior
466 choice from Bayesian MSC analyses (Supplementary Figures S18-S24). The pattern of
467 divergence and population structure formed by reduced gene flow is reflected in θ estimates
468 from BPP. MSC estimates suggested an increased population size proceeding divergence
469 events (Table 2), which likely explains why a larger prior on θ resulted in a higher posterior
470 probability and reduced uncertainty in the divergence time for the MRCA of CHS and eastern *M.*
471 *lehilahytsara*.
472
473 Second, the strong preference for an *n-islands* model from MIGRATE (Figure 4; Supplementary
474 Figure S12; Supplementary Table S5) also supported the demographic scenario of populations
475 that diverged simultaneously from a single larger distribution. Although there is genetic
476 structure, there is little discernable direction in gene flow, which may explain why models that
477 incorporated divergence did not perform as well as the *n-islands* model. The migration rates
478 inferred for our data are arguably high, between 0.8 and 1.2 migrants exchanged per
479 generation. Such migration rates can reduce fixation indices (e.g. Hartl and Clark, 1989) and
480 explain the relatively high proportion of admixture in Ambohitantely (Figure 2; Supplementary
481 Figure S11) and low F_{ST} despite the many generations that must have passed since the initial

divergence between the CHS and eastern forest populations (Supplementary Tables S3 and S4). Although, migration model parameters may be difficult to identify with many migration rate parameters (Beerli and Palczewski, 2010) and parameter identifiability is a concern for our short RAD loci (Supplementary Figures S14-S16) that individually contain very few informative coalescent events, we are nonetheless confident in these results. Given that we acknowledge the challenges of model testing and parameter estimation, the *n-islands* model agreed with the MSC results in that eastern *M. lehilahytsara* has a higher N_e than *M. lehilahytsara* in the CHS. Consistent with our model-based hypothesis testing, analyses of genetic variation with F_{ST} suggested that the mosaic of grassland and wooded savanna in the Central Highlands may have posed as much of a barrier to gene flow as the Mangoro river (Supplementary Table S3; Figure 2). Relatively little of the genetic variation is explained by grouping populations by the CHS and eastern forest compared to variation within populations (Supplementary Table S4).

The third piece of evidence for *M. lehilahytsara's* large ancestral distribution is that analyses of population size recovered similar demographic trends at all four sampling locations analyzed with EBSP methods. Both the CHS and eastern forests reflect a slight population decrease before the LGM followed by a recovery through the Holocene (Figure 6; Supplementary Figures S26-S30). Because we do not see changes in the CHS while the eastern forests remained stable, population size changes are likely due to global or Madagascar-wide effects rather than smaller-scale regional changes like agriculture. Although, there are large disparities in N_e estimates between EBSP and other methods, which raises some concern for interpreting absolute population size values from molecular data, patterns inferred from the different analyses are consistent.

Making conservation biology predictions from models of molecular evolution

507 Here we provided a range of present-day N_e estimates that are sensitive to model (i.e. isolation-
508 with-migration, MSC, Extended Bayesian Skyline) and prior choice, but all analyses suggest *M.*
509 *lehilahytsara* populations sizes are healthy (i.e. > 5000; Lynch and Lande, 1998) and have likely
510 increased through the Holocene. The most recent IUCN evaluation of *M. lehilahytsara* considers
511 the species to be vulnerable (Andriaholinirina et al., 2014c) while other mouse lemur species in
512 the eastern rainforest are considered endangered (Andriaholinirina et al., 2014a;
513 Andriaholinirina et al., 2014b; Baden et al., 2014). If *M. lehilahytsara* has adaptive traits that
514 allowed the species to expand its range into the CHS while related species were restricted to
515 lowland forests, *M. lehilahytsara* may have been more robust to Pleistocene climate change and
516 even more recent anthropogenic disturbances. *M. lehilahytsara* has been regarded as a
517 highland specialist (Radespiel et al., 2012) and has been documented to be able to enter a state
518 of prolonged torpor (Blanco et al., 2017; Andriambeloson et al., 2020). We hypothesize that the
519 capacity for extended torpor in *M. lehilahytsara* is an adaptive trait that has facilitated its
520 increased geographic range compared to most other mouse lemur species. Although individual
521 *M. lehilahytsara* populations in the CHS have maintained smaller population sizes compared to
522 those in the eastern forests (Table 2; Figure 6; Supplementary Figure S26), there is increasing
523 evidence that these populations are robust. For example, the Ankafobe forest patch is less than
524 one Km^2 yet the resident population is thriving (Blanco et al., in review).
525
526 We emphasize, however, that *M. lehilahytsara* in the CHS is not simply some sink population
527 from an eastern forest source. CHS populations have been evolutionarily independent from the
528 eastern escarpment for approximately 5000 generations if we consider the lower end of the 95%
529 HPD intervals of divergence times for nodes A and B in our BPP analyses (Table 2). Even so,
530 CHS populations have exchanged migrants (Figure 4) with the eastern forests despite the
531 historically open grassland-woodland mosaic (Solofondranohatra et al., 2018). Indeed, the
532 occasional migration to and from the east may have been important for sustaining large N_e

533 despite the small range sizes. Our best migration model estimated an N_e of 13,000 for both
534 Ankafobe and Ambohitantely and MSC estimates are 19,000 and 31,000 respectively. There
535 are of course discrepancies in the absolute N_e, but our results suggest that N_e has remained
536 sufficiently large such that the strength of selection has been strong enough to overcome
537 fixation of deleterious alleles by drift alone (Lynch & Lande, 1998). Nonetheless, the sustained
538 lower N_e in the CHS compared to the eastern forests may imply a negative consequence of
539 continued habitat loss and fragmentation if human activities are not mitigated.
540
541 *The importance of interpreting biodiversity through evolution*
542 The processes that shaped existing *M. lehilahytsara* population structure must have occurred
543 before human arrival and mid-Holocene droughts. Although charcoal records (Burney,1987a;
544 Burney 1987b), stratigraphic pollen occurrence (Gasse & Van Campo, 1998; Burns et al., 2016),
545 and megafauna subfossil extinctions (Hansford et al., 2018) all point to human-mediated large-
546 scale ecological changes in the Central Highlands (Virah-Sawmy et al., 2010), our analyses
547 implied that fundamental *M. lehilahytsara* population divergence coincides with Pleistocene
548 climate change near the LGM – not after. A similar narrative has been recovered in other
549 analyses of lemur populations (Quémére, Amelot, Pierson, Crouau-Roy, & Chikhi, 2012;
550 Salmona et al., 2017) and in an endemic olive tree (Salmona et al. 2019). In summary,
551 phylogeographic analysis of multiple species and localities agree with ecological (Bond et al.
552 2008) and evolutionary (Vorontsova et al., 2016; Hackel et al., 2018) studies of grasses that
553 demonstrate that Malagasy savannas are of ancient origin and have long persisted as a mix of
554 open grassland and woodland. Although fires due to human activity continue to ravage the
555 Central Highlands (Burney, 1987a; Gasse & Van Campo, 1998; Virah-Sawmy et al., 2010;
556 Burns et al., 2016; Samonds et al., 2019), leading to devastating forest loss in Madagascar
557 (Vieilledent et al., 2018), ongoing habitat fragmentation is occurring on a background of habitat
558 mosaicism that is likely associated with Pleistocene climate change.

559

560 **Ethics**

561 Research was conducted under permits N 0233/07/MINENV.EF/SG/DGEF/DPSAP/SSE
562 (Tsinjoarivo), N 229/14/MEEF/SG/DGF/DCB.SAP/SCB (Ambatovy), and
563 N225/15/MEEMF/SG/DGF/DAPT/SCBT (Ankafobe) issued by the Ministry of Environment,
564 Water and Forests of the Malagasy Government.

565

566 **Data Accessibility**

567 All new sequence data has been made available through NCBI BioProject PRJNA560399. All
568 individual SRA identifiers are available in Supplementary Table S7. Alignments and other data
569 for analyses are available from the Dryad Digital Repository (Tiley et al., 2020).

570

571 **Author Contributions**

572 GPT and ADY conceived the study. MBB, JMR, and RMR performed the fieldwork. ARS and
573 PAH constructed RADseq libraries. GPT performed analyses. GPT, MBB, and ADY wrote the
574 manuscript. All authors read and approved the final manuscript.

575

576 **Competing Interests**

577 The authors declare no competing interests.

578

579 **Funding**

580 The research was supported by Duke Tropical Conservation Initiative Grant to ADY, CI/Primate
581 Action Fund and Duke Lemur Center/SAVA Conservation research funds to MBB, and Duke
582 University research funds to ADY. ADY would also like to acknowledge support from the John
583 Simon Guggenheim Memorial Foundation and the Alexander von Humboldt Foundation.

584

Acknowledgements

This is Duke Lemur Center publication #XXXX.

References

Ali, O.A., O'Rourke, S.M., Amish, S.J., Meek, M.H., Luikart, G., Jeffres, C., & Miller, M.R. (2016). RAD Capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics* **202**, 389-400.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol.* **215**, 403-410.

Andriaholinirina, N., Baden, A., Blanco, M., Chikhi, L., Cooke, A., Davies, N., … Zaonarivelo, J. (2014). *Microcebus danfossi*. *The IUCN Red List of Threatened Species* 2014a: e.T136347A16112905. http://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T136347A16112905.en. Downloaded on 25 November 2018.

Andriaholinirina, N., Baden, A., Blanco, M., Chikhi, L., Cooke, A., Davies, N., … Zaonarivelo, J. (2014). *Microcebus bongolavensis*. *The IUCN Red List of Threatened Species* 2014b: e.T136537A16112834. http://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T136537A16112834.en. Downloaded on 25 November 2018.

Andriaholinirina, N., Baden, A., Blanco, M., Chikhi, L., Cooke, A., Davies, N., … Zaonarivelo, J. (2014). *Microcebus lehilahytsara*. *The IUCN Red List of Threatened Species* 2014c: e.T136199A16113109. http://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T136199A16113109.en. Downloaded on 28 November 2018.

Andriambeloson, J.B., Greene, L.K., & Blanco, M.B.. 2020. Prolonged torpor in Goodman's mouse Lemur (Microcebus lehilahytsara) from the high-altitude forest of Tsinjoarivo, central-eastern Madagascar. *Folia Primatol* (Basel) doi: 10.1159/000509102.

Baden, A,, Blanco, M., Chikhi, L., Cooke, A., Davies, N., Dolch, R., … Zaonarivelo, J. *Microcebus ravelobensis*. *The IUCN Red List of Threatened Species* 2014: e.T39751A16113536. http://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T39751A16113536.en. Downloaded on 25 November 2018.

Beerli, P., & Palczewski, M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* **185**, 313-326.

Blanco, M.B., Andriambeloson, J.B., Andriantsalohimisantatra, A.A., Rivoharison, T.V., Walker, N., Birkinshaw, C.R., & Yoder, A.D. (2020). Habitat constraints under natural conditions: a glimpse at the ecology of Goodman's mouse lemurs (*Microcebus lehilahytsara*) in the forest fragments of Ankafobe, central highlands, Madagascar. in review.

Blanco, M.B., Andriantsalohimisantatra, A.A., Rivoharison, T.V., & Andriambeloson, J-B. (2017). Evidence of prolonged torpor in Goodman's mouse lemurs at Ankafobe forest, central Madagascar. *Primates* **58**, 31-37.

Blanco, M.B., Dausmann, K.H., Faherty, S.L., & Yoder, A.D. (2018). Tropical heterothermy is "cool": The expression of daily torpor and hibernation in primates. *Evolutionary Anthropology* 2**7**, 147-161.

Bond, W.J., Silander Jr, J.A., Ranaivonasy, J., & Ratirarson, J. (2008). The antiquity of Madagascar's grasslands and the rise of C_4 grassy biomes. *J Biogeogr.* **35**, 1743-1758.

Bouckaert, R.R., Heled, J., Kuehnert, D., Vaughan, T.G., Wu, C.-H., Xie, D., … Drummond, S.J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput Biol.* **10**, e1003537. doi: 10.1371/journal.pcbi.1003537.

Bowman, D.M.J.S., & Franklin, D.C. (2005). Fire ecology. *Progress in Physical Geography* **29**, 248-255.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol Biol Evol.* **29**, 1917-1932.

Burney, D.A. (1987a). Late Holocene vegetational change in central Madagascar. *Quatern Res.* **28**, 130-143.

Burney, D.A. (1987b). Late Quaternary stratigraphic charcoal records from Madagascar. *Quatern Res.* **28**, 274-280.

Burney, D.A., Burney, L.P., Godfrey, L.R., Jungers, W.L., Goodman, S.M., Wright, H.T., & Jull, A.J.T. (2004). A chronology for late prehistoric Madagascar. *Journal of Human Evolution* **47**, 25-63.

Burns, S.J., Godfrey, L.R., Faina, P., McGee, D., Hardt, B., Ranivoharimanana, L., & Randrianasy, J. (2016). Rapid human-induced landscape transformation in Madagascar at the end of the first millennium of the common era. *Quaternary Science Reviews* **134**, 92-99.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T.L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421.

Campbell, C.R., Tiley, G.P., Poelstra, J.W., Hunnicutt, K.E., Larsen, P.A., Lee, H.-J., … Yoder, A.D. (2020). Substitution rate analysis validates surprising results from a pedigree-based measurement of the de novo mutation rate in gray mouse lemurs. *Heredity* in revision.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3* **1**, 171-182.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol.* **22**, 3124-3140.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**, 3317-3324.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., … McCabe, A.M. (2009). The last glacial maximum. *Science* 325, 710-714.

Dausmann, K.H., Glos, J., Ganzhorn, J.U., Heldmaier, G. (2005). Hibernation in the tropics: lessons from a primate. *Journal of Comparative Physiology B.* **175**, 147-155.

Dewar, R.E., & Wright, H.T. (1993). The culture history of Madagascar. *Journal of World Prehistory* **7**, 417-466.

Dewar, R.E., Radimilahy, C., Wright, H.T., Jacobs, Z., Kelly, G.O., & Berna, F. (2013). Stone tools and foraging in northern Madagascar challenge Holocene extinction models. *Proc Natl Acad Sci USA*. **110**,12583-12588.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Res.* **32**, 1792-1797.

Edwards, E.J., & Smith, S.A. (2010). Phylogenetic analyses reveal the shady history of C_4 grasses. *Proc Natl Acad Sci USA.* **107**, 2532-2537.

Excoffier, L., & Lischer, H.E.L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* **10**, 564-567.

Excoffier, L., Smouse, P.E., & Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* **14**, 2611-2620.

Farris, Z.J., Gerber, B.D., Valenta, K., Rafaliarison, R., Razafimahaimodison, J.C., Larney, E., … Chapman, C.A. (2017). Threats to a rainforest carnivore community: A multi-year assessment of occupancy and co-occurrence in Madagascar. *Biological Conservation* **210**, 116-124.

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Mol Biol Evol.* **35**, 2585-2593.

Gade, D.W. (1996). Deforestation and its effects in highland Madagascar. *Mountain Research and Development* **16**, 101-116.

Gasse, F., & Van Campo, E. (1998). A 40,000-yr pollen and diatom record from Lake Tritrivakely, Madagascar, in the southern tropics. *Quatern Res.* **49**, 299-311.

Goodman, S.M., Raherilalao, M.J., & Wohlauser, S. (2018). *The terrestrial protected areas of Madagascar: Their history, description, and biota*. Association Vahatra. Antananarivo, Madagascar.

Green, G.M., & Sussman, R.W. (1990). Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science* **248**, 212-215.

Hansford, J., Wright, P.C., Rasoamiaramanana, A., Pérez, V.R., Godfrey, L.R., Errickson, D., … Turvey, S.T. (2018). Early Holocene human presence in Madagascar evidenced by exploitation of avian megafauna. *Science Advances* **4**, eaat6925.

Harper, G.J., Steininger, M.K., Tucker, C.J., Juhn, D., & Hawkins, F. (2007). Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation* **34**, 325-333.

Hartl, D.L., & Clark, A.G. (1989). *Principles of population genetics* (2nd ed.). (pp. 310-314). Sunderland, MA: Sinauer Associates.

Heled, J., & Drummond, A.J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evol Biol*. **8**, 289.

Hotaling, S., Foley, M.E., Lawrence, N.M., Bocanegra, J., Blanco, M.B., Rasoloarison, R., … Weisrock, D.W. (2016). Species discovery and validation in a cryptic radiation of endangered primates: coalescent-based species delimitation in Madagascar's mouse lemurs. *Mol Ecol.* **25**, 2029-2045.

Humbert, H. (1927). La destruction d'une flore insulaire par le feu. Principaux aspects de la vegetation à Madagascar. *Mémoires de L'Académie Malgache* **5**, 1-78.

Kass, R.E., & Raferty, A.E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773-795.

Koechlin, J. (1993). *Natural grasslands: eastern hemisphere and resume*. (pp. 291-301.). In: Coupland RT editor. Grasslands of Madagascar. Amsterdam: Elsevier.

Kull, C.A. (2003). *Fire and the management of highland vegetation*. (pp. 153-157.). In: Goodman SM, Benstead JP, editors. The natural history of Madagascar. Chicago: The University of Chicago Press.

Lynch, M., & Lande, R. (1998). The critical effective size for a genetically secure population. *Animal Conservation* **1**,70-72.

Mitchell, P. (2019). Settling Madagascar: when did people first colonize the world's largest island? *Journal of Island and Coastal Archaeology* doi: 10.1080/15564894.2019.1582567.

Muldoon, K.M., Crowley, B.E., Godfrey, L.R., Rasoamiaramanana, A., Aronson, A., Jernvall, J., … Simons. E.L. (2012). Early Holocene fauna from a new subfossil site: A first assessment from Christmas River, south central Madagascar. *Madagascar Conservation and Development* **7**, 23-29.

Olivieri, G.L., Sousa, V., Chikhi, L., & Radespiel, U. (2008). From genetic diversity and structure to conservation: Genetic signature of recent population declines in three mouse lemur species (*Microcebus* spp.). *Biological Conservation* **141**, 1257-1271.

Perrier de la Bâthie, H. (1921). La vegetation Malgache. *Annales du Musée Colonial de Marseille* **9**, 1-226.

Poelstra, J.W., Salmona, J., Tiley, G.P., Schüßler, D., Blanco, M.B., Andriambeloson, J.B., … Yoder, A.D. (2020). Cryptic patterns of speciation in cryptic primates: microendemic mouse lemurs and the multispecies coalescent. *Systematic Biology* doi: https://doi.org/10.1093/sysbio/syaa053.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*. **6**, 7-11.

Pritchard, J.K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Quémére, E., Amelot, X., Pierson, J., Crouau-Roy, B., & Chikhi, L. (2012). Genetic data suggest a natural prehuman origin of open habitats in northern Madagascar and question the deforestation narrative in this region. *Proc Natl Acad Sci USA.* **109**, 13028-13033.

Radespiel, U., Lutermann, H., Schmelting, B., & Zimmermann, E. (2019): An empirical estimate of 1253 the generation time of mouse lemurs. *Am. J. Primatol.* https://doi.org/10.1002/ajp.23062.

Radespiel, U., Ratsimbazafy, J.H., Rasoloharijaona, S., Raveloson, H., Andriaholinirina, N., Rakotondravony, R., … Randrianambinina, B. (2012). First indications of a highland specialist among mouse lemurs (*Microcebus* spp.) and evidence for a new mouse lemur species from eastern Madagascar. *Primates* **53**, 157-170.

Railsback, L.B., Dupont, L.A., Liang, F., Brook, G.A., Burney, D.A., Cheng, H., & Edwards, R.L. (2020). Relationships between climate change, human environmental impact, and megafaunal extinction inferred from a 4000-year multi-proxy record from a stalagmite from northwestern Madagascar. *Quaternary Science Reviews* **234**, 106244.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., & Suchard, M.A. (2018). Posterior summarisation in Bayesian pylogenetics using Tracer 1.7. *Sys Biol.* doi: https://doi.org/10.1093/sysbio/syy032.

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* **164**, 1645-1656.

Robinson, D.R., & Foulds, L.R. (1981). Comparison of phylogenetic trees. *Math Biosci.* **53**, 131-147.

Salmona, J., Heller, R., Quémére, E., & Chikhi, L. (2017). Climate change and human colonization triggered habitat loss and fragmentation in Madagascar. *Mol Ecol.* **26**, 5203-5222.

Salmona, J., Olofsson, J.K., Hong-Wa, C., Razanatsoa, J., Rakotonasolo, F., Ralimanana, H., … Besnard, G. (2020). Late Miocene origin and recent population collapse of the Malagasy savanna olive tree. *Biol J Linn Soc.* **129**, 227-243.

Samonds, K.E., Crowley, B.E., Rasolofomanana, T.R.N., Andriambelomanana, M.C., Andrianavalona, H.T., Ramihangihajason, T.N., & Godrey, L.R. (2019). A new late Pleistocene subfossil site (Tsaramody, Sambaina basin, central Madagascar) with implications for the chronology of habitat and megafaunal community change on Madagascar's Central Highlands. *Journal of Quaternary Science* **34**, 379-392.

Schmid, J. (2001). Daily torpor in free-ranging gray mouse lemurs (*Microcebus murinus*) in Madagascar. *International Journal of Primatology* **22**, 1021-1031.

Scholz, C.A., Cohen, A.S., Johnson, T.C., King, J., Talbot, M.R., & Brown, E.T. (2011). Scientific drilling in the great rift valley: The 2005 Lake Malawi scientific drilling project – an overview of the past 145,000 years of climate variability in southern hemisphere East Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology* **303**, 3-19.

Sgarlata, G.M., Salmona, J., Alexio-Pais, I., Rakotonanahary, A., Sousa, A.P., Kun-Rodrigues, C., … Chikhi, L. (2018). Genetic differentiation and demographic history of the northern rufous

mouse lemur (*Microcebus tavaratra*) across a fragmented landscape in northern Madagascar. *International Journal of Primatology* **39**, 65-89.

Shimodaira, H. (2004). Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statistics* **32**, 2616-2641.

Shimodaira, H., & Hasegawa, M. (2001). CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246-1247.

Solofondranohatra, C.L., Vorontsova, M.S., Hackel, J., Besnard, G., Cable, S., Williams, J., … Lehmann, C.E.R. (2018). Grass functional traits differentiate forest and savanna in the Madagascar Central Highlands. *Frontiers in Ecology and Evolution* **6**, 184.

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

Styger, E., Rakotondramasy, H.M., Pfeffer, M.J., Fernandes, E.C.M., & Bates, D.M. (2007). Influence of slash-and-burn farming practices on fallow succession and land degradation in the rainforest region of Madagascar. *Agriculture, Ecosystems, and Environment* **119**, 257-269.

Swofford, D.L. (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tavaré, S. (1986). *Some probabilistic and statistical problems in the ananlysis of DNA sequences. In: Some mathematical questions in biology – DNA sequence analysis*. (pp. 57-86.). In Miura RM editor. Providence, RI: Amer Math Soc.

[dataset] Tiley G.P., Blanco, M.B., Ralision, J.M., Rasoloarison, R.M., Stalhke A.R., Hohenlohe, P.A., & Yoder, A.D. 2020. Dryad Data Repository. Version 2. doi: https://doi.org/10.5061/dryad.dncjsxkvz.

Trucchi, E., Gratton, P., Whittington, J.D., Cristofari, R., Le Maho, Y., Stenseth, N.C., & Le Bohec, C. (2014). King penguin demography since the last glaciation inferred from genome-wide data. *Proc R Soc B*. **281**, 20140528.

Vieilledent, G., Gardi, O., Grinand, C., Burren, C., Andriamanjato, M., Camara, C., … Rakotoarijaona, J-R. (2016). Bioclimatic envelope models predict a decrease in tropical forest carbon stocks with climate change in Madagascar. *Journal of Ecology* **104**, 703-715.

Vieilledent, G., Grinand, C., Rakotomalala, F.A., Ranaivosoa, R., Rakotoarijaona, J-R., Allnutt, T.F., & Achard, F. (2018). Combining global tree cover loss data with historical national forest cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. *Biological Conservation* **222**, 189-197.

Virah-Sawmy, M., Willis, K.J., & Gillson, L. (2009). Threshold response of Madagascar's littoral forest to sea-level rise. *Global Ecology and Biogeography* **18**, 98-110.

Virah-Sawmy, M., Willis, K.J., & Gillson, L. (2010). Evidence for drought and forest declines during the recent megafaunal extinctions in Madagascar. *J Biogeogr.* **37**, 506-519.

Vorontsova, M.S., Besnard, G., Forest, F., Malakasi, P., Morat, K., Clayton, W.D., … Linder, H.P. (2016). Madagascar's grasses and grasslands: anthropogenic or natural? *Proc Biol Sco.* **283**, 20152226.

Weir, B.S., & Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358-1370.

Wilmé, L., Goodman, S.M., & Ganzhorn, J.U. (2006). Biogeographic evolution of Madagascar's microendemic biota. *Science* **312**, 1063-1065.

Xie, W., Lewis, P.O., Fan, Y., Kuo, L., & Chen, M-H. (2011). Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst Biol*. **60**, 150-160.

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol.* **39**, 306-314.

Yoder, A.D., Campbell, C.R., Blanco, M.B., dos Reis, M., Ganzhorn, J.U., Goodman, S.M., … Weisrock, D.W. (2016). Geogenetic patterns in mouse lemurs (genus *Microcebus*) reveal the ghosts of Madagascar;s forests past. *Proc Natl Acad Sci USA.* **113**, 8049-8056.

Zohdy, S., Gerber, B.D., Tecot, S., Blanco, M.B., Winchester, J.M., Wright, P.C., & Jernvall, J. (2014). Teeth, sex, and testosterone: aging in the world's smallest primate. *PLoS One* 9, e109528.

915 **Tables**

Table 1- Marginal likelihood estimates and model selection with SNAPP

Steps	Topological Hypothesis	Marginal Likelihood	Bayes Factor†	Model Probability
20	RA	-41399.18	0.00	1
	TA	-41847.70	-897.05	1.61×10^{-195}
	RT	-41689.27	-580.19	1.03×10^{-126}
40	RA	-41531.26	0.00	1
	TA	-41848.94	-635.35	1.08×10^{-138}
	RT	-42060.55	-1058.58	1.36×10^{-230}
60	RA	-41328.81	0.00	1
	TA	-41618.87	-580.12	1.07×10^{-126}
	RT	-42057.39	-1457.16	0

†Bayes factors computed by marginal log-likelihoods with respect to the best model as $2 * (\ln L_i - \ln L_{Max})$

916

Table 2 - Absolute divergence times and population sizes estimated† from BPP.

Node	N_e (x 10^3)	95% HPD	Divergence time (KYA)	95% HPD
Ankafobe	19.2	17.3, 25.7	-	-
Ambohitantely	31.2	27.1, 40.9	-	-
Tsinjoarivo	23.1	14.0, 26.1	-	-
Ambatovy	55	33.5, 63.0	-	-
Riamalandy	204.5	193.4, 216.0	-	-
A	60.2	10.6, 201.3	26.7	15.9,30.4
B	22	15.1, 34.2	19	16.9, 25.2
C	875.1	776.9, 955.0	31	25.4, 42.7
D	108.3	106.6, 110.1	194.1	189.0, 198.8

†Estimates are based on combined posterior for priors $\theta = 0$ and $\tau = 1$.

917

918 **Figure Legends**

919 **Figure 1 – Recovery of RAD loci among *Microcebus lehilahytsara*.** The diagonal represents
920 the proportion of total loci recovered for each sample while the upper triangular represents the
921 proportion of total loci shared between two individuals. The lower triangular represents the
922 evolutionary distance between two individuals, which is the patristic distance from the
923 concatenated ML analysis.

924

925 **Figure 2 – Well-supported phylogenetic lineages of *Microcebus lehilahytsara* reflect**
926 **genetic structure.** The phylogeny is the maximum likelihood tree with bootstrap support at
927 nodes. Branch lengths are expected substitutions per site. Metadata for individual identifiers is
928 available in Table S1. STRUCTURE results for the one SNP per locus dataset is shown with an
929 optimal k = 3. The GPS coordinates of each sampling site are shown on the map, with
930 Riamalandy, Ambatovy, and Tsinjoarivo found in the eastern rainforest while Ambohitantely and
931 Ankafobe are fragmented forest patches in the Central Highland Savanna. Rivers are shown in
932 blue while forest cover is shown in green. The specific region of Madagascar studied here
933 shown in the boxed region on the country map. Major ecoregions of Madagascar are shown on
934 the country map. Green shows the moist forest while brown is dry forest and red is spiny forest.
935 Shape files of Madagascar's ecoregions were taken from Vieilledent et al. (2016). All *M.*
936 *lehilahytsara* sampling locations are considered wet forest.

937

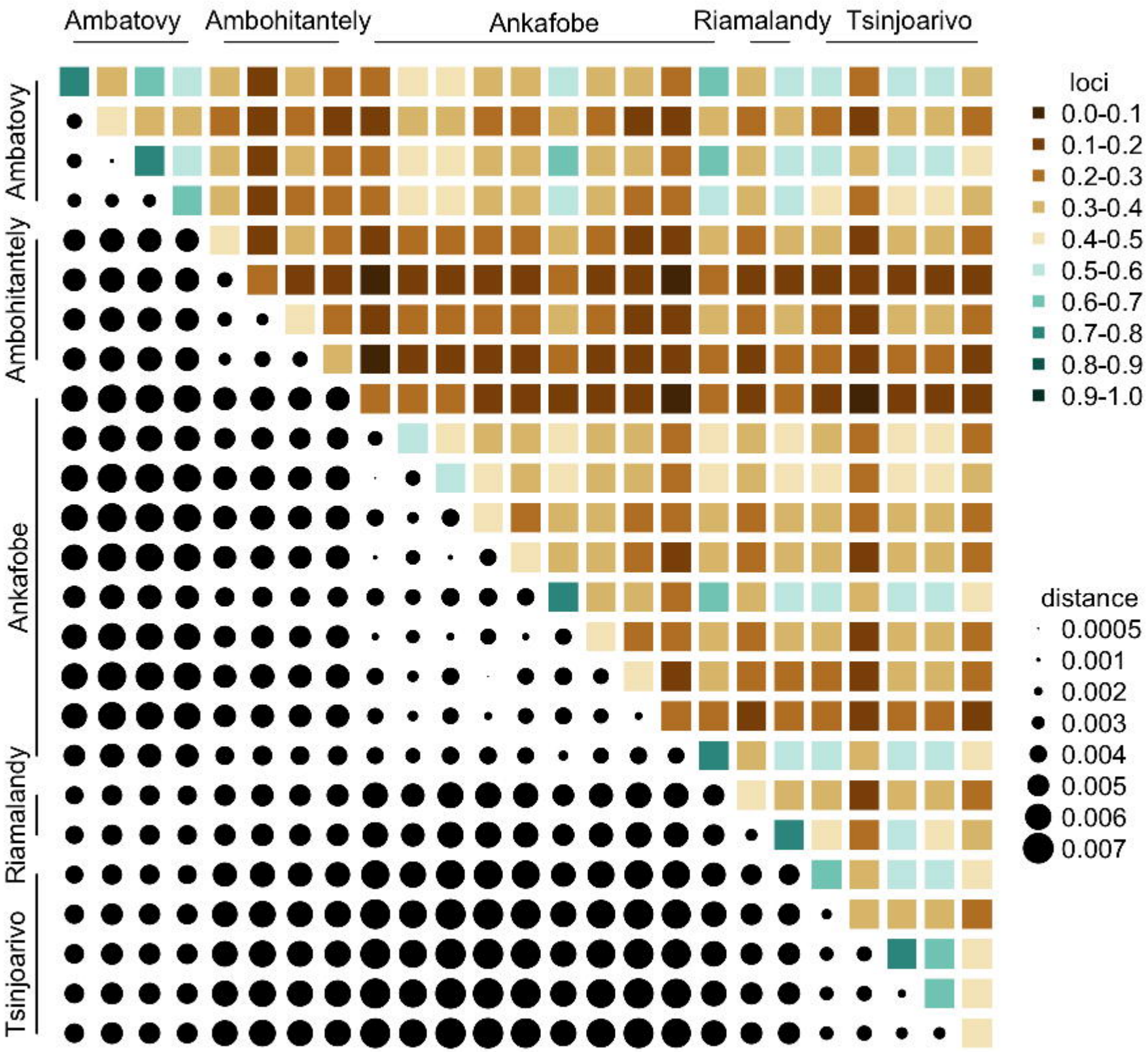
938 **Figure 3 – Phylogenetic hypothesis testing of eastern Madagascar.** Three competing
939 topological hypotheses were considered using both the AU test with site likelihoods and Bayes
940 factors from SNAPP. Posterior probabilities of topologies were obtained from marginal
941 likelihoods for SNAPP analyses and AU tests perform a multiscale bootstrap of site likelihoods
942 to obtain a p-value that should be interpreted as strength against a topology. No analogous
943 topology test exists for invariants, as in SVDquartets, but we did examine the RF distances

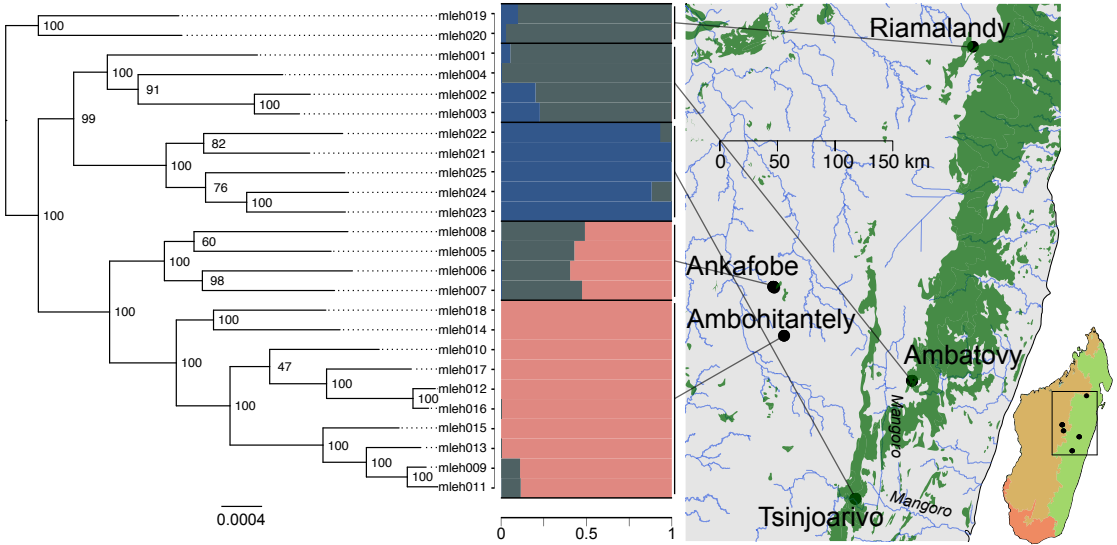
among bootstrap trees from ML and SVDquartets analyses to summarize phylogenetic discordance among RAD loci. Pairwise are all bootstrap trees for ML or SVDquartets analyses compared with themselves. ML are all bootstrap trees, from either ML or SVDquartets analyses, compared to the ML tree. SVDq are all bootstrap trees, from either ML or SVDquartets analyses, compared to the SVDquartets bootstrap consensus tree. Although RF distances among SVDquartets bootstraps are typically higher, they are also closer to being normally distributed, implying either less uncertainty or potential biases in ML analyses.

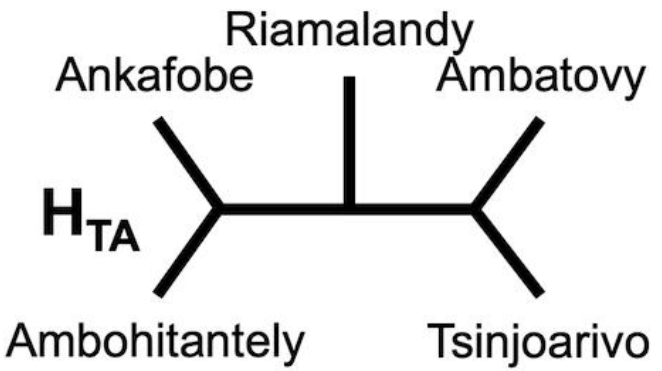
Figure 4 – Migration models reveal gene flow between the Central Highlands and eastern forests. Bayes factors were used to test migration hypotheses between the eastern rainforest, Ambatovy (grey), and two fragmented forest patches in the Central Highlands, Ambohitantely (pink and grey) and Ankafobe (pink). These three locations were selected because of admixture detected in Ambatovy based on STRUCTURE analyses, and geographic locations are colored as such. θ parameters are scaled as $4N_e\mu$ and mutation-scaled migration parameters are shown near arrows. Divergence times are shown for relevant models in substitutions per site.

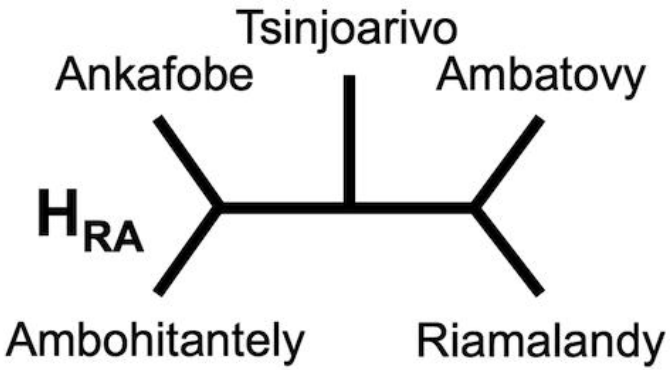
Figure 5 – Divergence time estimates with BPP. Posterior distributions of MSC parameters for the combined posterior of prior choice 1. Branch lengths are in coalescent units and error bars represent 95% highest posterior density intervals around the median node height. Violin plots are the posterior distributions of node heights for prior choices 1 though 4 to show that while divergence time estimates are sensitive to prior choice, our biological conclusions are not.

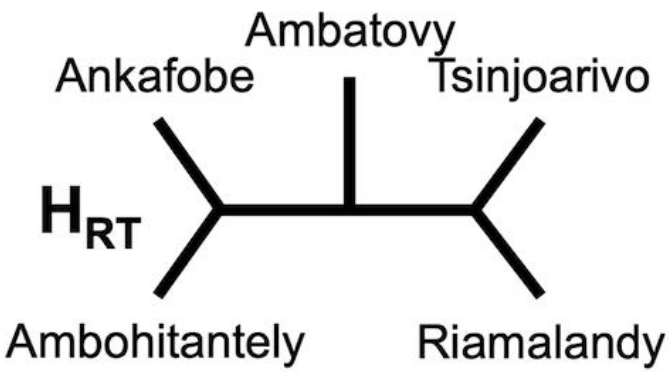
Figure 6 – Population size change over time at sampling locations. Posterior distributions of θ at discretized bins of τ determined by the number of change points. Only the first sampled dataset is shown for each population. The blue box shows the duration of the LGM.


Ambatovy
Ambohitantely
Ankafobe
Riamalandy
Tsinjoarivo
loci
0.0-0.1
0.1-0.2
0.2-0.3
0.3-0.4
0.4-0.5
0.5-0.6
0.6-0.7
0.7-0.8
0.8-0.9
0.9-1.0
distance
0.0005
0.001
0.002
0.003
0.004
0.005
0.006
0.007

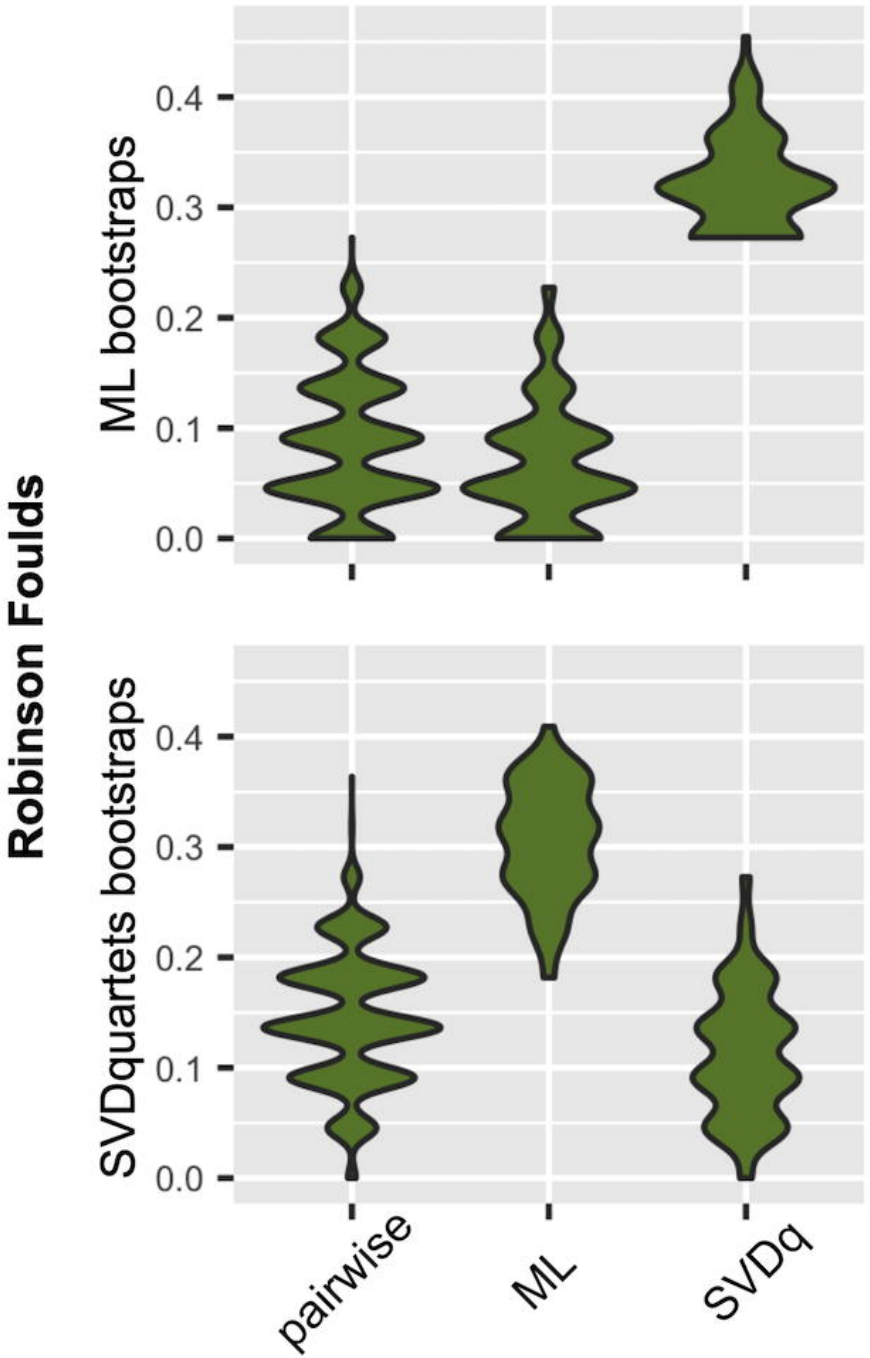

mleh019
mleh020
mleh001
mleh004
mleh002
mleh003
mleh022
mleh021
mleh025
mleh024
mleh023
mleh008
mleh005
mleh006
mleh007
mleh018
mleh014
mleh010
mleh017
mleh012
mleh016
mleh015
mleh013
mleh009
mleh011
100
91
99
82
76
60
98
47
0.0004
0
0.5
1
Riamalandy
0
50
100
150 km
Ankafobe
Ambohitantely
Ambatovy
Mangoro
Tsinjoarivo

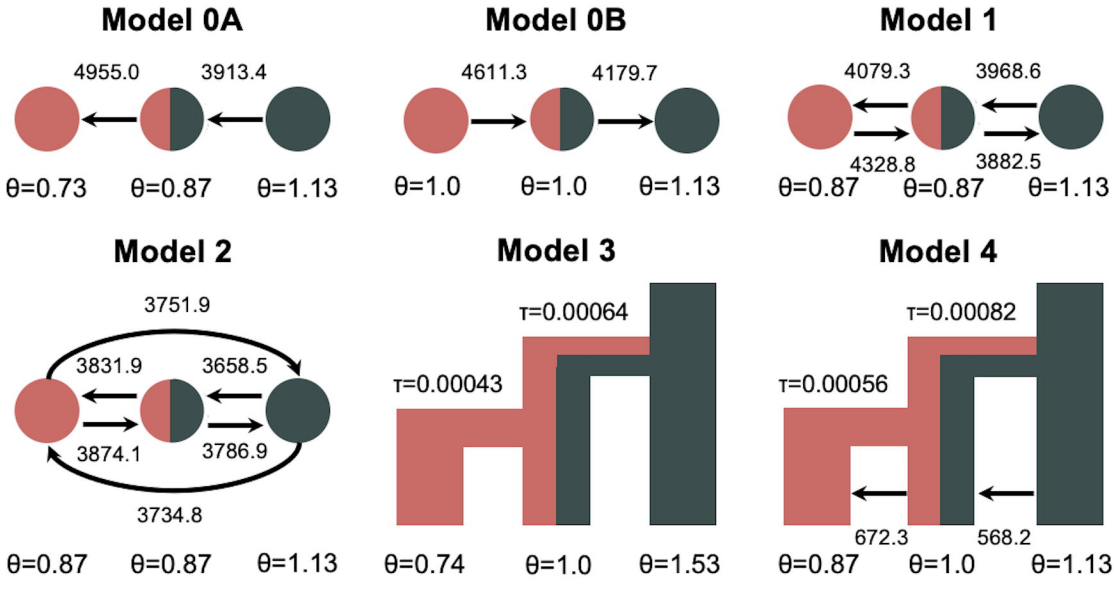






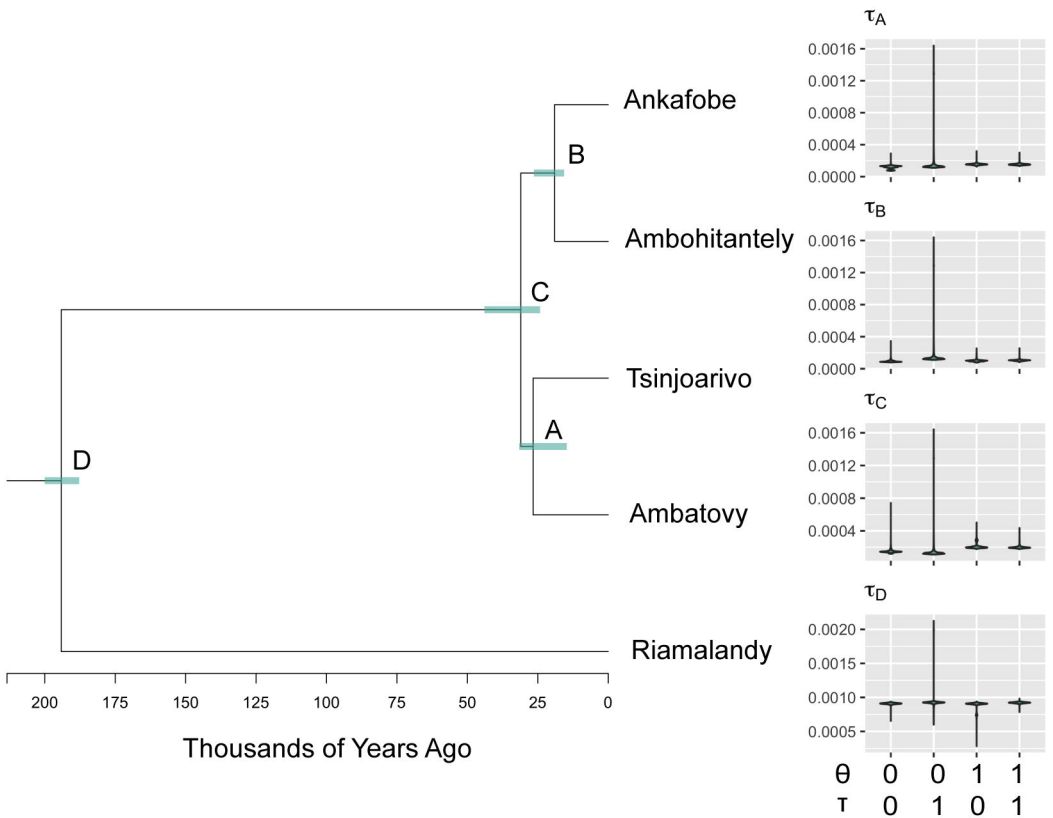
AU p-value
0.895
0.105
8e-99

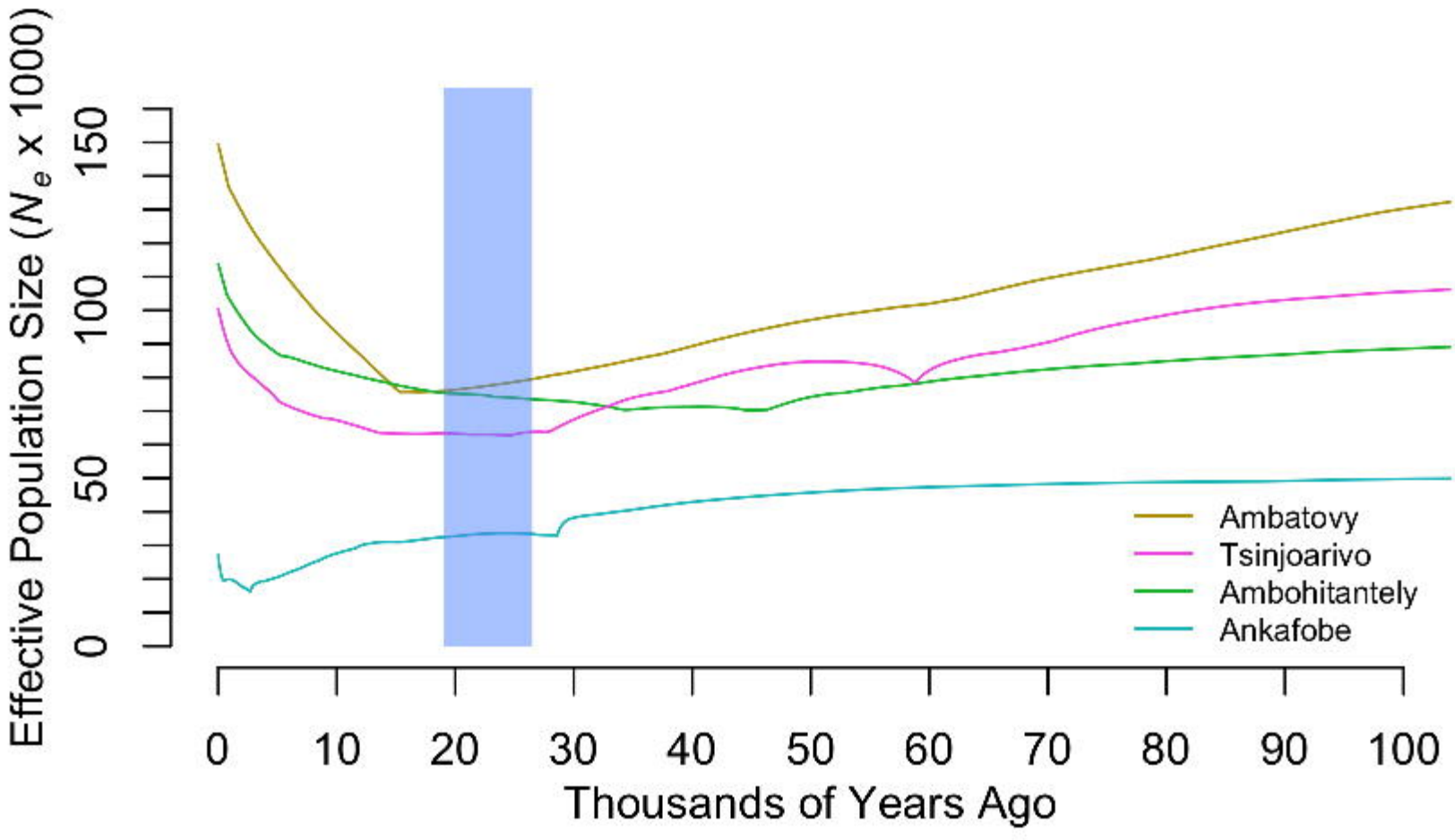




Model	Marginal Likelihood	Bayes Factor*	Model Probability
0A – East to west	-408766.23	NA	0
0B – West to east	-408169.12	1194.22	0
1 – Stepping-stone	-404252.47	9027.52	0
2 – n-islands	-402558.14	12416.18	1
3 – Divergence and isolation	-415036.4	-12540.34	0
4 – Divergence and migration	-407777.86	1976.74	0

*Bayes factors computed with natural logarithms of marginal likelihoods with respect to model 0A as $2 * (\ln L_i - \ln L_{0A})$
†θ values are presented as 10^{-3}


Ankafobe
Ambohitantely
Tsinjoarivo
Ambatovy
Riamalandy
A
B
C
D
200
175
150
125
100
75
50
25
0
Thousands of Years Ago
τA
τB
τC
τD
0.0016
0.0012
0.0008
0.0004
0.0000
0.0020
0.0015
0.0010
0.0005
θ 0 0 1 1
τ 0 1 0 1


Effective Population Size (N_e x 1000)
0
50
100
150
0
10
20
30
40
50
60
70
80
90
100
Thousands of Years Ago
Ambatovy
Tsinjoarivo
Ambohitantely
Ankafobe