nature ecology & evolution

Article

https://doi.org/10.1038/s41559-024-02596-1

Ecological and anthropogenic effects on the genomic diversity of lemurs in Madagascar

Received: 17 October 2023

Accepted: 1 November 2024

Published online: 27 December 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

Ecological variation and anthropogenic landscape modification have had key roles in the diversification and extinction of mammals in Madagascar. Lemurs represent a radiation with more than 100 species, constituting roughly one-fifth of the primate order. Almost all species of lemurs are threatened with extinction, but little is known about their genetic diversity and demographic history. Here, we analyse high-coverage genome-wide resequencing data from 162 unique individuals comprising 50 species of Lemuriformes, including multiple individuals from most species. Genomic diversity varies widely across the infraorder and yet is broadly consistent among individuals within species. We show widespread introgression in multiple genera and generally high levels of genomic diversity likely resulting from allele sharing that occurred during periods of connectivity and fragmentation during climatic shifts. We find distinct patterns of demographic history in lemurs across the ecogeographic regions of Madagascar within the last million years. Within the past 2,000 years, lemurs underwent major declines in effective population size that corresponded to the timing of human population expansion in Madagascar. In multiple regions of the island, we identified chronological trajectories of inbreeding that are consistent across genera and species, suggesting localized effects of human activity. Our results show how the extraordinary diversity of these long-neglected, endangered primates has been influenced by ecological and anthropogenic factors.

Despite its relatively small geographic area, Madagascar harbours an extraordinarily high level of biodiversity that evolved in the context of profound environmental heterogeneity[1,2]. The central highlands of Madagascar divide the island into western dry forest and eastern humid forest biomes, which are subdivided by multiple rivers into areas of endemism that often reflect the geographic ranges of distinct species[1,3–5]. The interaction of this environmental heterogeneity with periodic climatic fluctuations and the expansion and contraction of forested habitats along watersheds and altitudinal gradients is believed to have had an essential role in the formation of Madagascar's rich species diversity[1,3]. The effect of this ecogeographic patterning on speciation has been explored in a wide range of organisms[1,2,4–9], including many of the 103 recognized species of lemurs (infraorder Lemuriformes). The adaptive radiation of Lemuriformes in Madagascar followed their split with other strepsirrhine primates (~50 Ma)[10] and one or two overwater crossings from mainland Africa[11,12]. As such, the Lemuriformes provide an ideal system to examine the impacts of ecologically heterogeneous island biogeography on the population genomics of an expansive adaptive radiation of mammals.

In spite of this extraordinary species diversity, there has been no large-scale effort to resequence genomes of multiple individuals from a large number of Lemuriformes species to high coverage. Given the paucity of high-coverage genome-wide resequencing data, the effect of Madagascar's ecological heterogeneity on primate genomic diversity, interspecific allele sharing and demographic history remains largely unexplored at a system-wide level. The small number of population

e-mail: joseph.orkin@umontreal.ca; tomas.marques@upf.edu

genomic studies conducted on the Lemuriformes have led to important insights into individual clades[13–17] or broad patterns of diversity in relation to other primates when sequencing single individuals per species[10]. These studies have revealed that some species of lemurs appear to have notably high levels of genomic diversity[10,13]. Likewise, analyses of demographic history and historical effective population sizes have indicated a general pattern of declining populations that fluctuated during the Pleistocene[10,13,15–17]. However, substantial uncertainty surrounds the underlying causes of the extreme variation in genomic diversity across the Lemuriformes. It remains unclear whether the ecological variation in Madagascar has exerted a consistent influence on the genomic diversity and demographic history of the Lemuriformes.

Likewise, uncertainty remains about the timing and effect of late Holocene human activity on declining lemur populations[18–20]. Debate remains about when humans first arrived in Madagascar, but evidence points to a wide-scale human population expansion and landscape modification in the last 1,000 years[2,18–20]. The timing and extent of human impacts on the extinctions of several large-bodied species of subfossil lemurs and declining populations of extant lemurs remain uncertain and could be elucidated with population genomic data. Given the overwhelming threat of extinction for most Lemuriformes, there is an urgent need to leverage newly abundant sequence data for a robust understanding of the history of environmental and anthropogenic effects that have influenced the genomic diversity and survival of these species.

Here, we present the results of a large-scale high-coverage resequencing effort examining the genomic diversity and demographic history of the Lemuriformes in Madagascar. We demonstrate a pattern of widespread interspecific allele sharing and high levels of heterozygosity in many species that probably resulted from periodic population fragmentation and connectivity during climatic shifts. Additionally, we demonstrate regionally consistent patterns of demographic history among non-human primates throughout Madagascar that are consistent with both ecological variation and recent anthropogenic effects. Furthermore, these data substantially increase the number of publicly available high-coverage resequenced genomes for an understudied primate radiation under severe threat of extinction.

Results

Resequencing of the Lemuriformes

Our dataset comprises high-coverage resequenced whole genomes from half of all recognized Lemuriformes species, which encompasses nearly one-fifth of all primate species. We sequenced 166 samples, of which 162 were unique individuals from 50 out of 103 species of Lemuriformes. Of these unique resequenced genomes, data from 85 individuals are unpublished and newly sequenced and 77 were downloaded from publicly available sources (Supplementary Table 1), more than doubling the number of high-coverage resequenced genomes of Lemuriformes. Of these 50 species, five have not been sequenced to high coverage until now. Our sample represents all five families, 13 out of 15 genera and every species in the Lemuridae. We obtained more than one individual from 32 species and at least five individuals from 16 species (Supplementary Table 2). Among these samples, 121 (74%) are confirmed to be wild-born, 38 (23%) are of unknown origin and only 3 (2%) are known to be captive-born. Newly generated libraries were sequenced to a target mean depth of ~30× with PCR-free libraries using Illumina NovaSeq 6000 technology.

Genome-wide heterozygosity in the Lemuriformes

We observed high median values with wide ranges of variation in genome-wide heterozygosity per individual (het × bp^{-1}) (median, 0.0029, min., 0.0005, max., 0.0078) (Figs. 1a and 2). Among genera, the highest median heterozygosities were observed in *Indri* (0.0060), *Microcebus* (0.0049) and *Hapalemur* (0.0035). We confirm the placement of the aye-aye, *Daubentonia madagascariensis*, as an outlier with low median heterozygosity relative to the other members of the clade (het × bp^{-1} = 0.0006)[16,21]. Individual heterozygosity values tend to cluster in discrete ranges relative to their congeneric species (Supplementary Table 1 and Extended Data Figs. 1 and 2). We observed a clear break in heterozygosity among *Eulemur* species, segregating the northern and central members of the *fulvus* group from the non-*fulvus* group (Extended Data Figs. 2 and 3). Individuals within the *fulvus* group have among the highest levels of heterozygosity among the Lemuriformes, in contrast to their counterparts, who are among the least genomically diverse lemurs. To estimate relative levels of recent inbreeding and population bottlenecks, we identified runs of homozygosity (ROHs) in each individual using BCFtools[22]. We observed a wide range of values for the autozygous proportion of the genome (F_{ROH}) (Fig. 1b). In general, genera displayed an inverse relationship between heterozygosity per base pair and F_{ROH} (for example, *Lepilemur*, *Hapalemur*, *Microcebus*, *Daubentonia*). Compared to other members of the Lemuriformes, *Daubentonia* has a substantially elevated and highly variable callable fraction of its genome in ROHs (≥1 Mb) relative to the number of ROHs (Fig. 1c), suggesting a bottlenecked, consanguineous population[23].

Patterns of introgression among the Lemuriformes

To examine patterns of allele sharing (introgression) within strepsirrhine genera, we first generated a phylogenetic tree of the Lemuriformes using the complete sequences of 3,072 ultraconserved elements (UCEs) from each individual in our dataset using the PHYLUCE[24] pipeline (Fig. 2). Given the substantial evolutionary depth of the Lemuriformes radiation and the large scale of our dataset, UCEs allowed for a computationally tractable phylogenetic reconstruction from full resequencing data. The internal nodes of our species tree are well resolved with strong branch support (local posterior probability, ≥0.95), with the exception of the node placing *Eulemur coronatus* outside the clade containing *Eulemur mongoz* and the *Eulemur fulvus* species complex (local posterior probability, 0.79). All species resolved as monophyletic clades in the UCE tree. We used treePL[25] to estimate species divergence dates using a version of this tree based upon single individuals with a penalized maximum likelihood approach (Extended Data Fig. 4). We observed frequent speciation events occurring from the Early Pliocene through the Middle Pleistocene in all multi-species genera with the exception of *Cheirogaleus*.

To validate our UCE phylogenetic tree and resolve the placement of individuals within *Eulemur*, we also generated intrageneric species trees from 5,000 randomly chosen 10 kb windows for each Lemuriformes family (excluding the single species in Daubentoniidae) (Extended Data Figs. 5–8). Our window trees produced identical topologies to that of our UCE tree with strong branch support (local posterior probability, ≥0.95), with the exception of *E. coronatus* being placed outside the clade containing *Eulemur macaco* and *Eulemur flavifrons* (local posterior probability, 0.68). Additionally, we identified five samples that were probably mislabeled at some point during their curation (Supplementary Table 3 and Supplementary Text). In each case, the node supporting the species clade has strong support (local posterior probability, ≥0.95), but a single individual does not cluster with the other individuals in its designated species (Supplementary Text).

To assess the levels of excess sharing of derived alleles among species within each genus of non-anthropoid primate (limited to those genera with at least three species in our dataset), we used whole-genome resequencing data to generate f_4-ratios and f-branch statistics (f_b) with Dsuite[26]. The f_4-ratios were above 1%, 3% and 5% in 65, 24 and 15 out of 260 species triads, respectively (false discovery rate (FDR)-corrected $P < 0.01$, Z-score ≥ 2), indicating that pervasive allele sharing among species has been commonplace across non-human primates throughout Madagascar (Fig. 3a and Supplementary Tables 4 and 5). To account for the correlated structure of f_4-ratios that may have occurred during ancestral gene flow events, we generated matrices of f_b statistics, which

https://doi.org/10.1038/s41559-024-02596-1

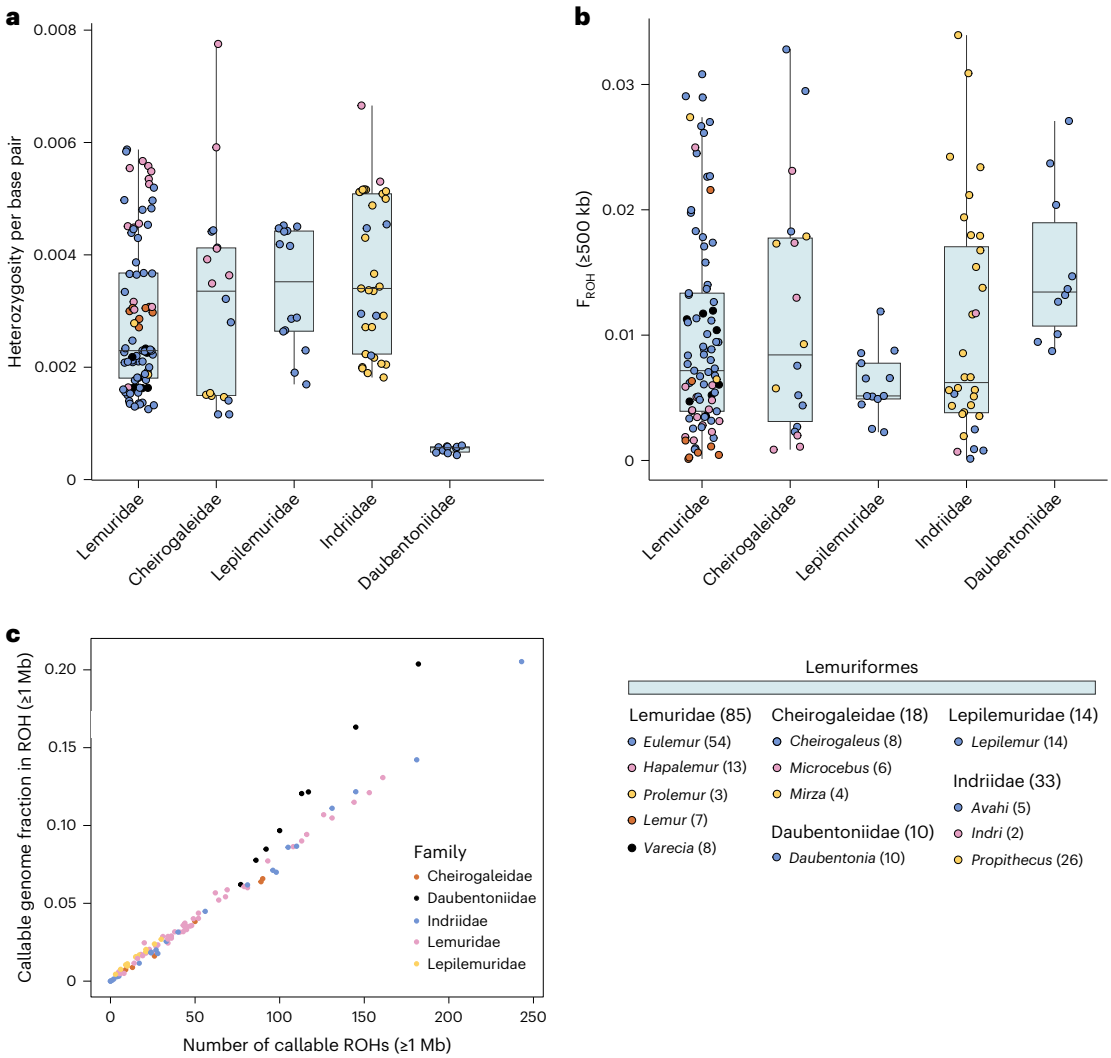


Fig. 1 | Genomic diversity of the Lemuriformes. a, Heterozygosity calculated at the base-pair level from the callable genome per individual. **b**, Proportion of the callable genome in ROHs that are at least 500 kb in length per individual. **c**, Proportion of the callable genome for each wild-born individual that is in ROHs of at least 1 Mb in length plotted against the number of callable ROHs per individual. Bottom right: colour scheme for the genera within families in **a** and **b**. For the boxplots in **a** and **b**, the box corresponds to the interquartile range (IQR), the horizontal line is the median value and the length of whiskers extends up to 1.5 times the IQR. The number of individuals used to derive each boxplot (*n*) is listed in parentheses following the taxonomic group in the legend, and each point represents a unique individual (biological replicate).

assign gene flow to both external and internal branches in a phylogeny (Fig. 3b). The f_b values are largely consistent with f_4-ratios, with the highest observed allele sharing between members of *Hapalemur*, *Eulemur*, *Lepilemur* and *Propithecus* (Table 1), particularly those species inhabiting multiple ecogeographic regions or distribution zones. In accordance with our dated phylogeny, these periods of introgression would have been likely to occur during the Pleistocene following species divergences that occurred during the Early Pliocene through the Middle Pleistocene.

Ecogeographic influence on genomic diversity and demographic history

Most lemur species occupy non-overlapping distribution zones (as previously defined in refs. 27,28 and ref. 4) within genera that are composed of one or more adjacent centres of endemism within larger ecogeographic regions (Fig. 4a,b). We observed a significant effect of ecogeographic region membership (wet forest, dry forest or both) on per base-pair heterozygosity (phylogenetic ANOVA, $F = 6.07$, $P = 0.0101$). The heterozygosity of species that inhabit both forest types was significantly higher than those that inhabit only the wet forest or dry forest (post hoc phylogenetic *t*-tests, $P = 0.0012$ and $P = 0.0424$, respectively). We also observed a significant effect on heterozygosity based on species inhabiting different biogeographic lemur distribution areas: phylogenetic ANOVA ($F = 5.00$, $P = 0.0186$). Although our analysis of these areas was limited to those for which we had data from five or more species (northern east, southern east, north, Sambirano and multiple areas), post hoc phylogenetic *t*-tests only identified significantly higher levels of median heterozygosity when comparing those taxa occupying multiple endemic areas to those in the southern east ($P = 0.0010$) and the Sambirano ($P = 0.0117$).

To assess the role of Madagascar's ecological variation on the Pleistocene demographic history of lemurs, we reconstructed effective population sizes (N_e) through time using SMC++[29] and categorized the resulting demographic history plots according to the presence of each species in wet forests, dry forests or both biomes (Fig. 4a,b and Supplementary Table 1). To minimize the effect of incomplete sampling distributions and major variation in life histories on the overlay of demographic history plots, we only included genera from larger-bodied species that are broadly distributed across Madagascar, for which we have high-coverage resequencing data from all or nearly

https://doi.org/10.1038/s41559-024-02596-1

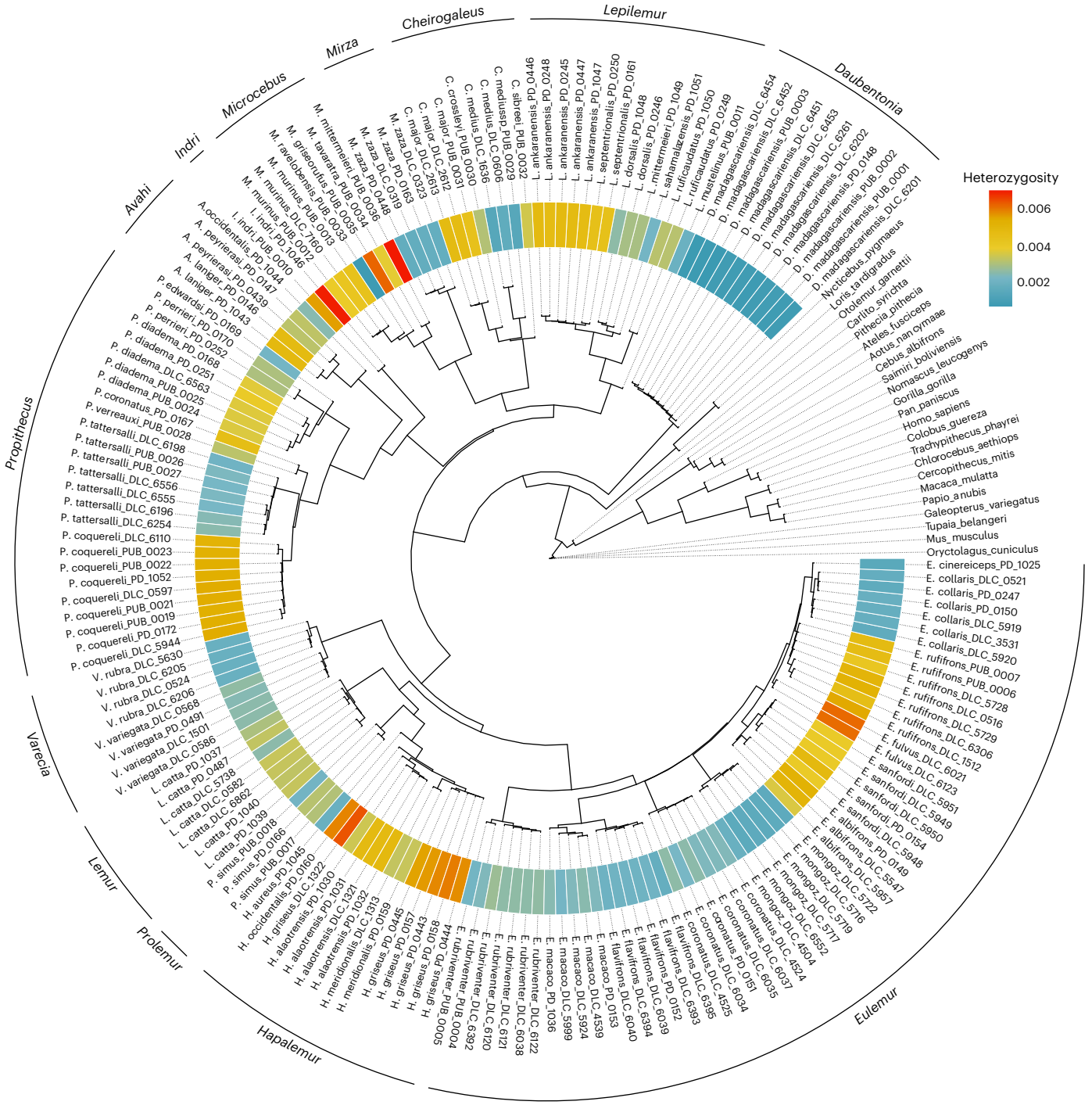


Fig. 2 | Phylogenetic tree of Lemuriformes individuals generated from the gene trees of 3,072 UCEs reconstructed with ASTRAL. Branch lengths are in units of coalescence distance, and dotted lines extending from tree tips are only to improve visualization. The ring of colour-coded values circumscribing the tree corresponds to the heterozygosity of the corresponding individual, ranging from 0.00036 to 0.006658 het × bp^{-1}. The outlier value of *M. mittermeieri* (0.007754 het x bp^{-1}) is coloured independently to retain the colour scale. Individuals without values are outgroup reference assemblies from which heterozygosity values were not calculated.

all species (*Eulemur*, *Hapalemur*, *Propithecus*). Distinct patterns of changing demographic history are visible in each ecogeographic category (Fig. 4c). Demographic history plots from species occupying the dry-forest biome overlap tightly and their N_e begins to collapse at roughly 100 ka. By contrast, species inhabiting the wet-forest biome do not display a consistent pattern throughout the last million years. Species inhabiting both wet-forest and dry-forest biomes also exhibit a consistent pattern of demographic history, which, while having declined from its highest level at 1 Ma, has remained relatively stable for the past 100 ka.

Anthropogenic population declines across Madagascar

To assess population declines in lemurs during the period of human arrival and population expansion across Madagascar, we reconstructed recent demographic histories using GONE[30]. GONE generates N_e estimates for 200 generations before present using patterns

https://doi.org/10.1038/s41559-024-02596-1

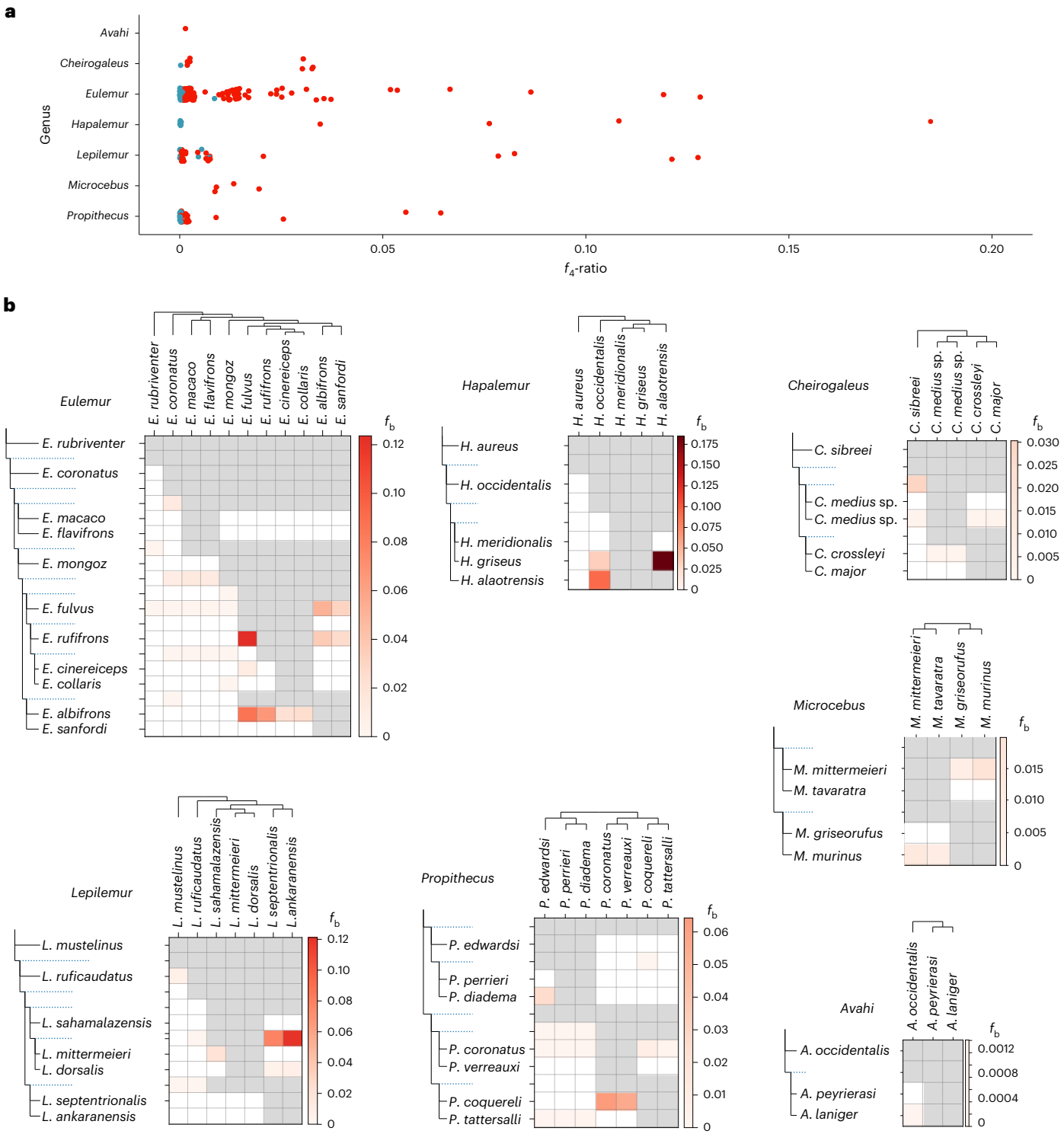


Fig. 3 | Patterns of interspecific allele sharing within genera of lemuriformes. **a**, f_4-ratio admixture proportions for each triad of species (((Species 1, Species 2), Species 3), outgroup) within each genus in our dataset that has at least three species. The significance of the f_4 values was tested with a block-jackknife procedure[26] to produce *Z*-scores and FDR-adjusted *P* values in Dsuite. Values with FDR-corrected *P* values of <0.01 and *Z*-values of ≥2 are depicted in red. **b**, f_b values in Lemuriformes genera. Each box indicates the amount of excess sharing of derived alleles between a species (*x* axis) and a branch (compared to its sister branch) on the UCE tree (*y* axis). The colour scale is shared across all genera and corresponds to the f_b value for a given pairing. Grey boxes represent comparisons that cannot be made by Dsuite as a result of the tree topology structure.

of linkage disequilibrium from unphased genomes. Given that our sample sizes are near the minimum required for GONE as well as the uncertainty of accurate generation times, we sought to test whether the general timing of population shifts is consistent with known anthropogenic activity in Madagascar rather than to obtain precise measures of N_e. For each of the four species with data sufficient to run GONE (Methods), we observed clear declines in N_e coinciding with the expansion of anthropogenic activity in Madagascar (Fig. 5a). Substantial effective population declines appear to have occurred within the last 1,000 years for *Lepilemur ankaranensis*,

Table 1 | Highest values of interspecific excess allele sharing among Lemuriformes species as measured f_b values from Dsuite[26]

Branch 1	Branch 2	f_b	Z-score
H. alaotrensis	***H. griseus***	0.185	36.1
E. fulvus	***E. rufifrons***	0.123	10.1
L. ankaranensis	(*L. mittermeieri*, *L. dorsalis*)	0.121	37.0
H. occidentalis	*H. alaotrensis*	0.092	57.6
E. fulvus	*E. albifrons*	0.087	13.1
L. septentrionalis	(*L. mittermeieri*, *L. dorsalis*)	0.078	36.1
E. albifrons	***E. rufifrons***	0.067	16.1
P. coronatus	*P. coquereli*	0.064	40.1
P. verrreauxi	*P. coquereli*	0.056	41.6
H. occidentalis	***H. griseus***	0.035	25.8
E. sanfordi	***E. fulvus***	0.032	13.2
C. sibreei	(*C. medius*, *C. medius* sp.)	0.030	53.9

The significance of the f_4 values was tested with a block-jackknifed procedure to produce *Z*-scores and FDR-adjusted *P*values. Pairs of species within parentheses represent allele sharing between the species in the first column and the ancestral branch leading to the descendent taxa (relative to its sister branch). In each case, FDR-corrected *P*values are below the minimum reported value of 1×10^{-19}. f_b values for all comparisons are available in Supplementary Table 5. Bold taxa inhabit multiple ecogeographic regions or distribution zones.

Propithecus coquereli and *E. flavifrons* and the past 1,400 years for *D. madagascariensis*.

Secondly, we used the number and length of ROHs identified in each species to calculate F_{ROH} values in 100-year bins and plotted time-specific F_{ROH} values for the past 2,000 years. These chronological ROH (chronROH) plots depict fluctuating levels of inbreeding in recent generations scaled to years before present using species generation times. For cross-validation of this method, we generated chronROH plots for all single and adjacent pairs of endemic areas inhabited by at least two lemur species. There is a clear effect of species distribution range on the F_{ROH} values in the past 2,000 years (Fig. 5b): in three regions, two examined species have parallel F_{ROH} trajectories and consistently overlapping or very similar F_{ROH} values (regions 1, 6, and 10 & 11); in three regions, two or more species have partially overlapping F_{ROH} values and/or parallel trajectories (regions 1 & 12, 2 and 10); and in two regions, the F_{ROH} values and trajectories of two species diverge (regions 2 & 3 and 9). Furthermore, in six regions, we observe an inflection point in the F_{ROH} trajectories at roughly 1,000–1,200 years before present (regions 1, 2 & 3, 9, 10, 1 & 12). This time period corresponds to a shift in human hunting practices away from dwindling populations of now-extinct large-bodied lemurs (and other megafauna) to the medium-bodied and smaller-bodied lemurs that remain extant today[31–33].

Discussion

We analysed 162 high-coverage resequenced genomes from the Lemuriformes. Our results suggest that (1) levels of genomic diversity are consistent within and widely variable among Lemuriformes species; (2) frequent interspecific hybridization, possibly resulting from environmental heterogeneity, is leading to high levels of heterozygosity in some species of lemurs; (3) the long-term demographic history of multiple radiations of lemurs has been shaped by ecological factors; and (4) lemurs have undergone major declines in N_e in the last 2,000 years, during which time inbreeding patterns have been shaped by anthropogenic forces in their local environments.

Our multiple-individual dataset provides robust evidence that the Lemuriformes generally have elevated levels of genomic diversity relative to that observed in other primates[10]. Furthermore, the range of genomic diversity within multiple species-rich genera of Lemuriformes (*Eulemur*, *Microcebus* and *Propithecus*) is comparable to or larger than that reported across the entire Lorisiformes infraorder[10]. We suggest that these high levels of genomic diversity are at least in part a consequence of historical biogeographic effects that have facilitated recurrent interspecific gene flow across the infraorder. Within Madagascar, periodic expansion and contraction of forests and watersheds during climatic shifts have resulted in a history of fluctuating habitat connectivity and isolation, resulting in discrete areas of endemism[1,3]. Previous research has demonstrated that speciation patterns within many lemuriform genera are broadly consistent with this retreat–dispersal watershed model[1,4,9]. Our results indicate that these same underlying ecological processes in Madagascar likely influenced the genomic diversity of lemurs. We suggest that the extreme geographic heterogeneity of Madagascar, coupled with the high density of species in close geographic proximity, facilitated interspecific allele sharing through recurrent interspecific hybridization that occurred when formerly fragmented habitats were reconnected across Madagascar during climatic shifts. In agreement with other recently dated phylogenetic trees[10,34], the divergence dates we estimated revealed frequent speciation events in many genera of Lemuriformes between the Early Pliocene and Middle Pleistocene (Extended Data Fig. 4). Divergences within this time range were particularly common among species within *Eulemur*, *Hapaplemur*, *Propithecus* and *Lepilemur*, for which we observed the strongest signals of allele sharing (Fig. 3 and Extended Data Fig. 4). Given that these speciation events would have pre-dated much of the Pleistocene climatic oscillations in Madagascar, there would have been sufficient time for introgression to have occurred among these species during periods of forest connectivity in the Pleistocene. However, we do caution interpretation of the results from *Lepilemur*; given that we were unable to obtain samples from most species, the signals and timing of introgression within this genus need additional investigation. Furthermore, we observed a significant positive effect on genetic diversity among species occupying multiple ecogeographic regions and/or distribution zones, and the results of our introgression analysis indicate that a disproportionate number of the lemurs with the highest values of interspecific excess allele sharing—*E. fulvus*, *Eulemur rufifrons*, *Hapalemur griseus* and *Hapalemur occidentalis*—also occupy multiple ecogeographic regions or distribution zones (Figs. 3 and 4 and Table 1). It is possible that the centrally distributed species, especially those with cross-island populations, are hybridizing with multiple species on the edges of their distributions, transferring variants across Madagascar through a pot-stirring effect. Accordingly, within the *E. fulvus* group, *Hapalemur* and *Propithecus*, in particular, there appears to be a higher level of genomic diversity in the central (and to a lesser extent some northern) species to the exclusion of the southern members (*Eulemur collaris*, *Eulemur cinereiceps*, *Hapalemur meridionalis*, *Hapalemur aureus*, *Propithecus verreauxi*). Substantial levels of introgression are also evident in some species of *Lepilemur* and *Propithecus*, which typically occupy microhabitats and contiguous habitats, respectively. In these cases, hybridization appears to be occurring between neighbouring species on opposite sides of rivers; however, as we do not have resequencing data from all members of these genera, the introgression patterns are more challenging to discern. Although we did observe some weaker signals of allele sharing between species with non-adjacent distributions, we caution that they could be the result of correlated allele frequencies that result from a stronger signal of introgression in related taxa. Alternatively, some species did formerly occupy broader ranges in the past[35], and gene flow occurring between multiple species in a stepping-stone process could also explain a weaker but still valid signal of introgression between non-adjacent taxa. Nonetheless, our results suggest that allele sharing during periods of population reconnection has been integral in maintaining genomic diversity in lemurs. Given the widespread deforestation and habitat fragmentation across Madagascar, the loss of contact or hybrid zones raises additional conservation concerns.

https://doi.org/10.1038/s41559-024-02596-1

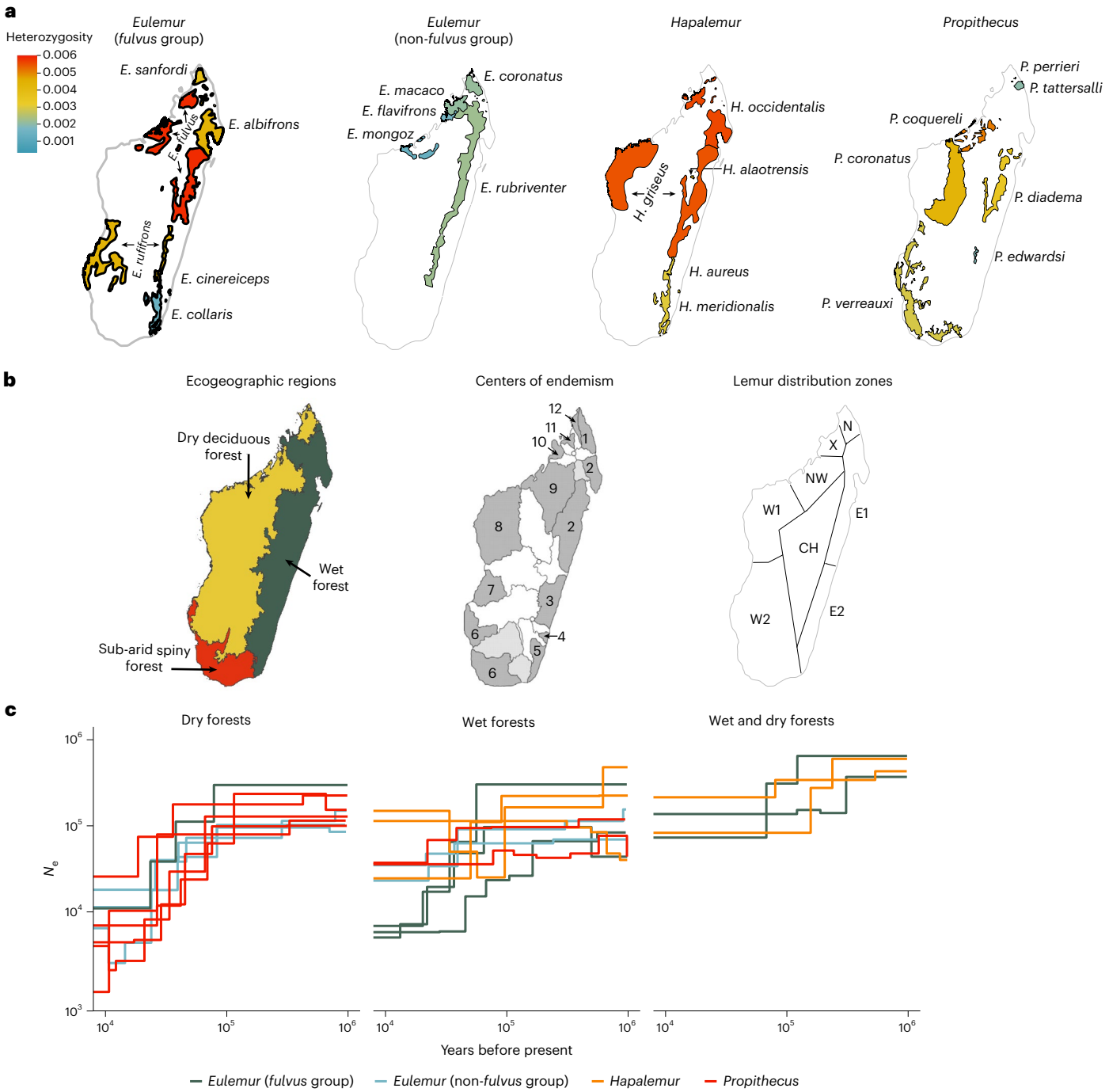


Fig. 4 | Ecogeographic distribution and demographic history of lemurs in Madagascar. a, Species range distributions in three genera of lemurs, colour-coded by per base-pair heterozygosity in the callable genome. Species with multiple distributions in the eastern wet and northern dry forests are labelled with arrows from the centre of the island. **b**, Biogeographic models of species diversity in Madagascar. Left: primary ecogeographic zones of climate variation. Data from ref. 70. Middle: centres of endemism corresponding to retreat or dispersal watersheds[3,71]. Right: biogeographic areas of lemur distribution. CH, central highlands region; E1, northern east region; E2, southern east region; N, north region; NW, northwest region; W1, central western region; W2, southern western region; X, Sambirano region. **c**, Log-transformed N_e from 1 Ma to 10 ka as reconstructed by SMC++ depicting demographic histories of *Eulemur*, *Hapalemur* and *Propithecus* grouped by the presence of each species in dry forest, wet forest or both. During the last 100 ka, dry-forest (dry-deciduous and sub-arid spiny forest) species follow a steep pattern of declining N_e, while those species with wet and dry distributions remain more stable. Species exclusively in the wet forest do not show a consistent pattern. Panel **b**, centre, adapted with permission from ref. 71, OUP. Panel **b**, right, adapted with permission from ref. 4, National Academy of Sciences.

We also recommend further work to identify not only patterns of admixture in the Lemuriformes but also the effect of introgression on their phylogeny. Both our UCE and 10 kb window tree topologies are largely consistent with the results of primate species trees derived from single-individual genome-scale data[10,34] and have high levels of branch support at almost all nodes. However, many published molecular phylogenies of the Lemuriformes (including ours) display topological differences with each other, in some cases with high branch support[9,10,34,36]. We note that different regions of the genome can have different histories, and complete topological consistency is

https://doi.org/10.1038/s41559-024-02596-1

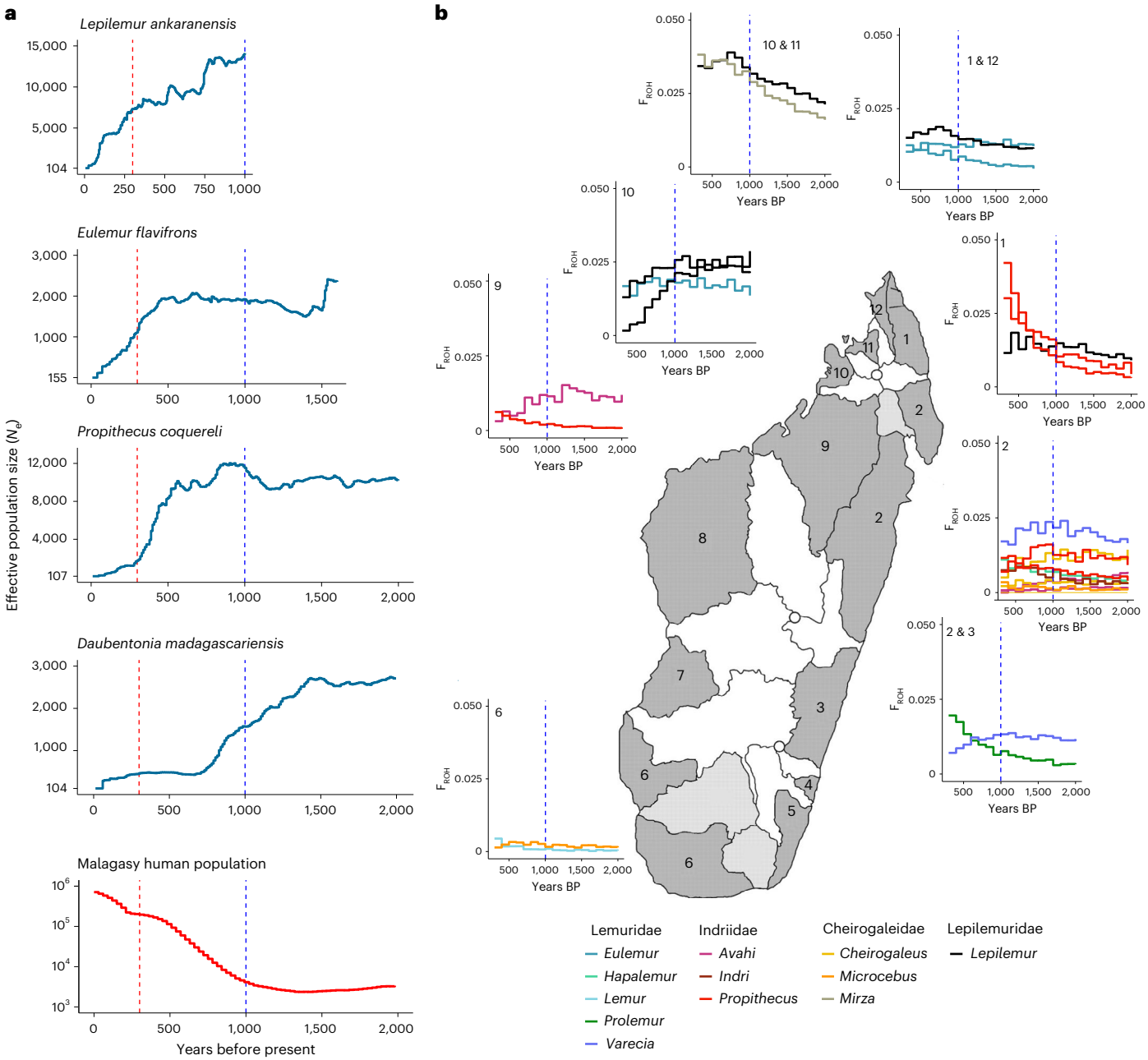


Fig. 5 | Declining effective population sizes in lemurs and regional patterns of inbreeding across Madagascar during the last 2,000 years. a, Historical N_e in the last 2,000 years (up to 200 generations) from four lemur species reconstructed using GONE[30] and the Malagasy human population N_e (from ref. 73). The variable time ranges of N_e plots result from the species generation times used to calculate time before present. The orange dotted line at 300 years before present indicates the rapid acceleration of anthropogenic deforestation in Madagascar. The blue dotted line at 1 ka corresponds to the beginning of human expansion and the change in lemur hunting patterns following the extinction of large-bodied subfossil species[31]. **b**, chronROH plots of the proportion of the callable genome per species in ROH between 300 and 2,000 years before present (BP) across endemic areas in Madagascar. In seven of nine endemic area groupings, where multiple species of lemurs are distributed, the recent temporal patterning of F_{ROH} either directly overlaps or follows a consistent pattern. An inflection point for inbreeding levels is visible at ~1 ka across multiple species and regions. Map of endemic areas adapted from previous publications[3,71].

not always expected[37]. Furthermore, this lack of topological consistency could result from the use of different combinations of loci that yield different gene trees as a result of interspecific allele sharing. Nonconformity among multiple trees is particularly notable in the genus *Eulemur*, which exhibits multiple hybrid zones and is unlikely to fit a strictly tree-like phylogenetic model[38,39]. Additionally, we cautiously acknowledge the phylogenetic position of *Microcebus ravelobensis* in our tree. This individual was sequenced at a substantially lower genomic depth of coverage (~5×) than other samples in our dataset, which could have influenced its placement relative to that observed in other phylogenies[34]. We also draw attention to the need for future studies using high-coverage genome-wide resequencing data from multiple specimens collected with known geographical provenance. Recent refinement of the taxonomy of the Lemuriformes has resulted in the splitting of many species, and this must be taken into account when working with biobanked samples that may have been collected before taxonomic subdivision. Broader geographic sampling will be essential to fully reveal the complex relationship between interspecific gene flow and phylogenetic history in the Lemuriformes.

We observed a clear ecological influence on N_e in multiple lemur genera throughout the last million years (Fig. 4c). This effect is most striking in the dry-forest biome, where all ten species demonstrate a tightly connected pattern of collapsing N_e around 100 ka. The dry-forest pattern is distinct from the more stable pattern observed in multiple-biome species and the lack of continuity among the wet-forest species. Although we acknowledge the potential effects of linked selection on SMC models[40] and the possibility that current distributions do not necessarily reflect historic ranges, the demographic patterns are nonetheless consistent. Recent studies have demonstrated that primates living in seasonal and dry forests must overcome strong ecological challenges to survive hotter and drier conditions[41–44] and, in some cases, have adapted to these environments[43]. We suggest that the harsher conditions in the dry-forest biomes have allowed for less ecological flexibility among the resident primates and exacted similar population stresses throughout the Pleistocene. By contrast, the wet-forest biomes are probably a less challenging environment, allowing for other local effects to have had greater influence on the lemurs' demographic history.

We identified a pattern of declining lemur N_e that corresponds with explosive human population growth and landscape modification in Madagascar during the last ~1,000 years[32,45]. This relationship is demonstrated by the striking chronological inversion of N_e between the four lemur species suitable for analysis with GONE and the Malagasy human population (Fig. 5a). We also observed a frequent inflection point of ~1,000 years before present in the trajectories of inbreeding levels in our chronROH plots (Fig. 5b), suggesting that previously stable lemur populations were disrupted by the expansion of cattle grazing and deforestation during the last 1,000 years[32,45]. Furthermore, the temporal pattern of cut-marked lemur bones in the archaeological record indicates that this time period coincides with a shift in hunting patterns toward extant lemurs following the collapse of megafauna populations, including the now-extinct larger-bodied subfossil lemurs[18,31,32,46]. Given the low number of samples we used to run GONE, the absolute N_e values should be treated with caution, although the trajectories and general timing of population-size shifts are consistent with archaeological and botanical evidence of human activity in Madagascar. Furthermore, we used an extremely cautious approach that only reconstructed demographic history for those species with the highest number of samples that did not display obvious artefacts.

We have provided direct genomic evidence that primates in localized geographic regions have undergone consistent patterns of demographic history. Our chronROH plotting approach revealed distinct patterns of population demographic history across multiple genera of lemurs throughout areas of sympatric distribution. The striking consistency of chronological F_{ROH} slopes and values in many distribution areas suggests that anthropogenic and natural disturbances may variably affect population sizes of lemurs across Madagascar, but uniformly within localities. Given the large variety of F_{ROH} trajectories and values in our dataset, it would seem implausible that both phylogenetically distinct taxa (*Lemur catta* and *Microcebus griseorufus* (region 6); *Lepilemur dorsalis* and *Mirza zaza* (regions 10 & 11) and closely related species (*Propithecus perrieri* and *Propithecus tattersalli* (regions 1 & 12); *Lepilemur mittermeieri* and *Lepilemur sahamalazensis* (region 10)) would have parallel slopes and nearly overlapping levels of F_{ROH}, were it not for the shared impact of their local environments and internal anthropogenic effects throughout time. Although several studies have used large bins of ROH length (for example, 1–5 Mb) to generate rough estimates of time-specific F_{ROH} values[23,47–51], the present study uses this approach to assess relatively fine-scale changes in inbreeding levels across time. Our goal was to use F_{ROH} inbreeding levels as a proxy for time-specific population declines that have resulted from human activity (hunting and deforestation) during the last ~2,000 years of human population expansion in Madagascar. Given that individual species from different genera of lemurs inhabit the same discrete regions of the island, we were able to use this natural patterning not only as a way to control for phylogenetic effects but also to cross-validate our chronROH plot approach.

Conclusions

We generated the most complete set of high-coverage resequenced genomes of the Lemuriformes to date, which includes 162 unique individuals from 50 species. Our results reveal consistent levels and broad ranges of genomic diversity in these primates. We identified a widespread pattern of introgression in most genera of lemurs. Levels of allele sharing are particularly high in species with multiple populations in multiple ecogeographic regions, which probably indicates that patterns of connectivity and fragmentation allow some species to share alleles during climatic shifts. We have also identified a shared pattern of demographic history in the different ecogeographic regions of Madagascar. Species occupying the dry-forest biome have undergone more uniform population declines than those lemurs in the wet forest and those inhabiting both wet and dry forests. During the last ~1,000 years, lemurs appear to have undergone rapid and intense declines in N_e that correspond to the expansion of human activity in Madagascar and the transition from the hunting of recently extinct large-bodied subfossil lemurs. Furthermore, recent historical inbreeding levels in lemurs are consistent across genera in discrete areas of endemism throughout Madagascar. We observed clear overlapping and parallel inbreeding trajectories across time in most areas of endemism. These results advance our understanding of how the genomic diversity and demographic history of the Lemuriformes have been influenced by ecological and anthropogenic factors. Furthermore, it is also possible that the recent dramatic population declines many species have experienced in the last few decades are not fully reflected in our data. We strongly encourage the development of resequencing projects that target broader panels of geolocalized individuals from multiple populations to circumvent the limitations that we encountered.

Methods

Sample collection

Samples of frozen tissue, blood and cell lines were obtained from previously collected specimens stored at The Duke Lemur Center, German Primate Center, The Barcelona Zoo, Penn State University, National Museums Scotland, The University of Toulouse and the European Association of Zoos and Aquaria (EAZA) Biobank. Detailed sample data are available in Supplementary Table 1. In accordance with the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) regulations, no sample was moved from outside its primary storage jurisdiction. Preparation and sequencing of samples from European collections occurred at UPF or CNAG in Barcelona, and the processing of samples from collections in the USA was performed at the Baylor College of Medicine (Houston, Texas). DNA extraction and library preparation methods are described in detail elsewhere[10]. In brief, DNA was extracted from tissue samples and prepared for sequencing with PCR-free libraries using the KAPA HyperPrep kit (Roche). All sequencing data were generated with the Illumina NovaSeq platform to produce 2 × 150 bp reads with a target depth of coverage of 30×. Additional short reads from 31 individual lemurs were downloaded from the NCBI Sequence Read Archive (Supplementary Table 1).

Variant calling and filtration

All sequenced individuals were mapped to the highest quality annotated reference genome assembly of the nearest phylogenetic relative (Supplementary Table 1), including *L. catta*[52], *P. coquereli*[13], *Microcebus murinus*[53] and *D. madagascariensis*[54]. Reference quality metrics are presented in Supplementary Table 6. Although some species are mapped to distant reference assemblies, estimates of genomic diversity and

heterozygosity are either not significantly or negligibly affected by the choice of and distance to these reference assemblies[10].

Reads were trimmed of adaptors using cutadapt (v.3.4) and mapped with BWA (v.0.7.15) mem[55]. Multiple sequencing libraries were merged into single BAM files with SAMtools (v.1.9)[56] merge and read group information was added with the Picard Toolkit (v.2.8.2). Reference assemblies were split into 30 Mb windows and variants were called using the GATK (v.4.1.7.0)[57] unified haplotype caller, following the GATK 'best practices' pipeline with the ERC BP_RESOLUTION option. Each sample was genotyped individually using the GATK haplotype caller. Joint genotyping approaches can be more accurate than individual genotyping when the number of individuals per species and sequenced depths of coverage are similar. However, because we sequenced a variable number of individuals per species (1–12 individuals) with uneven depths of coverage between newly sequenced genomes (~30 ×) and those publicly available on NCBI (~5× to ~30×), we concluded that joint genotyping could have introduced artefacts when making comparisons between species.

Variant call format (VCF) filtration followed a protocol implemented in a previous work[10]. First, we calculated the depth of coverage at each genomic position using mosdepth[58] and generated the modal coverage for each sequenced individual across its genome, assigning minimum and maximum coverage values to one-third and twice the modal values, respectively. Subsequently, we concatenated each 30 mb gVCF and removed uncalled (./.) and homozygous reference positions (0/0) using BCFtools (v.1.9)[56]. We then removed low-quality single-nucleotide polymorphisms (SNPs) and insertions and deletions (indels) using the following exclusive filters:

```
bcftools filter --threads 1 -e \"TYPE!='snp' |
(GT='het' & FMT/AD[*:*] < $MIN_HET_AD) | AC > 2 |
FMT/DP <= $MIN_COV | FMT/DP >= $MAX_COV | QD < 2 |
FS >60 | MQ < 40 | SOR > 3 | ReadPosRankSum < -8.0
| MQRankSum < -12.5
bcftools filter--threads 1 -e \"TYPE!='indel' |
(GT='het' & FMT/AD[*:*] < 3) | FMT/DP <= $min_
coverage | FMT/DP >= $max_coverage | QD < 2 | FS >
200 | MQ < 40 | SOR > 3 | ReadPosRankSum < -20.0
```

To generate individual callability maps, we excluded uncalled positions and those with aberrant depth values or low-quality genotypes.

```
bcftools filter -e \"(GT='./.') | (GT='het' & FMT/
AD[*:*] < $MIN_HET_AD) | FMT/DP <= $MIN_COV | FMT/
DP >= $MAX_COV | FMT/GQ <= 30
```

We also generated a secondary allele balance filter by intersecting the callability filters with a bed file of all heterozygous positions with allele balances of >0.75 and <0.25. With these parameters, we generated two sets of VCFs. First, we produced filtered SNP and indel VCFs for each sequenced individual, with its accompanying callability masks. Secondly, we generated merged SNP VCFs composed of each individual mapped to a given reference assembly. In the latter case, the order of operations varied slightly to allow for the variable individual sequencing depths across our dataset. Initially, each 30 Mb gVCF window was merged across corresponding individuals with BCFtools and only SNP positions across all samples were retained. Each window was then split by individual and filtered as above. We then re-merged each VCF window across individuals, concatenated them and removed any remaining invariant positions. Combined callability masks (for example, species-level or genus-level) were generated at the level of interest by intersecting individual callability masks with BEDTools (v2.29.0)[59] intersect and further filtered to exclude scaffolds that were non-autosomal or less than 1 Mb in length.

Relatedness

For each species, we identified potentially related individuals using NgsRelate[60] (v.2.0). All members of each species were selected from their corresponding joint VCFs and filtered to keep only variable positions with no more than 20% missing data, using VCFtools[61], and piped into NgsRelate:

```
vcftools--gzvcf $VCF--keep
$sampledir/${genus}_${sp}.all.txt--max-missing
0.8--maf 0.000001--recode--recode-INFO-all--stdout
| ngsRelate -h-T GT
```

The degree of relatedness among individuals was determined using the theta kinship coefficient[62]. We classified individuals with theta values of 0.5, 0.25 and 0.125 as identical, first-degree and second-degree, respectively. We removed identical individuals from all further analyses and those classified as first-degree or second-degree relatives from further population-level analyses (Supplementary Table 7). NgsRelate requires at least three individuals to identify relatedness coefficients within populations. For species represented by only two individuals, both were retained for downstream analyses and treated with caution.

Heterozygosity

We calculated genome-wide heterozygosity by identifying the levels of nucleotide diversity across the genome. For each individual, we calculated the total number of heterozygous single-nucleotide variant positions that passed filtration parameters and divided this sum by the total number of callable nucleotide positions in the individual's genome using custom perl scripts. Median heterozygosity was calculated for each species, and differences between superfamilies were tested for significance with Wilcoxon rank sum tests in R (v.4.2.2). We identified long ROHs in our samples, using a hidden Markov model classification programme in BCFtools and a constant recombination rate per base of 1 Mb per 1 cM. For optimal results, BCFtools recommends including population-specific allele frequencies from >20 individuals and a genetic map to model recombination hotspots. Given the unique composition of our dataset (many species, variable and limited numbers of individuals, mapping to non-specific reference assemblies and so on) it was not feasible to meet these strict guidelines. Nonetheless, BCFtools can still produce accurate results by calculating allele frequencies on the fly and in the absence of a genetic map, by using the default parameters as follows:

```
bcftools roh -G30--AF-dflt 0.4 -R $individual_
callablity_mask $vcf
```

From each output file, we selected ROHs of at least 500 kb in length with quality scores of ≥70.

Phylogenetic tree construction

We used the PHYLUCE[24] (v.1.2.1) UCE phylogenomics pipeline to identify the UCE sequences in our dataset. First, we downloaded the sequences of ~5,000 UCEs from a previous publication[63], which we aligned to our four Lemuriformes reference assemblies and a broad panel of outgroups, including 19 non-Lemuriformes primates, *Galeopterus variegatus*, *Tupaia belangeri*, *Oryctolagus cuniculus* and *Mus musculus*. For the alignments that passed default filtration parameters, the assembly-specific sequences plus 500 bp flanking regions were extracted. For each individual in our dataset, we overlaid indel and SNP variants onto the UCE sequence from the corresponding assembly and removed non-callable regions with BCFtools[56]. Subsequently, for each individual, any UCE that did not match the alignment according to default parameters in PHYLUCE was removed from the individual dataset. We then selected all UCEs that were present in at least 75%

of samples for downstream analysis. Multiple sequence alignments (MSAs) were constructed for each UCE using MAFFT[64] and internally trimmed with GBLOCKS[65]. Within each MSA, we used Treeshrink[66] to remove individuals that introduced extremely long (possibly artefactual) branches. Phylogenetic trees for each UCE were generated from the resulting MSAs with IQ-TREE 2 (ref. 67) using 1,000 ultrafast bootstrap replicates. Ultimately, we constructed a species tree from the individual UCE trees using ASTRAL[68]. We generated divergence dates for each species in our phylogenetic tree using a penalized maximum likelihood approach with treePL[25]. We selected the single individual from each species with the highest sequencing depth and concatenated all UCEs with 100% coverage across samples (and outgroups) using PHYLUCE. We concatenated these 646 UCEs and then used IQ-TREE 2 to identify the most probable mutation model (TVM + F + R7). Branch lengths were then estimated with IQ-TREE 2 as follows:

```
iqtree2 -s concatenated_UCE_alignment.phylip
-te astral_species_tree.tree -pre branch_length_
estimates -nt 4 -m TVM+F+R7
```

We augmented the topologically similar dated UCE tree from ref. 10 (with which we shared resequencing data and methodology) with the additional species in our dataset using treePL. Given the paucity of Lemuriformes in the fossil record, there are no available fossil calibrations within the infraorder (excluding the potential Chiromyiformes *Plesopithecus* and *Propotto* from mainland Africa[12]). As a result, we used the high-confidence divergence dates from the primate-wide MCMC-tree analysis in ref. 10 as calibration points for treePL (Supplementary Table 8). Following the treePL priming and cross-validation steps, we set the following parameters for the divergence dating analysis: opt = 3, optad = 2, optcvad = 2, smooth = 0.0001.

Additionally, in an effort to resolve ambiguous nodes within the Lemuriformes, we constructed species trees for each family using the sequences of 5,000 random genomic windows of 10 kb length. For each reference assembly, we generated 10,000 non-overlapping 10 kb windows using BEDTools makewindow. Nucleotide sequences for each window were generated as above by overlaying indel and SNP variants onto the reference sequence using BCFtools. For each reference assembly (and all species mapping to it), windows were ranked globally according to the average number of missing positions, and the 5,000 most complete windows were selected for phylogenetic analysis. As above, MSAs were generated with MAFFT, phylogenetic trees were inferred with IQ-TREE 2 and species trees were reconstructed with ASTRAL.

Interspecific allele sharing

We used Dsuite[26] to identify patterns of allele sharing between Lemuriformes species. Tests were run on all genera represented by at least four species in our dataset, including *Eulemur*, *Propithecus*, *Hapalemur*, *Microcebus*, *Lepilemur*, *Cheirogaleus* and *Avahi*. First, we generated filtered VCFs for each genus by using BCFtools and VCFtools with callability filters intersected for each genus. We excluded coding regions, scaffolds that were non-autosomal or less than 1 Mb in length and positions that were non-variable or had missing data. Subsequently, we calculated f_4-ratio statistics for all species trios within the phylogenetic tree of each genera using the Dsuite Dtrios command. We specified the species tree structure generated from our UCE or (where applicable) genomic window trees. To facilitate the interpretation of correlated f_4-ratios, we calculated f_b statistics using the Dsuite Fbranch command. Heatmaps of the resulting f_b matrices were constructed with dtools.py.

Ecogeographic–genetic associations

We downloaded species distribution polygons from the IUCN Red List of Threatened Species[69] and the ecoregions of Madagascar from ref. 70. Heterozygosity values for each species were merged with their respective distribution shapefiles and plotted in R (v.4.2.2) with a shared colour scale for heterozygosity. Ecoregion categories followed those of ref. 70, with the exception that the arid dry-deciduous forest and sub-arid spiny forest zones were both categorized as dry forest. As above, lemur species distributions were categorized into previously defined discrete centres of endemism[1,3,27,71] (Supplementary Table 1).

Significant differences in heterozygosity between ecoregions and centres of endemism were tested with phylogenetic ANOVAs and post hoc phylogenetic *t*-tests with Holm–Bonferroni-adjusted *P* values in phytools (v.1.5-1)[72]. In each case, species median heterozygosity values were square-root transformed to normalize the skewed distribution. We removed two species in our dataset that were extreme heterozygosity outliers (*D. madagascariensis* and *Microcebus mittermeieri*) and one species (*M. ravelobensis*) that had low coverage. Phylogenetic ANOVA and *t*-tests of different mean heterozygosity levels between ecoregions included all other species. For the tests of comparing distinct centres of endemism[3,71], we used a slightly coarser classification[4], which corresponds to the eight main centres where lemurs are distributed. To limit the statistical bias of small sample sizes, we restricted our tests to compare those regions that are inhabited by at least five species: northern east, southern east, north, Sambirano and multiple areas. Using the finer-grained classification depicted in Fig. 5b, these regions correspond to areas 2, 3–5, 1 & 12, 10 & 11 and multiple areas as previously defined[3,4,71].

Given our focus on recent changes in demographic history, we used SMC++[29] to reconstruct the N_e throughout the Pleistocene. Although MSC-based demographic history reconstructions can provide higher accuracy in recent years with larger sample sizes, we used SMC++ to take advantage of multiple samples where possible (Supplementary Table 1). For each species, we used SMC++ vcf2msc to convert our joint filtered VCFs to SMC format for each autosomal scaffold longer than 1 Mb, while iterating over each distinguished lineage in our VCFs. We then estimated demographic histories using SMC++ estimate with per genus mutation rates from ref. 10. SMC++ output files for each species were merged in R and categorized by species distribution (dry forest, wet forest or both). Ecogeographically categorized demographic history data was then overlaid and plotted with log transformation.

Holocene population declines and inbreeding levels

We used GONE[30] to estimate changes in the N_e of lemurs during the most recent 200 generations. In contrast to MSC-based approaches, GONE uses patterns of linkage disequilibrium to identify historical changes in N_e. GONE generates per-generation N_e values from the distribution of linkage disequilibrium levels between sampled SNP pairs across the genome. As such, it only requires SNP data as input, without the need for phased genomes, mutation rates or chromosome-level genome contiguity. Although GONE is optimally suited for large sample sizes, it can reliably reconstruct demographic history from fewer than ten individuals[30]. We converted our filtered masked VCFs to PLINK format using BCFtools and VCFtools[61] and used a custom script to convert non-standard chromosome or scaffold names. Given that GONE is not optimized for a very large number of variants, we randomly subset VCFs with more than five million variants before converting them to GONE format. We undertook several steps to account for the small sample sizes in our dataset. For each species with at least five individuals, we conducted three independent runs of GONE, using randomly generated SNP sets (50,000 SNPs per autosomal scaffold of ≥1 Mb) to confirm the absence of any major deviations in demographic history, then selected the median N_e value for each temporal bin. Given that GONE can be sensitive to migration and population structure, we removed individuals that were clear PCA outliers, reduced the default maximum recombination value to 0.01 and excluded any species with deviations from Hardy–Weinberg equilibrium of ≥0.03 or the presence of known artefacts (rapid increase and decrease of N_e). The time before present

for each generation was calculated with species generation times reported in ref. 10.

We used the length and number of ROH segments to calculate the changing levels of inbreeding for the past 2,000 years in Madagascar. We first calculated the age of each ROH segment in generations using the following equation: generation = 100 / (2 × ROH length)[51]. Each generation value was converted to an estimate of the number of years before present by multiplying it by the generation time of the corresponding species. Each ROH segment was then assigned to 100-year bins for the last 2,000 years. Subsequently, we calculated the average F_{ROH} value across each ROH segment in each bin for each species. We plotted binned ROH times against the time-specific F_{ROH} values to generate stairway plots of inbreeding (chronROH plots) for the last 2,000 years. We observed a sharp drop in the signal for time bins corresponding to the first 300 years before present and removed them from subsequent analyses. Based on IUCN species range distribution data, we classified each species as belonging to one or more areas of endemism[3,4,27,71]. To identify whether consistent patterns of historical inbreeding occurred across Madagascar, we overlaid chronROH plots for each distribution range that is occupied by two or more lemur species.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All newly generated sequencing data are deposited in the European Nucleotide Archive (PRJEB77609) and NCBI (PRJNA1156176). Additional accession numbers for previously published sequencing data to allow for a minimum dataset for reproducibility are available in Supplementary Table 1.

Code availability

Custom scripts are available at https://github.com/nomascus/Lemur_PopGen_NEE.

References

1. Vences, M., Wollenberg, K. C., Vieites, D. R. & Lees, D. C. Madagascar as a model region of species diversification. *Trends Ecol. Evol.* **24**, 456–465 (2009).
2. Antonelli, A. et al. Madagascar's extraordinary biodiversity: evolution, distribution, and use. *Science* **378**, eabf0869 (2022).
3. Wilmé, L., Goodman, S. M. & Ganzhorn, J. U. Biogeographic evolution of Madagascar's microendemic biota. *Science* **312**, 1063–1065 (2006).
4. Pastorini, J., Thalmann, U. & Martin, R. D. A molecular approach to comparative phylogeography of extant Malagasy lemurs. *Proc. Natl Acad. Sci. USA* **100**, 5879–5884 (2003).
5. Everson, K. M., Jansa, S. A., Goodman, S. M. & Olson, L. E. Montane regions shape patterns of diversification in small mammals and reptiles from Madagascar's moist evergreen forest. *J. Biogeogr.* **47**, 2059–2072 (2020).
6. Brown, J. L. et al. Spatial biodiversity patterns of Madagascar's amphibians and reptiles. *PLoS ONE* **11**, e0144076 (2016).
7. Yoder, A. D., Heckman, K. L., Lehman, S. M. & Fleagle, J. G. in *Primate Biogeography* (eds Lehman, S. M. & Fleagle, J. G.) 255–268 (Springer, 2006).
8. Lei, R. et al. Phylogenomic reconstruction of sportive lemurs (genus *Lepilemur*) recovered from mitogenomes with inferences for Madagascar biogeography. *J. Hered.* **108**, 107–119 (2017).
9. Markolf, M. & Kappeler, P. M. Phylogeographic analysis of the true lemurs (genus *Eulemur*) underlines the role of river catchments for the evolution of micro-endemism in Madagascar. *Front. Zool.* **10**, 70 (2013).
10. Kuderna, L. F. K. et al. A global catalog of whole-genome diversity from 233 primate species. *Science* **380**, 906–913 (2023).
11. Yoder, A. D., Cartmill, M., Ruvolo, M., Smith, K. & Vilgalys, R. Ancient single origin for Malagasy primates. *Proc. Natl Acad. Sci. USA* **93**, 5122–5126 (1996).
12. Gunnell, G. F. et al. Fossil lemurs from Egypt and Kenya suggest an African origin for Madagascar's aye-aye. *Nat. Commun.* **9**, 3193 (2018).
13. Guevara, E. E. et al. Comparative genomic analysis of sifakas (*Propithecus*) reveals selection for folivory and high heterozygosity despite endangered status. *Sci. Adv.* **7**, eabd2274 (2021).
14. Williams, R. C. et al. Conservation genomic analysis reveals ancient introgression and declining levels of genetic diversity in Madagascar's hibernating dwarf lemurs. *Heredity* **124**, 236–251 (2020).
15. Hawkins, M. T. R. et al. Genome sequence and population declines in the critically endangered greater bamboo lemur (*Prolemur simus*) and implications for conservation. *BMC Genomics* **19**, 445 (2018).
16. Perry, G. H. et al. Aye-aye population genomic analyses highlight an important center of endemism in northern Madagascar. *Proc. Natl Acad. Sci. USA* **110**, 5823–5828 (2013).
17. Hunnicutt, K. E. et al. Comparative genomic analysis of the pheromone receptor class 1 family (V1R) reveals extreme complexity in mouse lemurs (Genus, *Microcebus*) and a chromosomal hotspot across mammals. *Genome Biol. Evol.* **12**, 3562–3579 (2020).
18. Hansford, J. P., Lister, A. M., Weston, E. M. & Turvey, S. T. Simultaneous extinction of Madagascar's megaherbivores correlates with late Holocene human-caused landscape transformation. *Quat. Sci. Rev.* **263**, 106996 (2021).
19. Hixon, S. W. et al. Late Holocene spread of pastoralism coincides with endemic megafaunal extinction on Madagascar. *Proc. Biol. Sci.* **288**, 20211204 (2021).
20. Ralimanana, H. et al. Madagascar's extraordinary biodiversity: threats and opportunities. *Science* **378**, eadf1466 (2022).
21. Perry, G. H. et al. A genome sequence resource for the aye-aye (*Daubentonia madagascariensis*), a nocturnal lemur from Madagascar. *Genome Biol. Evol.* **4**, 126–135 (2012).
22. Narasimhan, V. et al. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* **32**, 1749–1751 (2016).
23. Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. & Wilson, J. F. Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* **19**, 220–234 (2018).
24. Faircloth, B. C. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* **32**, 786–788 (2016).
25. Smith, S. A. & O'Meara, B. C. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690 (2012).
26. Malinsky, M., Matschiner, M. & Svardal, H. Dsuite—fast *D*-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* **21**, 584–595 (2021).
27. Martin, R. D. Adaptive radiation and behaviour of the Malagasy lemurs. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **264**, 295–352 (1972).
28. Martin, R. D. in *Creatures in the Dark: The Nocturnal Prosimians* (eds Alterman, L. et al.) 535–563 (Plenum, 1995).
29. Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2017).
30. Santiago, E. et al. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol. Biol. Evol.* **37**, 3642–3653 (2020).

31. Godfrey, L. R. et al. A new interpretation of Madagascar's megafaunal decline: the 'Subsistence Shift Hypothesis'. *J. Hum. Evol.* **130**, 126–140 (2019).
32. Hixon, S. W. et al. Cutmarked bone of drought-tolerant extinct megafauna deposited with traces of fire, human foraging, and introduced animals in SW Madagascar. *Sci. Rep.* **12**, 18504 (2022).
33. Faina, P. et al. Comparing the paleoclimates of northwestern and southwestern Madagascar during the late Holocene: implications for the role of climate in megafaunal extinction. *Malagasy Nat.* **15**, 108–127 (2021).
34. Dos Reis, M. et al. Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. *Syst. Biol.* **67**, 594–615 (2018).
35. Britt, A., Axel, A. & Young, R. Brief surveys of two classified forests in Toamasina Province, eastern Madagascar. *Lemur News* **4**, 25–27 (1999).
36. Hawkins, M. T. R. et al. Nuclear and mitochondrial phylogenomics of the sifakas reveal cryptic variation in the diademed sifaka. *Genes* **13**, 1026 (2022).
37. Som, A. Causes, consequences and solutions of phylogenetic incongruence. *Brief. Bioinform.* **16**, 536–548 (2015).
38. Mittermeier, R. A. et al. Lemur diversity in Madagascar. *Int. J. Primatol.* **29**, 1607–1656 (2008).
39. Everson, K. M., Donohue, M. E. & Weisrock, D. W. A pervasive history of gene flow in Madagascar's true lemurs (genus *Eulemur*). *Genes* **14**, 1130 (2023).
40. Schrider, D. R., Shanku, A. G. & Kern, A. D. Effects of linked selective sweeps on demographic inference and model selection. *Genetics* **204**, 1207–1223 (2016).
41. Orkin, J. D. et al. Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. *ISME J.* **13**, 183–196 (2019).
42. Fan, P. F., Ni, Q. Y., Sun, G. Z., Huang, B. & Jiang, X. L. Gibbons under seasonal stress: the diet of the black crested gibbon (*Nomascus concolor*) on Mt. Wuliang, Central Yunnan, China. *Primates* **50**, 37–44 (2009).
43. Orkin, J. D. et al. The genomics of ecological flexibility, large brains, and long lives in capuchin monkeys revealed with fecalFACS. *Proc. Natl Acad. Sci. USA* **118**, e2010632118 (2021).
44. Campos, F. A. et al. Differential impact of severe drought on infant mortality in two sympatric neotropical primates. *R. Soc. Open Sci.* **7**, 200302 (2020).
45. Burney, D. A., Robinson, G. S. & Burney, L. P. Sporormiella and the late Holocene extinctions in Madagascar. *Proc. Natl Acad. Sci. USA* **100**, 10800–10805 (2003).
46. Sullivan, A. P. et al. Potential evolutionary body size reduction in a Malagasy primate (*Propithecus verreauxi*) in response to human size-selective hunting pressure. *Am. J. Biol. Anthropol.* **178**, 385–398 (2022).
47. Robinson, J. A. et al. Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci. Adv.* **5**, eaau0757 (2019).
48. Fontsere, C. et al. Population dynamics and genetic connectivity in recent chimpanzee history. *Cell Genom.* **2**, 100133 (2022).
49. Alvarez-Estape, M. et al. Past connectivity but recent inbreeding in Cross River gorillas determined using whole genomes from single hairs. *Genes* **14**, 743 (2023).
50. van der Valk, T. et al. The genome of the endangered Dryas monkey provides new insights into the evolutionary history of the vervets. *Mol. Biol. Evol.* **37**, 183–194 (2020).
51. Thompson, E. A. Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics* **194**, 301–326 (2013).
52. Palmada-Flores, M. et al. A high-quality, long-read genome assembly of the endangered ring-tailed lemur (*Lemur catta*). *Gigascience* **11**, giac026 (2022).
53. Larsen, P. A. et al. Hybrid de novo genome assembly and centromere characterization of the gray mouse lemur (*Microcebus murinus*). *BMC Biol.* **15**, 110 (2017).
54. Shao, Y. et al. Phylogenomic analyses provide insights into primate evolution. *Science* **380**, 913–924 (2023).
55. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
56. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
57. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics* **34**, 867–868 (2018).
59. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
60. Korneliussen, T. S. & Moltke, I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics* **31**, 4009–4011 (2015).
61. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
62. Jacquard, A. *The Genetic Structure of Populations* (Springer Science & Business Media, 2012).
63. Faircloth, B. C. et al. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).
64. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
65. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).
66. Mai, U. & Mirarab, S. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* **19**, 272 (2018).
67. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
68. Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).
69. IUCN. The IUCN Red List of Threatened Species, version 2022-2. https://www.iucnredlist.org (2022).
70. Vieilledent, G. et al. Combining global tree cover loss data with historical national forest cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. *Biol. Conserv.* **222**, 189–197 (2018).
71. Townsend, T. M., Vieites, D. R., Glaw, F. & Vences, M. Testing species-level diversification hypotheses in Madagascar: the case of microendemic Brookesia leaf chameleons. *Syst. Biol.* **58**, 641–656 (2009).
72. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods. Ecol. Evol.* **3**, 217–223 (2012).
73. Pierron, D. et al. Genomic landscape of human diversity across Madagascar. *Proc. Natl Acad. Sci. USA* **114**, E6498–E6506 (2017).

Acknowledgements

We acknowledge the Duke Lemur Center and the European Association of Zoos and Aquaria (EAZA) Biobank for collecting primate samples. This is Duke Lemur Center publication no. 1593. We thank D. Pierron for access to the Malagasy human population data. J.D.O. was supported by the 'La Caixa' Foundation (ID 100010434) and the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie

grant agreement no. 847648 (LCF/BQ/PI20/11760004) and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2023-04399, DGECR-2023-00272). Nous remercions le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) de son soutien. T.M.B. is supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 864203), PID2021-126004NB-100 (MICIIN/FEDER, UE) and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2021 SGR 00177). L.F.K.K. was supported by an EMBO STF 8286. M.C.J. was supported by the Natural Environment Research Council (NE/T000341/1). This research was supported by the National Science Foundation to L.P. (BCS 1926105) and M.E.B. (BCS 1926215). P.B. was supported by the Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) from the Centre National de la Recherche Scientifique (DEFI X-LiFE grant 'unknown living branches'). The research reported in this article was funded by the Vietnamese Ministry of Science and Technology's Program 562 (grant no. ĐTĐL.CN-64/19) to M.D.L. This project also received internal support from the Baylor College of Medicine. This research was enabled in part by support provided by Calcul Quebec (www.calculquebec.ca) and the Digital Research Alliance of Canada (https://alliancecan.ca).

Author contributions

J.D.O., L.F.K.K., N.H.-A., C.F., M.L.A. and M.C.J. conducted the analysis. J.D.O. wrote the paper with comments and edits from all authors. N.A., P.B., M.E.B., J.-L.F., K.K.-H.F., R.A.H., J.E.H., I.G.G., M.G., M.W.H., C.K., A.C.K., M.D.L., E.L., T.M.B., S.M., T.N., G.H.P., L.P., C.J.R., L.R., M.R., J.R., C.R., D.D.W., A.Z., G.Z. and D.Z. provided primate samples, sequencing data and laboratory resources. T.M.B., J.R. and K.K.-H.F. supervised the project.

Competing interests

L.F.K.K. and K.K.-H.F. are employees of Illumina as of the submission of the paper. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-024-02596-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-024-02596-1.

Correspondence and requests for materials should be addressed to Joseph D. Orkin or Tomas Marques Bonet.

Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

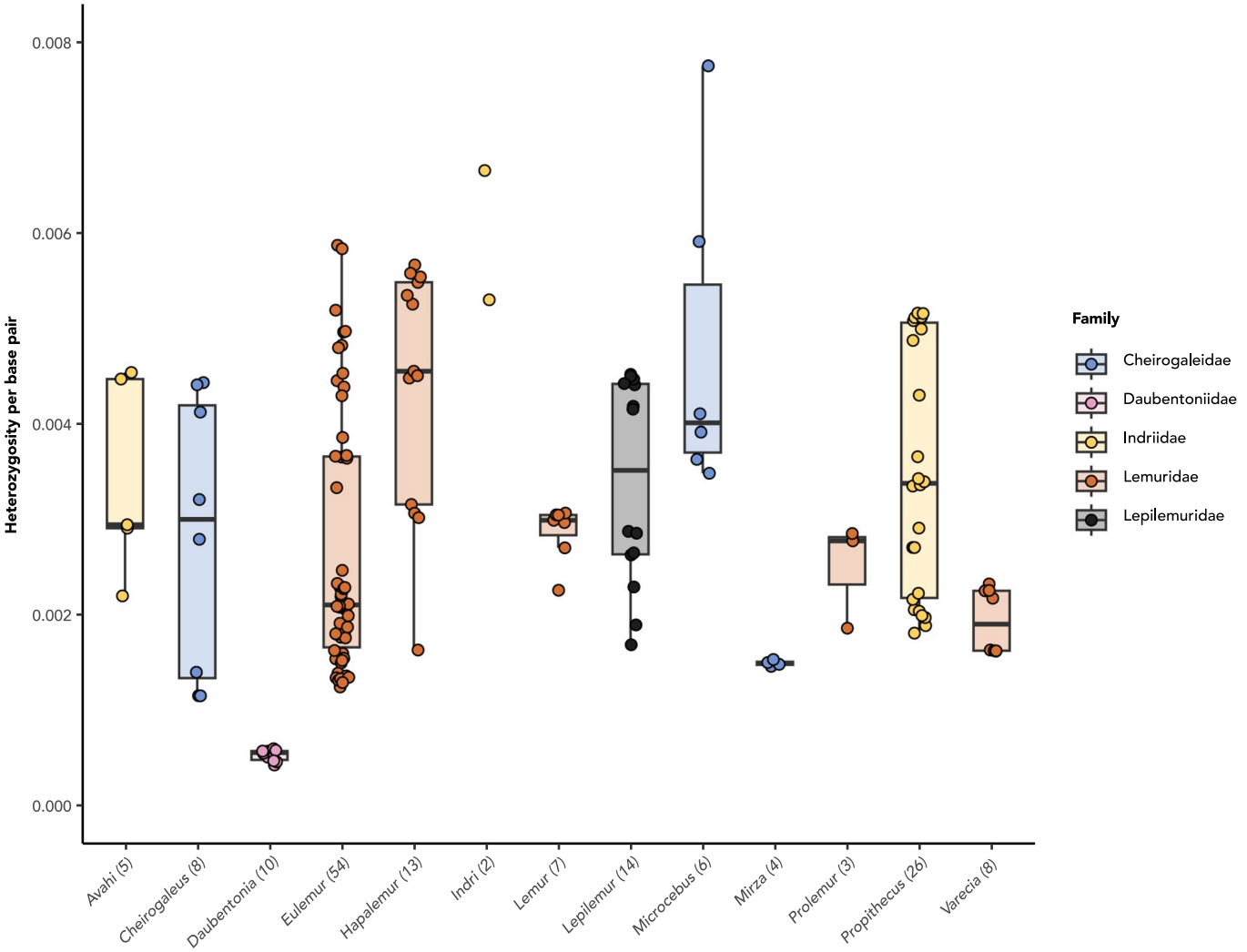
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Joseph D. Orkin[1,2,3] ✉, Lukas F. K. Kuderna[3,4], Núria Hermosilla-Albala[3], Claudia Fontsere[3,5], Megan L. Aylward[6], Mareike C. Janiak[7], Nicole Andriaholinirina[8], Patricia Balaresque[9], Mary E. Blair[10], Jean-Luc Fausser[11], Ivo Glynne Gut[12], Marta Gut[12], Matthew W. Hahn[13], R. Alan Harris[14], Julie E. Horvath[15,16,17,18,19], Christine Keyser[11], Andrew C. Kitchener[20,21], Minh D. Le[22], Esther Lizano[3,23,24], Stefan Merker[25], Tilo Nadler[26], George H. Perry[27,28], Clément J. Rabarivola[8,29], Linett Rasmussen[30], Muthuswamy Raveendran[14], Christian Roos[31], Dong Dong Wu[32], Alphonse Zaramody[8], Guojie Zhang[32,33,34,35,36], Dietmar Zinner[37,38,39], Luca Pozzi[40], Jeffrey Rogers[14], Kyle Kai-How Farh[4] & Tomas Marques Bonet[3,24,41,42] ✉

[1]Département d'anthropologie, Université de Montréal, Montréal, Québec, Canada. [2]Département de sciences biologiques, Université de Montréal, Montréal, Québec, Canada. [3]IBE, Institute of Evolutionary Biology (UPF–CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain. [4]Illumina Artificial Intelligence Laboratory, Illumina Inc, Foster City, CA, USA. [5]Center for Evolutionary Hologenomics, The Globe Institute, University of Copenhagen, Copenhagen, Denmark. [6]Department of Field and Conservation Science, Bristol Zoological Society, Bristol, UK. [7]School of Science, Engineering & Environment, University of Salford, Salford, UK. [8]Life Sciences and Environment, Technology and Environment of Mahajanga, University of Mahajanga, Mahajanga, Madagascar. [9]Centre de Recherche sur la Biodiversité et l'Environnement, CNRS UMR5300, Université Toulouse III, Université de Toulouse, CNRS IRD, Toulouse, France. [10]Center for Biodiversity and Conservation, American Museum of Natural History, New York, NY, USA. [11]Institut de Médecine Légale, Faculté de Médecine, Université de Strasbourg, Strasbourg, France. [12]Centro Nacional de Analisis Genomico (CNAG), Barcelona, Spain. [13]Department of Biology and Department of Computer Science, Indiana University, Bloomington, IN, USA. [14]Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. [15]Research & Collections, North Carolina Museum of Natural Sciences, Raleigh, NC, USA. [16]Department of Biological and Biomedical Sciences, North Carolina Central University, Durham, NC, USA. [17]Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [18]Department of Biological Sciences, North Carolina State University, Raleigh, NC, USA. [19]Department of Evolutionary Anthropology, Duke University, Durham, NC, USA. [20]Department of Natural Sciences, National Museums Scotland, Edinburgh, UK. [21]UK and School of Geosciences, University of Edinburgh, Edinburgh, UK. [22]Department of Environmental Ecology, Faculty of Environmental Sciences, University of Science and Central Institute for Natural Resources and Environmental Studies, Vietnam National University, Hanoi, Vietnam. [23]Unidad de Paleobiología, ICP-CERCA, Unidad Asociada al CSIC por el IBE UPF-CSIC, Cerdanyola del Vallès, Spain. [24]Institut Català de Paleontologia Miquel Crusafont (ICP-CERCA), Universitat Autònoma de Barcelona, Edifici ICTA-ICP, Cerdanyola del Vallès, Spain. [25]Department of Zoology, State Museum of Natural History Stuttgart, Stuttgart, Germany. [26]Cuc Phuong Commune, Ninh Binh Province, Vietnam. [27]Departments of Anthropology and Biology, Pennsylvania State University, University Park, PA, USA. [28]Huck Institutes of the Life Sciences, Pennsylvania

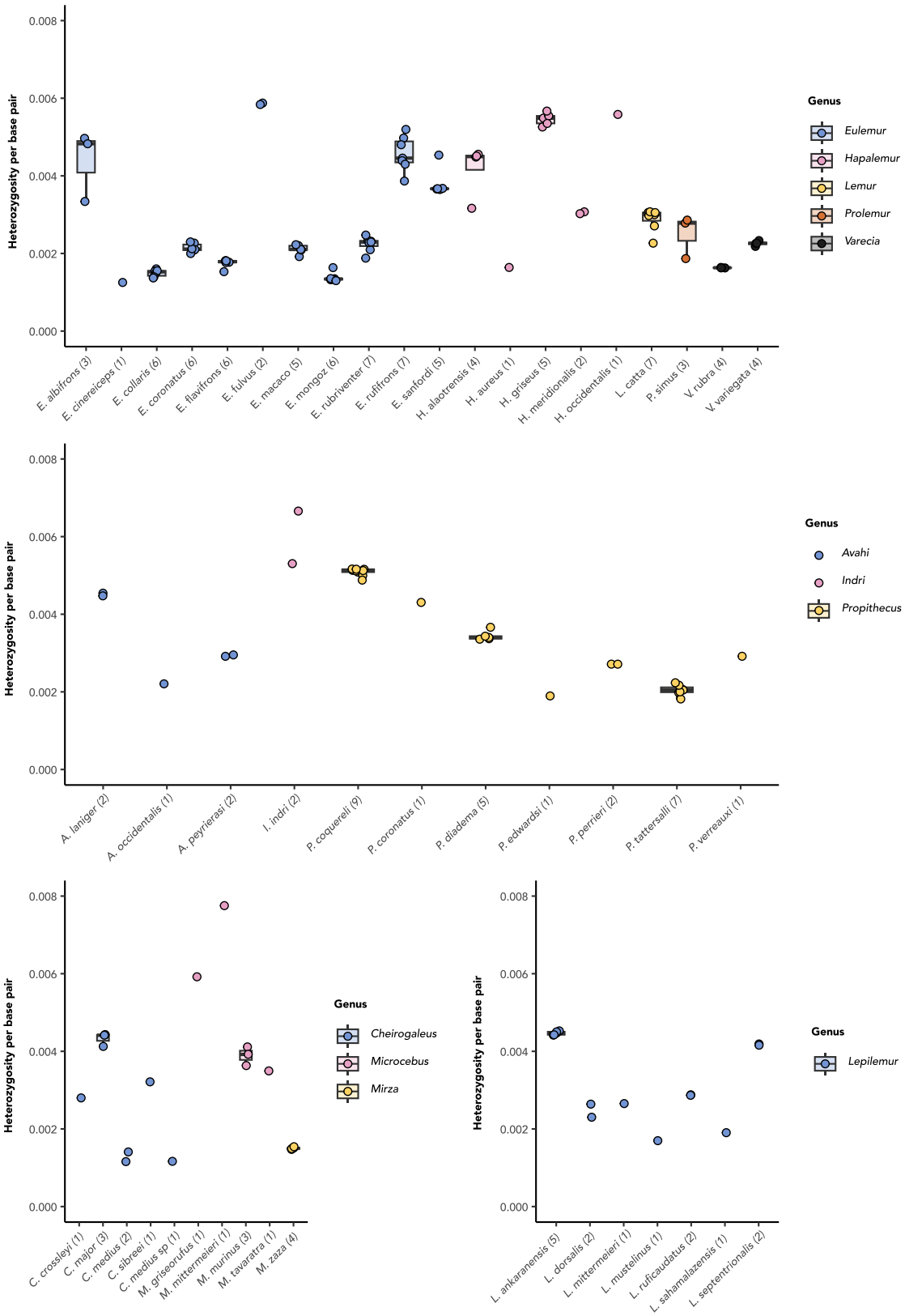
State University, University Park, PA, USA. [29]Université de l'Itasy, Antananarivo, Madagascar. [30]Copenhagen Zoo, Frederiksberg, Denmark. [31]Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany. [32]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China. [33]Center for Evolutionary and Organismal Biology, Zhejiang University School of Medicine, Hangzhou, China. [34]Villum Centre for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark. [35]Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou, China. [36]Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, China. [37]Cognitive Ethology Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany. [38]Department of Primate Cognition, Georg-August-University, Göttingen, Germany. [39]Leibniz-ScienceCampus Primate Cognition, Göttingen, Germany. [40]Department of Anthropology, University of Texas San Antonio, San Antonio, TX, USA. [41]CNAG–Centre for Genomic Analyses, Barcelona, Spain. [42]Institució Catalana de Recerca i Estudis Avançats (ICREA) and Universitat Pompeu Fabra, Barcelona, Spain. ✉e-mail: joseph.orkin@umontreal.ca; tomas.marques@upf.edu

https://doi.org/10.1038/s41559-024-02596-1



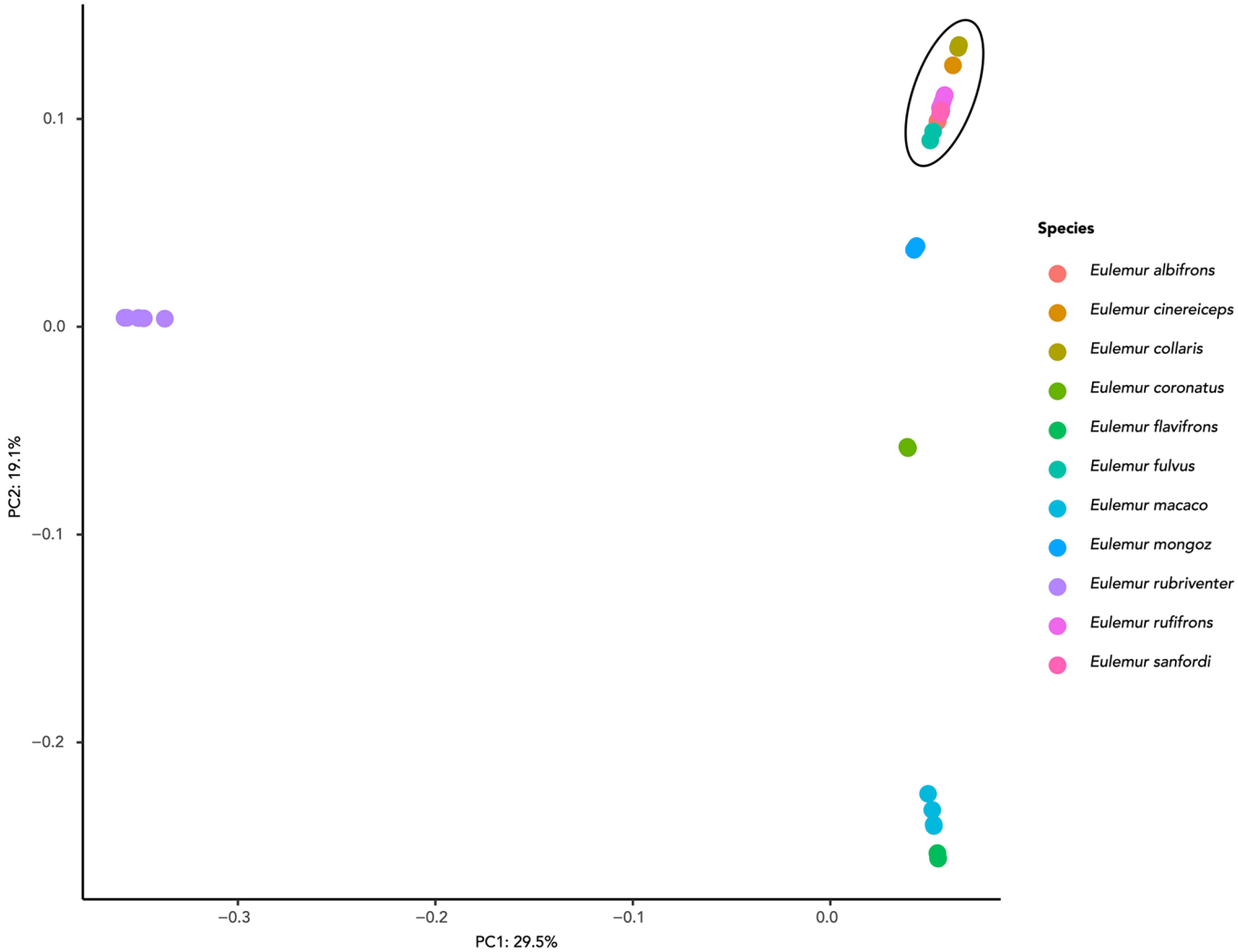
Extended Data Fig. 1 | Individual heterozygosity levels per genus. Heterozygosity calculated on the base-pair level per genus. Each point represents a unique individual (biological replicate), and boxplots are displayed for genera with at least three samples. The box corresponds to the interquartile range (IQR), the horizontal line is the median value, and the length of whiskers extends up to 1.5 times the IQR. The number of individuals per genus (n) is listed in parentheses following the taxonomic group on the X-axis.

https://doi.org/10.1038/s41559-024-02596-1



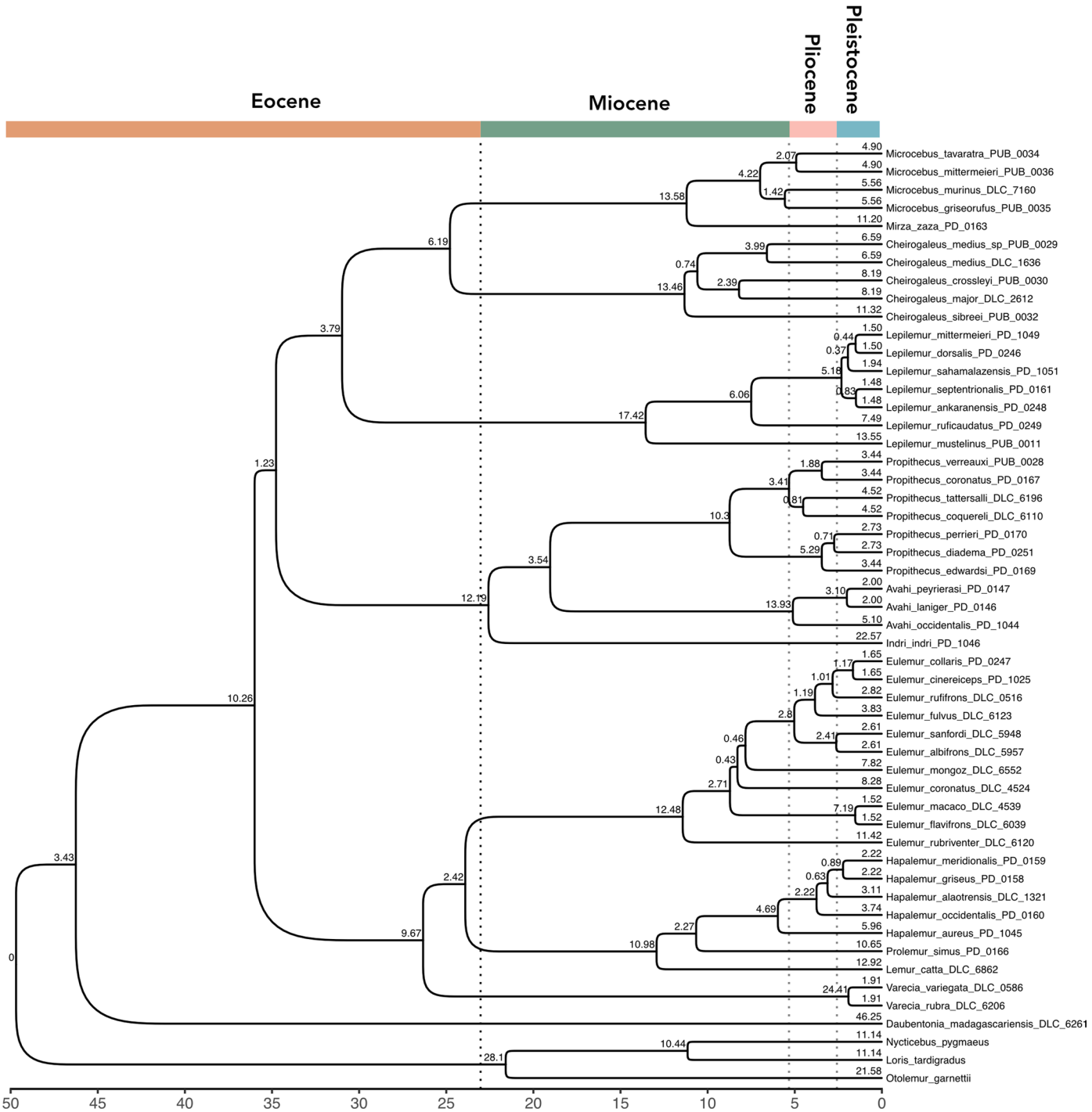
Extended Data Fig. 2 | Individual heterozygosity levels per species. Heterozygosity calculated on the base-pair level per species organized by family: **Top**) Lemuridae, **Middle**) Indriidae, **Bottom Left**) Cheirogaleidae, **Bottom Right**) Lepilemuridae. Each point represents a unique individual (biological replicate), and boxplots are displayed for species with at least three samples. The box corresponds to the interquartile range (IQR), the horizontal line is the median value, and the length of whiskers extends up to 1.5 times the IQR. The number of individuals per species (n) is listed in parentheses following the taxonomic group on the X-axis.

https://doi.org/10.1038/s41559-024-02596-1



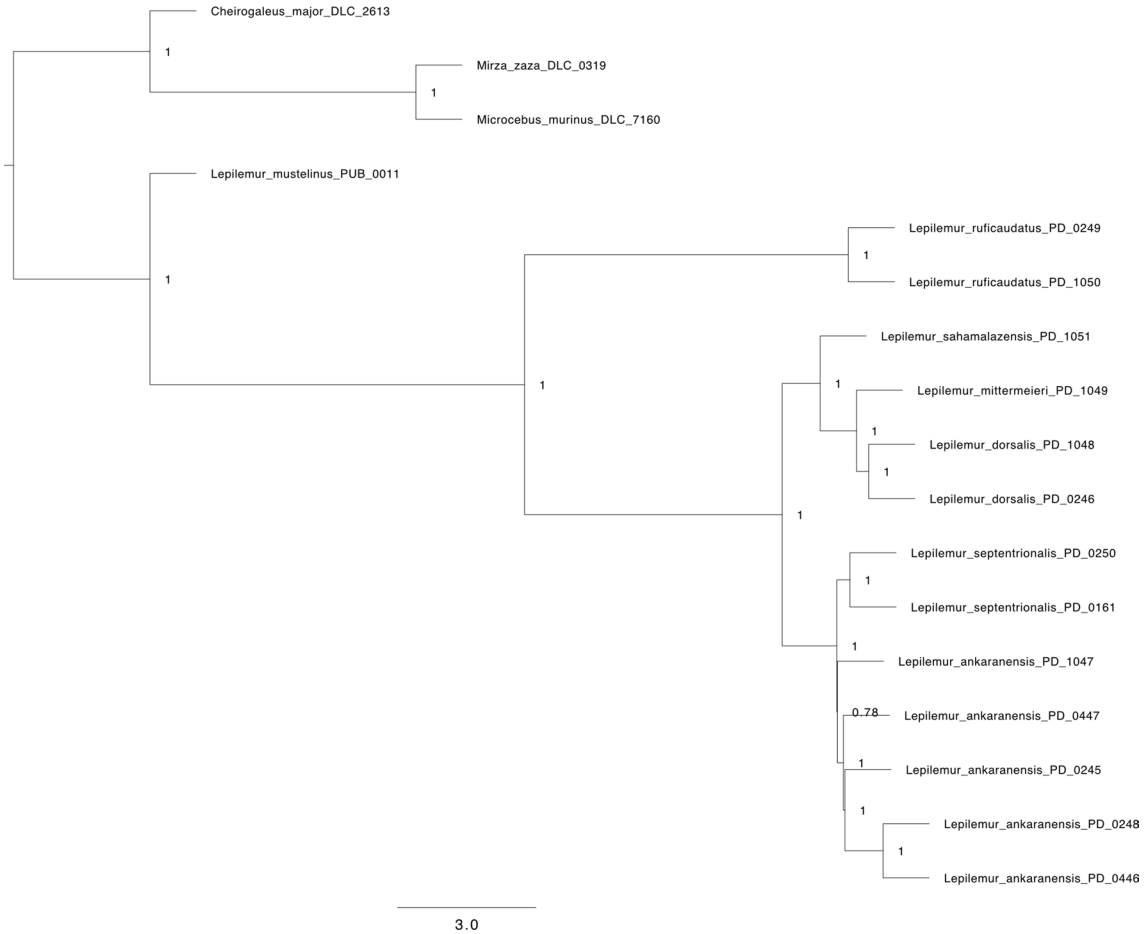
Extended Data Fig. 3 | Principal components analysis of all *Eulemur* individuals. PCA of *Eulemur* genus depicting fulvus and non-fulvus groups. Members of the *Eulemur* fulvus group are enclosed in the oval.

https://doi.org/10.1038/s41559-024-02596-1



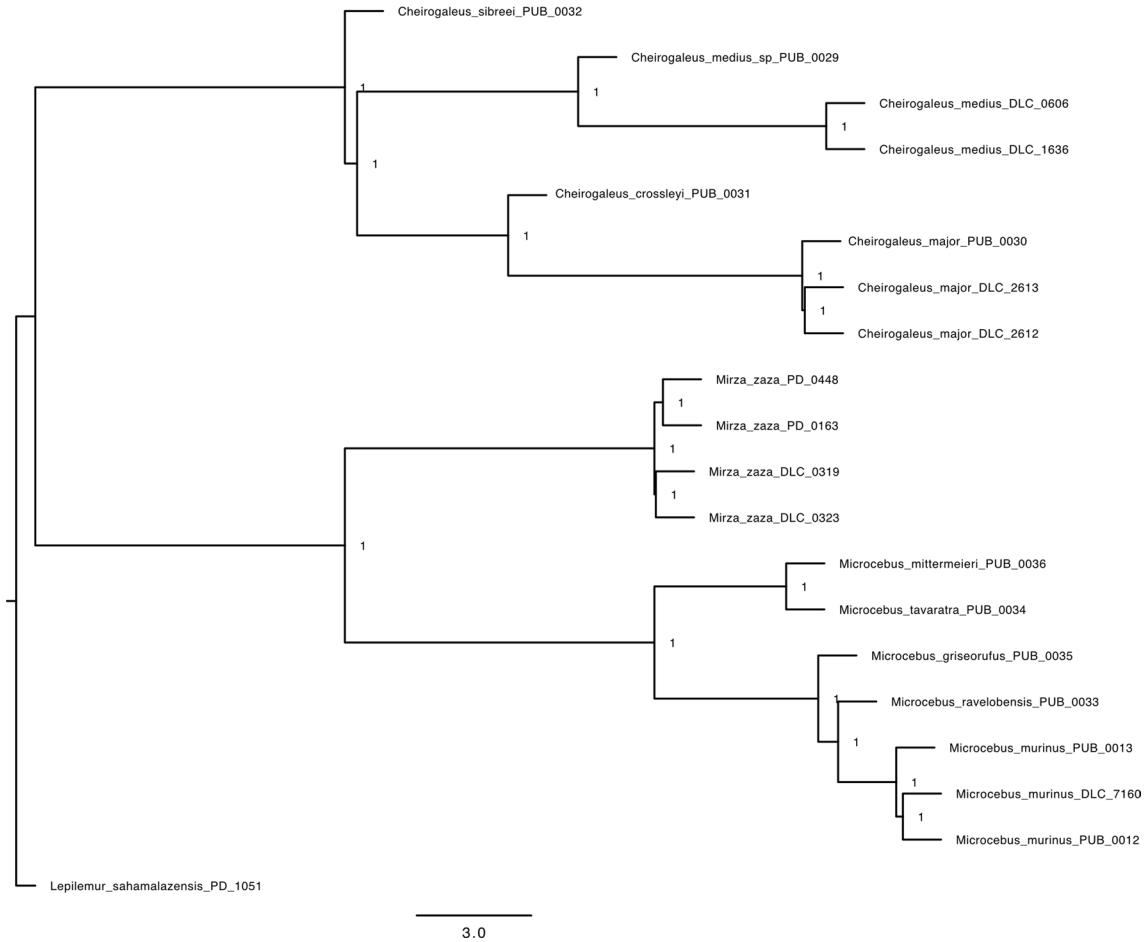
Extended Data Fig. 4 | Dated UCE phylogenetic tree of the Lemuriformes. Dates since split with sister taxon are listed as branch lengths in millions of years. Local posterior probability of all nodes is 1, with the exception of the node placing *E. coronatus* outside the clade containing *E. mongoz* and the *E. fulvus* species complex (local PP = 0.79).

https://doi.org/10.1038/s41559-024-02596-1



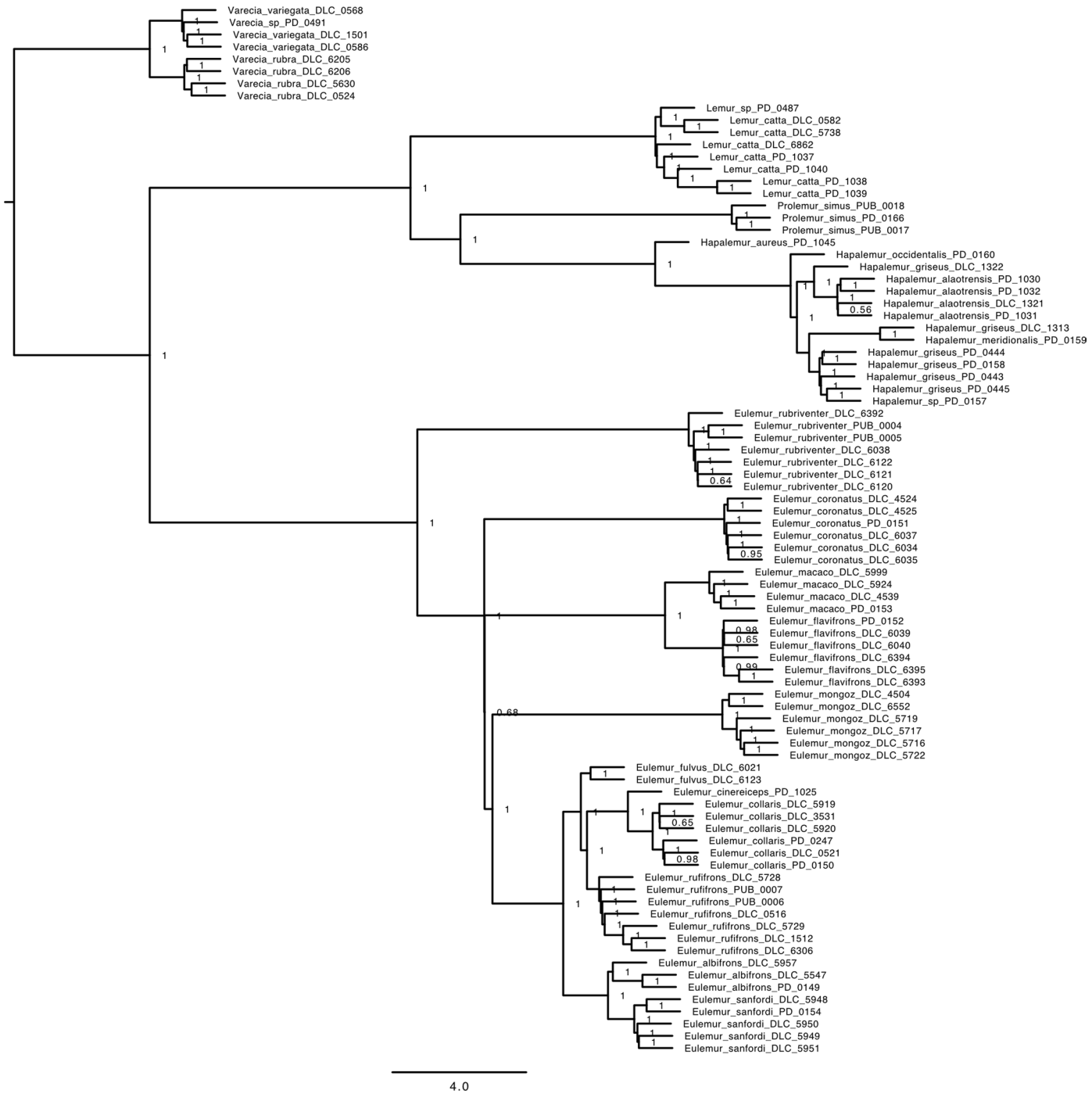
Extended Data Fig. 5 | Phylogenetic tree of the Lepilemuridae. 5000 10 kb window ASTRAL species tree for the Lepilemuridae. Branch node labels are ASTRAL local posterior probabilities.

https://doi.org/10.1038/s41559-024-02596-1



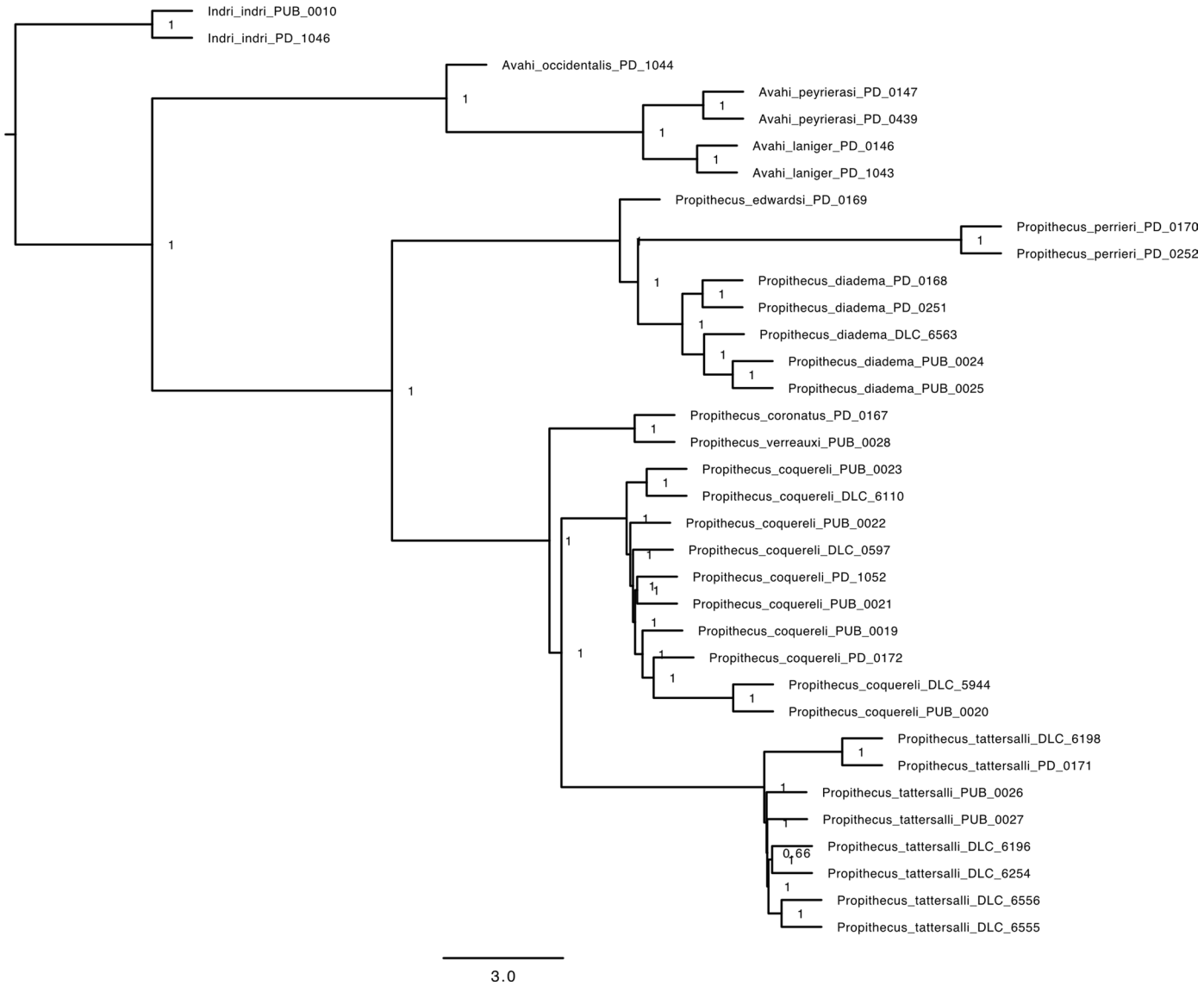
Extended Data Fig. 6 | Phylogenetic tree of the Cheirogaleidae. 5000 10 kb window ASTRAL species tree for the Cheirogaleidae. Branch node labels are ASTRAL local posterior probabilities.

https://doi.org/10.1038/s41559-024-02596-1



Extended Data Fig. 7 | Phylogenetic tree of the Lemuridae. 5000 10 kb window ASTRAL species tree for the Lemuridae. Branch node labels are ASTRAL local posterior probabilities.

https://doi.org/10.1038/s41559-024-02596-1



Extended Data Fig. 8 | Phylogenetic tree of the Indriidae. 5000 10 kb window ASTRAL species tree for the Indriidae. Branch node labels are ASTRAL local posterior probabilities.

nature portfolio

Corresponding author(s): Joseph Orkin, Tomas Marques Bonet

Last updated by author(s): Oct 31, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect data
Data analysis	ASTRAL 5.7.8, BCFtools 1.14, Bedtools 2.16, BWA 0.7.15, Dsuite v0.5, GONE, GATK 4.1.7.0, GBLOCKS 0.91, IQTree2, mafft, MosDepth 0.2.6, NGSrelate 2.0, PLINK 1.90, Phyluce 1.7.2, Phytools 1.5.1, picard 1.60, R 4.2.2, samtools 1.12 , SMC++ 1.15.2, Treeshrink, Vcftools 0.1.12

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All newly generated sequencing data is deposited in the European Nucleotide Archive (PRJEB77609) and NCBI (PRNJA1156176). Additional accession numbers for previously published sequencing data to allow for a minimally reproducible dataset are available in Table S1.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	No human participants, their data, or biological material were involved in this research
Reporting on race, ethnicity, or other socially relevant groupings	No human participants, their data, or biological material were involved in this research
Population characteristics	No human participants, their data, or biological material were involved in this research
Recruitment	No human participants, their data, or biological material were involved in this research
Ethics oversight	No human participants, their data, or biological material were involved in this research

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Our samples include a total of 162 unique individuals and 50 species of Lemuriformes primates.
Data exclusions	Several samples were excluded from the analysis based on quality control failures (inadequate quality/quantity of DNA for sequencing, low average genome coverage, etc.). First and second degree relatives were excluded from some analyses as described in the manuscript.
Replication	Replication was not relevant to this study, because each primate genome we used originates from an independent individual
Randomization	Randomization was not relevant to this study, because these no experimental conditions were used
Blinding	Blinding was not relevant to this study, because these no experimental conditions were used

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	One genome sequence was derived from a fibroblast cell line of a male lemur catta individual.

Authentication	The fibroblast cell line was generated at Universitat Pompeu Fabra from a skin biopsy of a known animal.
Mycoplasma contamination	The cell line was screened for mycoplasma using a MycoSPY® - RCR Mycoplasma Test Kit from Biontex.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Plants

Seed stocks	No plant materials were used in this study.
Novel plant genotypes	No plant materials were used in this study.
Authentication	No plant materials were used in this study.