

Title: A global catalog of whole-genome diversity from 233 primate species

Authors:

Lukas F.K. Kuderna^{1, †, ‡}, Hong Gao², Mareike C. Janiak³, Martin Kuhlwilm^{1,4,5}, Joseph D. Orkin^{1,6}, Thomas Bataillon⁷, Shivakumara Manu^{8,9}, Alejandro Valenzuela¹, Juraj Bergman^{7,10}, Marjolaine Rouselle⁷, Felipe Ennes Silva^{11,12}, Lidia Agueda¹³, Julie Blanc¹³, Marta Gut¹³, Dorien de Vries³, Ian Goodhead³, R. Alan Harris¹⁴, Muthuswamy Raveendran¹⁴, Axel Jensen¹⁵, Idriss S. Chuma¹⁶, Julie Horvath^{17,18,19,20}, Christina Hvilson²¹, David Juan¹, Peter Frandsen²¹, Joshua G. Schraiber², Fabiano R. de Melo²², Fabricio Bertuol²³, Hazel Byrne²⁴, Iracilda Sampaio²⁵, Izeni Farias²⁶, João Valsecchi do Amaral^{26,27,28}, Malu Messias²⁹, Maria N. F. da Silva³⁰, Mihir Trivedi⁹, Rogerio Rossi³¹, Tomas Hrbek^{23,32}, Nicole Andriaholinirina³³, Clément J. Rabarivola³³, Alphonse Zaramody³³, Clifford J. Jolly³⁴, Jane Phillips-Conroy³⁵, Gregory Wilkerson³⁶, Christian Abee³⁶, Joe H. Simmons³⁶, Eduardo Fernandez-Duque³⁷, Sree Kanthaswamy³⁸, Fekadu Shiferaw³⁹, Dongdong Wu⁴⁰, Long Zhou⁴¹, Yong Shao⁴⁰, Guojie Zhang^{42,43,40,41}, Julius D. Keyyu⁴⁴, Sascha Knauf⁴⁵, Minh D. Le⁴⁶, Esther Lizano^{1,47}, Stefan Merker⁴⁸, Arcadi Navarro^{1,49,50,51}, Tilo Nadler⁵², Chiea Chuen Khor⁵³, Jessica Lee⁵⁴, Patrick Tan^{55,56,53}, Weng Khong Lim^{56,55,57}, Andrew C. Kitchener⁵⁸, Dietmar Zinner^{59,60,61}, Ivo Gut¹³, Amanda Melin^{62,63}, Katerina Guschanski^{64,15}, Mikkel Heide Schierup⁷, Robin M. D. Beck³, Govindhaswamy Umapathy^{8,9}, Christian Roos⁶⁵, Jean P. Boubli³, Jeffrey Rogers^{14,†}, Kyle Farh^{2,†}, Tomas Marques Bonet^{1,47,13,49,†}

Affiliations:

¹Institute of Evolutionary Biology (UPF-CSIC), PRBB, Dr. Aiguader 88, 08003 Barcelona, Spain

²Illumina, Inc., CA, USA

³School of Science, Engineering & Environment, University of Salford, Salford M5 4WT, United Kingdom

⁴Department of Evolutionary Anthropology, University of Vienna, Djerassiplatz 1, 1030 Vienna, Austria

⁵Human Evolution and Archaeological Sciences (HEAS), University of Vienna, Austria

⁶Département d'anthropologie, Université de Montréal, 3150 Jean-Brillant, Montréal, QC, H3T 1N8, Canada

⁷Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark

⁸Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India

⁹Laboratory for the Conservation of Endangered Species, CSIR-Centre for Cellular and Molecular Biology, Hyderabad 500007, India

¹⁰Section for Ecoinformatics & Biodiversity, Department of Biology, Aarhus University, Aarhus, Denmark

¹¹Research Group on Primate Biology and Conservation, Mamirauá Institute for Sustainable Development, Estrada da Bexiga 2584, CEP 69553-225, Tefé, Amazonas, Brazil

5 ¹²Faculty of Sciences, Department of Organismal Biology, Unit of Evolutionary Biology and Ecology, Université Libre de Bruxelles (ULB), Avenue Franklin D. Roosevelt 50, 1050, Brussels, Belgium

¹³CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain

10 ¹⁴Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

¹⁵Department of Ecology and Genetics, Animal Ecology, Uppsala University, SE-75236 Uppsala, Sweden

¹⁶Tanzania National Parks, Arusha, Tanzania

15 ¹⁷North Carolina Museum of Natural Sciences, Raleigh, NC, 27601, USA

¹⁸Department of Biological and Biomedical Sciences, North Carolina Central University, Durham, NC, 27707, USA

¹⁹Department of Biological Sciences, North Carolina State University, Raleigh, NC, 27695, USA

20 ²⁰Department of Evolutionary Anthropology, Duke University, Durham, NC, 27708, USA

²¹Copenhagen Zoo, 2000 Frederiksberg, Denmark

²²Universidade Federal de Viçosa, Brazil

²³Universidade Federal do Amazonas, Departamento de Genética, Laboratório de Evolução Genética Animal (LEGAL), Manaus, Amazonas 69080-900, Brazil

25 ²⁴Department of Anthropology, University of Utah, Salt Lake City UT 84102, USA

²⁵Universidade Federal do Para, Brazil

²⁶Research Group on Terrestrial Vertebrate Ecology, Mamirauá Institute for Sustainable Development, Tefé, Amazonas, Brazil

30 ²⁷Rede de Pesquisa para Estudos sobre Diversidade, Conservação e Uso da Fauna na Amazônia – RedeFauna, Manaus, Amazonas, Brazil

²⁸Comunidad de Manejo de Fauna Silvestre en la Amazonía y en Latinoamérica – ComFauna, Iquitos, Loreto, Peru

²⁹Universidade Federal de Rondonia, Brazil

³⁰Instituto Nacional de Pesquisas da Amazonia, Brazil

35 ³¹Universidade Federal do Mato Grosso, Brazil

³²Department of Biology, Trinity University, San Antonio, TX 78212, USA

³³Life Sciences and Environment, Technology and Environment of Mahajanga, University of Mahajanga, Mahajanga, Madagascar

³⁴New York University, NY, USA

³⁵Washington University, MO, USA

³⁶Keeling Center for Comparative Medicine and Research, MD Anderson Cancer Center

³⁷Yale University, CT, USA

5 ³⁸Arizona State University, AZ, USA

³⁹Hawassa University, Hawassa, Ethiopia

⁴⁰State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

10 ⁴¹Evolutionary & Organismal Biology Research Center, Zhejiang University School of Medicine, Hangzhou, 310058, China

⁴²Villum Center for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Denmark

⁴³BGI-Shenzhen, Shenzhen 518083, China

15 ⁴⁴Tanzania Wildlife Research Institute (TAWIRI), Head Office, P.O.Box 661, Arusha, Tanzania

⁴⁵Institute of International Animal Health/One Health, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, 17493 Greifswald - Isle of Riems, Germany

20 ⁴⁶Department of Environmental Ecology, Faculty of Environmental Sciences, University of Science and Central Institute for Natural Resources and Environmental Studies, Vietnam National University, Hanoi

⁴⁷Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Barcelona, SpainCatalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain

⁴⁸Department of Zoology, State Museum of Natural History Stuttgart, Germany

25 ⁴⁹Institució Catalana de Recerca i Estudis Avançats (ICREA) and Universitat Pompeu Fabra. Pg. Luís Companys 23, 08010, Barcelona, Spain

⁵⁰Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Av. Doctor Aiguader, N88, 08003 Barcelona, Spain

30 ⁵¹BarcelonaBeta Brain Research Center, Pasqual Maragall Foundation, C. Wellington 30, 08005, Barcelona, Spain

⁵²Cuc Phuong Commune, Nho Quan District, Ninh Binh Province, Vietnam

⁵³Genome Institute of Singapore, Singapore

⁵⁴Mandai Nature, 80 Mandai Lake Road, Singapore 729826

⁵⁵SingHealth Duke-NUS Institute of Precision Medicine (PRISM), Singapore

35 ⁵⁶Cancer and Stem Cell Biology Program, Duke-NUS Medical School, Singapore

⁵⁷SingHealth Duke-NUS Genomic Medicine Centre, Singapore

⁵⁸Department of Natural Sciences, National Museums Scotland, Chambers Street, Edinburgh EH1 1JF, UK and School of Geosciences, Drummond Street, Edinburgh EH8 9XP

⁵⁹Cognitive Ethology Laboratory, Germany Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany

5 ⁶⁰Department of Primate Cognition, Georg-August-Universität Göttingen, 37077 Göttingen, Germany

⁶¹Leibniz ScienceCampus Primate Cognition, 37077 Göttingen, Germany

⁶²Department of Anthropology & Archaeology and Department of Medical Genetics

10 ⁶³Alberta Children's Hospital Research Institute; University of Calgary, 2500 University Dr NW T2N 1N4, Calgary, AB Canada

⁶⁴Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

⁶⁵Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany

15

† Corresponding authors. Email: lkuderna@illumina.com, jr13@bcm.edu, kfarh@illumina.com, tomas.marques@upf.edu

‡ Current affiliation: Illumina Inc.

Abstract: The rich diversity of morphology and behavior displayed across primate species provides an informative context in which to study the impact of genomic diversity on fundamental biological processes. Analysis of that diversity provides insight into long-standing questions in evolutionary and conservation biology, and is urgent given severe threats these species are facing. Here, we present high coverage whole-genome data from 233 primate species representing 86% of genera and all 16 families. This dataset was used, together with fossil calibration, to create a nuclear DNA phylogeny and to reassess evolutionary divergence times among primate clades. We found within-species genetic diversity across families and geographic regions to be associated with climate and sociality, but not with extinction risk. Furthermore, mutation rates differ across species, potentially influenced by effective population sizes. Lastly, we identified extensive recurrence of missense mutations previously thought to be human-specific. This study will open a wide range of research avenues for future primate genomic research.

One-Sentence Summary: The whole genome sequences of 233 primate species provide insight into the determinants of genetic diversity, phylogenomics, and human uniqueness.

Main Text:

The order Primates includes over 500 recognized species that display a remarkable array of morphological, physiological, and behavioral adaptations (1). Spanning a broad range of social systems, locomotory styles, dietary specializations, and habitat preferences, these species rightly attract attention from scientists with equally diverse research interests. Because humans are members of the order Primates, we also find many important and informative biological parallels between ourselves and other primates. The analysis of nonhuman primate genomes has long been motivated by a desire to understand human evolutionary origins, human health, and disease. However, past comparative genomic analyses have mainly focused on a relatively small number of species (2, 3), thus providing a limited understanding of genome variability in only a few key lineages, such as members of the great apes (4–10), or macaques (11–13). Furthermore, low numbers of wild-born individuals in these studies potentially result in assessments of diversity that may not reflect natural populations (3). To gain a more complete picture of how evolution has shaped genomic variation across primates, large-scale sequencing of many species and individuals is necessary, especially within previously neglected lineages such as strepsirrhines (lemurs, lorises, galagos, and relatives) and platyrrhines (monkeys of the Americas). The need for a more complete understanding of primate genetic diversity in the wild, and its determinants is urgent given the current extinction crisis driven by climate change, habitat loss, and illegal trading and hunting (14). At present, 60% of the world's primate species are threatened with extinction, and current trends are likely to exacerbate the rates of biodiversity loss in the near future (14, 15). The analysis of whole-genome sequences allows estimation of genetic diversity and evaluation of its association with ecological traits, degrees of inbreeding, and phylogenetic relationships, all metrics relevant to primate conservation genomics.

High coverage genome sequences of 233 primate species

We sequenced the genomes of 703 individuals from 211 primate species on the Illumina NovaSeq 6000 platform (16). For 78% of individuals, the available amount of DNA permitted us to generate PCR-free libraries. We sequenced paired-end reads of 151 bp to an average production target of at least 100 Gigabases (Gb), resulting in an average mapped coverage of 32.4 X per individual (15.3 - 77.6 X, see (16)). We expanded our dataset by including 106 individuals representing 29 species from previously published studies to maximize phylogenetic diversity (8, 17–24). Altogether, we compiled data from 809 individuals from 233 primate species, amounting to 47% of the 521 currently recognized species (14). Our sampling covers 86% of primate genera (69), and all 16 families. Over 72% of individuals in this study are wild-born. Furthermore, 58% of species in our dataset are classified as threatened with extinction by the IUCN (i.e., classified in the categories vulnerable (VU), endangered (EN), and critically endangered (CR), and 30 species are critically endangered. It is worth noting that among the species we sampled are some of the world's most endangered primates, which face an extremely high risk of extinction in the wild. Examples include the Western black crested gibbon (*Nomascus concolor*), with an estimated 1500 individuals left in the wild and scattered across an array of discontinuous habitats, and the northern sportive lemur (*Lepilemur septentrionalis*), with roughly 40 individuals estimated to remain in the wild, inhabiting an area potentially as small as 12 km² (25, 26).

For 100 species, we generated sequencing data from more than one individual, and for 36 species from five or more individuals, 29 of which belong to newly sequenced species. We thus gathered broad primate taxonomic coverage by compiling species from all major geographical regions

currently inhabited by primates, including the Americas, mainland Africa, Madagascar, and Asia (Fig. 1A). The data presented here provides the foundation for several additional studies in this issue, informing important and diverse topics including hybrid speciation and reticulation among primates (Sørensen et al.), the role of functional constraint in primate evolution (Rashid et al.), and predicting the landscape of tolerated mutations in the human genome (Gao et al.).

Owing to technical challenges inherent to short-read assembly, we aligned our data to a backbone of 32 reference genomes for further analyses, most of which are derived from long-read sequencing technologies (16). These references are well distributed across the primate phylogeny, and result in a median pairwise distance between the focal and reference species of 6.6×10^{-3} substitutions per site ($0-4.1 \times 10^{-2}$), which is within the range of previous projects using a similar approach (8). To ensure our estimates of genetic diversity over these phylogenetic distances are minimally biased, we compared pairs of diversity estimates where reads from one species were mapped to its own reference as well as mapped to another species reference. Across 19 species pairs that fully cover the phylogenetic distances between focal species and reference in our data, we find heterozygosity estimates to be highly correlated (Pearson's $r = 0.97$, $p = 6.8 \times 10^{-12}$). Overall, we find a median value of 2.4 Gb per individual to be callable across all references, thus enabling genome-wide comparisons.

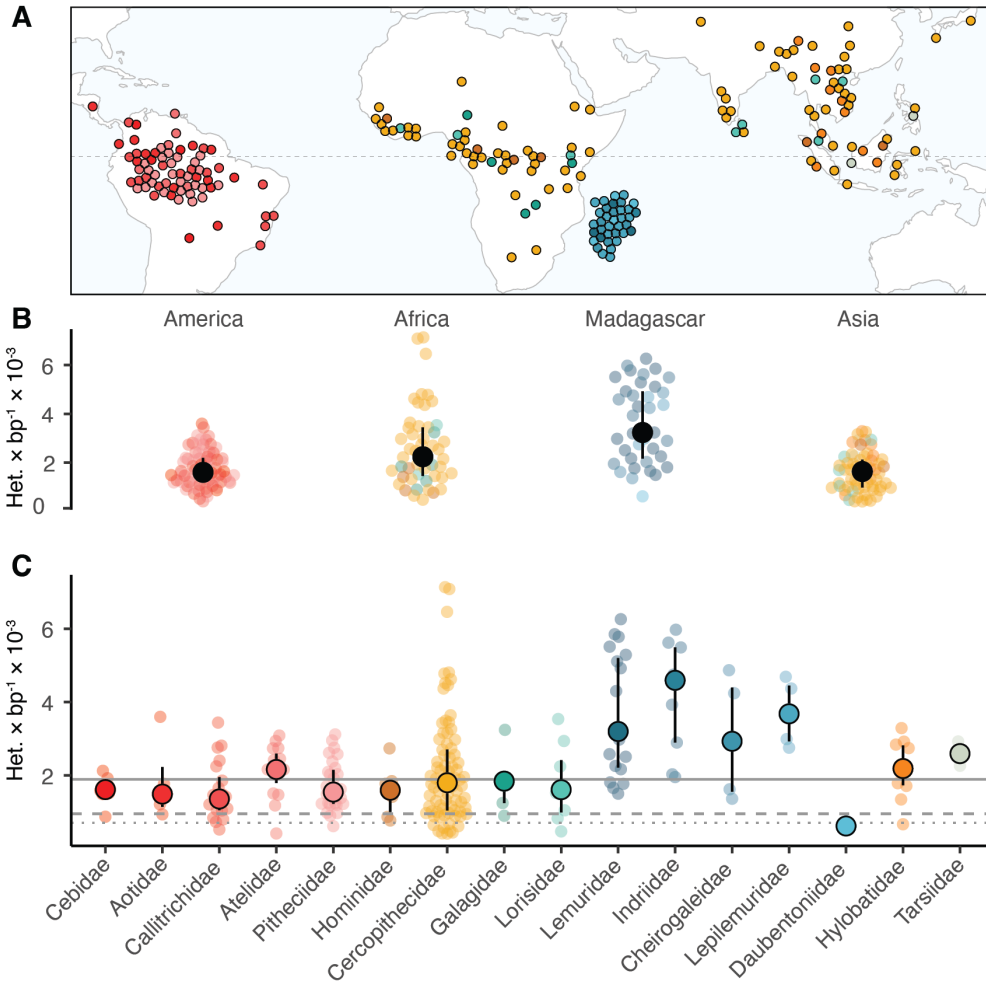


Fig. 1. Genetic diversity in primates across geographic regions and families (A) Sampling range of species analyzed in this project. Each point represents the approximate species range centroid of all sampled species with available ranges. Points are repelled to avoid overplotting. (B) Heterozygosity stratified by geographic region, solid black points and whiskers represent median values and interquartile range. (C) Median species heterozygosity by family. Solid circles and whiskers represent median and interquartile range. Solid gray line denotes primate-wide median heterozygosity, dashed and dotted lines denote human heterozygosity for African, and bottlenecked out-of-Africa populations, respectively. Points are colored according to the family a species belongs to, as denoted on the x-axis of panel C

Genetic diversity across primates

Heterozygosity in primates spans over an order of magnitude, with values ranging from $0.41 \cdot 10^{-3}$ heterozygotes per base pair ($\text{het} \times \text{bp}^{-1}$) to $7.14 \times 10^{-3} \text{ het} \times \text{bp}^{-1}$ (see Fig. 1C). We observe the lowest levels of diversity in the golden snub-nosed monkey (*Rhinopithecus roxellana*) at about one heterozygous position every 2400 bp. Interestingly, only 15 species have lower median genetic

diversity than humans, the primate with by far the largest census size. Among these are several Asian colobines, but also the aye-aye, the western hoolock gibbon, and the Guinea baboon. There are marked differences in genetic diversity across genera, families, and geographic regions, with high-diversity species found among cercopithecines from mainland Africa and lemurs in Madagascar (Fig. 1B). Among cercopithecines, guenons of the genus *Cercopithecus* are almost exclusively responsible for high diversity with a median value of 4.54×10^{-3} het•bp⁻¹, more than double the primate-wide median. Some members of this tribe also show large historical effective population sizes, and there are several known instances of past and present interspecific hybridization (27–30). We further observe high diversity across several genera of lemurs, which are among the most endangered primates, primarily due to rapid habitat loss and severe population decline. Examples include members of the true lemurs (*Eulemur* sp.), bamboo lemurs (*Haplemur* sp.), and sifakas (*Propithecus* sp.).

We investigated whether genetic diversity estimates are correlated with extinction risk in primates, a subject of previous debate (17, 31, 32). Despite our broad sampling, we find no global relationship between numerically coded IUCN extinction risk categories and estimated heterozygosity ($p > 0.05$ PGLS, see Fig. 2A and (16)). Since genetic diversity is strongly determined by long-term demographic history, rapid recent population declines such as those currently experienced by many primate species are unlikely to be detected in a cross-species comparison. Instead, temporal datasets within the same species are better suited to quantify recent changes in genetic diversity (33). Nevertheless, comparing genetic diversity for non-threatened (LC, NT) and threatened (VU, EN, CR) species within the same family consistently uncovers lower diversity among species in the threatened categories for all families with more than one species in both categories, although not all comparisons reach statistical significance ($p < 0.05$, MWU, see Fig. 2B). The only exception is Lorisidae, which showed no difference in genetic diversity between non-threatened and threatened species.

To further assess the potential impact of recent population decline, we analyzed runs of homozygosity (RoH) across species. We focused on tracts with a minimum length of one megabase (Mb), which in humans indicate recent inbreeding (8). The order-wide median fraction of the genome in RoH is 5.1%, and individual values vary substantially, reaching over 50%. We find critically endangered species, such as the white-headed langur (*Trachypithecus leucocephalus*), the eastern gorilla (*Gorilla beringei*), and mongoose lemur (*Eulemur mongoz*) among the species with the highest proportion of RoHs (see Fig. 2C). However, some species not currently classified as threatened, such as Azara's owl monkey (*Aotus azarae*) and the northern greater galago (*Otolemur garnettii*), also have a high fraction of the genome in RoHs. While the overall conservation status of these two species might not be worrisome, some individuals may belong to smaller local populations, which can exacerbate inbreeding. We find 13 critically endangered species with lower than the primate-wide average fractions of their genomes in RoHs, among them the three douc langur species (*Pygathrix cinerea*, *P. nemaus*, *P. nigripes*), red-tailed sportive lemur (*Lepilemur ruficaudatus*), and Verreaux's sifaka (*Propithecus verreauxi*). We find no overall relationship between extinction risk and degree of inbreeding deduced from the total fraction of the genome in RoHs (Pearson's $r = 0.03$, $p = 0.71$). This implies that RoHs are not a good predictor of extinction risk in primates, and suggests that many critically endangered species are threatened by non-genetic factors, likely reflecting population declines that have been too fast to be detectable on the genomic level. Given the potential importance of functional variation to conservation efforts, we sought to quantify the proportion of loss of functional variation in each lineage (32, 34). To this end, we quantified stop-gain and missense mutations and normalized them by the number of synonymous mutations to account for lineage-specific differences in

evolutionary rates. We found inverse relationships between the missense/synonymous ratios (Pearson's $r = -0.35$, $p = 9.3 \times 10^{-8}$) and, to a lesser extent, stop-gain/synonymous ratios and heterozygosity across primates, suggesting effects of purifying selection on deleterious variation, although the latter does not reach statistical significance (Pearson's $r = -0.12$, $p = 0.082$). We do not find deleterious variations as measured by the stop-gain/synonymous ratio to be correlated with extinction risk (Pearson's $r < 0.01$, $p = 0.94$). Nevertheless, we caution that the varying quality of the references and their annotations, together with potential changes in gene structure between the references and analyzed species, might add noise to the comparisons across our references.

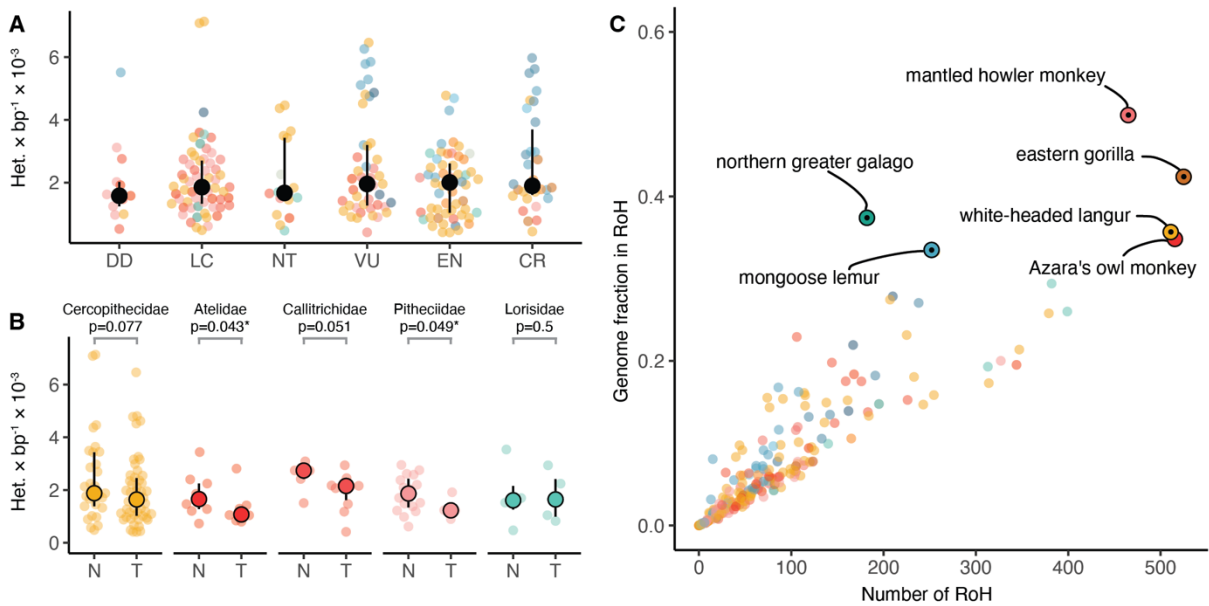


Fig. 2. Runs of homozygosity and impact of extinction risk on diversity (A) Relationship between IUCN extinction risk categories and heterozygosity. Solid black dots and bars denote median and IQR. **(B)** Partition into threatened (T: VU, EN, CR) and non-threatened (N: LC, NT) categories for all families with more than one species in either partition. Significant differences ($p < 0.05$, one-sided rank-sum test) are marked with an asterisk **(B)** Median number of tracts of homozygosity versus median proportion of the genome in runs of homozygosity per species. Species with a fraction over 1/3 are highlighted. Solid black dots within highlights denote threatened species (VU, EN, CR).

A time-calibrated nuclear phylogeny of primates

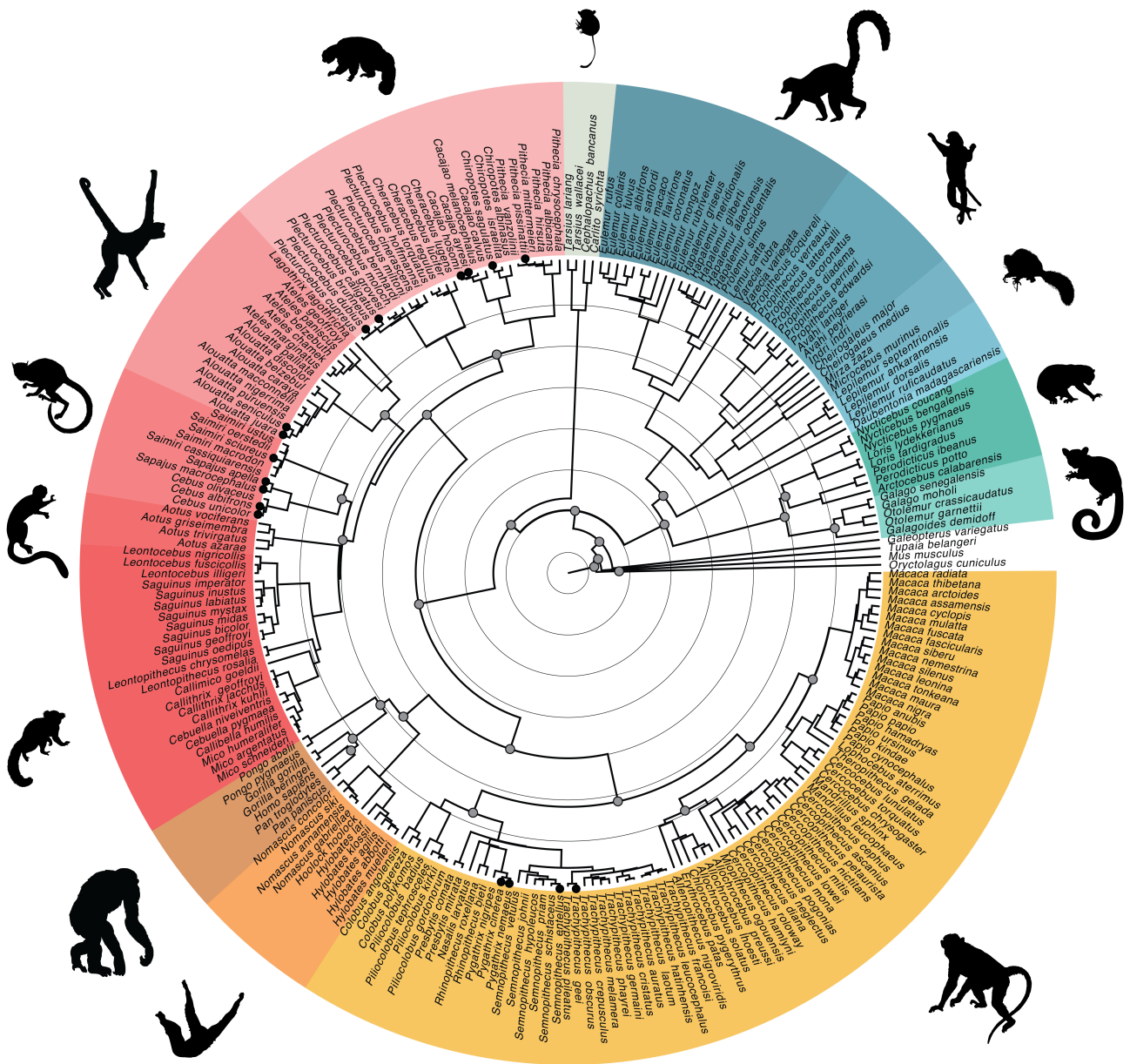


Fig. 3. Fossil calibrated nuclear time tree. Concentric background circles mark 10 million year intervals; solid gray circles in internal nodes show fossil calibration points (34); species marked with solid circles at tips show paraphyly or polyphyly when including additional individuals to estimate the topology.

We generated a genome-wide nuclear phylogeny of ultraconserved elements (UCEs) and 500 bp of their flanking regions, a widely used marker that enables easy detection of sequence orthologs across species (35). To this end, we identified the location of ~3500 UCE probes across all primate genomes and generated individual gene trees for each locus using a maximum likelihood approach (36–38). We used the resulting trees as input for a coalescent analysis to obtain the topology of the species tree, which has strong support at most nodes and recovers all currently recognized primate families, tribes, and genera as monophyletic (39–42). We used a newly established set of 27 well-justified fossil calibration points to constrain the timing of key phylogenetic divergences among different lineages (43). We estimate the split between Haplorhini and Strepsirrhini to have happened between 63.3–58.3 Ma ago, and thus the radiation of crown Primates is entirely within the Paleocene. We find the deepest divergence within tarsiers to be strikingly recent at 15.2–9.5 Ma, which, together with fossil evidence, implies considerable extinction along the long branch leading to extant tarsiers (44–47). All inter-familial relationships within our phylogeny receive strong support (posterior probability (PP) = 1), except for the position of Aotidae (owl monkeys), which is weakly supported as sister to Callitrichidae (marmosets and tamarins) rather than Cebidae (capuchin and squirrel monkeys) (PP = 0.56). We consider the precise relationship among these three families to remain uncertain. Lastly, we estimate the human-chimpanzee divergence between 9.0–6.9 Ma, and thus slightly older than other recent analyses, although these overlap our confidence intervals (39–41).

Taking advantage of our rich resequencing data, we generated a tree topology including two individuals per species for all species with more than one sequenced individual. We observe paraphyletic or polyphyletic placements of these individuals in 17 species, possibly calling several currently established species boundaries into question (see Fig. 3). These cases could result from genetic structure interpreted as species delimitation, incomplete lineage sorting, or hybridization, and most are also observed at the mitochondrial level (16, 48–51). While some instances of hybridization have previously been described, such as among different species of langurs (52), we find most of the paraphyletic or polyphyletic placements among platyrrhines. These include 13 species, among them capuchins, squirrel monkeys, howler monkeys, uakaris, sakis, and titis, and points to the need for more taxonomic studies using genomic data in this group (53). Finally, we retrieved previously unknown phylogenetic relationships for species that were sequenced for the first time in this study, such as different species of howler monkeys (e.g. *Alouatta puruensis*, or *A. juara*).

Determinants of diversity and mutation rate

We used the topology of the species tree and 614 UCE alignments for which we had full species coverage to estimate branch lengths as the number of substitutions per site. We combined this with our dated phylogeny and published estimates of generation times to estimate mutation rates per generation for all primate species via their substitution rates (16). While we caution that we cannot

rule out potential biases in these estimates, such as the effects of selection or uncertainties in fossil calibration, they agree well with published estimates for overlapping species based on trio sequencing (Spearman's $r=0.85$, $p = 0.02$, see Fig. 4C). Our estimated mutation rates (μ) per generation vary between 0.25×10^{-8} and 1.62×10^{-8} (see Fig. 4A), showing a considerably larger range than previously reported (54). We observe the lowest estimate per generation in Lemuridae and find highly variable estimates across some families such as Cebidae and Lorisidae, which also have variable generation times (8-17 and 4.6-9 years per generation, respectively). The highest estimates of μ are in great apes. We find a significant and positive correlation between μ per generation and the generation time (Spearman's $r=0.36$, $p=1.89 \times 10^{-8}$), which partly counteracts a generation time effect on the yearly mutation rate. The latter is therefore larger in species with a shorter generation time (see Fig. 4E). Together, variation in effective population size (N_e) and generation time explain roughly half of the observed variation in mutation rates among extant species.

We used our estimates of μ and estimates of genetic diversity π via median heterozygosity to get an estimate of the effective population sizes $N_e = \pi / (4 \times \mu)$. We find multiple species belonging to different families of lemurs, as well as several species of guenons within the Cercopithecidae, with the largest N_e estimates, often exceeding 2×10^5 (see Fig. 4B). For several critically endangered lemur species, e.g. the northern sportive lemur (*Lepilemur septentrionalis*), the red-tailed sportive lemur (*Lepilemur ruficaudatus*), or the Alaotra reed lemur (*Hapalemur alaotrensis*), these likely surpass census sizes by a significant margin. We find multiple members of the genera *Cercopithecus* and *Eulemur* exhibiting high N_e values, which may be driven by interspecific hybridization observed in these species. Conversely, we observe comparatively low N_e estimates in great apes, lorises and platyrrhines (see Fig. 4B and (16)).

The drift-barrier hypothesis (55, 56) predicts that μ per generation should decrease with N_e , because new mutations affecting fitness are predominantly deleterious, and the ability to select for lower mutation rate increases with the population size. We tested for a relationship between μ and N_e , while controlling for the relationship between μ and generation time in a phylogenetic generalized least squares (PGLS) model, and observed a significantly lower mutation rate for species with higher N_e . We find around 45% of the variation in μ to be explained by N_e , thus lending apparent support to the drift-barrier hypothesis (57). However, we caution that while this pattern is consistent with the drift-barrier hypothesis, N_e is estimated via the division of π by μ , which at least partially explains the negative relationship, and complicates a formal test. Additionally, our estimates of μ assume homogeneous levels of evolutionary constraint on the UCEs and flanking regions used to estimate divergence time and substitution rate. Should there be a strong covariation between substitution rates in these regions and effective size in branches, underlying variation in N_e along the branches of the phylogeny can act as a confounder of apparent variation in mutation rates, and thus further complicate a formal test of the drift-barrier hypothesis.

To further disentangle what factors might contribute to the levels of genetic diversity and mutation rates, we compiled a list of 32 traits that can be summarized into the broader categories of body mass, life history, activity budget, ranging patterns, climatic niche, social organization, sexual selection, diet composition, social systems, mating systems and natal dispersal mode (58–60). To account for potential phylogenetic inertia in trait evolution, we generated PGLS models using either genetic diversity or mutation rate as the response variable, and individual traits as the predictors. We find traits within mating systems, activity budget, climatic niche, ranging patterns, and life history to be significant predictors of diversity ($p < 0.05$), and traits within the former three categories remaining so after accounting for multiple testing (BH correction, FDR=0.05). Species

organized in single-male polygynous mating systems show lower diversity than the background ($r^2_{\text{pred}}=0.11$, $p_{\text{corr}}=1.53\times 10^{-2}$), consistent with expectations of reduced contribution of allelic diversity from males (61). Within the climatic niche, we observe a gradient of diversity declining from south to north ($r^2_{\text{pred}}=0.28$, $p_{\text{corr}}=1.45\times 10^{-5}$), which is driven by highly diverse lemur species in the southern hemisphere. We also find a significant correlation with mean temperature and amount of precipitation ($r^2_{\text{pred}}=0.33$, $p_{\text{corr}}=1.97\times 10^{-4}$). It is worth noting that these measurements are not highly correlated with each other (Pearson's r -0.27 – 0.17) and the relationships are thus at least partly independent. Lastly, within the activity budget, we find the amount of time spent socializing to be correlated with diversity ($r^2_{\text{pred}}=0.11$, $p_{\text{corr}}=5.56\times 10^{-3}$). However, we caution that the measurement of activity budget is difficult to standardize across species and interpreting this relationship thus challenging. We find no significant impact of life-history traits such as body mass or longevity on genetic diversity within primates, although body mass is significant before accounting for multiple testing. These relationships have been previously described, albeit for broader evolutionary distances including a wider range of genetic diversity and body mass (62, 63). We additionally calculated the relationship of the traits above to our mutation rate estimates. After correcting for multiple testing, we did not find any significant predictors of μ .

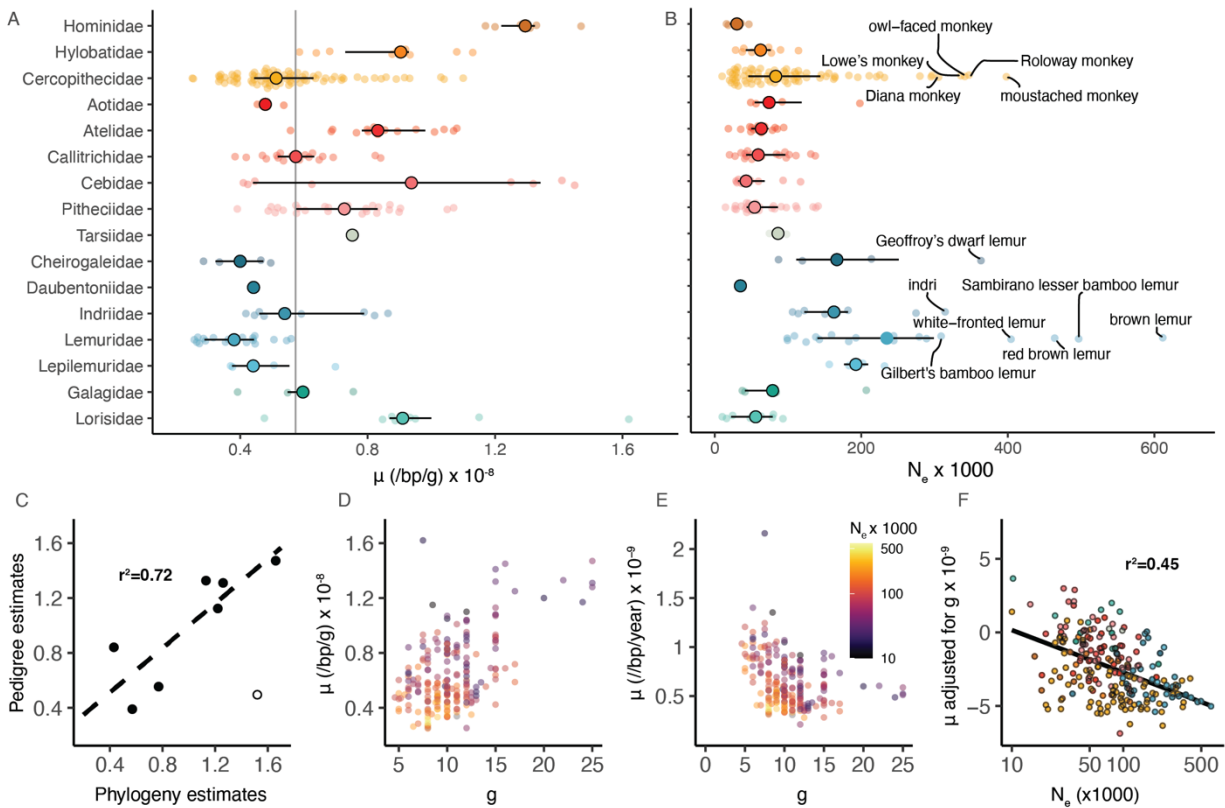


Fig. 4. Estimates of mutation rates and effective population size (A) Distribution of estimates of the per generation mutation rate across primate families (μ). Large solid circles denote median and horizontal bars denote the interquartile range. The gray line denotes the primate-wide median.

(B) Distribution of N_e estimates across primate families. Species with effective population size above 3×10^5 are highlighted. (C) Comparison of pedigree-based estimates of μ for great apes (64, 65), olive baboon (66), rhesus macaque (67), common marmoset (68) show a high correlation between the two estimates (Spearman's $r=0.85$ $p=0.02$). The hollow circle denotes the estimate for the mouse lemur (69), which was excluded from the comparison as an outlier (16). Data for trio estimates was derived from (70). (D) Positive correlation between estimates of per generation mutation rates and generation times (g) (Pearson's $r = 0.53$, $p = 2.1 \times 10^{-17}$). (E) Inverse relationship between yearly mutation rate and generation time. Circles in D and E are colored by the effective population size N_e (Pearson's $r = -0.34$, $p = 3.1 \times 10^{-7}$). (F) Relationship between per generation mutation rate, adjusted by first regressing the effects of generation time, and effective population size. The relationship is highly significant after phylogenetic correction ($r^2 = 0.45$, $p < 0.001$).

Unique variants in the human lineage

Finally, we revisited a previously published catalog of 647 high-frequency human-specific missense changes, i.e., amino-acid altering variants that putatively emerged specifically in the human lineage and quickly rose to high frequency or fixation (71). This catalog was mainly defined by looking at derived sites segregating at high frequency in anatomically modern humans, at which archaic hominins (Neandertals and Denisovans) carry the ancestral allele. While insufficient to explain the whole spectrum of human uniqueness, such a catalog should contain prime candidates for some of its molecular underpinnings. We sought to determine how often the putatively human-specific derived allele occurs at orthologous positions across the genomes of other primate species analyzed in this study. We find 63% (406) of high-frequency human-specific missense changes to occur in at least one other primate species and 55% in more than two, segregating at high frequency (>0.9) within the sampled individuals of a species (Fig. 5). This suggests that mutational recurrence generally might be widespread across primates. We find mutation pairs in recurrent high-frequency human-specific missense changes enriched in T-C and A-G mutations, and to a lesser extent in C-T and G-A compared to non-recurrent ones.

We leveraged our data to generate a more stringent picture of the mutations that arose specifically in the human lineage and have not emerged elsewhere in primates. We identified alleles present in anatomically modern humans at a frequency of at least 99.9% that differ in state from a set of four high coverage archaic hominins genomes (72–75). We ensured the human allele represents the derived state by requiring the ancestral allele to be present at a frequency of over 99% in a genetic diversity panel of 139 previously published great ape genomes (8, 9, 76, 77). The resulting 24,374 candidates include a conservative set of 124 missense coding mutations affecting 107 different genes, among which are 17 previously undescribed changes affecting 12 genes (71).

We further sought to detect which genes have not shown frequent allele recurrence in other primate species. To this end, we removed variants which we found to reoccur in $>1\%$ of species at a frequency of $>0.1\%$. In this set, we find 89 missense changes, affecting 80 distinct genes. We observe no enrichment for functional categories or association to diseases among them. Within our catalog, we also find the two amino-acid differences with demonstrated functional differences between humans and Neandertals: The ancestral allele in *NOVA1* (Neuro-oncological ventral antigen 1) leads to a slower development of cortical organoids and modifies synaptic protein interactions (78); the human derived allele of the adenylosuccinate lyase gene (*ADSL*) leads to a reduced *de novo* synthesis of purines in the brain (79). Furthermore, changes in mitotic spindle-associated genes previously reported to be under positive selection (*SPAG5*, *KIF18A*), maintain

their status as uniquely human (80). This may have had an impact on neurogenesis during development (81), although this hypothesis has not been experimentally validated. Interestingly, we find a uniquely human change in *TMPRSS2*, a main factor in the response to SARS-CoV-2 infection with known functional variants that have possibly been under selection in some human populations (82).

Analogous to the above, we additionally generated a catalog of sites that are fixed across great apes, but differ from rhesus macaque (*Macaca mulatta*). Among these 11.2M variants, we find 1M without observed recurrences beyond apes, corresponding to mutations specific to the great ape lineage. These contain 3,792 missense variants affecting 2,970 different genes that are significantly enriched for multiple cilia-related functional categories, such as axoneme assembly, motile cilium assembly, non-motile cilium assembly, cilium-dependent cell motility, and epithelial cilium movement involved in extracellular fluid movement, suggesting that the evolution of ape-specific features cilia has been important in shaping the lineage leading to our own species. The disruption of normally functioning cilia can lead to an array of heterogeneous pathologies in humans, collectively known as ciliopathies. Among 187 genes with established links to different ciliopathies, we find 30% to be affected by ape-specific missense changes (83) ($p < 0.01$, Fisher's Exact). More generally, we also find an overall significant enrichment of genes with non-recurrent ape-specific missense changes among genes with disease association in OMIM ($p < 0.01$, Fisher's Exact), suggesting that – to some degree – variants that give rise to the ape-specific phenotype, and thus ultimately also to the human one, affect a greater proportion of the genes that make us susceptible to diseases than would be expected by chance.

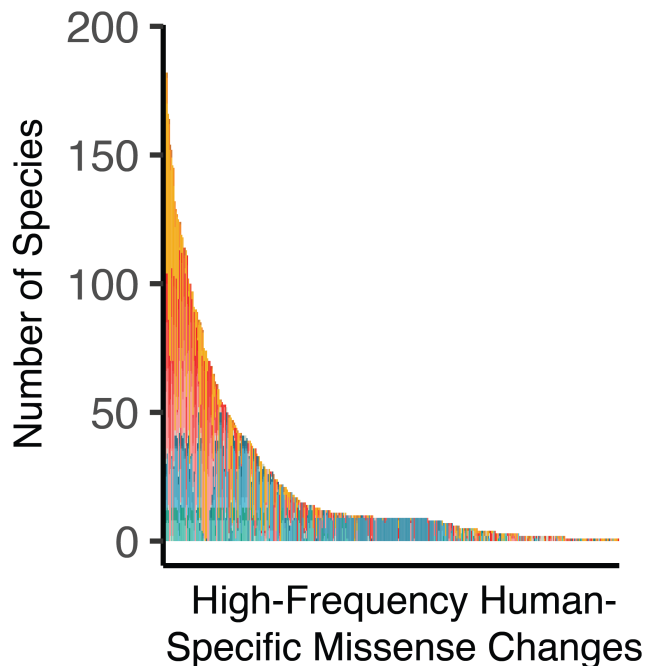


Fig. 5. Recurrent putative high-frequency human-specific missense changes. Each bar on the x-axis represents a high-frequency human-specific missense changes with the same allele found in a different species. Color schemes follow the same as presented in Fig. 1 and 2

References and Notes

1. Rylands, A. B., Mittermeier, R. A., "Primate names: Working Taxonomic List of Primates of the IUCN SSC Primate Specialist Group" in *Primate Behavioral Ecology, 6th edition*, K. B. Strier, Ed. (Routledge, New York, 2021), pp. 407–428.
2. L. F. Kuderna, P. Esteller-Cucala, T. Marques-Bonet, Branching out: what omics can tell us about primate evolution. *Curr. Opin. Genet. Dev.* **62**, 65–71 (2020).
3. J. D. Orkin, L. F. K. Kuderna, T. Marques-Bonet, The Diversity of Primates: From Biomedicine to Conservation Genomics. *Annu Rev Anim Biosci.* **9**, 103–124 (2021).
4. Chimpanzee Sequencing and Analysis Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature.* **437**, 69–87 (2005).
5. D. P. Locke, L. W. Hillier, W. C. Warren, K. C. Worley, L. V. Nazareth, D. M. Muzny, S.-P. Yang, Z. Wang, A. T. Chinwalla, P. Minx, M. Mitreva, L. Cook, K. D. Delehaunty, C. Fronick, H. Schmidt, L. A. Fulton, R. S. Fulton, J. O. Nelson, V. Magrini, C. Pohl, T. A. Graves, C. Markovic, A. Cree, H. H. Dinh, J. Hume, C. L. Kovar, G. R. Fowler, G. Lunter, S. Meader, A. Heger, C. P. Ponting, T. Marques-Bonet, C. Alkan, L. Chen, Z. Cheng, J. M. Kidd, E. E. Eichler, S. White, S. Searle, A. J. Vilella, Y. Chen, P. Flicek, J. Ma, B. Raney, B. Suh, R. Burhans, J. Herrero, D. Haussler, R. Faria, O. Fernando, F. Darré, D. Farré, E. Gazave, M. Oliva, A. Navarro, R. Roberto, O. Capozzi, N. Archidiacono, G. Della Valle, S. Purgato, M. Rocchi, M. K. Konkel, J. A. Walker, B. Ullmer, M. A. Batzer, A. F. A. Smit, R. Hubley, C. Casola, D. R. Schrider, M. W. Hahn, V. Quesada, X. S. Puente, G. R. Ordoñez, C. López-Otín, T. Vinar, B. Brejova, A. Ratan, R. S. Harris, W. Miller, C. Kosiol, H. A. Lawson, V. Taliwal, A. L. Martins, A. Siepel, A. Roychoudhury, X. Ma, J. Degenhardt, C. D. Bustamante, R. N. Gutenkunst, T. Mailund, J. Y. Dutheil, A. Hobolth, M. H. Schierup, O. A. Ryder, Y. Yoshinaga, P. J. de Jong, G. M. Weinstock, J. Rogers, E. R. Mardis, R. A. Gibbs, R. K. Wilson, Comparative and demographic analysis of orang-utan genomes. *Nature.* **469**, 529–533 (2011).
6. A. Scally, J. Y. Dutheil, L. W. Hillier, G. E. Jordan, I. Goodhead, J. Herrero, A. Hobolth, T. Lappalainen, T. Mailund, T. Marques-Bonet, S. McCarthy, S. H. Montgomery, P. C. Schwalie, Y. A. Tang, M. C. Ward, Y. Xue, B. Yngvadottir, C. Alkan, L. N. Andersen, Q. Ayub, E. V. Ball, K. Beal, B. J. Bradley, Y. Chen, C. M. Clee, S. Fitzgerald, T. A. Graves, Y. Gu, P. Heath, A. Heger, E. Karakoc, A. Kolb-Kokocinski, G. K. Laird, G. Lunter, S. Meader, M. Mort, J. C. Mullikin, K. Munch, T. D. O'Connor, A. D. Phillips, J. Prado-Martinez, A. S. Rogers, S. Sajjadian, D. Schmidt, K. Shaw, J. T. Simpson, P. D. Stenson, D. J. Turner, L. Vigilant, A. J. Vilella, W. Whitener, B. Zhu, D. N. Cooper, P. de Jong, E. T. Dermitzakis, E. E. Eichler, P. Flicek, N. Goldman, N. I. Mundy, Z. Ning, D. T. Odom, C. P. Ponting, M. A. Quail, O. A. Ryder, S. M. Searle, W. C. Warren, R. K. Wilson, M. H. Schierup, J. Rogers, C. Tyler-Smith, R. Durbin, Insights into hominid evolution from the gorilla genome sequence. *Nature.* **483**, 169–175 (2012).
7. K. Prüfer, K. Munch, I. Hellmann, K. Akagi, J. R. Miller, B. Walenz, S. Koren, G. Sutton, C. Kodira, R. Winer, J. R. Knight, J. C. Mullikin, S. J. Meader, C. P. Ponting, G. Lunter, S. Higashino, A. Hobolth, J. Dutheil, E. Karakoç, C. Alkan, S. Sajjadian, C. R. Catacchio, M. Ventura, T. Marques-Bonet, E. E. Eichler, C. André, R. Atencia, L. Mugisha, J. Junhold, N. Patterson, M. Siebauer, J. M. Good, A. Fischer, S. E. Ptak, M. Lachmann, D. E. Symer, T.

Mailund, M. H. Schierup, A. M. Andrés, J. Kelso, S. Pääbo, The bonobo genome compared with the chimpanzee and human genomes. *Nature*. **486**, 527–531 (2012).

8. J. Prado-Martinez, P. H. Sudmant, J. M. Kidd, H. Li, J. L. Kelley, B. Lorente-Galdos, K. R. Veeramah, A. E. Woerner, T. D. O'Connor, G. Santpere, A. Cagan, C. Theunert, F. Casals, H. Laayouni, K. Munch, A. Hobolth, A. E. Halager, M. Malig, J. Hernandez-Rodriguez, I. Hernando-Herraez, K. Prüfer, M. Pybus, L. Johnstone, M. Lachmann, C. Alkan, D. Twigg, N. Petit, C. Baker, F. Hormozdiari, M. Fernandez-Callejo, M. Dabad, M. L. Wilson, L. Stevison, C. Camprubí, T. Carvalho, A. Ruiz-Herrera, L. Vives, M. Mele, T. Abello, I. Kondova, R. E. Bontrop, A. Pusey, F. Lankester, J. A. Kiyang, R. A. Bergl, E. Lonsdorf, S. Myers, M. Ventura, P. Gagneux, D. Comas, H. Siegismund, J. Blanc, L. Agueda-Calpena, M. Gut, L. Fulton, S. A. Tishkoff, J. C. Mullikin, R. K. Wilson, I. G. Gut, M. K. Gonder, O. A. Ryder, B. H. Hahn, A. Navarro, J. M. Akey, J. Bertranpetit, D. Reich, T. Mailund, M. H. Schierup, C. Hvilsom, A. M. Andrés, J. D. Wall, C. D. Bustamante, M. F. Hammer, E. E. Eichler, T. Marques-Bonet, Great ape genetic diversity and population history. *Nature*. **499**, 471–475 (2013).
9. A. Nater, M. P. Mattle-Greminger, A. Nurcahyo, M. G. Nowak, M. de Manuel, T. Desai, C. Groves, M. Pybus, T. B. Sonay, C. Roos, A. R. Lameira, S. A. Wich, J. Askew, M. Davila-Ross, G. Fredriksson, G. de Valles, F. Casals, J. Prado-Martinez, B. Goossens, E. J. Verschoor, K. S. Warren, I. Singleton, D. A. Marques, J. Pamungkas, D. Perwitasari-Farajallah, P. Rianti, A. Tuuga, I. G. Gut, M. Gut, P. Orozco-terWengel, C. P. van Schaik, J. Bertranpetit, M. Anisimova, A. Scally, T. Marques-Bonet, E. Meijaard, M. Krützen, Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species. *Curr. Biol.* **27**, 3576–3577 (2017).
10. Z. N. Kronenberg, I. T. Fiddes, D. Gordon, S. Murali, S. Cantsilieris, O. S. Meyerson, J. G. Underwood, B. J. Nelson, M. J. P. Chaisson, M. L. Dougherty, K. M. Munson, A. R. Hastie, M. Diekhans, F. Hormozdiari, N. Lorusso, K. Hoekzema, R. Qiu, K. Clark, A. Raja, A. E. Welch, M. Sorensen, C. Baker, R. S. Fulton, J. Armstrong, T. A. Graves-Lindsay, A. M. Denli, E. R. Hoppe, P. Hsieh, C. M. Hill, A. W. C. Pang, J. Lee, E. T. Lam, S. K. Dutcher, F. H. Gage, W. C. Warren, J. Shendure, D. Haussler, V. A. Schneider, H. Cao, M. Ventura, R. K. Wilson, B. Paten, A. Pollen, E. E. Eichler, High-resolution comparative analysis of great ape genomes. *Science*. **360** (2018), doi:10.1126/science.aar6343.
11. Rhesus Macaque Genome Sequencing and Analysis Consortium, R. A. Gibbs, J. Rogers, M. G. Katze, R. Bumgarner, G. M. Weinstock, E. R. Mardis, K. A. Remington, R. L. Strausberg, J. C. Venter, R. K. Wilson, M. A. Batzer, C. D. Bustamante, E. E. Eichler, M. W. Hahn, R. C. Hardison, K. D. Makova, W. Miller, A. Milosavljevic, R. E. Palermo, A. Siepel, J. M. Sikela, T. Attaway, S. Bell, K. E. Bernard, C. J. Buhay, M. N. Chandrabose, M. Dao, C. Davis, K. D. Delehaunty, Y. Ding, H. H. Dinh, S. Dugan-Rocha, L. A. Fulton, R. A. Gabisi, T. T. Garner, J. Godfrey, A. C. Hawes, J. Hernandez, S. Hines, M. Holder, J. Hume, S. N. Jhangiani, V. Joshi, Z. M. Khan, E. F. Kirkness, A. Cree, R. G. Fowler, S. Lee, L. R. Lewis, Z. Li, Y.-S. Liu, S. M. Moore, D. Muzny, L. V. Nazareth, D. N. Ngo, G. O. Okwuonu, G. Pai, D. Parker, H. A. Paul, C. Pfannkoch, C. S. Pohl, Y.-H. Rogers, S. J. Ruiz, A. Sabo, J. Santibanez, B. W. Schneider, S. M. Smith, E. Sodergren, A. F. Svatek, T. R. Utterback, S. Vattathil, W. Warren, C. S. White, A. T. Chinwalla, Y. Feng, A. L. Halpern, L. W. Hillier, X. Huang, P. Minx, J. O. Nelson, K. H. Pepin, X. Qin, G. G. Sutton, E. Venter, B. P. Walenz, J. W. Wallis, K. C. Worley, S.-P. Yang, S. M. Jones, M. A. Marra, M. Rocchi, J. E. Schein, R. Baertsch, L. Clarke, M. Csürös, J. Glasscock, R. A. Harris, P. Havlak, A. R. Jackson, H. Jiang, Y. Liu, D. N.

- 5
10
15
20
25
30
35
40
- Messina, Y. Shen, H. X.-Z. Song, T. Wylie, L. Zhang, E. Birney, K. Han, M. K. Konkel, J. Lee, A. F. A. Smit, B. Ullmer, H. Wang, J. Xing, R. Burhans, Z. Cheng, J. E. Karro, J. Ma, B. Raney, X. She, M. J. Cox, J. P. Demuth, L. J. Dumas, S.-G. Han, J. Hopkins, A. Karimpour-Fard, Y. H. Kim, J. R. Pollack, T. Vinar, C. Addo-Quaye, J. Degenhardt, A. Denby, M. J. Hubisz, A. Indap, C. Kosiol, B. T. Lahn, H. A. Lawson, A. Marklein, R. Nielsen, E. J. Vallender, A. G. Clark, B. Ferguson, R. D. Hernandez, K. Hirani, H. Kehrer-Sawatzki, J. Kolb, S. Patil, L.-L. Pu, Y. Ren, D. G. Smith, D. A. Wheeler, I. Schenck, E. V. Ball, R. Chen, D. N. Cooper, B. Giardine, F. Hsu, W. J. Kent, A. Lesk, D. L. Nelson, W. E. O'brien, K. Prüfer, P. D. Stenson, J. C. Wallace, H. Ke, X.-M. Liu, P. Wang, A. P. Xiang, F. Yang, G. P. Barber, D. Haussler, D. Karolchik, A. D. Kern, R. M. Kuhn, K. E. Smith, A. S. Zwiag, Evolutionary and biomedical insights from the rhesus macaque genome. *Science*. **316**, 222–234 (2007).
12. W. C. Warren, R. A. Harris, M. Haukness, I. T. Fiddes, S. C. Murali, J. Fernandes, P. C. Dishuck, J. M. Storer, M. Raveendran, L. W. Hillier, D. Porubsky, Y. Mao, D. Gordon, M. R. Vollger, A. P. Lewis, K. M. Munson, E. DeVogelaere, J. Armstrong, M. Diekhans, J. A. Walker, C. Tomlinson, T. A. Graves-Lindsay, M. Kremitzki, S. R. Salama, P. A. Audano, M. Escalona, N. W. Maurer, F. Antonacci, L. Mercuri, F. A. M. Maggiolini, C. R. Catacchio, J. G. Underwood, D. H. O'Connor, A. D. Sanders, J. O. Korbel, B. Ferguson, H. M. Kubisch, L. Picker, N. H. Kalin, D. Rosene, J. Levine, D. H. Abbott, S. B. Gray, M. M. Sanchez, Z. A. Kovacs-Balint, J. W. Kemnitz, S. M. Thomasy, J. A. Roberts, E. L. Kinnally, J. P. Capitano, J. H. P. Skene, M. Platt, S. A. Cole, R. E. Green, M. Ventura, R. W. Wiseman, B. Paten, M. A. Batzer, J. Rogers, E. E. Eichler, Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science*. **370** (2020), doi:10.1126/science.abc6617.
13. C. Xue, M. Raveendran, R. A. Harris, G. L. Fawcett, X. Liu, S. White, M. Dahdouli, D. Rio Deiros, J. E. Below, W. Salerno, L. Cox, G. Fan, B. Ferguson, J. Horvath, Z. Johnson, S. Kanthaswamy, H. M. Kubisch, D. Liu, M. Platt, D. G. Smith, B. Sun, E. J. Vallender, F. Wang, R. W. Wiseman, R. Chen, D. M. Muzny, R. A. Gibbs, F. Yu, J. Rogers, The population genomics of rhesus macaques (*Macaca mulatta*) based on whole-genome sequences. *Genome Res*. **26**, 1651–1662 (2016).
14. A. Estrada, P. A. Garber, A. B. Rylands, C. Roos, E. Fernandez-Duque, A. Di Fiore, K. A.-I. Nekaris, V. Nijman, E. W. Heymann, J. E. Lambert, F. Rovero, C. Barelli, J. M. Setchell, T. R. Gillespie, R. A. Mittermeier, L. V. Arregoitia, M. de Guinea, S. Gouveia, R. Dobrovolski, S. Shanee, N. Shanee, S. A. Boyle, A. Fuentes, K. C. MacKinnon, K. R. Amato, A. L. S. Meyer, S. Wich, R. W. Sussman, R. Pan, I. Kone, B. Li, Impending extinction crisis of the world's primates: Why primates matter. *Sci. Adv.* **3**, e1600946 (2017).
15. C. S. Mantyka-Pringle, P. Visconti, M. Di Marco, Climate change modifies risk of global biodiversity loss due to land-cover change. *Biologicals* (2015) (available at <https://www.sciencedirect.com/science/article/pii/S0006320715001615>).
16. see Supplementary Material.
17. Zoonomia Consortium, A comparative genomics multitool for scientific discovery and conservation. *Nature*. **587**, 240–245 (2020).

18. L. Wang, J. Wu, X. Liu, D. Di, Y. Liang, Y. Feng, S. Zhang, B. Li, X.-G. Qi, A high-quality genome assembly for the endangered golden snub-nosed monkey (*Rhinopithecus roxellana*). *Gigascience*. **8** (2019), doi:10.1093/gigascience/giz098.
19. Z. Liu, L. Zhang, Z. Yan, Z. Ren, F. Han, X. Tan, Z. Xiang, F. Dong, Z. Yang, G. Liu, Z. Wang, J. Zhang, T. Que, C. Tang, Y. Li, S. Wang, J. Wu, L. Li, C. Huang, C. Roos, M. Li, Genomic Mechanisms of Physiological and Morphological Adaptations of Limestone Langurs to Karst Habitats. *Mol. Biol. Evol.* **37**, 952–968 (2020).
20. B. J. Evans, A. J. Tosi, K. Zeng, J. Dushoff, A. Corvelo, D. J. Melnick, Speciation over the edge: gene flow among non-human primate species across a formidable biogeographic barrier. *R Soc Open Sci.* **4**, 170351 (2017).
21. Z. Fan, A. Zhou, N. Osada, J. Yu, J. Jiang, P. Li, L. Du, L. Niu, J. Deng, H. Xu, J. Xing, B. Yue, J. Li, Ancient hybridization and admixture in macaques (genus *Macaca*) inferred from whole genome sequences. *Mol. Phylogenet. Evol.* **127**, 376–386 (2018).
22. D. Vanderpool, B. Q. Minh, R. Lanfear, D. Hughes, S. Murali, R. A. Harris, M. Raveendran, D. M. Muzny, M. S. Hibbins, R. J. Williamson, R. A. Gibbs, K. C. Worley, J. Rogers, M. W. Hahn, Primate phylogenomics uncovers multiple rapid radiations and ancient interspecific introgression. *PLoS Biol.* **18**, e3000954 (2020).
23. L. Yu, G.-D. Wang, J. Ruan, Y.-B. Chen, C.-P. Yang, X. Cao, H. Wu, Y.-H. Liu, Z.-L. Du, X.-P. Wang, J. Yang, S.-C. Cheng, L. Zhong, L. Wang, X. Wang, J.-Y. Hu, L. Fang, B. Bai, K.-L. Wang, N. Yuan, S.-F. Wu, B.-G. Li, J.-G. Zhang, Y.-Q. Yang, C.-L. Zhang, Y.-C. Long, H.-S. Li, J.-Y. Yang, D. M. Irwin, O. A. Ryder, Y. Li, C.-I. Wu, Y.-P. Zhang, Genomic analysis of snub-nosed monkeys (*Rhinopithecus*) identifies genes and processes related to high-altitude adaptation. *Nat. Genet.* **48**, 947–952 (2016).
24. N. Osada, K. Matsudaira, Y. Hamada, S. Malaivijitnond, Testing Sex-Biased Admixture Origin of Macaque Species Using Autosomal and X-Chromosomal Genomic Sequences. *Genome Biol. Evol.* **13** (2021), doi:10.1093/gbe/evaa209.
25. C. Schwitzer, ;, Edward, E. Louis, "Jonah Ratsimbazafy (Durrell Wildlife Conservation Trust-Madagascar Programme)" in *Madagascar Biodiversity Partnership, Henry Doorly Zoo, A. A. Unknown-L. Madagascar, L. R. Unknown-L. Madagascar, -. L. Madagascar*, Eds. (Clifton, Bristol; Omaha, USA; Leslie Wilmet (TERRA Research Center; Gembloux, Belgium; Omaha, NE, 68107, USA), Timothy Sefczek; Omaha, Nebraska, USA; Omaha, Nebraska, USA; Mary Dinsmore; Wisconsin, USA).
26. C. Coudrat, B. Rawson, P. Phiaphalath, F. Pengfei, C. Roos, M. H. Nguyen, IUCN Red List of Threatened Species: *Nomascus concolor*. *IUCN Red List of Threatened Species* (2015) (available at <https://www.iucnredlist.org/species/39775/17968556>).
27. T. van der Valk, C. M. Gonda, H. Silegowa, S. Almanza, I. Sifuentes-Romero, T. B. Hart, J. A. Hart, K. M. Detwiler, K. Guschanski, The Genome of the Endangered Dryas Monkey Provides New Insights into the Evolutionary History of the Vervets. *Mol. Biol. Evol.* **37**, 183–194 (2020).

28. K. M. Detwiler, Mitochondrial DNA Analyses of Cercopithecus Monkeys Reveal a Localized Hybrid Origin for *C. mitis doggetti* in Gombe National Park, Tanzania. *Int. J. Primatol.* **40**, 28–52 (2019).
29. Y. A. de Jong, T. M. Butynski, Three Sykes's Monkey *Cercopithecus mitis* × Vervet Monkey *Chlorocebus pygerythrus* Hybrids in Kenya. *Primate Conserv.* **25**, 43–56 (2010).
30. H. Svardal, A. J. Jasinska, C. Apetrei, G. Coppola, Y. Huang, C. A. Schmitt, B. Jacquelin, V. Ramensky, M. Müller-Trutwin, M. Antonio, G. Weinstock, J. P. Grobler, K. Dewar, R. K. Wilson, T. R. Turner, W. C. Warren, N. B. Freimer, M. Nordborg, Ancient hybridization and strong adaptation to viruses across African vervet monkey populations. *Nat. Genet.* **49**, 1705–1713 (2017).
31. D. Spielman, B. W. Brook, R. Frankham, Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15261–15264 (2004).
32. J. C. Teixeira, C. D. Huber, The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021), doi:10.1073/pnas.2015096118.
33. T. van der Valk, D. Díez-Del-Molino, T. Marques-Bonet, K. Guschanski, L. Dalén, Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. *Curr. Biol.* **29**, 165-170.e6 (2019).
34. J. A. Robinson, J. Rääkkönen, L. M. Vucetich, J. A. Vucetich, R. O. Peterson, K. E. Lohmueller, R. K. Wayne, Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci Adv.* **5**, eaau0757 (2019).
35. B. C. Faircloth, J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, T. C. Glenn, Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).
36. S. Naser-Khdour, B. Q. Minh, W. Zhang, E. A. Stone, R. Lanfear, The Prevalence and Impact of Model Violations in Phylogenetic Analysis. *Genome Biol. Evol.* **11**, 3341–3352 (2019).
37. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
38. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods.* **14**, 587–589 (2017).
39. P. Perelman, W. E. Johnson, C. Roos, H. N. Seuánez, J. E. Horvath, M. A. M. Moreira, B. Kessing, J. Pontius, M. Roelke, Y. Rumpler, M. P. C. Schneider, A. Silva, S. J. O'Brien, J. Pecon-Slattery, A molecular phylogeny of living primates. *PLoS Genet.* **7**, e1001342 (2011).
40. M. S. Springer, R. W. Meredith, J. Gatesy, C. A. Emerling, J. Park, D. L. Rabosky, T. Stadler, C. Steiner, O. A. Ryder, J. E. Janečka, C. A. Fisher, W. J. Murphy, Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. *PLoS One.* **7**, e49521 (2012).

41. M. D. Reis, G. F. Gunnell, J. Barba-Montoya, A. Wilkins, Z. Yang, A. D. Yoder, 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).
42. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics.* **19**, 153 (2018).
43. D. de Vries, R. M. D. Beck, Total evidence tip-dating phylogeny of platyrrhine primates and 27 well-justified fossil calibrations for primate divergences. *bioRxiv* (2021) (available at <https://www.biorxiv.org/content/10.1101/2021.10.21.465342.abstract>).
44. J. B. Rossie, X. Ni, K. C. Beard, Cranial remains of an Eocene tarsier. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4381–4385 (2006).
45. J. S. Zijlstra, L. J. Flynn, W. Wessels, The westernmost tarsier: a new genus and species from the Miocene of Pakistan. *J. Hum. Evol.* **65**, 544–550 (2013).
46. Y. Chaimanee, R. Lebrun, C. Yamee, J.-J. Jaeger, A new Middle Miocene tarsier from Thailand and the reconstruction of its orbital morphology using a geometric–morphometric method. *Proceedings of the Royal Society B: Biological Sciences.* **278**, 1956–1963 (2011).
47. X. Ni, Q. Li, L. Li, K. C. Beard, Oligocene primates from China reveal divergence between African and Asian primate evolution. *Science.* **352**, 673–677 (2016).
48. J. Tung, L. B. Barreiro, The contribution of admixture to primate evolution. *Curr. Opin. Genet. Dev.* **47**, 61–68 (2017).
49. C. Fontserè, M. de Manuel, T. Marques-Bonet, M. Kuhlwilm, Admixture in mammals and how to understand its functional implications: On the abundance of gene flow in mammalian species, its impact on the genome, and roads into a functional understanding. *Bioessays.* **41**, e1900123 (2019).
50. J. Sukumaran, L. L. Knowles, Multispecies coalescent delimits structure, not species. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1607–1612 (2017).
51. M. C. Janiak, F. E. Silva, R. M. D. Beck, D. de Vries, L. F. K. Kuderna, N. S. Torosin, A. D. Melin, T. Marquès-Bonet, I. B. Goodhead, M. Messias, M. N. F. da Silva, I. Sampaio, I. P. Farias, R. Rossi, F. R. de Melo, J. Valsecchi, T. Hrbek, J. P. Boubli, Two hundred and five newly assembled mitogenomes provide mixed evidence for rivers as drivers of speciation for Amazonian primates. *Mol. Ecol.* **31**, 3888–3902 (2022).
52. C. Roos, D. Zinner, L. S. Kubatko, C. Schwarz, M. Yang, D. Meyer, S. D. Nash, J. Xing, M. A. Batzer, M. Brameier, F. H. Leendertz, T. Ziegler, D. Perwitasari-Farajallah, T. Nadler, L. Walter, M. Osterholz, Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. *BMC Evol. Biol.* **11**, 77 (2011).
53. M. G. M. Lima, J. de S. E. Silva-Júnior, D. Černý, J. C. Buckner, A. Aleixo, J. Chang, J. Zheng, M. E. Alfaro, A. Martins, A. Di Fiore, J. P. Boubli, J. W. Lynch Alfaro, A phylogenomic perspective on the robust capuchin monkey (*Sapajus*) radiation: First evidence for extensive population admixture across South America. *Mol. Phylogenet. Evol.* **124**, 137–150 (2018).

54. M. Chintalapati, P. Moorjani, Evolution of the mutation rate across primates. *Curr. Opin. Genet. Dev.* **62**, 58–64 (2020).
55. T. Ohta, Slightly deleterious mutant substitutions in evolution. *Nature.* **246**, 96–98 (1973).
56. W. Sung, M. S. Ackerman, S. F. Miller, T. G. Doak, M. Lynch, Drift-barrier hypothesis and mutation-rate evolution. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18488–18492 (2012).
57. M. Lynch, Evolution of the mutation rate. *Trends Genet.* **26**, 345–352 (2010).
58. J. M. Kamilari, N. Cooper, Phylogenetic signal in primate behaviour, ecology and life history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120341 (2013).
59. A. R. DeCasien, S. A. Williams, J. P. Higham, Primate brain size is predicted by diet but not sociality. *Nat Ecol Evol.* **1**, 112 (2017).
60. S. Shultz, C. Opie, Q. D. Atkinson, Stepwise evolution of stable sociality in primates. *Nature.* **479** (2011), pp. 219–222.
61. B. Charlesworth, Effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* **10**, 195–205 (2009).
62. H. Ellegren, N. Galtier, Determinants of genetic diversity. *Nat. Rev. Genet.* **17**, 422–433 (2016).
63. J. Romiguier, P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, Y. Chiari, R. Derrat, L. Duret, N. Faivre, E. Loire, J. M. Lourenco, B. Nabholz, C. Roux, G. Tsagkogeorga, A. A.-T. Weber, L. A. Weinert, K. Belkhir, N. Bierne, S. Glémin, N. Galtier, Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature.* **515**, 261–263 (2014).
64. S. Besenbacher, C. Hvilsom, T. Marques-Bonet, T. Mailund, M. H. Schierup, Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nat Ecol Evol.* **3**, 286–292 (2019).
65. M. D. Kessler, D. P. Loesch, J. A. Perry, N. L. Heard-Costa, D. Taliun, B. E. Cade, H. Wang, M. Daya, J. Ziniti, S. Datta, J. C. Celedón, M. E. Soto-Quiros, L. Avila, S. T. Weiss, K. Barnes, S. S. Redline, R. S. Vasan, A. D. Johnson, R. A. Mathias, R. Hernandez, J. G. Wilson, D. A. Nickerson, G. Abecasis, S. R. Browning, S. Zöllner, J. R. O’Connell, B. D. Mitchell, National Heart, Lung, and Blood Institute Trans-Omics for Precision Medicine (TOPMed) Consortium, TOPMed Population Genetics Working Group, T. D. O’Connor, De novo mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 2560–2569 (2020).
66. F. L. Wu, A. I. Strand, L. A. Cox, C. Ober, J. D. Wall, P. Moorjani, M. Przeworski, A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. *PLoS Biol.* **18**, e3000838 (2020).
67. L. A. Bergeron, S. Besenbacher, J. Bakker, J. Zheng, P. Li, G. Pacheco, M.-H. S. Sinding, M. Kamilari, M. T. P. Gilbert, M. H. Schierup, G. Zhang, The germline mutational process in

rhesus macaque and its implications for phylogenetic dating. *Gigascience*. **10** (2021), doi:10.1093/gigascience/giab029.

- 5 68. C. Yang, Y. Zhou, S. Marcus, G. Formenti, L. A. Bergeron, Z. Song, X. Bi, J. Bergman, M. M. C. Rousselle, C. Zhou, L. Zhou, Y. Deng, M. Fang, D. Xie, Y. Zhu, S. Tan, J. Mountcastle, B. Haase, J. Balacco, J. Wood, W. Chow, A. Rhie, M. Pippel, M. M. Fabiszak, S. Koren, O. Fedrigo, W. A. Freiwald, K. Howe, H. Yang, A. M. Phillippy, M. H. Schierup, E. D. Jarvis, G. Zhang, Evolutionary and biomedical insights from a marmoset diploid genome assembly. *Nature*. **594**, 227–233 (2021).
- 10 69. C. Ryan Campbell, G. P. Tiley, J. W. Poelstra, K. E. Hunnicutt, P. A. Larsen, H.-J. Lee, J. L. Thorne, M. dos Reis, A. D. Yoder, Pedigree-based measurement of the de novo mutation rate in the gray mouse lemur reveals a high mutation rate, few mutations in CpG sites, and a weak sex bias. *bioRxiv* (2020), p. 724880.
- 15 70. L. A. Bergeron, S. Besenbacher, T. Turner, C. J. Versoza, R. J. Wang, A. L. Price, E. Armstrong, M. Riera, J. Carlson, H.-Y. Chen, M. W. Hahn, K. Harris, A. S. Kleppe, E. H. López-Nandam, P. Moorjani, S. P. Pfeifer, G. P. Tiley, A. D. Yoder, G. Zhang, M. H. Schierup, The Mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *Elife*. **11** (2022), doi:10.7554/eLife.73577.
- 20 71. M. Kuhlwilm, C. Boeckx, A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. *Sci. Rep.* **9**, 8463 (2019).
- 25 72. K. Prüfer, C. de Filippo, S. Grote, F. Mafessoni, P. Korlević, M. Hajdinjak, B. Vernot, L. Skov, P. Hsieh, S. Peyrégne, D. Reher, C. Hopfe, S. Nagel, T. Maricic, Q. Fu, C. Theunert, R. Rogers, P. Skoglund, M. Chintalapati, M. Dannemann, B. J. Nelson, F. M. Key, P. Rudan, Ž. Kučan, I. Gušić, L. V. Golovanova, V. B. Doronichev, N. Patterson, D. Reich, E. E. Eichler, M. Slatkin, M. H. Schierup, A. M. Andrés, J. Kelso, M. Meyer, S. Pääbo, A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science*. **358**, 655–658 (2017).
- 30 73. F. Mafessoni, S. Grote, C. de Filippo, V. Slon, K. A. Kolobova, B. Viola, S. V. Markin, M. Chintalapati, S. Peyrégne, L. Skov, P. Skoglund, A. I. Krivoschapkin, A. P. Derevianko, M. Meyer, J. Kelso, B. Peter, K. Prüfer, S. Pääbo, A high-coverage Neandertal genome from Chagyrskaya Cave. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15132–15136 (2020).
- 35 74. R. E. Green, J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Ž. Kucan, I. Gušić, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Pääbo, A draft sequence of the Neandertal genome. *Science*. **328**, 710–722 (2010).
- 40 75. D. Reich, R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. F. Johnson, T. Maricic, J. M. Good, T. Marques-Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V.

Shunkov, A. P. Derevianko, J.-J. Hublin, J. Kelso, M. Slatkin, S. Pääbo, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*. **468**, 1053–1060 (2010).

76. Y. Xue, J. Prado-Martinez, P. H. Sudmant, V. Narasimhan, Q. Ayub, M. Szpak, P. Frandsen, Y. Chen, B. Yngvadottir, D. N. Cooper, M. de Manuel, J. Hernandez-Rodriguez, I. Lobon, H. R. Siegismund, L. Pagani, M. A. Quail, C. Hvilsom, A. Mudakikwa, E. E. Eichler, M. R. Cranfield, T. Marques-Bonet, C. Tyler-Smith, A. Scally, Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* (2015) (available at <https://science.sciencemag.org/content/348/6231/242.abstract>).
77. M. de Manuel, M. Kuhlwilm, P. Frandsen, V. C. Sousa, T. Desai, J. Prado-Martinez, J. Hernandez-Rodriguez, I. Dupanloup, O. Lao, P. Hallast, J. M. Schmidt, J. M. Heredia-Genestar, A. Benazzo, G. Barbujani, B. M. Peter, L. F. K. Kuderna, F. Casals, S. Angedakin, M. Arandjelovic, C. Boesch, H. Kühl, L. Vigilant, K. Langergraber, J. Novembre, M. Gut, I. Gut, A. Navarro, F. Carlsen, A. M. Andrés, H. R. Siegismund, A. Scally, L. Excoffier, C. Tyler-Smith, S. Castellano, Y. Xue, C. Hvilsom, T. Marques-Bonet, Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*. **354**, 477–481 (2016).
78. C. A. Trujillo, E. S. Rice, N. K. Schaefer, I. A. Chaim, E. C. Wheeler, A. A. Madrigal, J. Buchanan, S. Preissl, A. Wang, P. D. Negraes, R. A. Szeto, R. H. Herai, A. Huseynov, M. S. A. Ferraz, F. S. Borges, A. H. Kihara, A. Byrne, M. Marin, C. Vollmers, A. N. Brooks, J. D. Lautz, K. Semendeferi, B. Shapiro, G. W. Yeo, S. E. P. Smith, R. E. Green, A. R. Muotri, Reintroduction of the archaic variant of NOVA1 in cortical organoids alters neurodevelopment. *Science*. **371** (2021), doi:10.1126/science.aax2537.
79. V. Stepanova, K. E. Moczulska, G. N. Vacano, I. Kurochkin, X. Ju, S. Riesenberger, D. Macak, T. Maricic, L. Dombrowski, M. Schörning, K. Anastasiadis, O. Baker, R. Naumann, E. Khrameeva, A. Vanushkina, E. Stekolshchikova, A. Egorova, A. Tkachev, R. Mazzarino, N. Duval, D. Zubkov, P. Giavalisco, T. G. Wilkinson, D. Patterson, P. Khaitovich, S. Pääbo, Reduced purine biosynthesis in humans after their divergence from Neandertals. *Elife*. **10** (2021), doi:10.7554/eLife.58741.
80. S. Peyrégne, J. Kelso, B. M. Peter, S. Pääbo, The evolutionary history of human spindle genes includes back-and-forth gene flow with Neandertals. *Elife*. **11** (2022), doi:10.7554/elife.75464.
81. S. Pääbo, The human condition—a molecular approach. *Cell*. **157**, 216–226 (2014).
82. S. Jeon, A. Blazyte, C. Yoon, H. Ryu, Y. Jeon, Y. Bhak, D. Bolser, A. Manica, E.-S. Shin, Y. S. Cho, B. C. Kim, N. Ryoo, H. Choi, J. Bhak, Regional TMPRSS2 V197M Allele Frequencies Are Correlated with COVID-19 Case Fatality Rates. *Mol. Cells*. **44**, 680–687 (2021).
83. J. F. Reiter, M. R. Leroux, Genes and molecular pathways underpinning ciliopathies. *Nat. Rev. Mol. Cell Biol.* **18**, 533–547 (2017).
84. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014).
85. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature*. **475**, 493–496 (2011).

86. G. Gower, J. Tuke, A. B. Rohrlach, J. Soubrier, B. Llamas, N. Bean, A. Cooper, Population size history from short genomic scaffolds: how short is too short? *bioRxiv* (2018), p. 382036.
87. T. Günther, C. Nettelblad, The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genet.* **15**, e1008302 (2019).
- 5 88. D. R. Schrider, A. G. Shanku, A. D. Kern, Effects of Linked Selective Sweeps on Demographic Inference and Model Selection. *Genetics.* **204**, 1207–1223 (2016).
89. M. Pacifici, L. Santini, M. Di Marco, D. Baisero, L. Francucci, G. Grotto Marasini, P. Visconti, C. Rondinini, Generation length for mammals. *NC.* **5**, 89–94 (2013).
- 10 90. C. D. Soria, M. Pacifici, M. Di Marco, S. M. Stephen, C. Rondinini, COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. *Ecology.* **102**, e03344 (2021).
91. S. Schiffels, K. Wang, "MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent" in *Statistical Population Genomics*, J. Y. Dutheil, Ed. (Springer US, New York, NY, 2020), pp. 147–166.
- 15 92. B. C. Faircloth, M. G. Branstetter, N. D. White, S. G. Brady, Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Mol. Ecol. Resour.* **15**, 489–501 (2015).
93. B. C. Faircloth, PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics.* **32**, 786–788 (2016).
- 20 94. D. M. Portik, J. J. Wiens, Do Alignment and Trimming Methods Matter for Phylogenomic (UCE) Analyses? *Syst. Biol.* **70**, 440–462 (2021).
95. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
- 25 96. U. Mai, S. Mirarab, TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics.* **19** (2018), , doi:10.1186/s12864-018-4620-2.
97. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
- 30 98. M. dos Reis, J. Inoue, M. Hasegawa, R. J. Asher, P. C. J. Donoghue, Z. Yang, Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. Biol. Sci.* **279**, 3491–3500 (2012).
- 35 99. M. R. McGowen, G. Tsagkogeorga, S. Álvarez-Carretero, M. Dos Reis, M. Struebig, R. Deaville, P. D. Jepson, S. Jarman, A. Polanowski, P. A. Morin, S. J. Rossiter, Phylogenomic Resolution of the Cetacean Tree of Life Using Target Sequence Capture. *Syst. Biol.* **69**, 479–501 (2020).
100. S. Y. W. Ho, M. J. Phillips, Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* **58**, 367–380 (2009).

101. Z. Yang, B. Rannala, Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226 (2006).
102. M. N. Puttick, MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics.* **35**, 5321–5322 (2019).
- 5 103. R. F. S. Computing, R: A language and environment for statistical computing. *Vienna: R Core Team* (2013) (available at <https://www.yumpu.com/en/document/view/6853895/r-a-language-and-environment-for-statistical-computing>).
104. M. dos Reis, Z. Yang, Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Mol. Biol. Evol.* **28**, 2161–2172 (2011).
- 10 105. J. Gabry, D. Simpson, A. Vehtari, M. Betancourt, A. Gelman, Visualization in Bayesian workflow. *J. R. Stat. Soc. Ser. A Stat. Soc.* **182**, 389–402 (2019).
106. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* **35**, 526–528 (2019).
107. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and
15 Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995).
108. A. R. Ives, R² for Correlated Data: Phylogenetic Models, LMMs, and GLMMs. *Syst. Biol.* **68**, 234–251 (2019).
109. M. Meyer, M. Kircher, M.-T. Gansauge, H. Li, F. Racimo, S. Mallick, J. G. Schraiber, F.
20 Jay, K. Prüfer, C. de Filippo, P. H. Sudmant, C. Alkan, Q. Fu, R. Do, N. Rohland, A. Tandon, M. Siebauer, R. E. Green, K. Bryc, A. W. Briggs, U. Stenzel, J. Dabney, J. Shendure, J. Kitzman, M. F. Hammer, M. V. Shunkov, A. P. Derevianko, N. Patterson, A. M. Andrés, E. E. Eichler, M. Slatkin, D. Reich, J. Kelso, S. Pääbo, A high-coverage genome sequence from an archaic Denisovan individual. *Science.* **338**, 222–226 (2012).
- 25 110. K. Prüfer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinze, G. Renaud, P. H. Sudmant, C. de Filippo, H. Li, S. Mallick, M. Dannemann, Q. Fu, M. Kircher, M. Kuhlwilm, M. Lachmann, M. Meyer, M. Ongyerth, M. Siebauer, C. Theunert, A. Tandon, P. Moorjani, J. Pickrell, J. C. Mullikin, S. H. Vohr, R. E. Green, I. Hellmann, P. L. F. Johnson, H. Blanche, H. Cann, J. O. Kitzman, J. Shendure, E. E. Eichler, E. S. Lein, T. E. Bakken, L.
30 V. Golovanova, V. B. Doronichev, M. V. Shunkov, A. P. Derevianko, B. Viola, M. Slatkin, D. Reich, J. Kelso, S. Pääbo, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature.* **505**, 43–49 (2014).
111. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience.* **10** (2021), doi:10.1093/gigascience/giab008.
- 35 112. ALFA: Allele Frequency Aggregator, (available at <http://www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/>).
113. M. Lawrence, R. Gentleman, V. Carey, rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics.* **25**, 1841–1842 (2009).

114. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**, 841–842 (2010).
115. M. Kuhlwilm, C. Fontseré, S. Han, M. Alvarez-Estape, T. Marques-Bonet, HuConTest: Testing Human Contamination in Great Ape Samples. *Genome Biol. Evol.* **13** (2021), doi:10.1093/gbe/evab117.
116. A. B. Rylands, R. A. Mittermeier, "Primate names: Working taxonomic list of primates of the IUCN SSC Primate Specialist Group" in *Primate Behavioral Ecology, 6th edition*, K. B. Strier, Ed. (Routledge, New York, ed. 6th, 2021), pp. 407–428.
117. C. J. Burgin, J. P. Colella, P. L. Kahn, N. S. Upham, How many species of mammals are there? *J. Mammal.* **99**, 1–14 (2018).
118. Mammal Diversity Database, Mammal Diversity Database (2020), doi:10.5281/ZENODO.4139818.
119. E. Sayyari, S. Mirarab, Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Mol. Biol. Evol.* **33**, 1654–1668 (2016).
120. R. C. M. Warnock, Z. Yang, P. C. J. Donoghue, Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc. Biol. Sci.* **284**, 20170227 (2017).
121. R. W. Meredith, J. E. Janečka, J. Gatesy, O. A. Ryder, C. A. Fisher, E. C. Teeling, A. Goodbla, E. Eizirik, T. L. L. Simão, T. Stadler, D. L. Rabosky, R. L. Honeycutt, J. J. Flynn, C. M. Ingram, C. Steiner, T. L. Williams, T. J. Robinson, A. Burk-Herrick, M. Westerman, N. A. Ayoub, M. S. Springer, W. J. Murphy, Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science*. **334**, 521–524 (2011).
122. W. J. Murphy, Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science*. **294**, 2348–2351 (2001).
123. J. E. Tarver, M. Dos Reis, S. Mirarab, R. J. Moran, S. Parker, J. E. O'Reilly, B. L. King, M. J. O'Connell, R. J. Asher, T. Warnow, K. J. Peterson, P. C. J. Donoghue, D. Pisani, The Interrelationships of Placental Mammals and the Limits of Phylogenetic Inference. *Genome Biol. Evol.* **8**, 330–344 (2016).
124. L. Liu, J. Zhang, F. E. Rheindt, F. Lei, Y. Qu, Y. Wang, Y. Zhang, C. Sullivan, W. Nie, J. Wang, F. Yang, J. Chen, S. V. Edwards, J. Meng, S. Wu, Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E7282–E7290 (2017).
125. J. A. Esselstyn, C. H. Oliveros, M. T. Swanson, B. C. Faircloth, Investigating Difficult Nodes in the Placental Mammal Tree with Expanded Taxon Sampling and Thousands of Ultraconserved Elements. *Genome Biol. Evol.* **9**, 2308–2321 (2017).
126. W. J. Murphy, N. M. Foley, K. R. Bredemeyer, J. Gatesy, M. S. Springer, Phylogenomics and the genetic architecture of the placental mammal radiation. *Annu Rev Anim Biosci.* **9**, 29–53 (2021).

127. M. A. O’Leary, J. I. Bloch, J. J. Flynn, T. J. Gaudin, A. Giallombardo, N. P. Giannini, S. L. Goldberg, B. P. Kraatz, Z.-X. Luo, J. Meng, X. Ni, M. J. Novacek, F. A. Perini, Z. S. Randall, G. W. Rougier, E. J. Sargis, M. T. Silcox, N. B. Simmons, M. Spaulding, P. M. Velazco, M. Weksler, J. R. Wible, A. L. Cirranello, The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science*. **339**, 662–667 (2013).
128. R. J. Asher, N. Bennett, T. Lehmann, The new framework for understanding placental mammal evolution. *Bioessays*. **31**, 853–864 (2009).
129. R. J. Asher, M. R. Smith, A. Rankin, R. J. Emry, Congruence, fossils and the evolutionary tree of rodents and lagomorphs. *R Soc Open Sci*. **6**, 190387 (2019).
130. V. C. Mason, G. Li, P. Minx, J. Schmitz, G. Churakov, L. Doronina, A. D. Melin, N. J. Dominy, N. T.-L. Lim, M. S. Springer, R. K. Wilson, W. C. Warren, K. M. Helgen, W. J. Murphy, Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates. *Sci Adv*. **2**, e1600633 (2016).
131. J. E. Janecka, W. Miller, T. H. Pringle, F. Wiens, A. Zitzmann, K. M. Helgen, M. S. Springer, W. J. Murphy, Molecular and genomic data identify the closest living relative of primates. *Science*. **318**, 792–794 (2007).
132. J. Lin, G. Chen, L. Gu, Y. Shen, M. Zheng, W. Zheng, X. Hu, X. Zhang, Y. Qiu, X. Liu, C. Jiang, Phylogenetic affinity of tree shrews to Glires is attributed to fast evolution rate. *Mol. Phylogenet. Evol.* **71**, 193–200 (2014).
133. X. Zhou, F. Sun, S. Xu, G. Yang, M. Li, The position of tree shrews in the mammalian tree: Comparing multi-gene analyses with phylogenomic results leaves monophyly of Euarchonta doubtful. *Integr. Zool.* **10**, 186–198 (2015).
134. Z. Kielan-Jaworowska, R. L. Cifelli, Z.-X. Luo, *Mammals from the Age of Dinosaurs: Origins, Evolution, and Structure* (Columbia University Press, 2005).
135. G. P. Wilson Mantilla, S. G. B. Chester, W. A. Clemens, J. R. Moore, C. J. Sprain, B. T. Hovatter, W. S. Mitchell, W. W. Mans, R. Mundil, P. R. Renne, Earliest Palaeocene purgatoriid and the initial radiation of stem primates. *R Soc Open Sci*. **8**, 210050 (2021).
136. M. J. Phillips, C. Fruciano, The soft explosive model of placental mammal evolution. *BMC Evol. Biol.* **18**, 104 (2018).
137. D. Marjanović, The Making of Calibration Sausage Exemplified by Recalibrating the Transcriptomic Timetree of Jawed Vertebrates. *Front. Genet.* **12**, 521693 (2021).
138. W. J. Murphy, E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, S. J. O’Brien, Molecular phylogenetics and the origins of placental mammals. *Nature*. **409**, 614–618 (2001).
139. H. J. Chatterjee, S. Y. W. Ho, I. Barnes, C. Groves, Estimating the phylogeny and divergence times of primates using a supermatrix approach. *BMC Evol. Biol.* **9**, 259 (2009).
140. L. Pozzi, J. A. Hodgson, A. S. Burrell, K. N. Sterner, R. L. Raaum, T. R. Disotell, Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Mol. Phylogenet. Evol.* **75**, 165–183 (2014).

141. K. Finstermeier, D. Zinner, M. Brameier, M. Meyer, E. Kreuz, M. Hofreiter, C. Roos, A mitogenomic phylogeny of living primates. *PLoS One*. **8**, e69504 (2013).
142. G. Hartig, G. Churakov, W. C. Warren, J. Brosius, W. Makałowski, J. Schmitz, Retrophylogenomics place tarsiers on the evolutionary branch of anthropoids. *Sci. Rep.* **3**, 1756 (2013).
143. N. M. Jameson, Z.-C. Hou, K. N. Sterner, A. Weckle, M. Goodman, M. E. Steiper, D. E. Wildman, Genomic data reject the hypothesis of a prosimian primate clade. *J. Hum. Evol.* **61**, 295–305 (2011).
144. P. E. Morse, S. G. B. Chester, D. M. Boyer, T. Smith, R. Smith, P. Gigase, J. I. Bloch, New fossils, systematics, and biogeography of the oldest known crown primate *Teilhardina* from the earliest Eocene of Asia, Europe, and North America. *J. Hum. Evol.* **128**, 103–131 (2019).
145. R. L. Anemone, B. Nachman, North American Primate Fossil Record. *The International Encyclopedia of Primatology*, 1–5 (2017).
146. M. Vianey-Liaud, R. Minwer-Barakat, L. Marivaux, The rodents from Santpedor-2 (Ebro Basin, NE Spain) confirm the Oligocene age of the latest primates from the Paleogene of Europe. *Geobios Mem. Spec.* **55**, 77–88 (2019).
147. K. C. Beard, A new genus of Tarsiidae (Mammalia: Primates) from the middle Eocene of Shanxi Province, China, with notes on the historical biogeography of tarsiers. *Bull. Carnegie Mus. Nat. Hist.* **34**, 260–277 (1998).
148. E. L. Simons, "The fossil record of tarsier evolution" in *Tarsiers: Past, present and future*, P. C. Wright, E. L. Simons, S. Gursky, Eds. (Rutgers University Press, New Brunswick, 2003), pp. 9–34.
149. X. Ni, Q. Li, C. Zhang, K. Samiullah, L. Zhang, Y. Yang, W. Cao, Paleogene mammalian fauna exchanges and the paleogeographic pattern in Asia. *Sci. China Earth Sci.* **63**, 202–211 (2020).
150. G. F. Gunnell, D. M. Boyer, A. R. Friscia, S. Heritage, F. K. Manthi, E. R. Miller, H. M. Sallam, N. B. Simmons, N. J. Stevens, E. R. Seiffert, Fossil lemurs from Egypt and Kenya suggest an African origin for Madagascar's aye-aye. *Nat. Commun.* **9**, 3193 (2018).
151. J. R. Ali, M. Huber, Mammalian biodiversity on Madagascar controlled by ocean currents. *Nature*. **463**, 653–656 (2010).
152. K. E. Samonds, L. R. Godfrey, J. R. Ali, S. M. Goodman, M. Vences, M. R. Sutherland, M. T. Irwin, D. W. Krause, Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor type. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5352–5357 (2012).
153. J. C. Masters, F. Génin, Y. Zhang, R. Pellen, T. Huck, P. P. A. Mazza, M. Rabineau, M. Doucouré, D. Aslanian, Biogeographic mechanisms involved in the colonization of Madagascar by African vertebrates: Rifting, rafting and runways. *J. Biogeogr.* **48**, 492–510 (2021).

154. J. C. Masters, F. Génin, D. Silvestro, A. M. Lister, M. DelPero, The red island and the seven dwarfs: body size reduction in Cheirogaleidae. *J. Biogeogr.* **41**, 1833–1847 (2014).
155. F. Génin, J. C. Masters, "The physiology of phyletic dwarfism in Cheirogaleidae" in *The Dwarf and Mouse Lemurs of Madagascar: Biology, Behavior and Conservation Biogeography of the Cheirogaleidae*, S. M. Lehman, U. Radespiel, E. Zimmermann, Eds. (Cambridge University Press, Cambridge, 2016), pp. 317–344.
156. S. H. Montgomery, N. I. Mundy, Parallel episodes of phyletic dwarfism in callitrichid and cheirogaleid primates. *J. Evol. Biol.* **26**, 810–819 (2013).
157. G. J. Slater, L. J. Harmon, M. E. Alfaro, Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution.* **66**, 3931–3944 (2012).
158. K. E. Samonds, L. R. Godfrey, J. R. Ali, S. M. Goodman, M. Vences, M. R. Sutherland, M. T. Irwin, D. W. Krause, Imperfect Isolation: Factors and Filters Shaping Madagascar's Extant Vertebrate Fauna. *PLoS ONE.* **8**, e62086 (2013).
159. E. R. Seiffert, E. L. Simons, Y. Attia, Fossil evidence for an ancient divergence of lorises and galagos. *Nature.* **422**, 421–424 (2003).
160. J.-J. Jaeger, K. C. Beard, Y. Chaimanee, M. Salem, M. Benammi, O. Hlal, P. Coster, A. A. Bilal, P. Düringer, M. Schuster, X. Valentin, B. Marandat, L. Marivaux, E. Métais, O. Hammuda, M. Brunet, Late middle Eocene epoch of Libya yields earliest known radiation of African anthropoids. *Nature.* **467**, 1095–1098 (2010).
161. S. López-Torres, M. T. Silcox, "What We Know (and Don't Know) About the Fossil Records of Lorises" in *Evolution, Ecology and Conservation of Lorises and Pottos*, K. A. I. Nekaris, A. M. Burrows, Eds. (Cambridge University Press, 2020), pp. 33–46.
162. S. López-Torres, K. R. Selig, A. M. Burrows, M. T. Silcox, "The toothcomb of *Karanisia clarki*" in *Evolution, Ecology and Conservation of Lorises and Pottos*, Nekaris, K. A. I., Burrows, A. M., Ed. (Cambridge University Press, 2020), pp. 67–75.
163. E. R. Seiffert, E. L. Simons, T. M. Ryan, Y. Attia, Additional remains of *Wadilemur elegans*, a primitive stem galagid from the late Eocene of Egypt. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11396–11401 (2005).
164. M. J. Phillips, Geomolecular Dating and the Origin of Placental Mammals. *Syst. Biol.* **65**, 546–557 (2016).
165. T. Harrison, "Later Tertiary Lorisiformes" in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (University of California Press, 2010), pp. 333–350.
166. L. J. Flynn, M. E. Morgan, New lower primates from the Miocene Siwaliks of Pakistan. *Interpreting the past: essays on human, primate, and mammal evolution in honor of David Pilbeam*, 81–102 (2005).
167. P. Mein, L. Ginsburg, Les mammifères du gisement miocène inférieur de Li Mae Long, Thaïlande: systématique, biostratigraphie et paléoenvironnement. *Geodiversitas.* **19**, 783–844 (1997).

168. M. Bond, M. F. Tejedor, K. E. Campbell Jr, L. Chornogubsky, N. Novo, F. Goin, Eocene primates of South America and the African origins of New World monkeys. *Nature*. **520**, 538–541 (2015).
169. K. E. Campbell Jr, P. B. O’Sullivan, J. G. Fleagle, D. de Vries, E. R. Seiffert, An Early Oligocene age for the oldest known monkeys and rodents of South America. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2105956118 (2021).
170. P.-O. Antoine, J. Yans, A. A. Castillo, N. Stutz, M. A. Abello, S. Adnet, M. A. Custódio, A. Benites-Palomino, G. Billet, M. Boivin, F. Herrera, C. Jaramillo, C. Martínez, F. Moreno, R. E. Navarrete, F. R. Negri, F. Parra, F. Pujos, J.-C. Rage, A. M. Ribeiro, C. Robinet, M. Roddaz, J. V. Tejada-Lara, R. Varas-Malca, R. Ventura Santos, R. Salas-Gismondi, L. Marivaux, Biotic community and landscape changes around the Eocene–Oligocene transition at Shapaja, Peruvian Amazonia: Regional or global drivers? *Glob. Planet. Change*. **202**, 103512 (2021).
171. N. M. Jameson Kiesling, S. V. Yi, K. Xu, F. Gianluca Sperone, D. E. Wildman, The tempo and mode of New World monkey evolution and biogeography in the context of phylogenomic analysis. *Mol. Phylogenet. Evol.* **82**, 386–399 (2015).
172. H. Schneider, I. Sampaio, The systematics and evolution of New World primates - A review. *Molecular Phylogenetics and Evolution*. **82** (2015), , doi:10.1016/j.ympev.2013.10.017.
173. C. G. Schrago, H. N. Seuánez, Large ancestral effective population size explains the difficult phylogenetic placement of owl monkeys. *Am. J. Primatol.* **81**, e22955 (2019).
174. A. L. Rosenberger, *New World Monkeys: The Evolutionary Odyssey* (Princeton University Press, 2020).
175. S. M. Ford, Callitrichids as phyletic dwarfs, and the place of the callitrichidae in platyrrhini. *Primates*. **21**, 31–43 (1980).
176. D. Silvestro, M. F. Tejedor, M. L. Serrano-Serrano, O. Loiseau, V. Rossier, J. Rolland, A. Zizka, S. Höhna, A. Antonelli, N. Salamin, Early Arrival and Climatically-Linked Geographic Expansion of New World Monkeys from Tiny African Ancestors. *Syst. Biol.* **68**, 78–92 (2019).
177. M. E. Morgan, K. L. Lewton, J. Kelley, E. Otárola-Castillo, J. C. Barry, L. J. Flynn, D. Pilbeam, A partial hominoid innominate from the Miocene of Pakistan: description and preliminary analyses. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 82–87 (2015).
178. I. Nengo, P. Tafforeau, C. C. Gilbert, J. G. Fleagle, E. R. Miller, C. Feibel, D. L. Fox, J. Feinberg, K. D. Pugh, C. Berruyer, S. Mana, Z. Engle, F. Spoor, New infant cranium from the African Miocene sheds light on ape evolution. *Nature*. **548**, 169–174 (2017).
179. C. C. Gilbert, A. Ortiz, K. D. Pugh, C. J. Campisano, B. A. Patel, N. P. Singh, J. G. Fleagle, R. Patnaik, New Middle Miocene Ape (Primates: Hylobatidae) from Ramnagar, India fills major gaps in the hominoid fossil record. *Proc. Biol. Sci.* **287**, 20201655 (2020).

180. M. Pickford, Y. Sawada, R. Tayama, Y.-K. Matsuda, T. Itaya, H. Hyodo, B. Senut, Refinement of the age of the Middle Miocene Fort Ternan Beds, Western Kenya, and its implications for Old World biochronology. *C. R. Geosci.* **338**, 545–555 (2006).
- 5 181. N. M. Young, L. MacLatchy, The phylogenetic position of *Morotopithecus*. *J. Hum. Evol.* **46**, 163–184 (2004).
182. D. R. Begun, M. C. Nargolwalla, L. Kordos, European Miocene hominids and the origin of the African ape and human clade. *Evol. Anthropol.* **21**, 10–23 (2012).
183. S. Worthington, thesis, New York University (2012).
184. J. B. Rossie, C. C. Gilbert, A. Hill, Early cercopithecoid monkeys from the Tugen Hills, Kenya. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 5818–5822 (2013).
- 10 185. M. Pickford, B. Senut, Hominoid teeth with chimpanzee- and gorilla-like features from the Miocene of Kenya: implications for the chronology of ape-human divergence and biogeography of Miocene hominoids. *Anthropol. Sci.* **113**, 95–102 (2005).
186. C. Roos, M. Kothe, D. M. Alba, E. Delson, D. Zinner, The radiation of macaques out of Africa: Evidence from mitogenome divergence times and the fossil record. *J. Hum. Evol.* **133**, 114–132 (2019).
- 15 187. M. J. Benton, P. C. J. Donoghue, R. J. Asher, M. Friedman, T. J. Near, J. Vinther, Constraints on the timescale of animal evolutionary history. *Palaeontol. Electronica.* **18**, 1–106 (2015).
- 20 188. H. P. Püschel, O. C. Bertrand, J. E. O'Reilly, R. Bobe, T. A. Püschel, Divergence-time estimates for hominins provide insight into encephalization and body mass trends in human evolution. *Nat Ecol Evol.* **5**, 808–819 (2021).

Acknowledgments:

Funding: LFKK was supported by an EMBO STF 8286. MCJ was supported by (NERC) NE/T000341/1. MK was supported by “la Caixa” Foundation (ID 100010434), fellowship code LCF/BQ/PR19/11700002, and by the Vienna Science and Technology Fund (WWTF) and the City of Vienna through project VRG20-001. JDO was supported by “la Caixa” Foundation (ID 100010434) and the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847648. The fellowship code is LCF/BQ/PI20/11760004. FES was supported by Brazilian National Council for Scientific and Technological Development (CNPq) (Process nos.: 303286/2014-8, 303579/2014-5, 200502/2015-8, 302140/2020-4, 300365/2021-7, 301407/2021-5, 301925/2021-6), and received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801505. Fieldwork for samples collected in the Brazilian Amazon was funded by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/SISBIOTA Program #563348/2010-0), Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM/SISBIOTA #2317/2011), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES AUX # 3261/2013) to IPF. Sampling of nonhuman primates in Tanzania was funded by the German Research Foundation (KN1097/3-1 to SK and RO3055/2-1 to CR) and by the US National Science Foundation (BNS83-03506 to JPC). No animals in Tanzania were sampled purposely for this study. Details of the original study on *Treponema pallidum* infection can be requested from SK. Sampling of baboons in Zambia was funded by US NSF grant BCS-1029451 to JPC, CJJ and JR. The research reported in this manuscript was also funded by the Vietnamese Ministry of Science and Technology’s Program 562 (grant no. ĐTDL.CN-64/19). ANC is supported by AEI-PGC2018-101927-BI00 704 (FEDER/UE), FEDER (Fondo Europeo de Desarrollo Regional)/FSE (Fondo Social Europeo), “Unidad de Excelencia María de Maeztu”, funded by the AEI (CEX2018-000792-M) and Secretaria d’Universitats i Recerca and CERCA Programme del Departament d’Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 880). ADM was supported by the National Sciences and Engineering Research Council of Canada and Canada Research Chairs program. The authors would like to thank the Veterinary and Zoology staff at Wildlife Reserves Singapore for their help in obtaining the tissue samples, as well as the Lee Kong Chian Natural History Museum for storage and provision of the tissue samples. We wish to thank H. Doddapaneni, D.M. Muzny and M.C. Gingras for their support of sequencing at the Baylor College of Medicine Human Genome Sequencing Center. We greatly appreciate the support of Richard Gibbs, Director of HGSC for this project and thank Baylor College of Medicine for internal funding. TMB 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), BFU2017-86471-P (MINECO/FEDER, UE), “Unidad de Excelencia María de Maeztu”, funded by the AEI (CEX2018-000792-M), Howard Hughes International Early Career, Obra Social “La Caixa” and Secretaria d’Universitats i Recerca and CERCA Programme del Departament d’Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 880). JPB, RMDB, IG and DV were supported by a UKRI Grant NERC (NE/T000341/1). We thank Dr. Praveen Karanth (IISc), Dr. H.N. Kumara (SACON) for collecting and providing us with some of the samples from India. SMA was supported by a BINC fellowship from the Department of Biotechnology (DBT), India. We acknowledge the

support provided by the Council of Scientific and Industrial Research (CSIR), India to GU for the sequencing at the Centre for Cellular and Molecular Biology (CCMB), India.

Author contributions:

- 5 Conceptualization: TMB, KF, JR
- Methodology & Analysis: LFKK, HG, MCJ, MK, JDO, SMA, AV, JB, MRO, RMDB, TMB, KF, JR, TB, YS, LZ, JGS, DDV, IGO, AJ, JPB, MR, RAH
- Fieldwork & Sample acquisition: JPB, CR, GU, KG, FES, FRDM, FB, HB, IS, IF, JVDS, MM, MNFDS, MT, RR, TH, AM, DZ, ACK, WKL, CCK, PT, JL, SME, MDL, SKN, JDK, FS, EFD, JHS, CA, GW, JPC, CJJ, AZ, CJR, NA, CH, PF, ISC, JH, JR
- 10 Topic Section Leaders: LFKK, JPB, MHS, RMDB, KG, CR, GU, AM, TMB
- Sequencing: LA, MG, JH, JB, IGU, EL, RAH, MRA
- Supervision: TMB, KF, JR, MHS, RMDB, GZ, DW, DJ, JPB
- Writing – original draft: LFKK, TMB
- 15 Writing – review & editing: All authors

Competing interests: LFKK, HG, JGS, and KF are employees of Illumina Inc. as of the submission of this manuscript.

- 20 **Data and materials availability:** All sequencing data has been deposited at the European Nucleotide Archive under the accession number PRJEB49549

Supplementary Materials

- Materials and Methods
- Supplementary Text
- 25 Figs. S1 to S111
- Tables S1 to S30
- Data S1 to S6



Supplementary Materials for

A global catalog of whole-genome diversity from 233 primate species

Kuderna et al.

Corresponding authors: Lukas F.K. Kuderna, lkuderna@illumina.com; Jeffrey Rogers, jr13@bcm.edu;
Kyle Farh, kfarh@illumina.com; Tomas Marques Bonet, tomas.marques@upf.edu

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S111
Tables S1 to S30
Captions for Data S1 to S6

Other Supplementary Materials for this manuscript include the following:

Data S1 – S6

Materials and Methods.....	1
Library preparation and whole-genome sequencing.....	1
Sampling.....	2
Mapping and Genotyping.....	3
Runs of Homozygosity.....	6
Loss of function variation.....	6
Inferences of Demographic History.....	7
UCE identification and alignment generation.....	9
Gene tree and species tree generation.....	10
Mitochondria assembly and phylogeny.....	11
Divergence Dating Analysis.....	12
Estimates of mutation rates and effective population sizes.....	13
PGLS models for traits as predictors of diversity and mutation rates.....	14
Correlates of mutation rate variation among primates.....	15
Human-specific and ape-specific variants.....	17
Effects of species cross mappings on estimates of diversity.....	19
Phylogenetic analysis.....	21
Supplementary Figures.....	26
Supplementary Tables.....	137
Data S1. (separate file).....	171
Data S2. (separate file).....	171
Data S3. (separate file).....	171
Data S4. (separate file).....	171
Data S5. (separate file).....	171
Data S6. (separate file).....	171

Materials and Methods

Library preparation and whole-genome sequencing

All samples were gathered in accordance with and abiding by local laws and regulations. The short-inserted paired-end libraries for the whole genome sequencing were prepared with PCR-free protocol using KAPA HyperPrep kit (Roche), with some modifications. In short, depending on the available material 0.1 - 1.0 microgram of genomic DNA was sheared on a Covaris™ LE220-Plus (Covaris) in order to reach the average fragment size of ~300bp. The fragmented DNA was size-selected for the fragment size of 220-550bp with AMPure XP beads (Agencourt, Beckman Coulter). The size selected genomic DNA fragments were end-repaired, adenylated and Illumina platform compatible adaptors with unique dual indexes and unique molecular identifiers (UMI, integrated DNA Technologies) were ligated. The libraries were quality controlled on an Agilent 2100 Bioanalyzer with the DNA 7500 assay (Agilent) for size and quantified by Kapa Library Quantification Kit for Illumina platforms (Roche). Library with final molarity below 3nM underwent PCR amplification of 6 - 10 cycles using KAPA Library Amp Primer Mix (Roche) and KAPA HiFi HotStart ReadyMix PCR Kit (Roche).

The libraries were sequenced on NovaSeq6000 (Illumina) in paired-end mode with a read length of 2x151+17+8bp following the manufacturer's protocol for dual indexing. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 3.4.4) and followed by generation of FASTQ sequence files.

Sampling

Research permissions and institutional approvals

The DNA samples analyzed in this project were obtained by various independent researchers over decades. Some samples were collected in the field or from zoos as long ago as the mid-1960s, with a substantial fraction of biomaterials collected prior to 2000. All necessary permits and approvals required at the time of collection and export were obtained from the relevant government agencies. Approval for both the collection and sequencing of samples was obtained by individual investigators from their home institutions, as required by the policies of each host institution. Some samples were obtained from live animals housed in zoos or other captive facilities, and in these cases approval for collection was also obtained from the institution with responsibility for care of the animal.

The habitat countries that generously granted permission for collecting and exporting primate samples were: Argentina, Brazil, Costa Rica, Ethiopia, Indonesia, Madagascar, Senegal, Tanzania, Vietnam and Zambia. Samples from primates in India were obtained under all appropriate permissions and approvals, and all of them were sequenced in India.

Procedures for sample collection

All samples from live animals were obtained using recognized and widely accepted procedures for the safe collection of biomaterials, following appropriate veterinary practice for the particular species and its research context (i.e., field, zoo or other collection circumstance). Wherever required, animal handling and biomaterials collection procedures for live animals were reviewed and approved by the home institution of the researcher conducting the study. Depending on the species, age and circumstances of each animal subject (wild-caught, captive in zoo, captive in research institution, etc.) collection of biomaterials involved either collection of blood from femoral vein, or collection of skin biopsy from ear, axilla, or other appropriate site. All samples were gathered in accordance with and abiding by local, national and international law and regulations, and established veterinary standards. Many of the samples analyzed in this project consisted of archived material collected from dead animals and provided by zoos, universities or other institutions with relevant biobanks.

Mapping and Genotyping

Given the relative scarcity of primate reference assemblies and the challenges of short read genome assemblies, we opted for a cross species reference mapping approach to analyze the resequencing data, using newly assembled and publicly available genomes (see Table S1). To this end, we first removed potential remaining adaptor sequences from the read pairs with cutadapt (v1.18) after interleaving with seqtk (v 1.2-r95-dirty) and removed all fragments with a length below 30 bp after trimming:

```
seqtk mergepe read_1.fastq.gz read_2.fastq.gz | \  
cutadapt -j 8 --interleaved -m 30 -a  
AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -A  
AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCA  
TT -
```

We used bwa mem (v 0.7.15-r1140) to map the trimmed data to appropriate reference genomes (see Table S1.), and the samtools suite (v 1.9) to sort and process the mappings:

```
bwa mem -pt 8 reference.fasta - | \  
samtools sort -@ 8 -m 3G -O bam -T $TMPDIR/part.map.sorted.TMP -o  
part.mapped.sorted.bam
```

We then merged all individual sequencing parts belonging to a given sample using samtools merge, and marked duplicate sequences using bammarkduplicates from the biobambam suite (v 2.0.35). We additionally generated indices for the joint mappings using samtools index:

```
samtools merge -@8 -b file_of_bam_paths - | \  
bammarkduplicates O=markdup.merged.bam markthreads=8 tmpfile=$TMPDIR  
samtools index markdup.merged.bam
```

For downstream compatibility, we added read groups to the mappings with AddOrReplaceReadGroups from the picard suite (v 2.8.2), and re-generated indices with samtools index:

```
java -jar picard.jar AddOrReplaceReadGroups I=markdup.merged.bam  
O=markdup.merged.addRG.bam RGID=id RGLB=lib1 RGPL=illumina RGPU=unit1 RGSM=id  
;  
samtools index markdup.merged.addRG.bam;
```

We used the resulting mappings as input for GATK HaplotypeCaller (v 4.1.6.0) to call variants. To this end, we generated bed files that partition the references into chunks of roughly 30 Mb to parallelize the computation. We then called variants for all chunks and samples on a per-sample basis with base-pair resolution in the following way:

```
java -jar gatk-package-4.1.6.0-local.jar HaplotypeCaller -R reference.fasta -I
markdup.merged.addRG.bam -ERC BP_RESOLUTION -L chunk.bed -O
chunk.raw.snps.indels.g.vcf.gz
```

We then genotyped the resulting gVCFs using GATK's GenotypeGVCFs:

```
java -jar gatk-package-4.1.6.0-local.jar GenotypeGVCFs --include-non-variant-sites -R
chunk.raw.snps.indels.g.vcf.gz --variant chunk.raw.snps.indels.g.vcf.gz -O
chunk.genotyped.g.vcf.gz"
```

We calculated the coverage distribution of mapped bases to estimate coverage-based filtering cutoffs on an individual basis (see Fig. S1 and Fig. S2). To this end, we used mosdepth (v 0.2.9) to generate a base-pair resolution coverage mask, from which we calculated the coverage histogram. We used the mode coverage values (excluding zero covered bases) to generate minimum and maximum coverage cutoffs, by removing regions that had below $\frac{1}{3}$ x mode or above 2 x mode coverage in each sample. We additionally required a hard cutoff of at least 3 reads supporting each allele in heterozygous calls. We further used the following set of hard filtering cutoffs to exclude single nucleotide variants as recommended by the developers of GATK:

```
QD < 2
FS > 60
MQ < 40
SOR > 3
ReadPosRankSum < -8.0
MQRankSum < -12.5
```

For indel calls, we used the following set of hard filter cutoffs:

```
QD < 2
FS > 200
MQ < 40
SOR > 3
ReadPosRankSum < -20.0
```

We applied the filters using bcftools (v 1.9) for SNPs and indels in the following way (for SNPs and indels, respectively):

```
bcftools filter -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" -O z -o variable.filtered.HF.snps.vcf.gz
```

```
bcftools filter -e "TYPE!='indel' | (GT='het' & FMT/AD[*:*] < $MIN_HET_AD ) | FMT/DP <= $MIN_COV | FMT/DP >= $MAX_COV | QD < 2 | FS > 200 | MQ < 40 | SOR > 3 | ReadPosRankSum < -20.0" -O z -o variable.filtered.HF.indels.vcf.gz
```

Lastly, we generated individual callability masks by applying the following coverage and quality cutoffs to all variants, including monomorphic ones:

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

For all estimates of diversity, we additionally filtered heterozygous positions with an allele balance below 0.25 or above 0.75. For our final estimates of genome-wide heterozygosity, we calculated the number of heterozygous sites and divided them by the callable portion of the genome.

Runs of Homozygosity

To detect signatures of recent inbreeding, we aimed to estimate the proportion of the genome in segments of homozygosity of at least 1Mb in length. To this end, we generated sliding windows of 1Mb size with an offset of 200 kb and calculated the number of heterozygous calls falling within each of them, as described in (8). We only retained windows with at least 50% callable bases for further analyses. The density distribution of heterozygosity across these windows has variable shapes across species and individuals, which might be influenced by demographic history, distance to the reference assemblies, or potential genotyping errors. We sought to establish cutoffs to differentiate windows with an excess of low diversity, putatively derived from the effects of inbreeding. To this end, we calculated local minima in the density distribution of heterozygosity across 1Mb windows and classified all windows below the cutoff as autozygous. In cases where such a minimum could not be established, for example, due to an excess of inbreeding, we used the genus wide median threshold for that individual. We then merged overlapping windows together to generate the final RoH set. We caution that differences in assembly qualities will influence the proportion of the genomes in large tracts of autozygosity that are detectable, i.e. more fragmented references will potentially have fewer possible 1Mb windows due to limitations of contig sizes.

Loss of function variation

As differences in gene structure between the focal and reference species might result in potentially faulty calls of stop gain variants, we applied a stringent set of filtering criteria as described in Gao et al in this same issue. We quantified only homozygous alternative calls within positions that passed all filtering steps until random forest.

Inferences of Demographic History

We reconstructed the demographic history for each species through coalescent-based analysis using the Multiple Sequentially Markovian Coalescent (MSMC v2.1.2, <https://github.com/stschiff/msmc2>) (84). Since many species had only one sample sequenced, we chose one representative sample for each species when multiple samples were available, prioritizing samples passing the quality checks for genome-wide coverage and allele balance distribution. When MSMC2 is provided with a single unphased diploid sample, it computationally reduces to a variant of the Pairwise Sequentially Markovian Coalescent known as PSMC' but performs better than PSMC (83), especially at recent timescales (85). As small scaffolds and sex-chromosomes affect the inference of demographic history (86), we filtered out scaffolds shorter than 1Mb and retained only the autosomal variants. For each species, we then generated the input files per scaffold containing the positions of heterozygous sites and distance from the last observed heterozygous site using the MSMC tools package (<https://github.com/stschiff/msmc-tools>). We accounted for all the uncalled positions in the genome due to abnormal coverage (87) by providing the callability masks for each sample and filtered out scaffolds with less than 100Kb of high-quality genotyped sites. To alleviate the confounding effects of selection on demographic inference (88), we applied a negative mask of the genomic coordinates of the protein-coding sequences (CDS) in the respective reference genomes to filter out sites that could be potentially under selection. To maintain computational tractability for hundreds of species, we ran MSMC2 for ten iterations to maximize the likelihood using the Baum-Welch algorithm. We specified 30 time segments with 28 free parameters by merging the first two and the penultimate two time segments using the MSMC2 input pattern (1*2+25*1+1*2+1*1). To obtain the variance of the estimated coalescent rates, we generated three bootstrapped replicates per sample by creating 20 artificial chromosomes, each with 100 randomly sampled 1Mb blocks with replacement. We analyzed the block-bootstrap replicates with MSMC2 using the same parameters. The resulting time and coalescent rate estimates are scaled in units of per-generation mutation rates as described below.

To convert the scaled time and coalescent rates into time in years and effective population size, we assembled a dataset of best-known generation length for each species from the literature and estimated the per generation mutation rates (see supplementary data S1). We adopted the definition of generation length proposed by IUCN as “the average age of parents of the current cohort”. As it is one of the important parameters calculated and published in the IUCN Red List assessment of species, we first collected the data directly from the IUCN portal (<https://www.iucnredlist.org>). We then obtained the generation length for species with missing data from other sources where it was either calculated from age at first reproduction and reproductive lifespan (89) or statistically imputed for species with no data on reproduction (90). Finally, we assigned the congeneric median generation length for species without any estimates in the sources mentioned above, and retained the same generation length for species that were previously considered subspecies. We used estimate of mutation rates for each species as described below. Using the estimates of generation length and per-generation mutation rates, we converted the scaled time and coalescent estimates from MSMC2 into years and effective population size following the standard protocol (91). We excluded the last infinite time segment and inspected the plots of demographic history for any inconsistencies. In cases where there were clear signs of

overfitting characterized by extremely large estimates in the most recent time segment, we reduced the number of free parameters by merging an additional time segment and repeated the analysis.

UCE identification and alignment generation

We identified the locations of 5472 UCE tetrapod-specific probes detected with a probe length of 120 bp derived from a previous publication across all reference genomes used in this study (92). These regions are useful for phylogenetic inference, as their high conservation enables their detection confidently across a broad range of evolutionary distances. In addition to the primate references used throughout the project, we included genome assemblies of several outgroup species, namely Colugo, Treeshrew, Mouse, and Rabbit (NCBI assembly versions GCA_000696425.1, GCA_000181375.1, GCA_000001635.9, GCA_000003625.1). To detect the coordinates of the UCE probes, we aligned them to the reference assemblies using last-z through the parser implemented in the phyluce package (version v1.7.1) with the following parameters (34, 92):

```
phyluce_probe_run_multiple_lastzs_sqlite --db $assembly_uce_mappings.sqlite --  
output $assembly_uce_mappings --scaffoldlist $assembly --genome-base-path  
$assembly_base_path --probefile uce-5k-probes.fasta
```

We then filtered the resulting alignments and retained only mapping where at least 90% of the probe aligned with an average identity to the reference of at least 90%. We additionally removed all probes with more than one mapping across a given reference to avoid potentially wrong placements. We subsetted the original probes to retain only probes which we could confidently identify across at least 75% of all reference assemblies, excluding outgroups, to minimize issues with missing data. By this means, we were left with 3516 UCE probes for downstream analysis. As there is likely to be reduced phylogenetic signal within the ultraconserved regions themselves, we padded them with 500 bp on either side to define the loci for further analysis. We then generated the private UCE sequences for resequenced individuals by including all variants within the loci's regions, and masking out non-callable bases by replacing them with Ns. We generated multiple sequence alignments of all sequences belonging to a given probe using MAFFT (v 7.471) using the following parameters:

```
mafft mafft --auto --adjustdirection $sequences
```

We trimmed the resulting alignments using trimAl (v 1.2rev59) to remove any alignment column with more than 20% gaps across individual sequences.

```
trimal -gt 0.8 -in $alignments
```

The parameters and cutoffs above were chosen in accordance with recent benchmarks to reconstruct UCE phylogenies from similarly sized datasets (93)

Gene tree and species tree generation

For the calculations of gene- and species trees, we included one representative of each species from the resequenced individuals, together with sequences from the reference genomes in the alignments. For species with more than one individual available, we additionally included one further individual to detect potential paraphyletic placements in the resulting phylogenies.

We used the trimmed alignments to generate individual gene trees using a maximum likelihood approach implemented in *iqtree2* (v 2.1.2) (94). To this end, we identified the best-fitting substitution model for each locus using the *ModelFinder* algorithm implemented in *iqtree* and kept it for subsequent analysis (37). We assessed branch support within gene trees by calculating the SH-like approximate likelihood ratio test and performed 1000 ultrafast bootstrap replicates (36). We used the following parameters for this analysis:

```
iqtree2 -s $trimmed_alignment.fa -m MFP --alrt 1000 -B 1000 -pre iqt_output
```

We used the resulting 3516 consensus gene trees as input for a coalescent analysis to infer the underlying species tree by using the “accurate species tree algorithm” implemented in *Astral* (v. 5.15.1) (41). To minimize the potential impact or errors in a given gene tree, we calculated the distribution of branch lengths for all taxa across gene trees, and removed taxa from trees in which they have abnormally long branch lengths, as these are likely to represent errors. This was done with the algorithm implemented in *TreeShrink* (99), at a false positive rate of 0.05:

```
run_treeshrink.py -t input_gene_trees.tree > output_gene_trees.tree
```

We used the resulting pruned gene trees as input for *Astral*:

```
java -jar astral.jar -i input_gene_trees -o astral_tree.tree
```

We checked for paraphyletic placements of oversampled species within the resulting species topology. To obtain our final species tree topology, we removed the reference sequences with the exception of Human, Colugo, Treeshrew, Mouse, and Rabbit, and randomly retained one of the resequenced individuals for oversampled species (see Table S5).

As *Astral* only outputs branch length estimates in coalescent units, we additionally sought to generate a phylogeny with branch length estimates as substitutions per site. To this end, we gathered all UCE loci with full coverage across species, leaving 614. We used *iqtree2* to estimate a substitution model on this data and calculated the branch lengths on the fixed final species tree topology under the inferred GTR+F+R7 model in the following way:

```
iqtree2 -s alignments/ -te species_tree.tree -pre bl_estimates -m GTR+F+R7
```


Mitochondria assembly and phylogeny

We generated assemblies and annotations for the mitochondria of all individuals included in our nuclear phylogeny using MitoFinder. For individuals in our study that were included in Janiak et al., we used the mitochondrial assemblies and annotations generate therein. For all remaining ones, we ran MitoFinder with the same parameters as Janiak et al. (described below). We identified publicly available mitochondrial assemblies for different clades, that are used by MitoFinder to guide the assembly process (see Table S4). We randomly subsampled 5,000,000 genomic read pairs from each individual, and provided them as the input for MitoFinder using metaspades as the assembly engine, and the following parameters:

```
Mitofinder -j $prefix --metaspades --max-contig 1 -1 read_pair_1.fastq -2 read_pair_2.fastq  
-r mitochondrial_reference_bait.fasta -o 2 -m 80
```

Wherever this did not result in a full reconstruction, we increased the number of read pairs to 1,000,000,000, and set `--max-contig 0`. From the resulting mitochondrial assemblies we then generated individual multiple sequence alignments from each of the 13 protein coding genes using `mafft`, and trimmed the resulting alignment with `trimal`:

```
mafft --auto --adjustdirection unaligned.fasta > aligned.fasta  
trimal -gappyout -in aligned.fasta > aligned.trimmed.fasta
```

The resulting trimmed alignments were used as a concatenated input for `iqtree`. This included an evolutionary model that identified `mtVer+R6` as the best fit, and both 1000 bootstrap replicates and 1000 Shimodaira–Hasegawa approximate likelihood ratio test to estimate clade support:

```
iqtree -s input_proteins/ --alrt 1000 -B 1000 -pre output_prefix
```

Within the oversampled tree, we observe polyphyletic or paraphyletic placements for 17 species (see Table S5), 9 of which we also find to be polyphyletic or paraphyletic in the nuclear phylogeny. After sub setting to a single individual per species to obtain a species tree, we find all but 4 genera to cluster monophyletically: *Cercocebus*, *Mandrillus*, *Cercopithecus*, and *Allochrocebus* show polyphyletic relationships. All families cluster monophyletically, except for the Lorisidae. In this family, the genera *Perodicticus* and *Arctocebus* cluster as a sister group to the Galagidae. However, this relationship does not receive full bootstrap support (bootstrap support = 93%). The remaining relationships between families follow the order expected based on the nuclear phylogeny.

Divergence Dating Analysis

To date the divergences of the primate tree, we used the program MCMCTree, implemented in PAML 4.9j (100), on a subset of the UCEs used to generate the tree topology. We retained only UCEs for which we had full representation of all samples and all outgroups with at least 75% of the sequence complete. The final dataset consisted of 614 different UCEs, totaling 676,475 bp. To reduce the computational burdens of analyzing the entire alignment as a single partition or each element as a separate partition, we partitioned the dataset into 10 separate partitions, grouping the UCEs into the same partition according to similar relative rates of divergence, following (101). The pairwise distance between one haplorrhine (*Pan paniscus*) and one strepsirrhine (*Lemur catta*) - representing the deepest split within crown-clade Primates (see below) - was estimated for each UCE with the baseml program in PAML 4.9j, and UCEs were grouped into 10 partitions of 61-62 elements each (mean length of partitions = 67,647.5 bp), based on similar rates of divergence.

We calibrated our phylogenetic tree with 27 node calibrations based on the fossil record (42). These calibrations were defined using the following prior distributions (103): “skew-T”, where we are confident that the calibrating fossil is close to the time of divergence (this is approximately equivalent to the offset exponential distribution implemented in the programs BEAST and MrBayes); bound” (uniform), where the fossil record is poor, and we cannot be confident that the calibrating fossil is close to the time of divergence); and “lower” minimum bound only, where a maximum bound cannot be confidently assigned (see Table S3). As a default, MCMCTree treats bounds for lower and bound (uniform) distributions as “soft,” allowing a 2.5% chance of the age going beyond the bounds of the calibration (98). For skew-T distributions minimum ages are hard bounds, but we allowed for a 5% probability of exceeding the soft maximum age. A full justification for all 27 node calibrations, their associated distributions, and the specific calibrating fossils used is given in (42). Prior distributions for the node calibrations were calculated and added to the tree topology with the package MCMCTreeR (105), modified by us to accept minimum age only calibrations) in R v. 4.1.0 (106) .

Divergence dating within MCMCTree was done with the approximate likelihood calculation method detailed in (107). We used the HKY85+G5 substitution model, as this is the most complex substitution model implemented by MCMCTree. For the clock model, we used the autocorrelated rates model, as this has been found previously to show a better fit than the independent rates clock model to molecular sequence data for primates (40) and other mammals (109). We used default settings for the other parameters and priors, adjusted to a 1-million-year time scale. We carried out 10 replicates of 1 million generations in MCMCTree, sampling every 50 generations for a total of 20,000 samples from each run. We checked for convergence with the R package bayesplot (110) and then combined all 10 runs, discarding the first 10% as burn-in, for a total of 180,010 samples. Each of the 10 replicates supported very similar divergence times for the nodes (see Fig. S32.), both when measured as dispersion from the mean (mean variance = 0.005, range = 1.13e-6 - 0.121) and as the difference between the highest and lowest age estimate per node relative to that node’s age (mean = 1.96%, range = 0.2-3.8%); combining these replicates was sufficient to give an Effective Sample Size (ESS) of >200 for every node (mean ESS = 4025.4, range = 522.8-80,813.4).

Estimates of mutation rates and effective population sizes

We sought to infer the mutation rates of all species via their substitution rates derived from the phylogeny. To this end, we used both the fossil calibrated phylogeny, and our estimates of branch lengths in substitution per site to estimate the per generation substitution rates across all species. We calculated the terminal branch lengths on both phylogenies, and estimated the per generation mutation rate (μ) in the following way:

$$\mu \approx K = D \times g/2 \times T$$

D is the number of pairwise nucleotide differences on the terminal branch, T is the divergence time (in years) on the terminal branch, and g is the generation length in years. We compiled estimates of g for all species from the IUCN red list and the literature (see Supplementary Data S2). We caution that our estimates of μ are likely to be biased, as by using the substitution rates to approximate them we are not accounting for the potential effects of selection. However, we also note that other sources of biases, such as uncertainties in the true generation time, are likely to swamp the potential effects of selection not captured herein. To estimate the amount of variance in substitution rates explained by mutation rates, we compared our estimates to all species for which de-novo mutation rates estimates based trio whole genome sequencing have been generated and find 72% of variance in substitution rates to be explained by de-novo mutation rates (Spearman's $r=0.85$, $p=0.02$, see Table S6). We note that our estimates are not systematically biased towards lower values, as we find them to disperse on either side of the $x=y$ line (see Fig. S38).

We have excluded the estimates for *Microcebus murinus* from this comparison, as we find the trio-based estimates to be an outlier compared to the remaining ones. Formally, the studentized residuals for this species are more than two standard deviations below the mean predicted value in a linear regression for both estimates. However, beyond this formal criterium we also note that the estimates for this species are outliers in other ways: The study reports de-novo mutation spectrum that differs from the spectra reported for all other primate estimates, such as e.g. a Ti:Tv ratio of 0.96, or around half the ratio reported across other species. This is likely the result of a lower observed number of C>T transitions at CpG sites, whose estimated enrichment is less than half that observed across other species (e.g. Thomas et al. 2018, Besenbacher et al. 2019). Furthermore, the authors do not find a paternal bias in de novo mutations, which has been consistently uncovered across mammalian species. Importantly, we note that as the total number of mutations to estimate the rates is small, the addition or exclusion of single events can significantly change them, underlining the importance of consistent analytical choices for accurate cross-species comparisons of mutation rates (see Bergeron et al., 2022).

We used estimates of μ and pairwise heterozygosity to calculate estimates of effective population sizes across species, by estimating nucleotide diversity (π) via the median pairwise heterozygosity:

$$N_e = \pi / (4 \times \mu)$$

PGLS models for traits as predictors of diversity and mutation rates

To determine potential predictors of genetic diversity (median pairwise heterozygosity) and mutation rates, we gather quantitative trait annotations for several broad trait categories from (57), as well as social- and mating systems (58), and data on Natal Dispersal from (59) (see Table S9). For categorical variables (mating system, social system, natal dispersal) we generated dummy variables for each of the categories therein. To ease potential issues due to multiple testing, we sought to remove highly correlated traits within each of the broad trait categories. To this end, we calculated the correlation coefficients of all traits within a category, including only species without any missing data across the focal traits. Starting at the traits with the highest number of initial observations across species, we then remove all traits with a correlation coefficient above 0.7, leaving 33 traits for final comparison. In case two correlated traits had an equal number of observations, we arbitrarily chose one for subsequent analyses. The final list of analyzed traits can be found in Table S10.

To account for phylogenetic non-independence between species, we used phylogenetic generalized least squares models and modeled the expected covariance between observations. We conservatively modeled the covariance under a Brownian motion model as implemented in the `corBrownian` function of the `APE` package (116). We generated a model for each individual trait rather than using broader categories to facilitate the interpretation of the results. We additionally log-transformed the measurements of several traits based on the shape of their distribution, as well as heterozygosity and mutation rate as the response variable (see Table S9). To determine violations of PGLS heterozygosity model assumptions, we generated diagnostic plots (bivariate scatterplots, Q-Q plots, residuals vs fits, and studentized residual histograms) and removed outliers as defined by studentized residuals greater than ± 3). Additionally, we performed an allometric correction for traits that scale with body mass, to remove its potential confounding effects. To this end, we regressed female body mass onto each allometric trait in a PGLS model under Brownian motion as described above, and kept the residuals of the regression for subsequent analysis. To account for multiple testing, we used a false discovery rate control using the Benjamini-Hochberg procedure (117). We report the significance of results by setting a false discovery rate of 0.05. Diagnostic plots of significant results are presented in Fig. S39 - Fig. S42.

We additionally ran OLS and PGLS models to test the influence of genetic diversity (median pairwise heterozygosity) on levels of inbreeding (as measured by RoH) using Brownian motion. To this end we numerically coded extinction risk from 1-5. None of these models yielded statistically significant results ($p > 0.05$).

Correlates of mutation rate variation among primates

We used a series of linear models to test how much the variation in mutation rates among species can be explained by variation in generation time and effective population size. We used the $\log_{10}(\mu)$ as the response variable and three continuous predictors: generation length (in years), the effective population size estimated $\log_{10}(N_e)$, and terminal branch length (TBL). We use TBL as a technical covariate to account for the possible biasing effects of terminal branch length on our inference of mutation rates.

Models were fitted by either ordinary least squares (OLS) or using phylogenetic generalized least squares (PGLS) models. PGLS models were assuming either a Brownian motion model or Pagel's phylogenetic regression model. Fitting using OLS amounts to assuming zero phylogenetic inertia and therefore that values in the primate species can be treated as independent data points. This is defensible if selection can change mutation rates on a time scale that is instantaneous on the phylogenetic scale considered here. Using the Brownian motion model amounts to assuming that mutation rates evolve neutrally and change at a rate directly proportional to the branch length estimated using UCEs. These fitting assumptions represent extremes in the range of assumptions one can make about how rapidly mutation rates can evolve. Pagel's model is more agnostic and is fitting jointly a parameter that tunes the amount of phylogenetic inertia in the model, we reported the AIC of each model, the R^2 of regression models as a measure of absolute goodness of fit and checked that these R^2 were relatively insensitive to our model assumptions. When fitting PGLS, we use a prediction R^2 , using the square of the correlation coefficient between observed and fitted values (118–123). As predicted R^2 is not computed via ratios of Deviance, in rare instances a model with a worse AIC might still exhibit a better R^2 .

All models were fitted in R by Maximum likelihood using the `gls()` function in the package `nlme`. We checked for an internal correlation between predictors (using the variance inflation function `vif()` implemented in the `car` package), checked visually for the homoscedasticity of model residuals as well as their normality (using a Shapiro-Wilk test). Statistical significance for the terms in the models (log effective size, generation length, and terminal branch length) was obtained by computing p-values for each effect using type II sum of squares (as implemented in the `car` package (`Anova()` R function)). We further checked that alternative linear models relying on different specifications of the effect of phylogenetic inertia (OLS assuming either zero phylogenetic inertia or Grafen's model that jointly estimates a magnitude of phylogenetic inertia) gave qualitatively similar results in terms of slopes estimates and overall proportion of variance explained (see Table S11).

Lastly, we also identified a conservative subset of 88 species that were deemed distant enough to minimize the effect of shared polymorphism. To do so, we iteratively removed one species out of any pair that was separated by less than $4N_e$ generations in our fossil calibrated phylogeny, i.e. the average time it takes for a new allele to be fixed due to genetic drift, until no more species pairs met that condition. We checked that the analysis we present using the full dataset - and phylogenetic least square linear models- is qualitatively robust to both model specification and the presence of closely related species in the original dataset. To do so, we refitted different regression models and investigated how much the presence or absence of closely related species, and choice of phylogenetic model could sway the proportion of variance explained by our model, as well as the effect size, i.e. the slopes estimates, inferred for each potential explanatory variable, i.e. the numerical regression covariate. We find our conclusions to be qualitatively remarkably robust to both these potential caveats. To illustrate this robustness, we compare the inference of mutation

rates using 3 different models below: one ignoring non independence due to shared phylogeny (OLS), a phylogenetic least square model using a Brownian motion model to specify the level of correlation between response variables (the model choice used throughout our phylogenetic least square analysis in the manuscript), as well PGLS Grafen, a more flexible model where the influence of shared phylogeny is not strictly proportional to phylogenetic distance as in the Brownian model, but is governed by a phylogenetic inertia parameter that can vary with each response variable. We find that, while the goodness of fit as measured by AIC varies considerably depending on the choice of models, the overall magnitude of R^2 and slopes estimates associated with $\log_{10}(N_e)$, Generation length, and terminal branch length are qualitatively remarkably stable across different model choices (see Table S11 - Table S12).

Human-specific and ape-specific variants

We strove to obtain the most complete set of differences between modern humans and archaic hominins to date. We used the genotypes of the three high-coverage Neandertals and the high-coverage Denisovan as published (cnda.eva.mpg.de, (71, 72, 108, 109)), and merged the four individuals using bcftools (v1.11, (110)). We also retrieved the filter bed files for the same individuals to obtain a high-confidence set.

Instead of relying on the human reference base as the human state, we determined the human majority allele at all positions where any genotype was called in the archaic individuals, making this approach robust to cases where the human reference allele is the minority allele in humans, but prevalent in archaics. We used the ALFA aggregate database (version 20201027095038XX, (111)) to determine the human majority allele. Since this database is based on the reference assembly GRCh38, while the archaic genotypes are based on the reference assembly hg19, we performed a liftover of all coordinates in archaics using the rtracklayer package (112), and retrieved the reference base using bedtools getfasta (113). After removing indels, we merged the data with the archaic genotypes, and transformed them from the vcf files to bases with bcftools query. We then retrieved sites that were showing >90% majority allele frequency in modern humans and <20% frequency in archaic hominins (i.e. tolerating singletons across 8 chromosomes). This resulted in a total of 445,981 sites across the 22 autosomes and the X chromosome.

This dataset includes differences between modern humans and archaic hominins, regardless of which of the two branches acquired a derived mutation. We then intersected these positions with the diversity of all great apes from published data (8, 76), processed as described previously (115–117), precisely, 43 gorillas, 10 bonobos, 59 chimpanzees and 27 orangutans. We retrieved the genotypes of these individuals at all positions passing the above filters, and applied the following filters on an individual level to the great ape genotypes: sequencing coverage between 6 and 100, mapping quality >20, reads with mapping quality of 0 at less than 10%. We then calculated the maximum allele frequency across all apes and in each species, as well as the human allele frequencies, and also provided a majority allele string of *Homo;Pan;Gorilla;Pongo*. We defined human lineage-specific changes as those changes where the human allele frequency was lower than 1% across all great apes, while the great apes carried the ancestral allele to more than 99% (to remove polymorphic positions across the great ape clade) and omitting sites where less than 10% of individuals had genotype data. This resulted in 166,917 positions that likely arose on the human lineage, among which 24,374 were almost fully fixed (>99.9%) in modern humans.

Finally, we intersected the data with the newly generated primate dataset. We retrieved the subset of sites for which at least 50% of individuals had genotype data, and where the majority allele was observed across 99% of individuals, while the human allele was observed at less than 1% or fixed.

Analogous to the above, we additionally generated a catalog of variants that are fixed across all great ape species (>99% frequency in at least 2/3 of individuals) and differ in state from rhesus macaque as a non-ape outgroup. Among the approximately 11.2M variants identified in this way, we find 32,033 missense variants affecting 11,541 different genes. While these coding variants should contain some of the substrate for ape specific evolution, the number of genes involved is too large to draw meaningful conclusions. In a functional enrichment analysis against the genome wide background, we observe only one general functional class enriched within them: “cellular process”. We furthermore conservatively excluded any variant that recurred in more than 1% of

callable species in other parts of the primate phylogeny, and was found at more than 1% frequency in our sampled individuals, which removed over 90% of sites. The remaining ~1M sites thus contain variants that have specifically changed at the root of all apes and did not remerge elsewhere in primates sampled in our data. Among them, we observe 3792 missense variants affecting 2970 different genes, provided in the supplementary Data S6. We performed a GO-term enrichment analyses including these genes and find that 30 terms contained in 10 hierarchies are significantly enriched (Fishers Exact test $p < 0.05$, BH-corrected FDR=0.05, see Table S13).

Effects of species cross mappings on estimates of diversity

To estimate the potential biases introduced by using reference assemblies based on species different from the focal species, we used the great ape clade as a test case. This family has a divergence time of ~20 Mya and has been characterized extensively on the genomic level, including multiple resequenced individuals from all species. The median pairwise distance between all great ape species and human is 9.1×10^{-3} substitutions per base pair, and thus 38% larger than the median pairwise distance between all species pairs included in our study (6.6×10^{-3} substitutions per base pair). It's worth pointing out that previous resequencing studies have relied on the human reference genome assembly to analyze different great apes, whose estimates of diversity are highly similar to others based on references from the same species (8, 9, 75, 76).

We used previously generated data from (8), and mapped and genotyped all individuals to both the human reference genome, and reference assemblies from the focal genus or species. To this end, we followed the same approach for each individual as described in the methods above using both references. We analyzed 25 Chimpanzees and 13 Bonobos mapped to the Chimpanzee reference, 31 Gorillas mapped to the Gorilla reference, and 5 Bornean and Sumatran Orangutans mapped to their respective species-specific reference. We observe an expected reduction in the proportion of callable sites in the human reference compared to the species-specific one, which ranges from 0.90 - 0.98 fold, with the lowest proportion in the Sumatran Orangutan, and the highest in Chimpanzees, concordant with the phylogenetic distance of these species to humans. We observed some fluctuations in the absolute number of heterozygous sites depending on the reference used, however, after accounting for differences in callable sequence space, we find estimates of heterozygosity to be highly correlated, at a Pearson's $r^2 = 0.993$ (see Fig. S3 - Fig. S4).

To further ensure our genome-wide estimates of diversity are not impacted by the choice of and distance to the reference genome, we additionally identified 19 species that were included in the sequencing data we have generated, and for which we also have a reference genome available (see Table S2. For each of these species, we identified an appropriate phylogenetically related species to formally test the effect of the distance to the reference species on our diversity estimates. We calculated pairwise distances between species pairs as the sum of branch lengths between them in our phylogeny, measured as substitutions per site. Across the 19 species pairs, these distances fully encompass the range observed across all species pairs of sequenced individuals and the respective reference genomes to which they were mapped in our analyses (see Fig. S6 - Fig. S7).

We performed variant calling and filtering, and identified the callable sequence space with identical parameters on both references as described in the Materials and Methods section. We find the resulting estimates of heterozygosity per base pair between the two references to be highly correlated with each other (Pearson's $r = 0.97$, $p = 6.8e-12$, see Fig. S8).

To test the influence of the distance to the reference genome, we performed a linear regression of the estimates based on heterozygosity estimated using distant reference versus heterozygosity using same species reference, and checked the correlation of the resulting residual values against the pairwise distance between the species. We do not find residual heterozygosity to be significantly influenced by the distance to the reference genome ($p = 0.25$, Spearman's r), and neither are the distant ($p = 0.5$ Spearman's r) and same species estimates of heterozygosity ($p = 0.46$, Spearman's r) estimates (see Fig. S9).

We also find no significant influence of the assembly quality on residual heterozygosity, when reference quality is evaluated from contig and scaffold N50 values. This holds for both close and distant assemblies.

The high correlation of heterozygosity estimates despite using references from a different species is attributable to our analytical choice of establishing a stringent callability mask for each species (see Materials and Methods). This base-pair resolution mask excludes genomic regions in the reference that fall outside the expected depth of coverage for the sequencing depth of a given individual, and serves as the denominator to calculate the per base heterozygosity. Therefore, particularly divergent and poorly mappable regions are excluded from the downstream analyses. The proportion of callable bases within each individual is highly correlated to the pairwise distances between the focal species and reference genome (Persons' $r = -0.92$, $p = 3.7e-8$, see Fig. S10), and thus correcting for it appropriately accounts for the effects this distance might have on estimates of heterozygosity. In an analysis that is uncorrected for callability, we find the genome-wide heterozygosity to exhibit poorer correlation across species pairs (Pearson's $r = 0.89$, $p = 3.2e-7$), and residual uncorrected heterozygosity to be significantly predicted by pairwise distance (Pearson's $r = -0.67$, $p = 0.0016$, see Fig. S11).

To obtain a more granular picture of potential confounding factors, we additionally ran window-wise comparisons across the aforementioned species pairs. To this end, we generated whole genome alignments across references of the species pairs, and lifted 100 kb sliding windows from the focal species reference to the distant one. To quantify the potential impact of variables beyond the distance to the reference genome, we regressed the estimates of heterozygosity of the distant genome onto the close one for each species, and calculated the correlation coefficients of the resulting residuals against GC-content and divergence. We find significant correlations ($p < 0.01$, Spearman's r) between residual heterozygosity and GC content in 7 species pairs, and between residual heterozygosity and divergence in 11 species pairs. However, while the relationships are significant, the effect sizes are comparatively small. For significant correlations, the median variance of heterozygosity calculated on distant species explained by GC content is 0.181% (0.003% - 1.057%), and the median variance explained by divergence is 0.254% (0.002% - 3.506%). We regard these influences as negligible for our genome-wide analyses comparisons.

Phylogenetic analysis

For our discussion of topology, we focus on relationships at the subfamily and above, although we note here that relationships at lower taxonomic levels are congruent with current classifications, with all currently recognized primate tribes and genera recovered as monophyletic (120–125). ASTRAL support values for nodes are local posterior probabilities (PP), and are calculated from gene tree quartet frequencies (120–124, 126–128). For our discussions of divergence dates, we focus on the 95% highest posterior density (HPD) estimates, as point estimates (e.g., means or medians) may not be particularly good estimates of true divergence dates, whereas 95% HPDs are more likely to encompass the true age of divergence (125, 129, 130).

Numerous large-scale molecular analyses carried out over the last twenty years (125, 131, 132) clearly indicate that the order Primates is a member of the superordinal clade Euarchontoglires, together with four other extant orders: tree shrews (Scandentia), colugos or flying lemurs (*Galeopterus*, Dermoptera), rodents (Rodentia), and lagomorphs (Lagomorpha). We included one member of each of these four other orders to act as outgroup taxa, namely the tree shrew *Tupaia belangeri*, the colugo *Galeopterus variegatus*, the rodent *Mus musculus* (the house mouse), and the lagomorph *Oryctolagus cuniculus* (the domestic rabbit). Monophyly of Glires has been strongly supported by previous molecular, morphological, and total evidence analyses (125). Colugos have been consistently found to be more closely related to Primates than to Glires (126), but the position of tree shrews within Euarchontoglires has proved difficult to resolve confidently (125, 129, 130); nevertheless, the majority of molecular evidence (133) supports monophyly of Euarchonta (Primates+Scandentia+Dermoptera), which is congruent with morphological data (134). We therefore rooted our ASTRAL phylogeny between Glires (= *Mus*+*Oryctolagus*) and Euarchonta (= *Tupaia*+*Galeopterus*+Primates).

Within Euarchonta, our tree places *Galeopterus* closer to Primates than to *Tupaia* with strong support (PP = 1), congruent with the majority of recent molecular studies that support colugos as the closest living relatives of primates (135); the orders Dermoptera and Primates together form the clade Primatomorpha. Our estimated ages for the interordinal divergences within Euarchonta - namely between Scandentia and Primatomorpha (95% HPD: 70.6-65.7 Ma), and between Dermoptera and Primates (95% HPD: 67.8-62.9 Ma) - overlap the K-Pg boundary, which is congruent with the absence of definitive euarchontans from Cretaceous fossil sites (136) but the presence of “plesiadapiforms” (plesiomorphic euarchontans, at least some of which appear to be members of Primatomorpha; see review by (42)) in the earliest Paleocene (139, 140).

Within crown-clade Primates (= Euprimates), our tree supports relationships that are highly congruent with other recent large-scale molecular analyses of primate phylogeny (21, 38–40). The deepest split within Primates is between Strepsirrhini (the so-called “wet-nosed” primates, which includes lorises, galagos, lemurs, and the aye-aye) and Haplorhini (the so-called “dry-nosed” primates, which include tarsiers, monkeys, and apes). The 95% HPD for this first divergence is 63.3-58.3 Ma, i.e. entirely within the Paleocene. This is younger than several other molecular clock analyses using large sequence datasets, which place the confidence interval for this divergence partially or entirely within the Late Cretaceous (38, 39, e.g., 40, 120). The younger date found here may be due, at least in part, to our use of a Skew-T calibration for the age of crown-clade Primates, with a soft maximum bound of 66 Ma (the K-Pg boundary) and only a 5% probability for a divergence older than this (see Table S3 and (42)). Nevertheless, we consider this calibration to be appropriate given the total lack of euprimates, stem-primates, or other definitive members of

Euarchontoglires (the clade that includes primates, dermopterans, scandentians, rodents, and lagomorphs) from Cretaceous deposits, as discussed by (42); see also (144) and (145). Vanderpool et al. (21) applied a similarly restrictive prior on the age of crown-clade Primates, with a soft maximum of 65.8 Ma, resulting in a mean estimate for the age of this node of 61.7 Ma, similar to that found here (mean = 60.7 Ma). A Paleocene age for crown-clade primates is in good agreement with the fossil record, as the oldest known crown-clade primates are latest Paleocene in age: the oldest well-supported haplorhine, *Teilhardina brandti*, dates to 55.985 ± 0.05 Ma, and the oldest well-supported strepsirrhine, *Donrussellia provincialis*, dates to 55.12–55.8 Ma (42).

Some previous analyses that used mitochondrial sequence data placed tarsiers (Tarsiiformes) sister to strepsirrhines (the “Prosimii” hypothesis) (149). However, some mitogenomic studies (146), have instead supported tarsiers as sister to monkeys and apes (Anthropoidea or Simiiformes), forming the clade Haplorhini, albeit with weak support. By contrast, retroposon insertions (150–152) and nuclear sequence data (21, 38, 120, 142) provide robust support for haplorhine monophyly. Our tree likewise strongly supports monophyly of Haplorhini (PP = 1), with Tarsiiformes as the sister-clade of Anthropoidea. We estimate the Tarsiiformes-Anthropoidea split to have occurred during the Paleocene or earliest Eocene (95% HPD: 60.4–55.4 Ma), broadly congruent with the oldest omomyiforms such as *Teilhardina* (the oldest records of which date to the latest Paleocene; see above) being early tarsiiforms (143).

Strikingly, the deepest divergence within crown Tarsiiformes, between *Tarsius* and *Cephalopachus*+*Carlito*, is very recent, dating to the middle-to-late Miocene (95% HPD: 15.2–9.5 Ma). No calibrations were specified for divergences within Tarsiiformes, and so this result is entirely due to the limited molecular divergence between the sampled members of this clade. Springer et al. (39) found a similarly young date in their molecular clock analysis when using an independent rates clock model, but not when using an autocorrelated rates model like the one used here. If this young date is correct, then the long branch leading to crown Tarsiiformes implies either very limited diversification or considerable extinction along the stem lineage. The latter interpretation seems more likely based on evidence from the fossil record: omomyiforms (which are probably stem tarsiiforms (144)) were diverse during the Eocene (145) but are last recorded in the early Oligocene (146), and representatives of the modern family Tarsiidae were once much more widely distributed across Asia (43–46, 146, 147), occurring as far west as southern Pakistan in the Miocene (44). The current restricted distribution of living tarsiers, on islands in southeast Asia, is therefore probably relictual. The young age of crown Tarsiiformes also indicates that fossil tarsiid remains from the middle Eocene (149) Shanghuang fissure fills in China that have been referred to the extant genus *Tarsius* (43) cannot belong to the crown clade, and so warrant referral to a separate genus (44, see also 146, 147).

Within Strepsirrhini, the Malagasy primates - the lemurs (Lemuriformes) and aye-aye (Chiromyiformes) - form a clade to the exclusion of the mainland African and Asian lorisiforms, which we estimate to have occurred during the early Eocene (95% HPD = 52.3–47.2 Ma). Within the Malagasy clade, the aye-aye (*Daubentonia madagascariensis*, the only living chiromyiform) is sister to the remaining taxa, which comprise Lemuriformes. *Daubentonia* and definitive lemuriforms are known only from Madagascar, but stem chiromyiforms (*Plesiopithecus* and *Propotto*) are known from mainland Africa (149), which implies that the Chiromyiformes-Lemuriformes split (95% HPD = 48.8–43.8 Ma) occurred in Africa, followed by independent dispersals of these lineages to Madagascar. Crown lemuriforms have not been found in Africa, and so our divergence estimates suggest that the time of dispersal to Madagascar by lemuriforms probably occurred after 52.3 Ma (the maximum age of the Chiromyiformes-Lemuriformes split),

but before 33.7 Ma (the minimum age of crown Lemuriformes). This is congruent with evidence for favorable ocean currents flowing eastward from mainland Africa to Madagascar during the Eocene (153–155). However, Masters et al. (156) argued that it is more likely that primates and other mammals dispersed between Africa and Madagascar via short-lived land bridges. Again based on the maximum age of the Chiromyiformes-Lemuriformes split, the timing of dispersal of the chiromyiform lineage to Madagascar probably occurred after 52.3 Ma, but cannot be further constrained by our divergence estimates. However, the total evidence clock analysis of Gunnell et al. (157) estimated the divergence between *Propotto* (the youngest known fossil chiromyiform from Africa) and *Daubentonia* to have occurred 35.8–20.7 Ma, in which case 20.7 Ma may be a plausible minimum age for the dispersal of the chiromyiform lineage to Madagascar.

Relationships within Lemuriformes are congruent with other recent molecular studies in supporting monophyly of the currently recognised families Cheirogaleidae (dwarf and mouse lemurs), Indriidae (woolly lemurs, sifakas and the indri), Lemuridae (the ring-tailed lemur, brown lemurs, ruffed lemurs, bamboo lemurs and relatives), and Lepilemuridae (sportive lemurs). Cheirogaleidae and Lepilemuridae form a clade, with Indriidae and Lemuridae successive outgroups to this. This topology is notable because it suggests that the small size of some dwarf and mouse lemurs (<100g, with a few species as small as 30g) may be secondary, the result of “phyletic dwarfing”, rather than a plesiomorphic retention (158, 159). However, robust testing of this hypothesis requires evidence from the fossil record (160), and no pre-Pleistocene primate fossils have been discovered in Madagascar to date (135, see also 161).

Within Lorisiformes, reciprocal monophyly of the two extant families - Galagidae (the galagos or bushbabies of Africa) and Lorisidae (lorises and pottos) - is supported. Several molecular clock studies have estimated a relatively ancient divergence time for the Galagidae-Lorisidae split, dating to the late Eocene (21, 38–40, 120). However, most of these studies assumed that the late Eocene (~37.5–37.0 Ma) *Saharagalago* and/or *Karanisia* (163, 164) are crown lorisiforms and have used this fossil evidence to calibrate this node, when in fact the position of these fossil taxa is quite unstable in published phylogenetic analyses, sometimes being placed outside the crown-clade Lorisiformes (42, 160, see summaries in 161). Use of the more conservative calibration preferred by de Vries and Beck (42) results in a much younger estimate for this divergence, dating to the latest Oligocene or early Miocene (95% HPD: 23.4–19.9 Ma). In turn, this young date provides additional evidence that the late Eocene *Saharagalago* and *Karanisia*, and also the slightly younger (earliest Oligocene, ~33.4 Ma) *Wadilemur* (135, see also 163), are not crown lorisiforms (160, 164).

Within Lorisidae, the subfamilies Perodictinae (the African pottos and angwantibos) and Lorisinae (the Asian slow and slender lorises) are reciprocally monophyletic. As noted above, the oldest definitive lorisiform appears to be the late Eocene *Saharagalago* from Africa, which is often placed outside the crown-clade in published phylogenetic analyses. Various crown lorisiforms are known from African fossil sites dating to the early Miocene onwards (160, 165), but remains of lorises have not been found in Africa. By contrast, the oldest well-preserved lorisiforms described from Asia date to the middle Miocene, and represent lorises only (166). Collectively, our tree and the evidence from the fossil record therefore suggest that the presence of lorises in Asia is the result of a single dispersal event from Africa. If so, then our divergence estimates indicate a time of dispersal after 21.7 Ma (the maximum age of the Perodictinae-Lorisinae split) but before 10.0 Ma (the minimum age of crown Lorisinae). Fragmentary remains of putative lorises from the early or middle Miocene of Thailand (?*Nycticebus linglom* (165)) and the middle

Miocene of Pakistan (*Nycticeboides* sp. (167)) suggest that the actual time of dispersal was probably somewhat earlier than 10.0 Ma.

The first split within Anthrooidea is between the platyrrhines of South and Central America (the so-called “New World monkeys”), and the catarrhines of Africa and Eurasia (the apes and so-called “Old World monkeys”) and is estimated to have occurred during the middle-to-late Eocene (95% HPD: 41.1-36.7 Ma). We estimate crown Platyrrhini to have begun diversifying during the late Oligocene or earliest Miocene (95% HPD: 26.5-22.7 Ma). If so, it is clear that the earliest fossil record of crown platyrrhines is missing, as the oldest definitive crown platyrrhines are middle Miocene (Laventan South American Land Mammal Age), with older fossils falling outside the crown clade (42). Given that crown platyrrhines are known only from South and Central America, our divergence estimates constrain the timing of dispersal of the platyrrhine lineage to the Americas to between 41.1 Ma (the maximum age of the Platyrrhini-Catarrhini split) and 22.7 Ma (the minimum age of crown Platyrrhini). However, this minimum dispersal time can be revised upwards to 29.52 Ma based on fossil evidence, specifically *Perupithecus* from the Santa Rosa site in Amazonian Peru (167), which has been robustly dated as 29.6 +/- 0.08 Ma (169), and which appears to be a stem platyrrhine (168). Potentially older South American primate fossils that may also represent stem platyrrhines have been reported from TAR-21 site in Shapaja, also in Amazonian Peru, which has been proposed to be ~34.3 Ma, i.e. latest Eocene (170, 171); if correct, this implies that the platyrrhine dispersal event occurred prior to 34.3 Ma. However, the reported age of TAR-21 has been questioned, and “tip-dating” analyses of fossil rodents appear to support an Oligocene age for this and other Shapaja sites (172).

Relationships within Platyrrhini are again congruent with other large-scale molecular studies (172), supporting a basal split between Pitheciidae (uakaris and titi monkeys) and a clade comprising the remaining families: Atelidae (howler, spider and woolly monkeys), Cebidae (squirrel and capuchin monkeys), Callitrichidae (tamarins and marmosets), and Aotidae (owl or night monkeys of the genus *Aotus*). Within the latter clade, the position of owl monkeys has proved difficult to resolve in previous studies (21, 172). Here, Aotidae is placed sister to Callitrichidae, in agreement with Schrago & Seuánez (155, 173, 174), albeit with low support (PP = 0.56); in fact, this is the only interfamilial relationship in our tree with PP < 1. By contrast, Vanderpool et al. (21) suggested that a closer relationship to Cebidae might be more likely, based on gene tree concordance factors, but the evidence in support of this alternative topology was not unequivocal. Based on this evidence, we consider that the precise relationship of Aotidae to Callitrichidae and Cebidae remains uncertain, the only interfamilial relationship within Primates for which this is the case. It is tempting to view this uncertainty as indicating a hard polytomy comprising these three families, but Schrago & Seuánez (176) and Vanderpool et al. (21) presented the results of statistical tests that appear to reject a hard polytomy, even though their analyses reached different conclusions as to the exact branching pattern. Similarly to cheirogaleids (see above), the position of the small-bodied callitrichids (the smallest of which, *Cebuella pygmaea*, weighs ~110g (178)) nested within a clade that otherwise comprises markedly larger species suggests that the small size of the former is the result of phyletic dwarfing (179), although this has been questioned (e.g., 180–182).

Within Catarrhini, the deepest split - between Cercopithecoidea (the so-called “Old World monkeys”) and Hominoidea (apes) - dates to the early Oligocene (95% HPD: 33.1-29.2 Ma). Relationships within Cercopithecoidea are again congruent with other recent large-scale molecular analyses in supporting reciprocal monophyly of the subfamilies Colobinae and Cercopithecinae, with the divergence between these two subfamilies estimated to have occurred during the early

Miocene (95% HPD: 20.6-18.0 Ma). Likewise, relationships within Hominoidea are as expected: there is a basal split between Hylobatidae (the gibbons or “lesser apes”), and Hominidae (the “great apes”), dated here to the latest Oligocene or early Miocene (95% HPD: 24.8-21.0 Ma). The first split within Hominidae, between Ponginae (orangutans) and Homininae (gorillas, chimps and humans), is dated to the early Miocene (95% HPD: 22.2-18.6 Ma). Given the rich record of African stem-hominids and stem-hominoids (reviewed by 42), the Ponginae-Homininae split presumably occurred in Africa. The oldest robustly-dated specimens of the earliest well-supported pongine, *Sivapithecus indicus*, from the Chinji Formation of Pakistan, are only 12.3 Ma old (184). It is therefore possible that the earliest stages of pongine evolution took place in Africa, an inference further supported by the phylogenetic analyses of Nengo et al. (185) and Gilbert et al. (182) that recovered *Kenyapithecus wickeri* from the middle Miocene (13.7 ± 0.3 Ma; (183) of Kenya as a stem-pongine; however, other analyses have placed *Kenyapithecus* as a stem-hominid (184); see summary by (42). Regardless, our divergence date estimate provides a maximum age for the timing of the dispersal of pongines to Asia, namely 22.2 Ma (our maximum estimate for the Ponginae-Homininae split).

Our estimate for the *Homo-Pan* divergence is slightly older than that of some other molecular clock analyses, with a 95% HPD of 9.0-6.9 Ma. This may be because of our conservative approach to calibrating this node: following de Vries and Beck (42), we specified a relatively young minimum bound of 4.631 Ma, based on the minimum age of well-dated specimens of the oldest well-supported member of the *Homo* lineage, *Ardipithecus ramidus*, but a relatively old maximum bound of 15 Ma, based on the poor record of fossil primates in Africa 15-6 Ma (123), and Pickford and Senut’s (124) suggestion that a ~12.5 Ma isolated lower molar from the Ngorora Formation may belong to the *Pan* lineage. This calibration was also modeled as a uniform distribution, again reflecting the poor African primate record 15-6 Ma (see 42). By contrast, some other molecular clock studies of primates (38, 40, 185, see also 186) have used more tightly constrained maximum bounds to calibrate this node, which will tend to favor younger dates, and which were rejected by de Vries and Beck (42). Regardless, our 95% HPD for the *Homo-Pan* split overlaps with those from other recent molecular clock (38–40) and total evidence clock (127) analyses.

Supplementary Figures

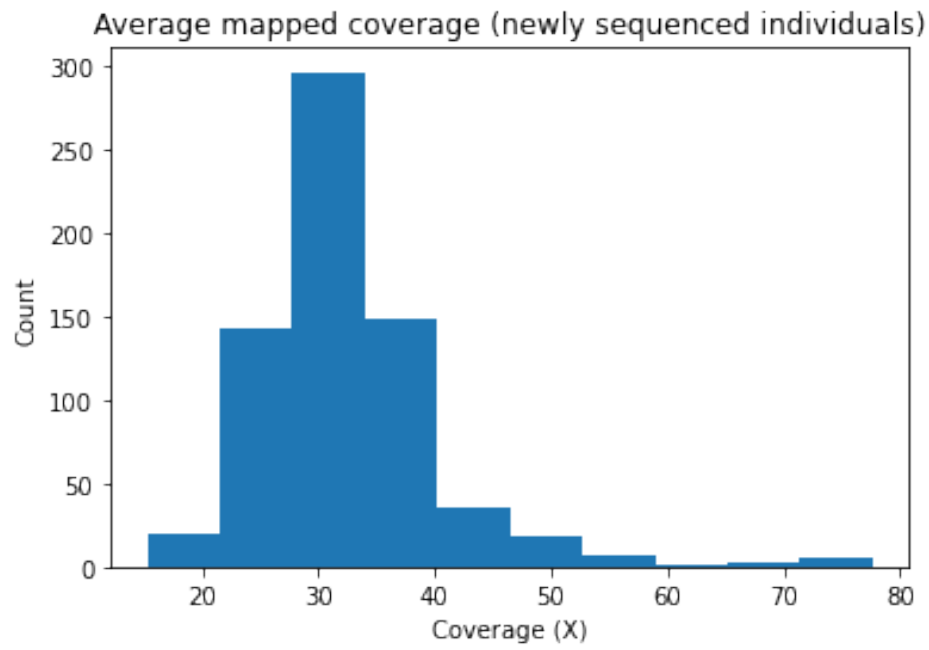


Fig. S1. Distribution of average mapped coverage of newly sequenced individuals in this study.

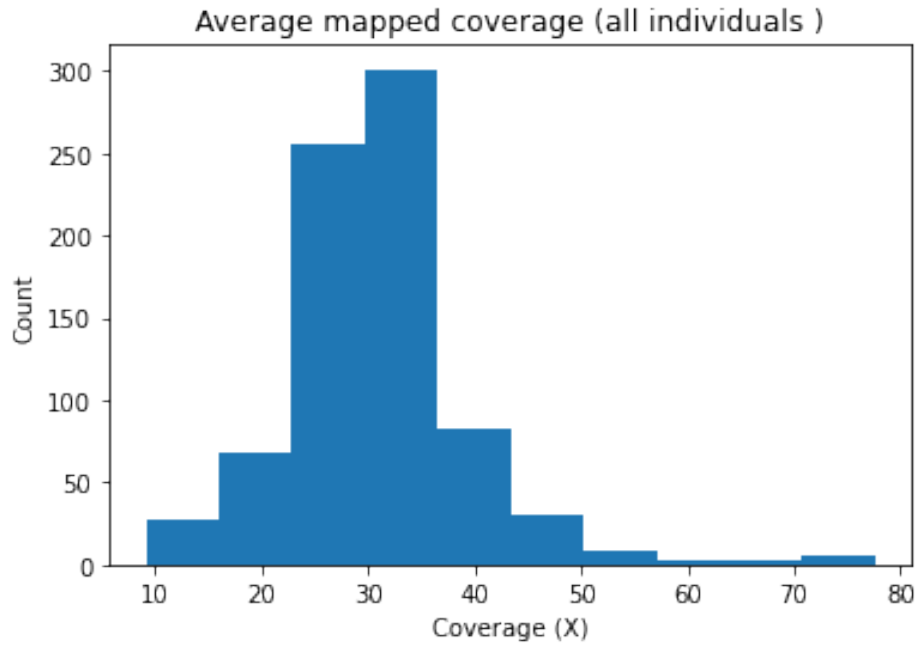


Fig. S2.

Distribution of average mapped coverage of all individuals in this study

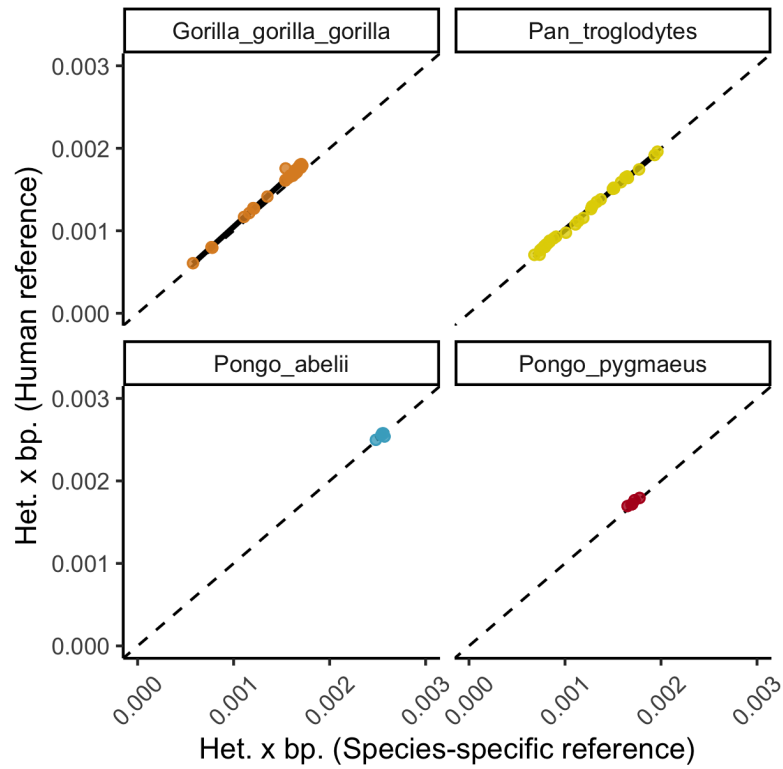


Fig. S3.

Concordance of heterozygosity estimates based on species specific and distant references split by species. Dashed lines denote theoretical $x=y$ slope. The left plot is split up by the different reference assemblies used for the great apes.

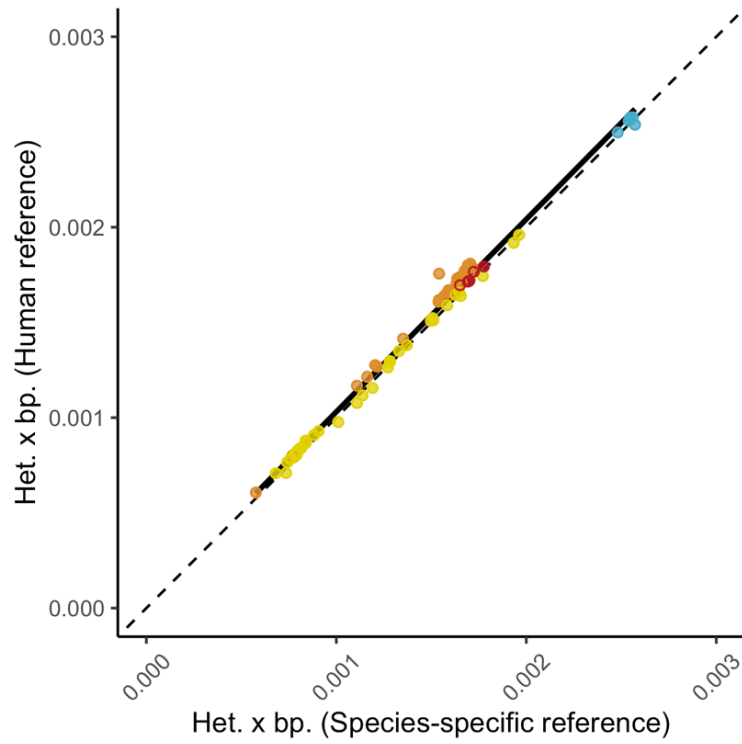


Fig. S4: Concordance of heterozygosity estimates based on species specific and distant references. Dashed lines denote theoretical $x=y$ slope.

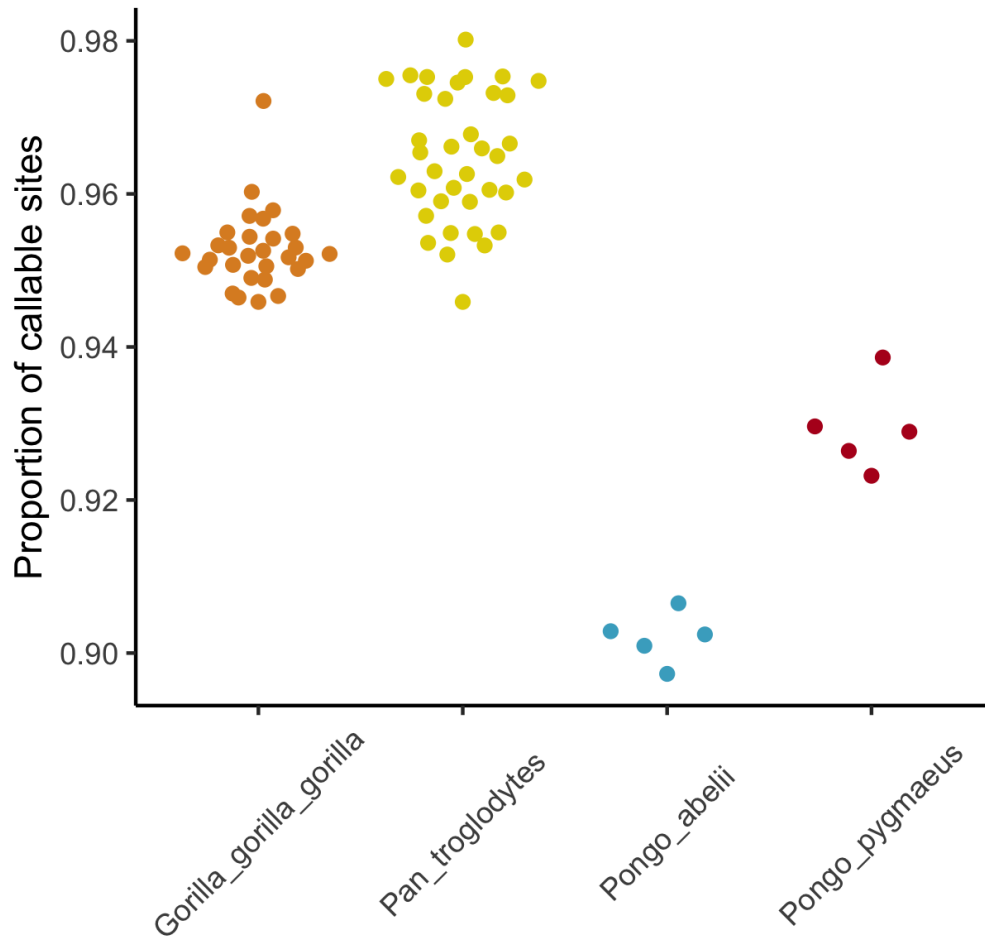


Fig. S5.

Proportion of callable sites calculated as the ratio of callable sites using the human reference genome versus the species-specific reference genome.

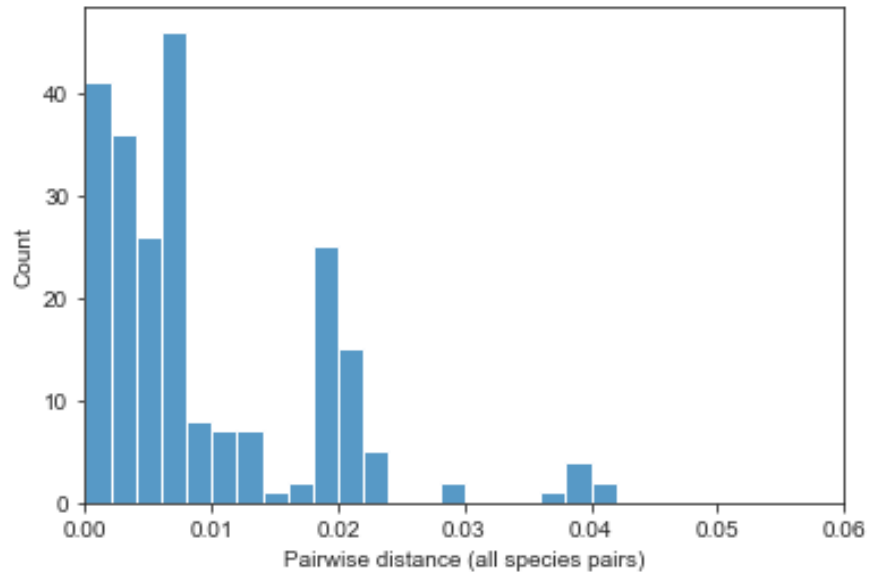


Fig. S6. Pairwise distances measured as substitutions per site between all focal species and reference species used in this study

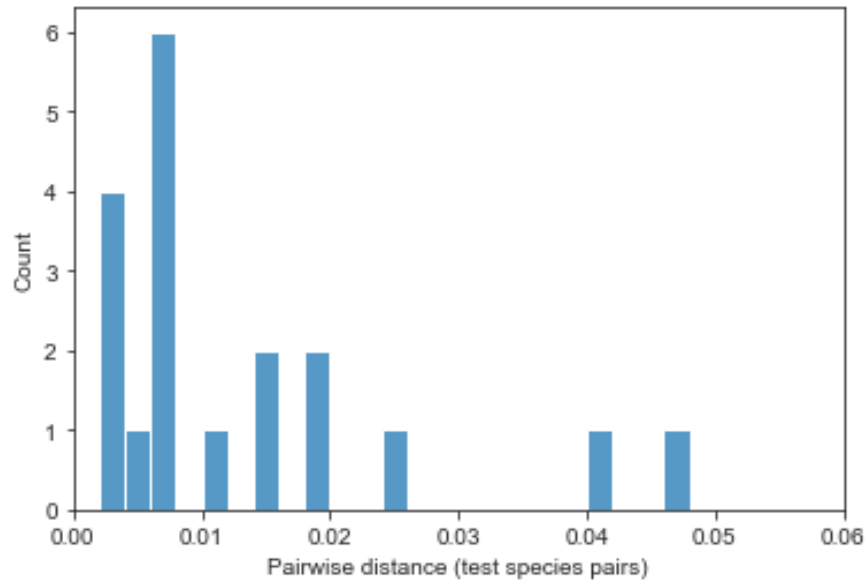


Fig. S7. Pairwise distances measured as substitutions per site between all species for which both a reference genome and resequencing data was available, and an outgroup species used to measure the effect of cross-reference species mapping.

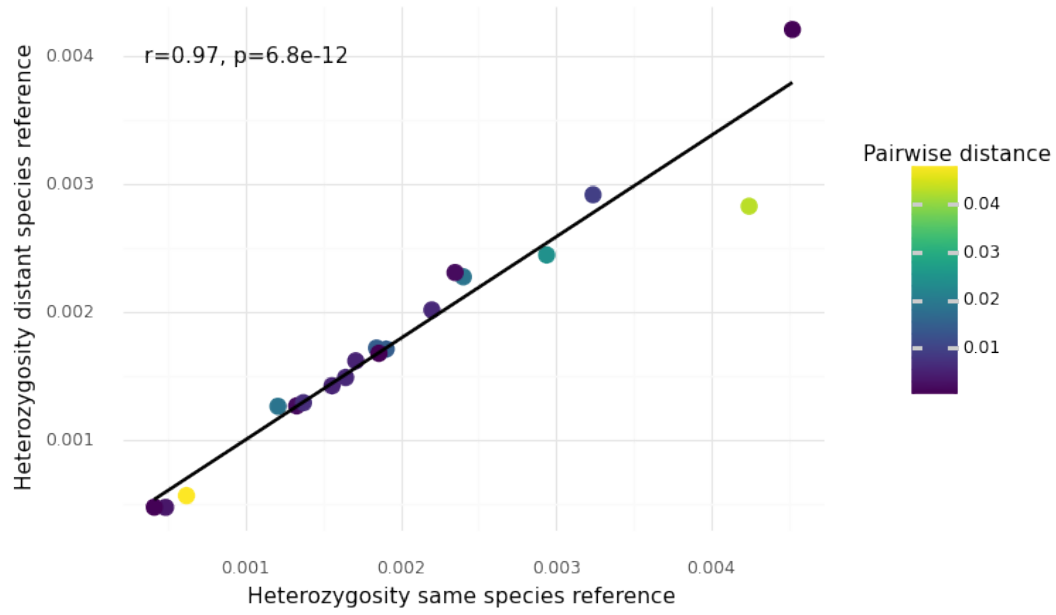


Fig. S8.

Estimates of heterozygosity based on reference genomes from the same species, and a distant one.

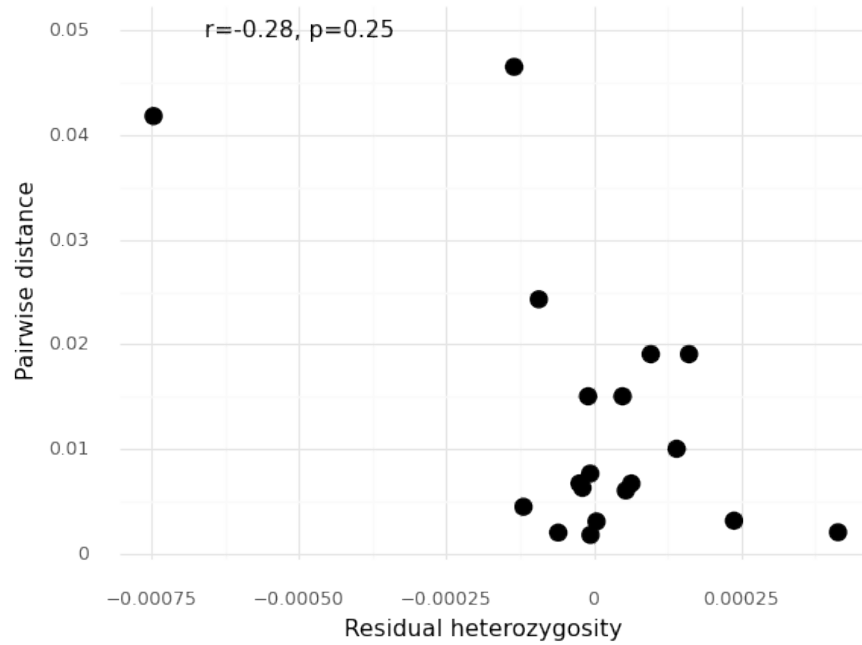


Fig. S9.

Estimates of residual heterozygosity after regressing heterozygosity estimates based on a distant reference genome onto heterozygosity based on a close reference genome, versus the pairwise distance between these two reference genomes.

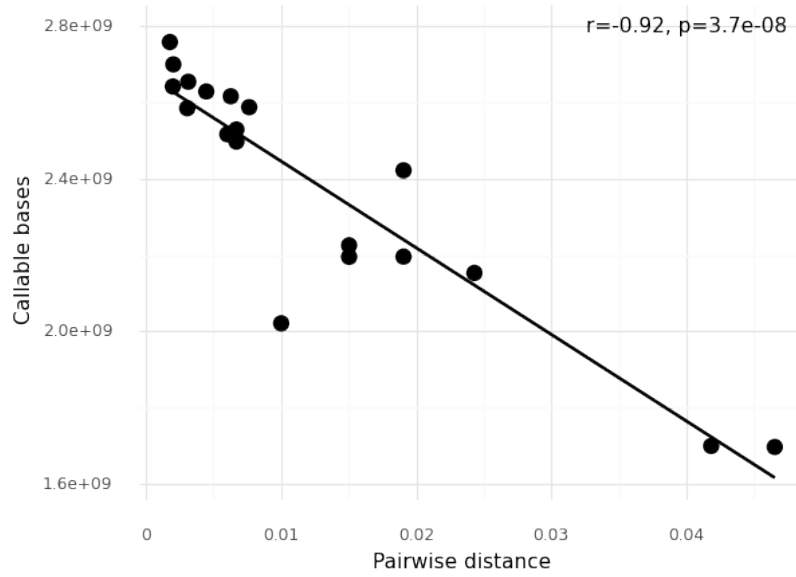


Fig. S10. Number of callable bases versus the pairwise distance between reference genomes used to estimate effects of cross-species mapping.

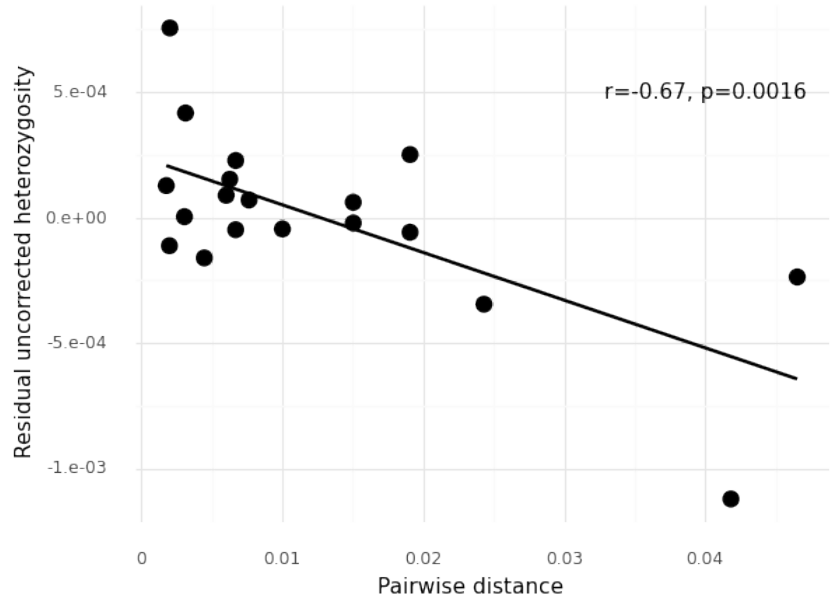


Fig. S11.

Estimates of residual heterozygosity after regressing heterozygosity estimates based on a distant reference genome onto heterozygosity based on a close reference genome without correcting for callability, versus the pairwise distance between these two reference genomes.

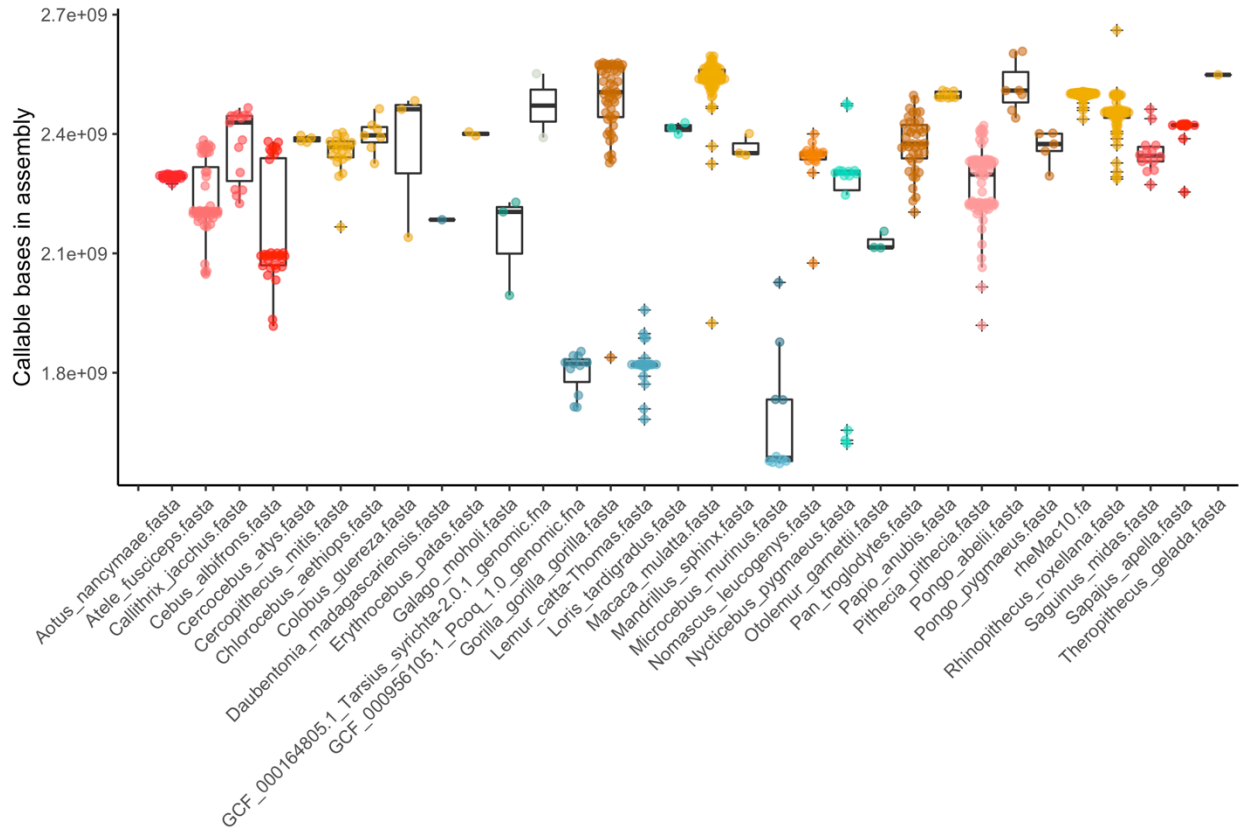


Fig. S12.

Total number of callable bases for all individuals and reference assemblies.

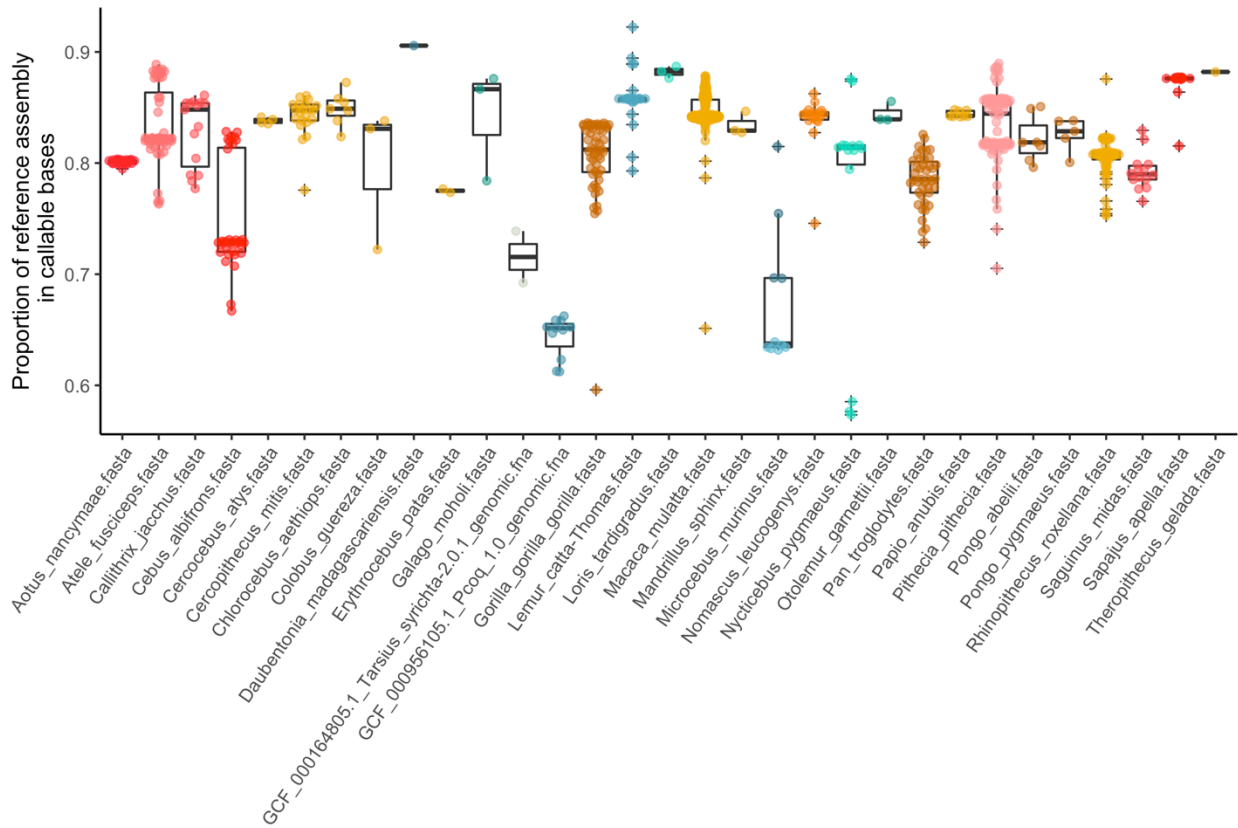


Fig. S13.

Callable bases as proportion of the total length of the reference genome for all individuals and reference assemblies.

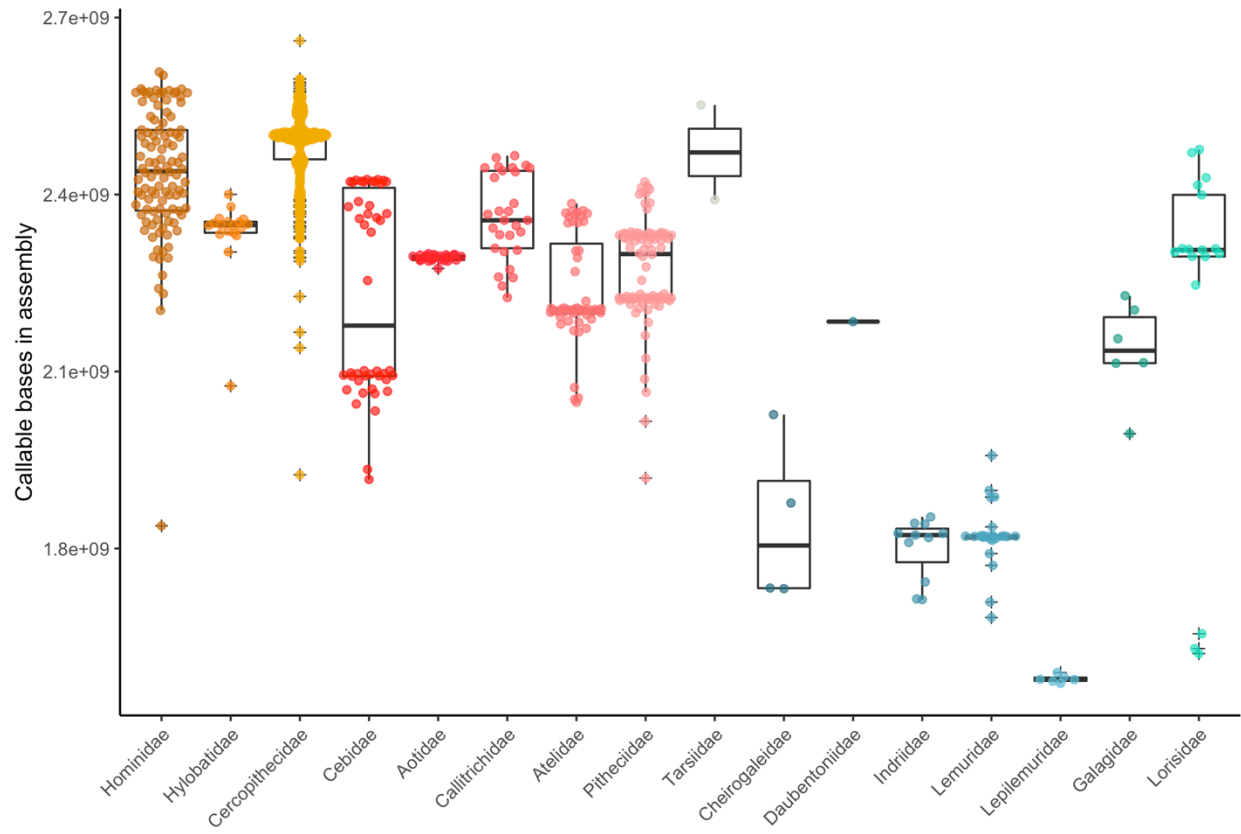


Fig. S14.
Total number of callable bases per genus

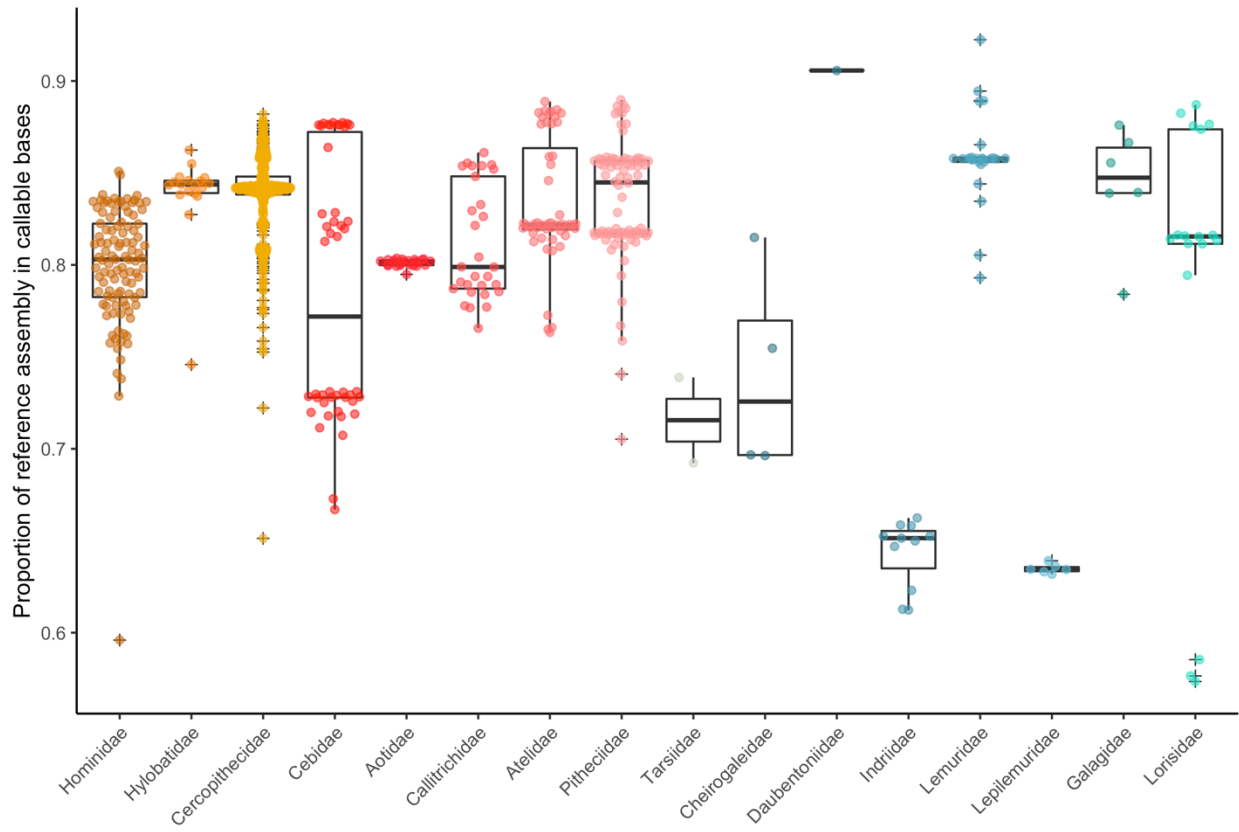


Fig. S 15.
Proportion of callable bases per genus

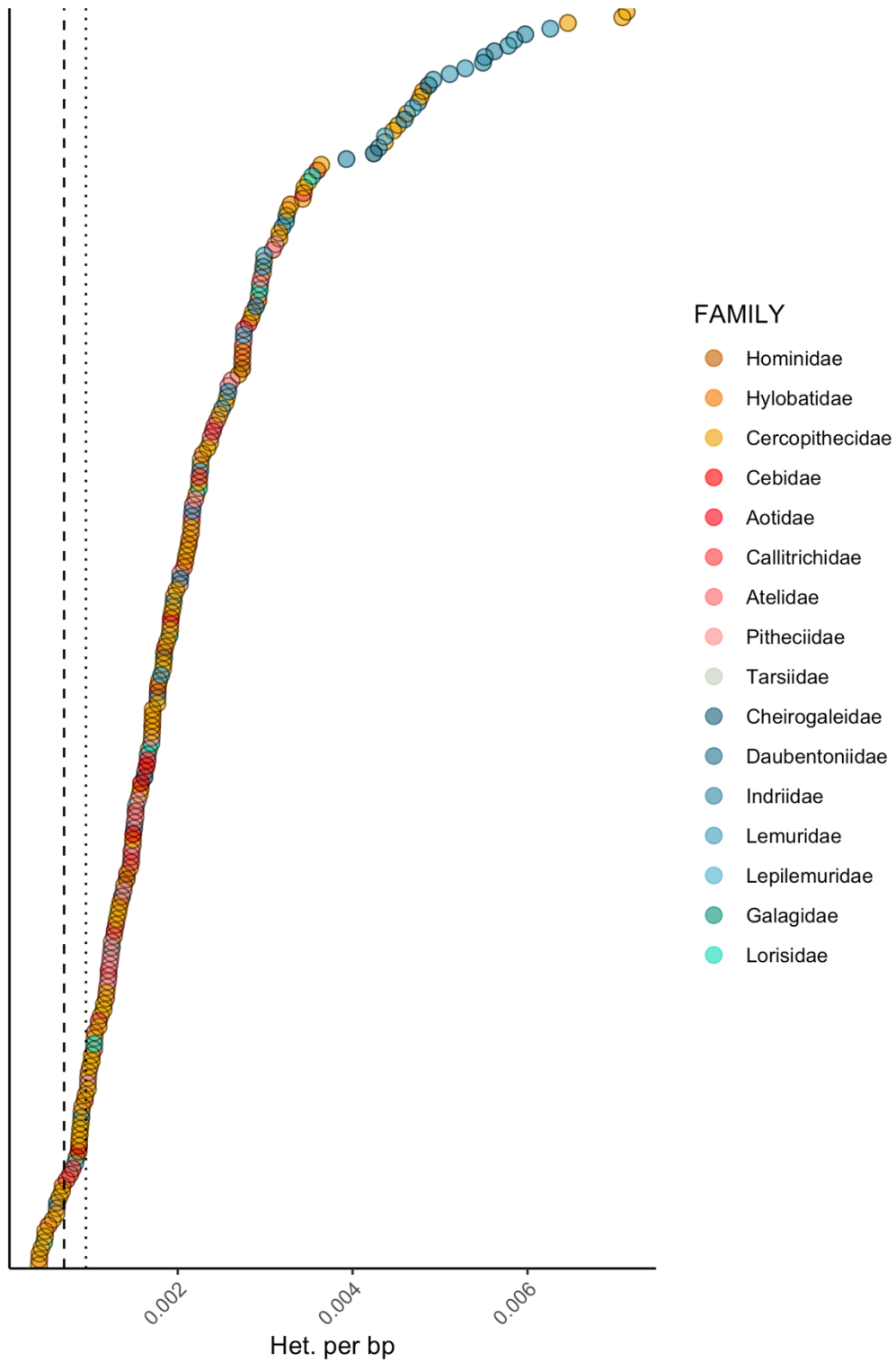


Fig. S16. Median heterozygosity for all species. Dotted and dashed lines denote median values for sub-Saharan and out of Africa human populations.

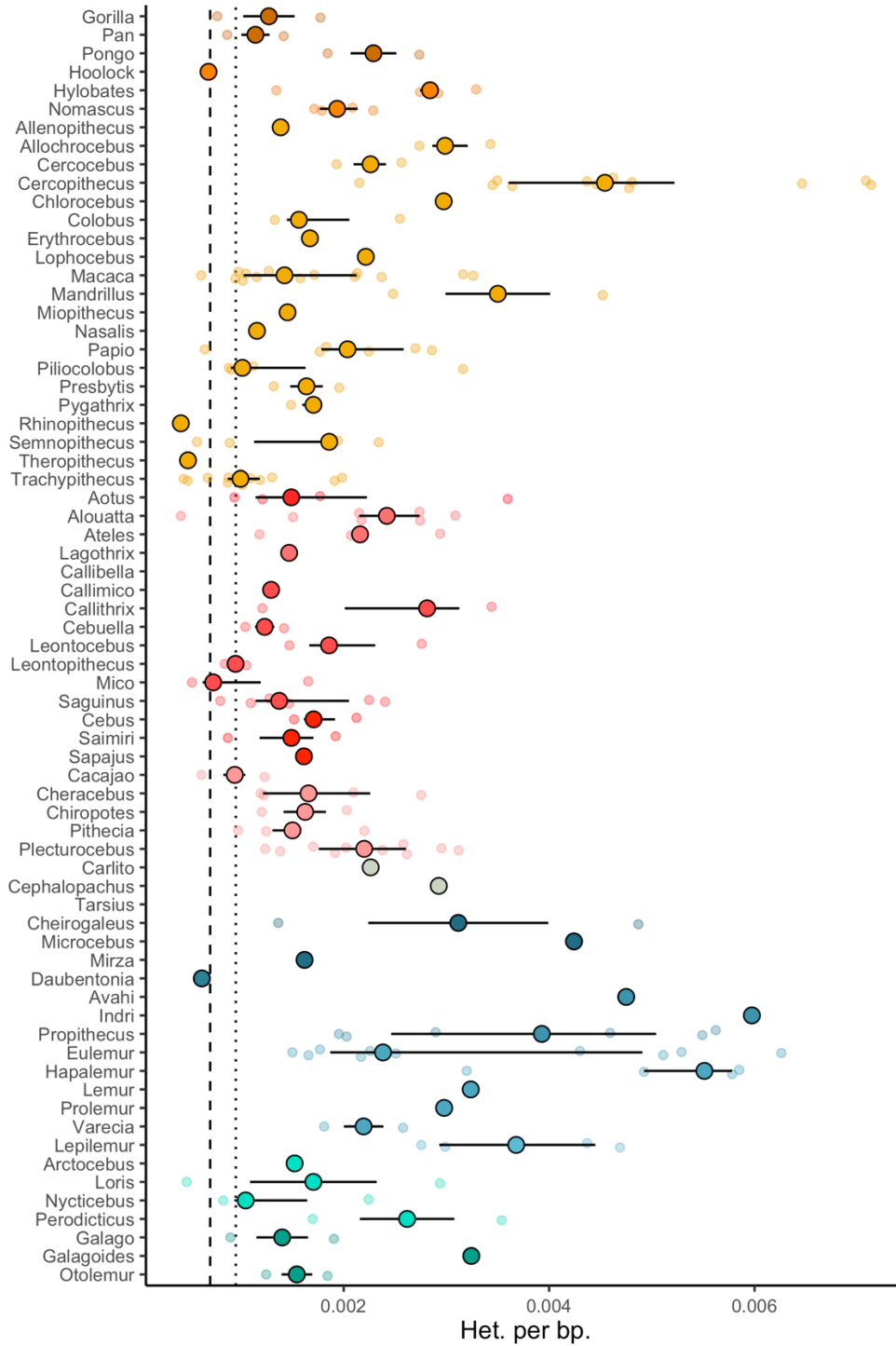


Fig. S17.

Distribution of heterozygosity across genera. Large solid circles and bars denote genus-wide median and interquartile range. Distribution of heterozygosity across genera. Large solid circles and bars denote genus-wide median and interquartile range.

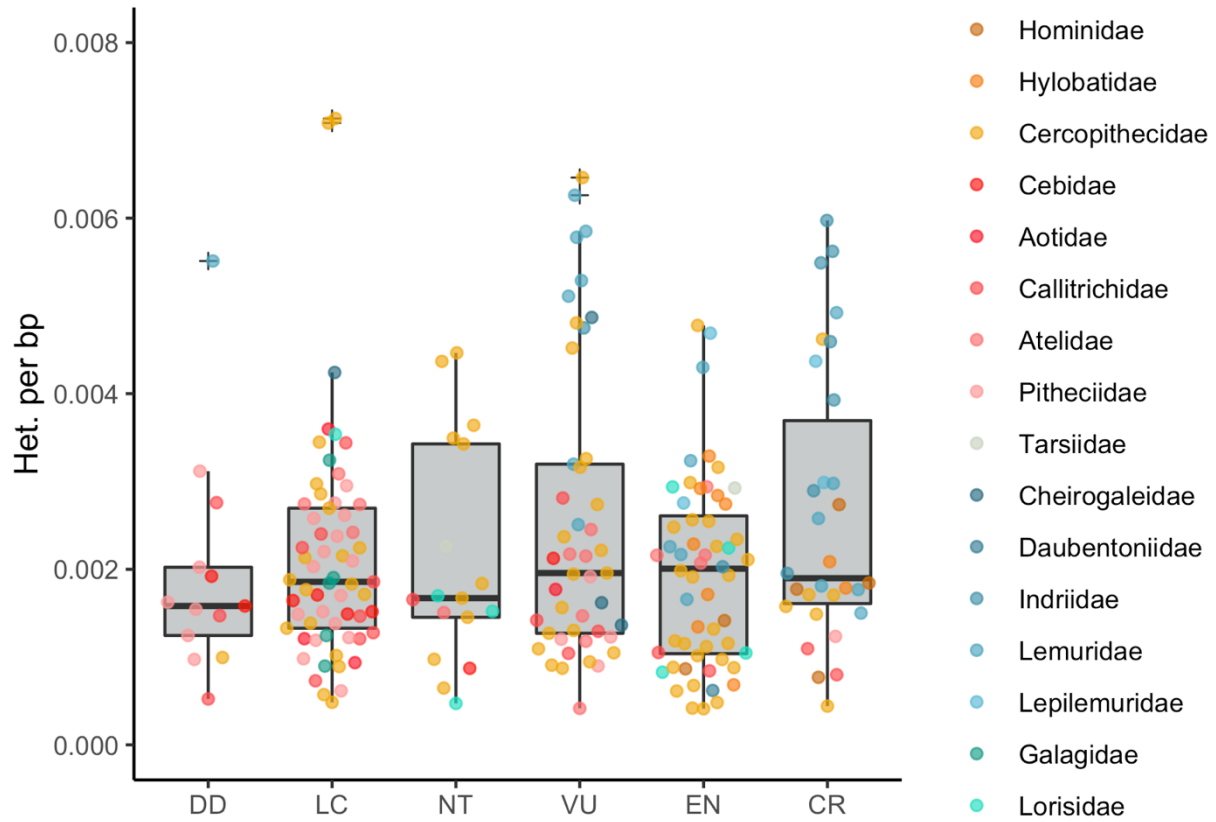


Fig. S18. Distribution of heterozygosity versus IUCN extinction risk category for all primate families, DD: Data deficient, LC: Least concern, NT: Near threatened, VU: Vulnerable, EN: Endangered, CR: Critically endangered

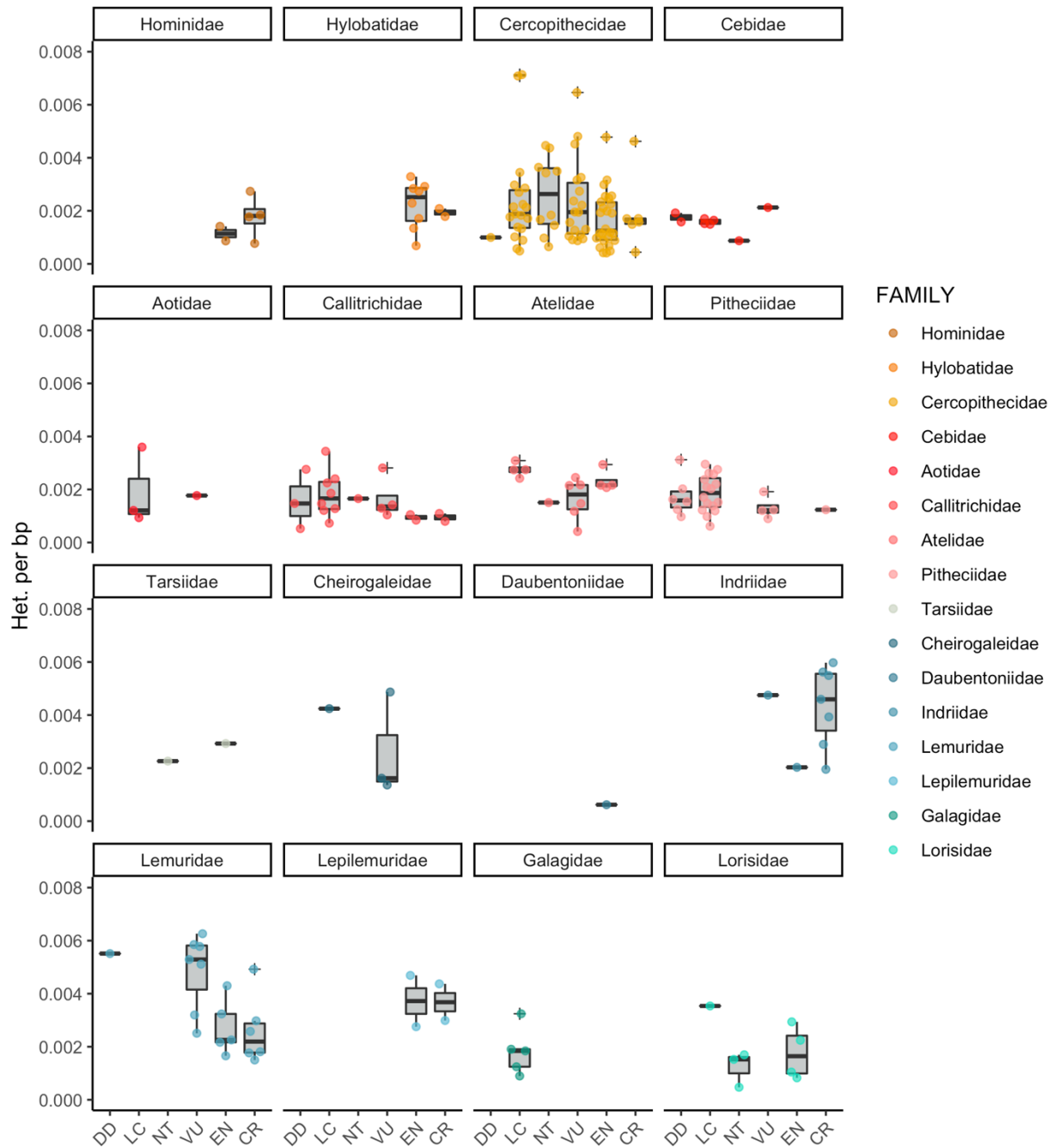


Fig. S19. Distribution of heterozygosity versus IUCN extinction risk category for all primate families, DD: Data deficient, LC: Least concern, NT: Near threatened, VU: Vulnerable, EN: Endangered, CR: Critically endangered

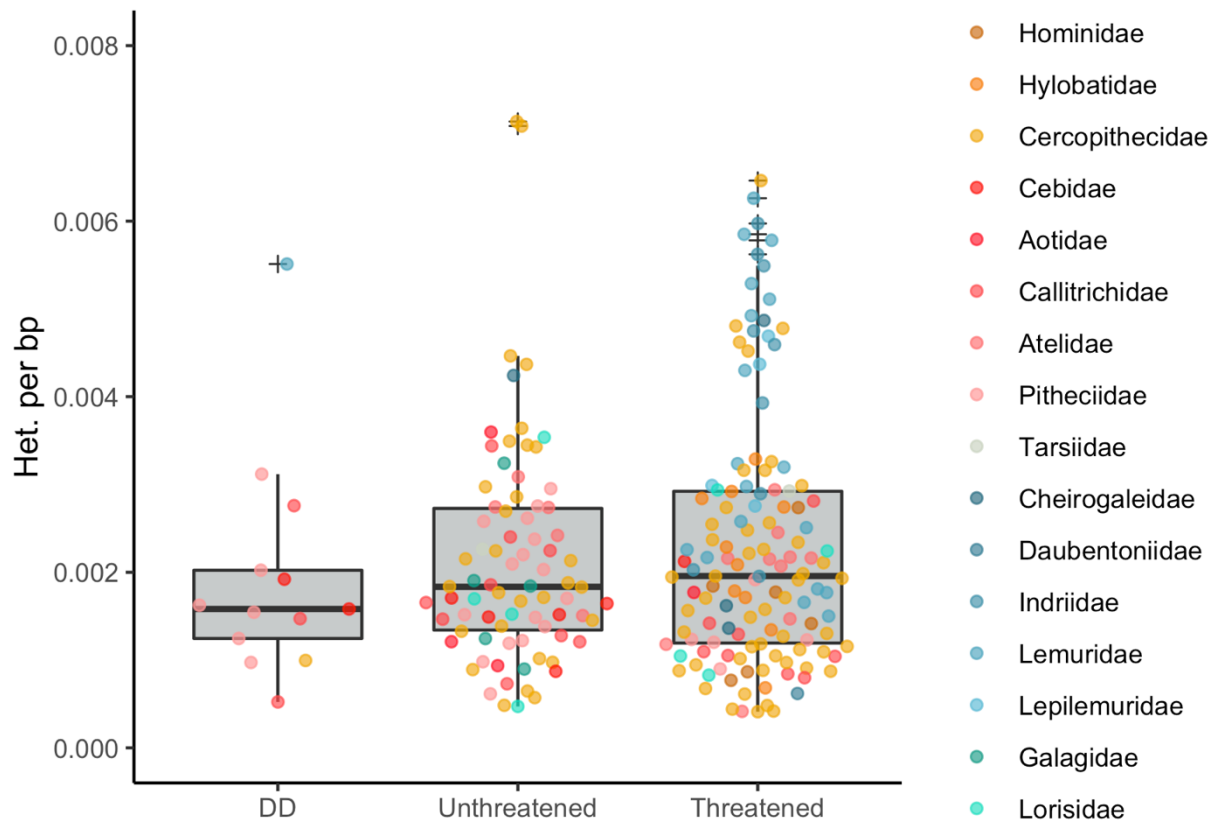


Fig. S20.

Distribution of heterozygosity versus IUCN extinction risk category for all primate families, DD: Data deficient, Unthreatened: LC , NT; Threatened: VU, EN, CR

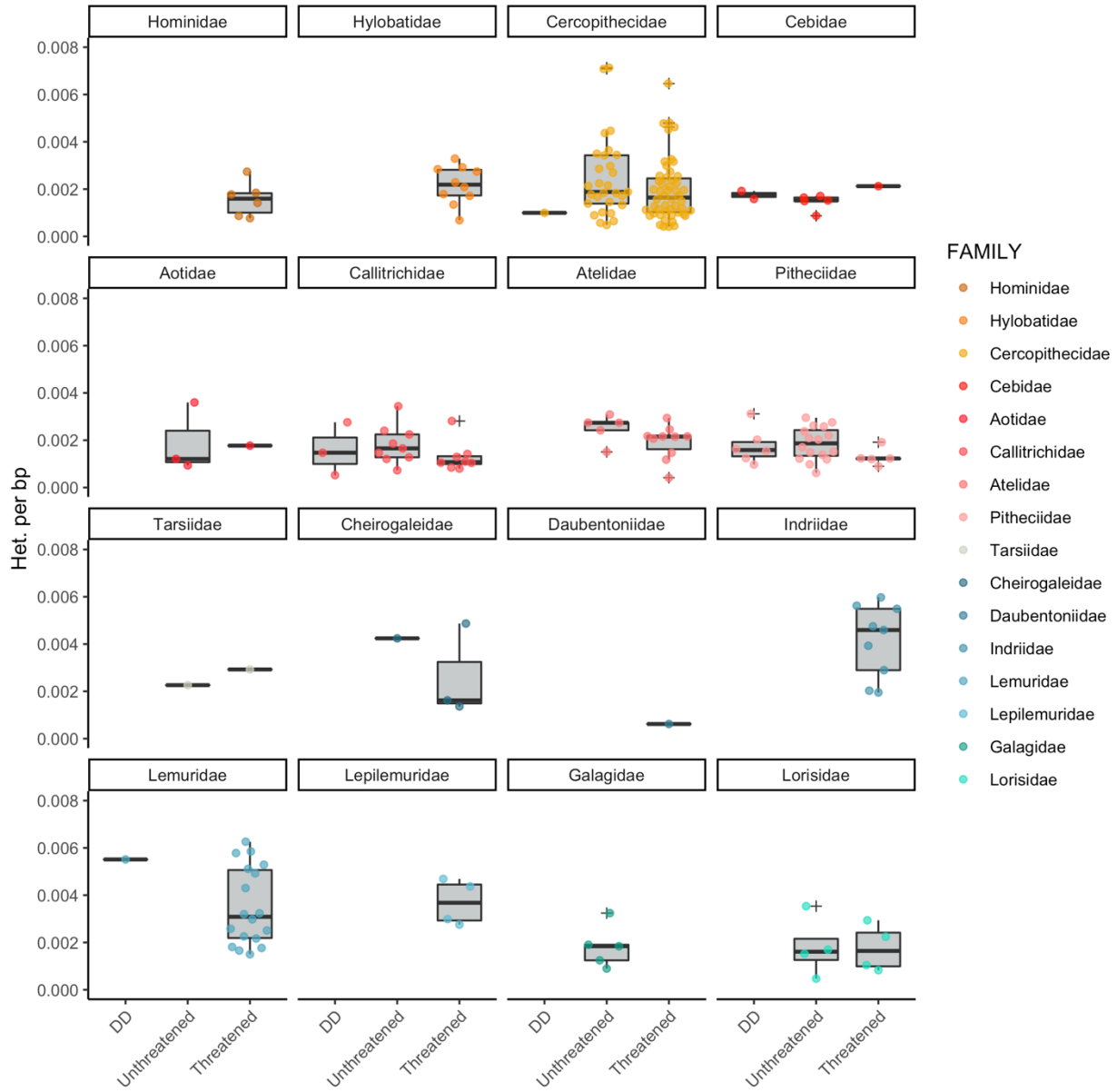


Fig. S21.

Distribution of heterozygosity versus IUCN extinction risk category for all primate families, DD: Data deficient, Unthreatened: LC , NT; Threatened: VU, EN, CR

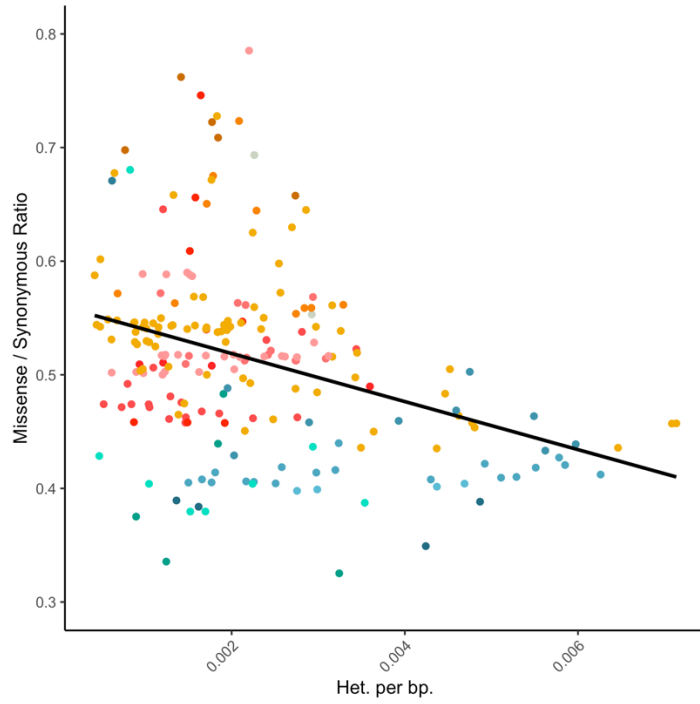


Fig. S22. Relationship between missense/synonymous ratio and heterozygosity. The negative slope is consistent with the effects of purifying selection.

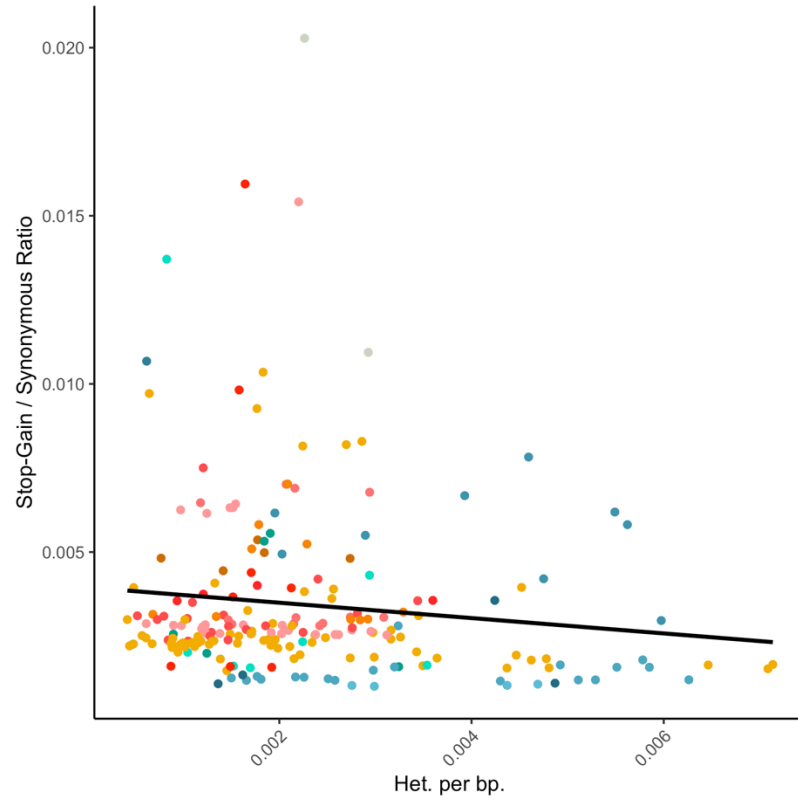


Fig. S23.
Relationship between stop-gain / synonymous ratio and heterozygosity.

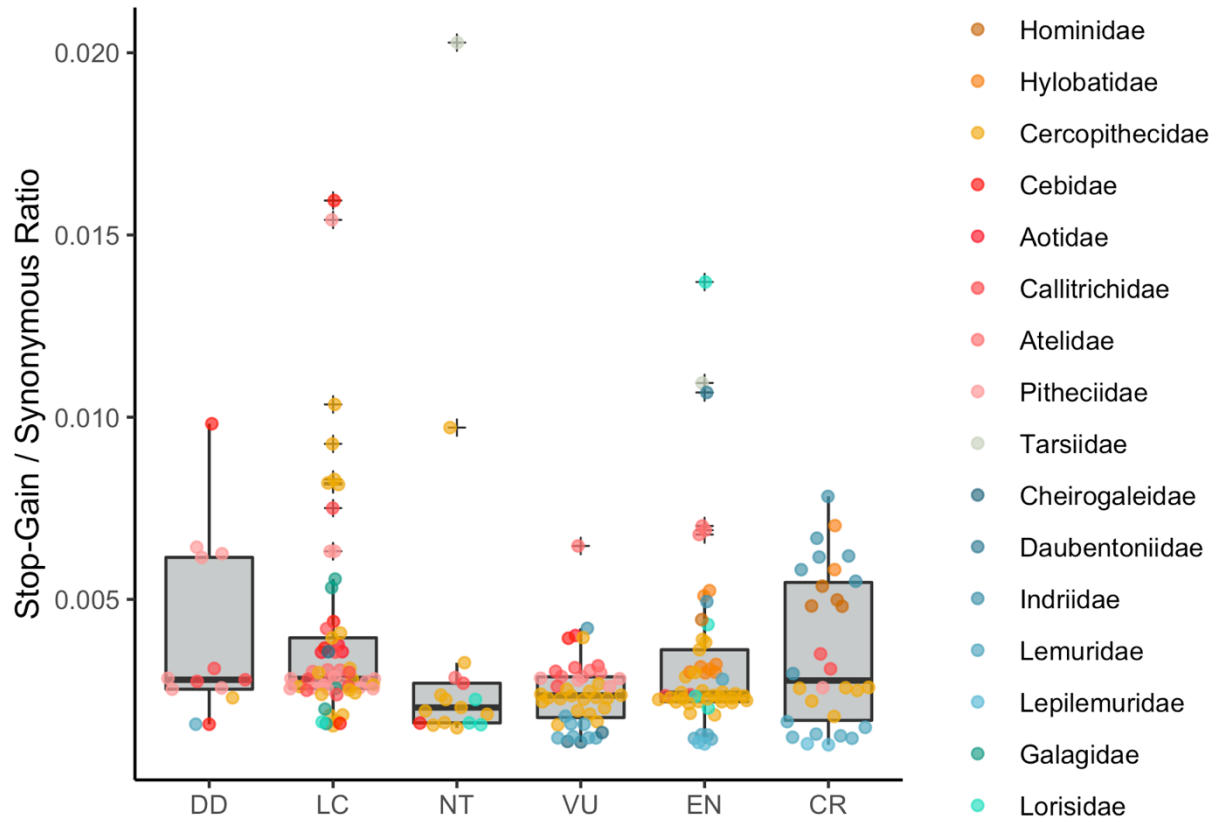


Fig. S24.

Loss of function variation as measured by stop-gain / synonymous ratio versus IUCN extinction risk category. DD: Data deficient, LC: Least concern, NT: Near threatened, VU: Vulnerable, EN: Endangered, CR: Critically endangered

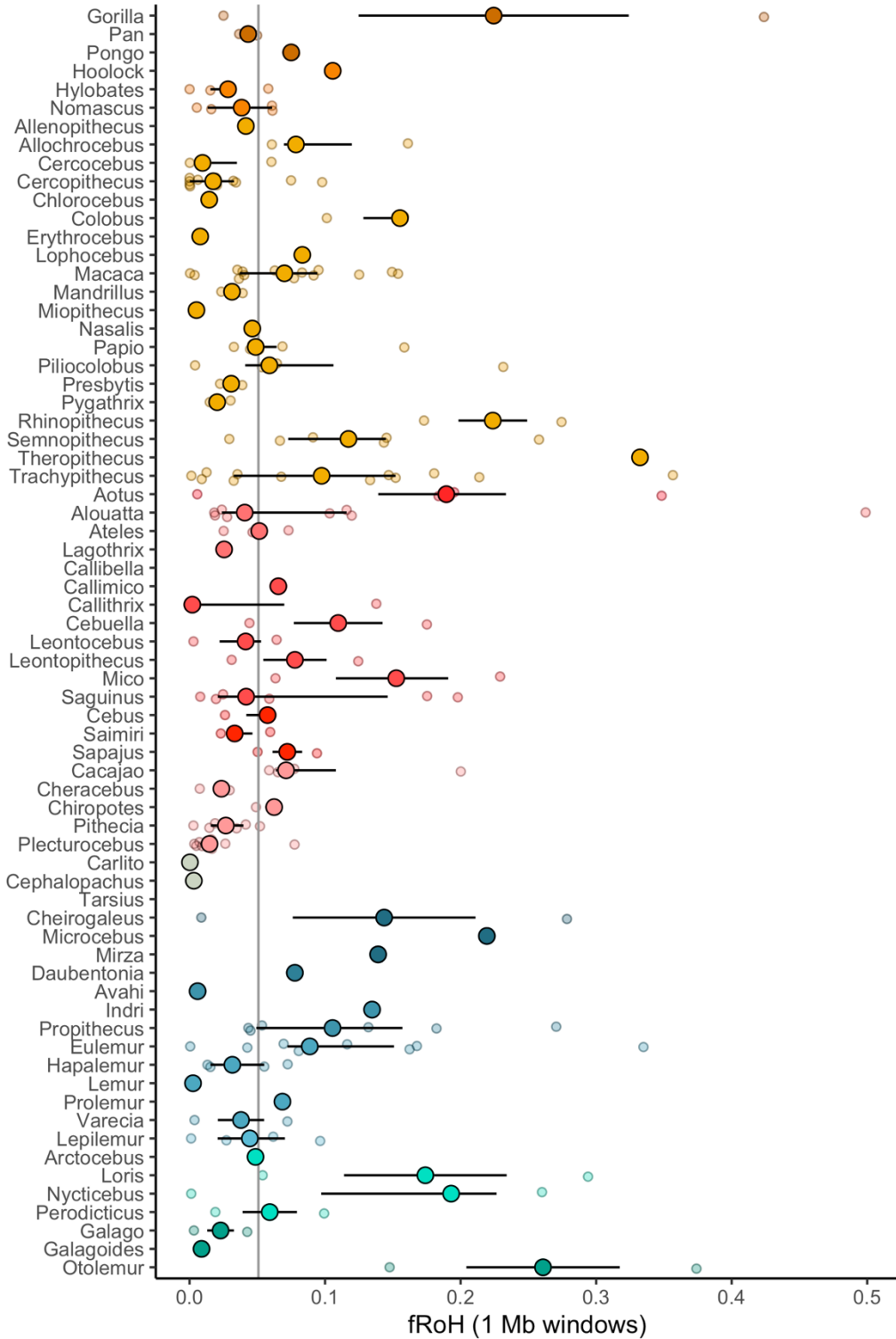


Fig. S25.

Fraction of genome in runs of homozygosity of at least 1Mb in size. Large solid circles and bars denote genus-wide median and interquartile range.

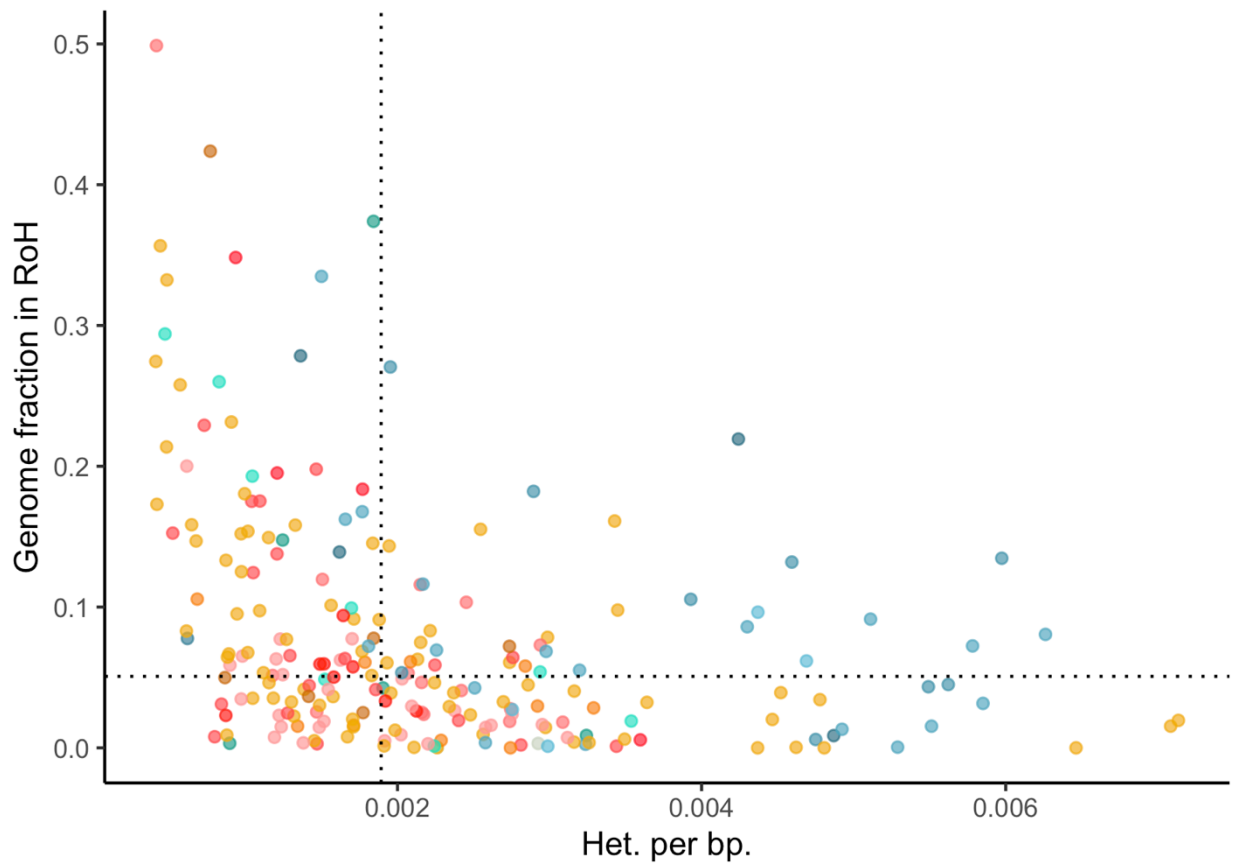


Fig. S26.

Relationship between median heterozygosity and fraction of the genome in runs of homozygosity of at least 1mb. Dotted lines correspond to median values of each axis.

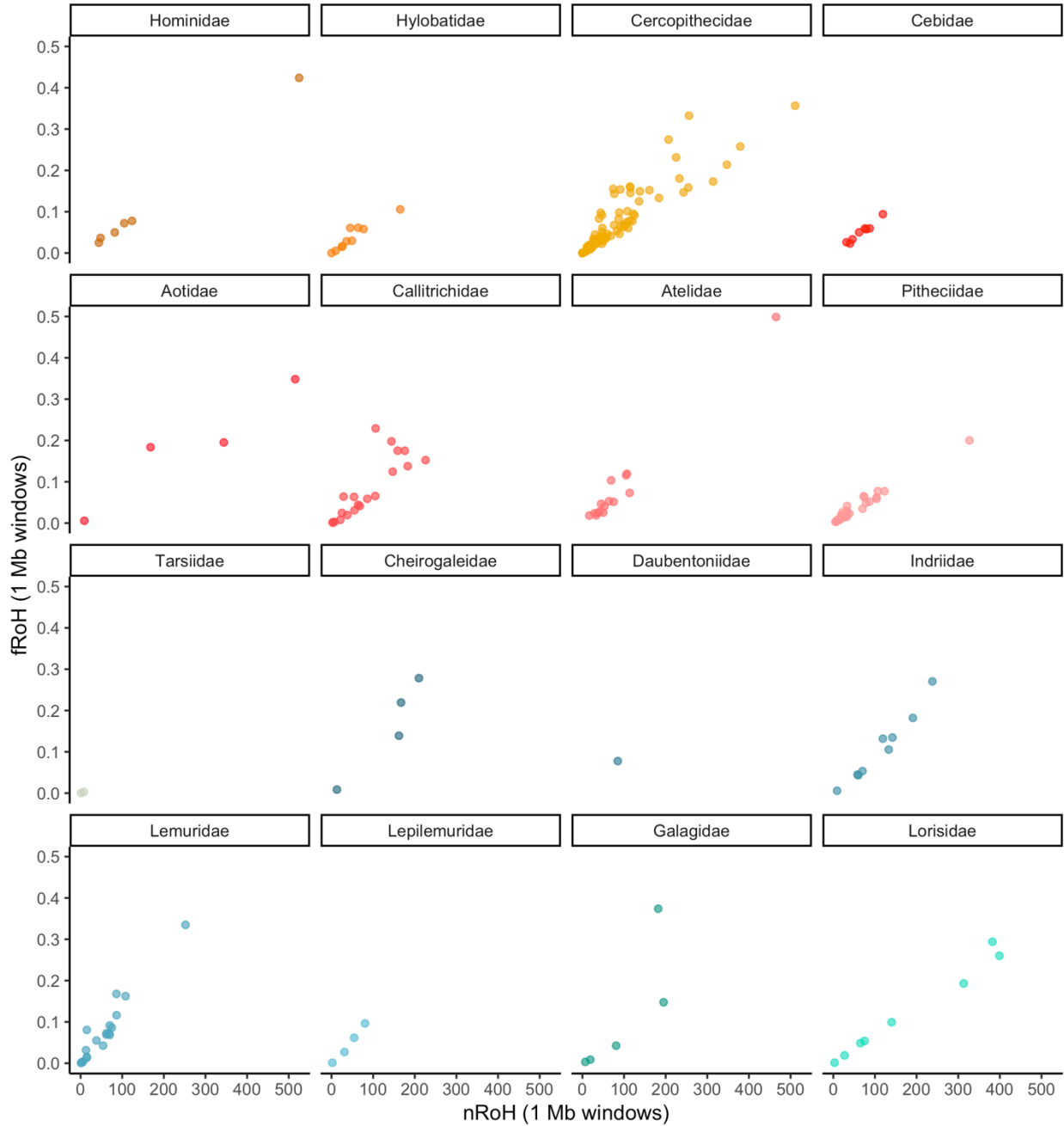


Fig. S27. Median number of runs of homozygosity versus median proportion of the genome in runs of homozygosity per species, stratified by families.

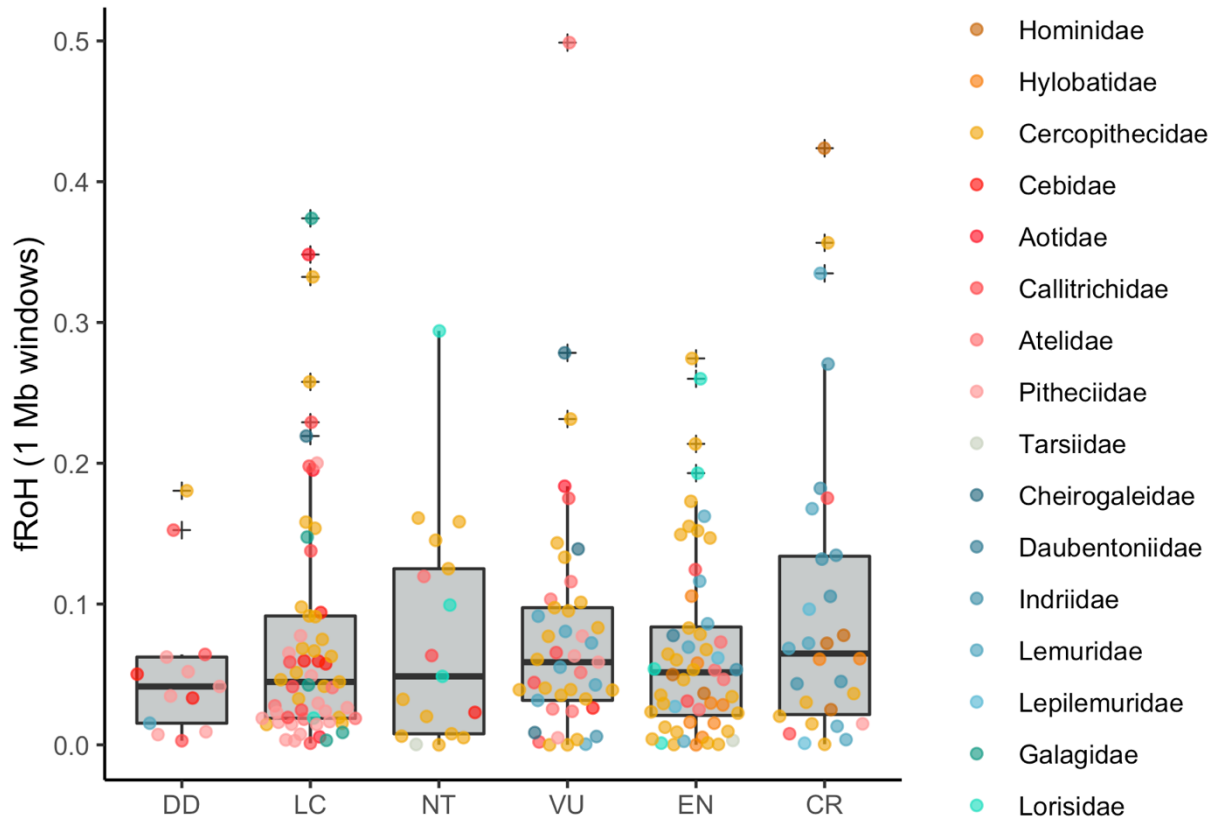


Fig. S28.

Median fraction of genome in runs of homozygosity per species IUCN extinction risk category. DD: Data deficient, LC: Least concern, NT: Near threatened, VU: Vulnerable, EN: Endangered, CR: Critically endangered

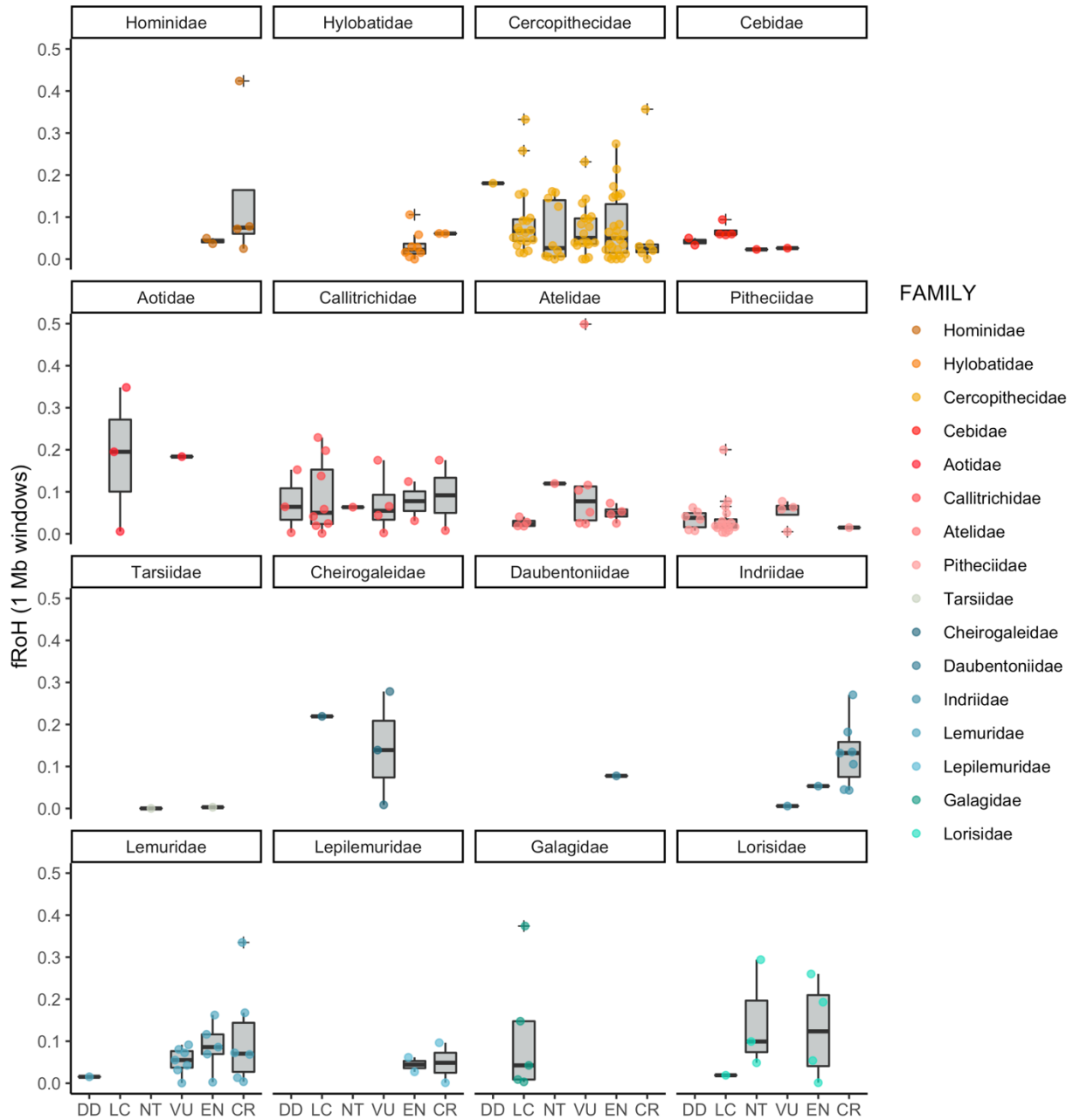


Fig. S29.

Median fraction of genome in runs of homozygosity per species versus IUCN extinction risk category. DD: Data deficient, LC: Least concern, NT: Near threatened, VU: Vulnerable, EN: Endangered, CR: Critically endangered

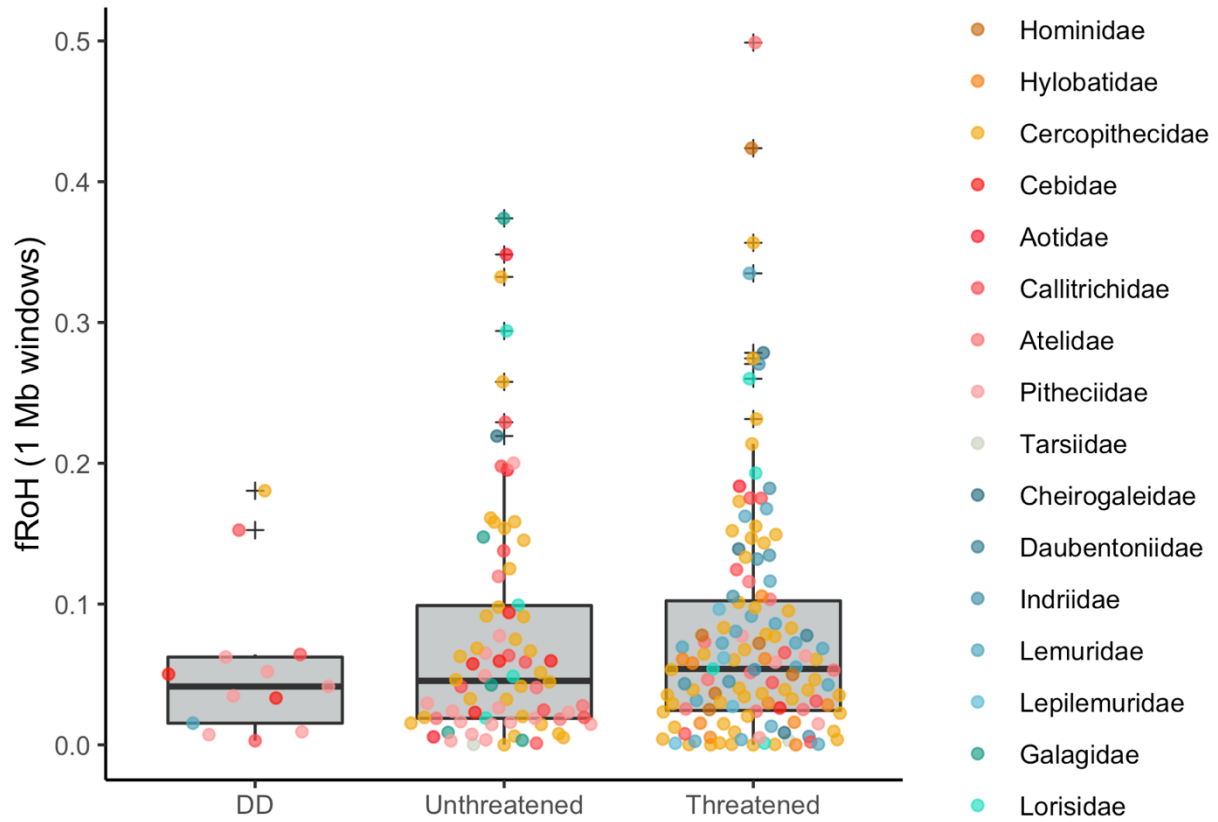


Fig. S 30.
 Median fraction of genome in runs of homozygosity per species versus threatened status

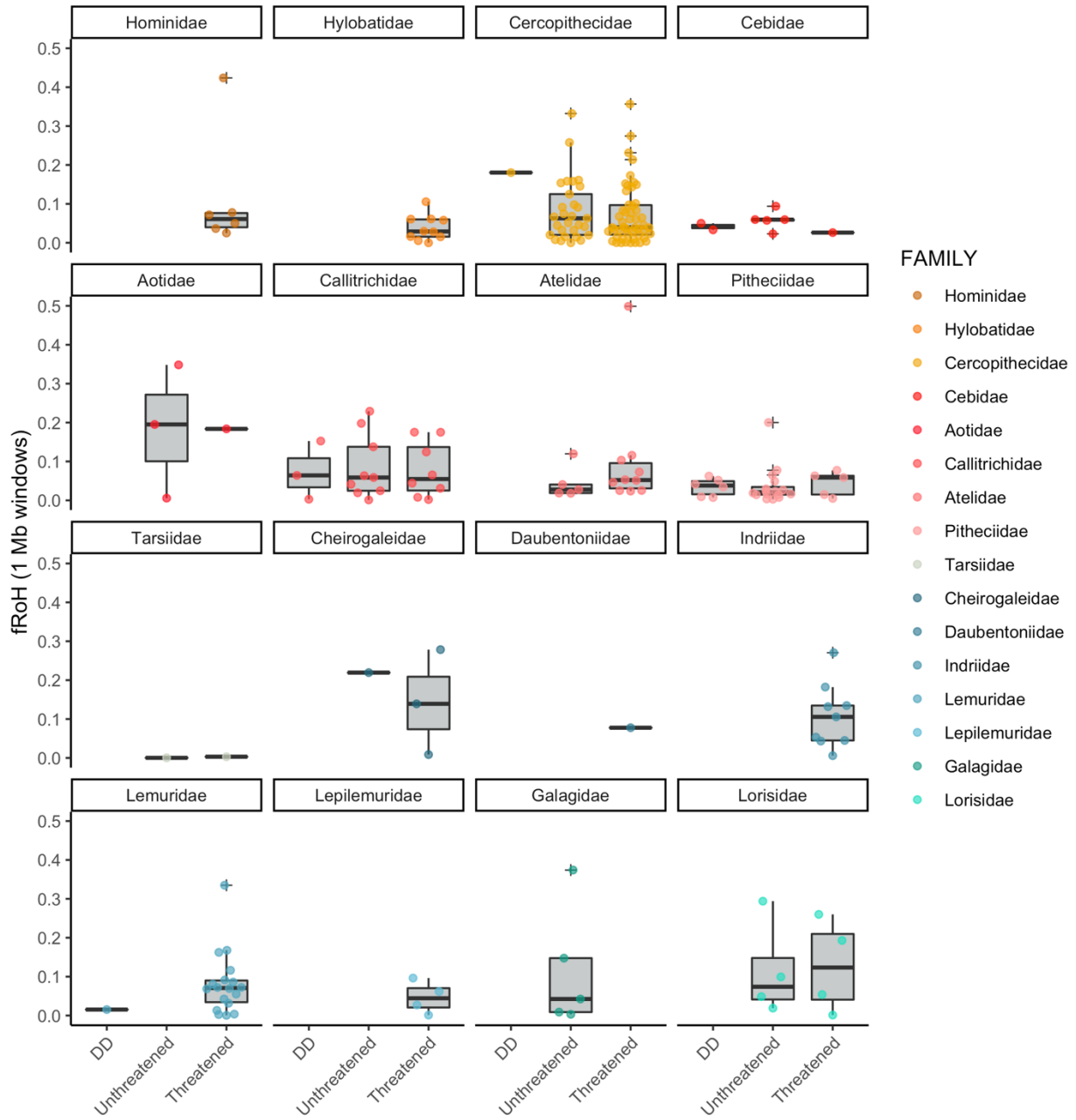


Fig. S31. Median fraction of genome in runs of homozygosity per species versus threatened status

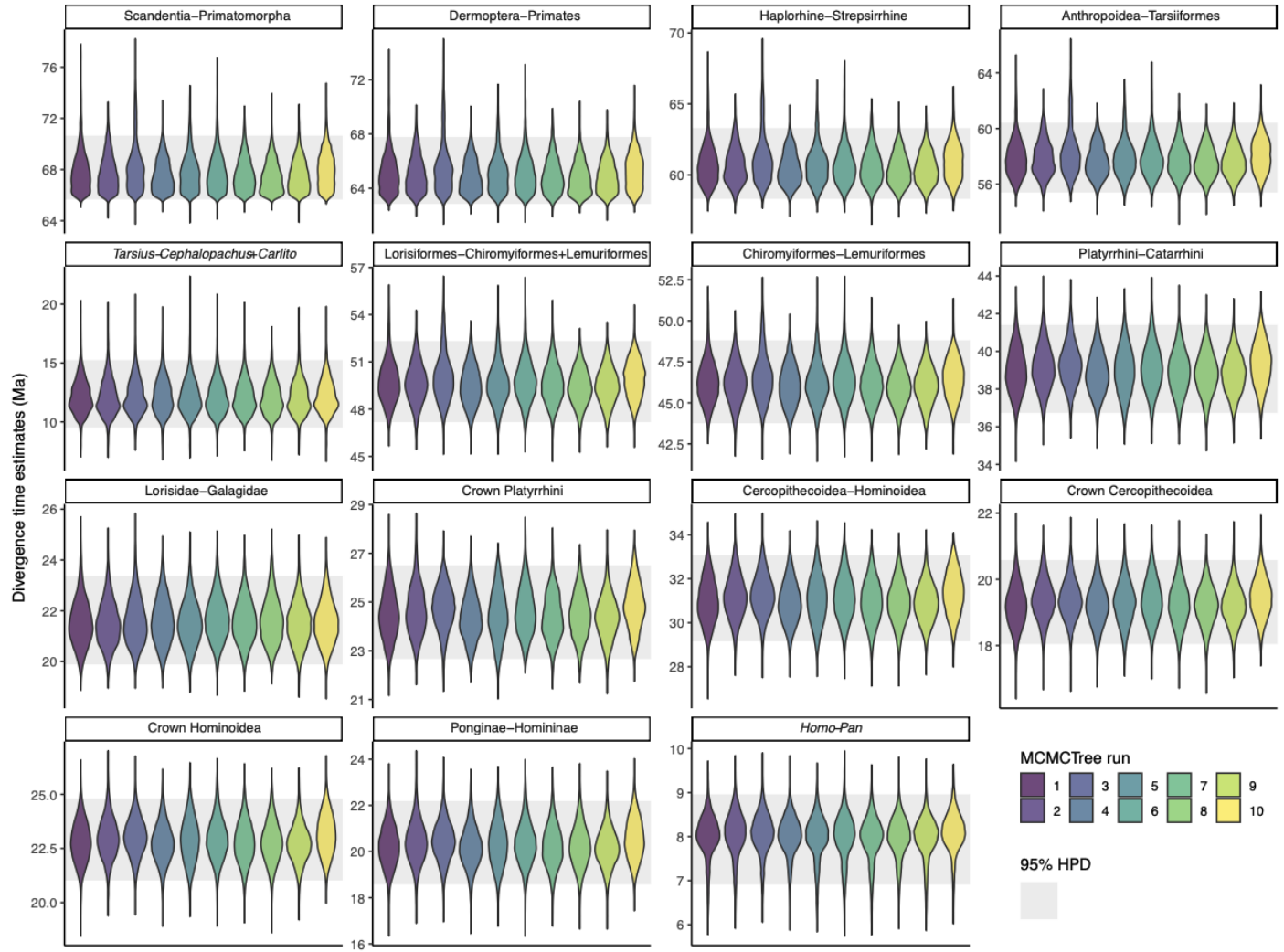


Fig. S32.

Divergence times of nodes discussed here. Violin plots show densities of the node ages estimated by separate MCMCTree replicates, gray shading indicates 95% highest posterior density of the combined runs.

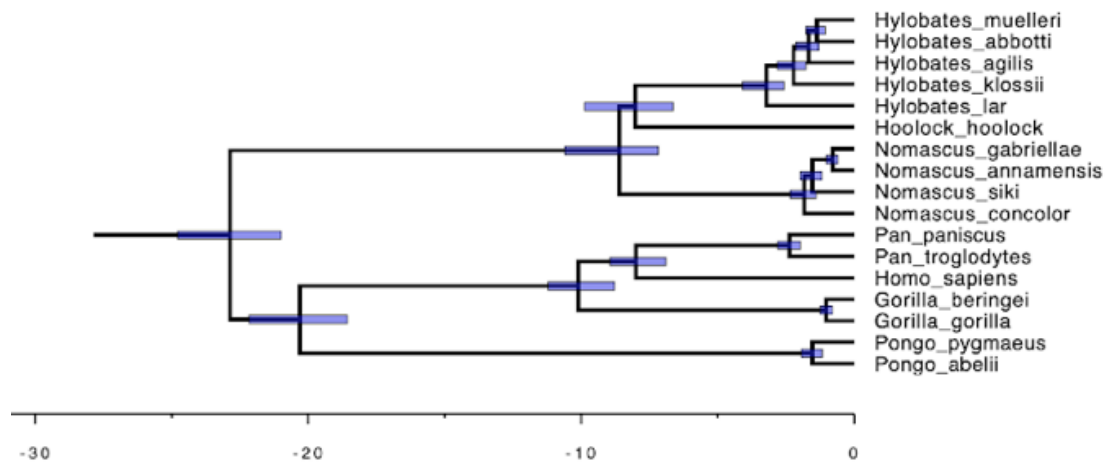


Fig. S 33

Hominidae and Hylobatidae subtree of the fossil-calibrated time tree. Scaling units are million years. Blue bars on nodes denote 95%

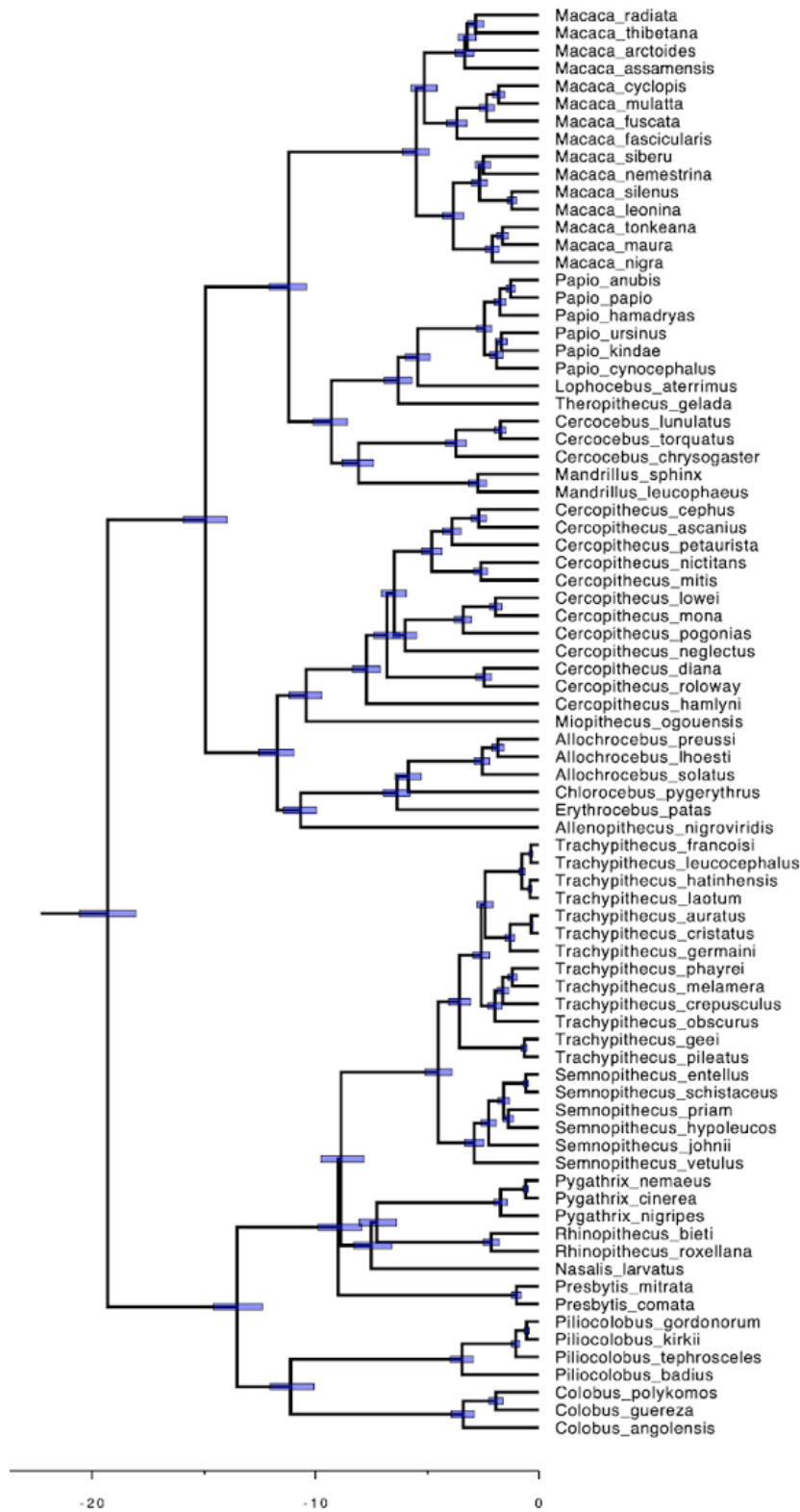


Fig. S 34

Cercopithecidae subtree of the fossil-calibrated time tree. Scaling units are million years, blue bars denote 95% highest posterior density.

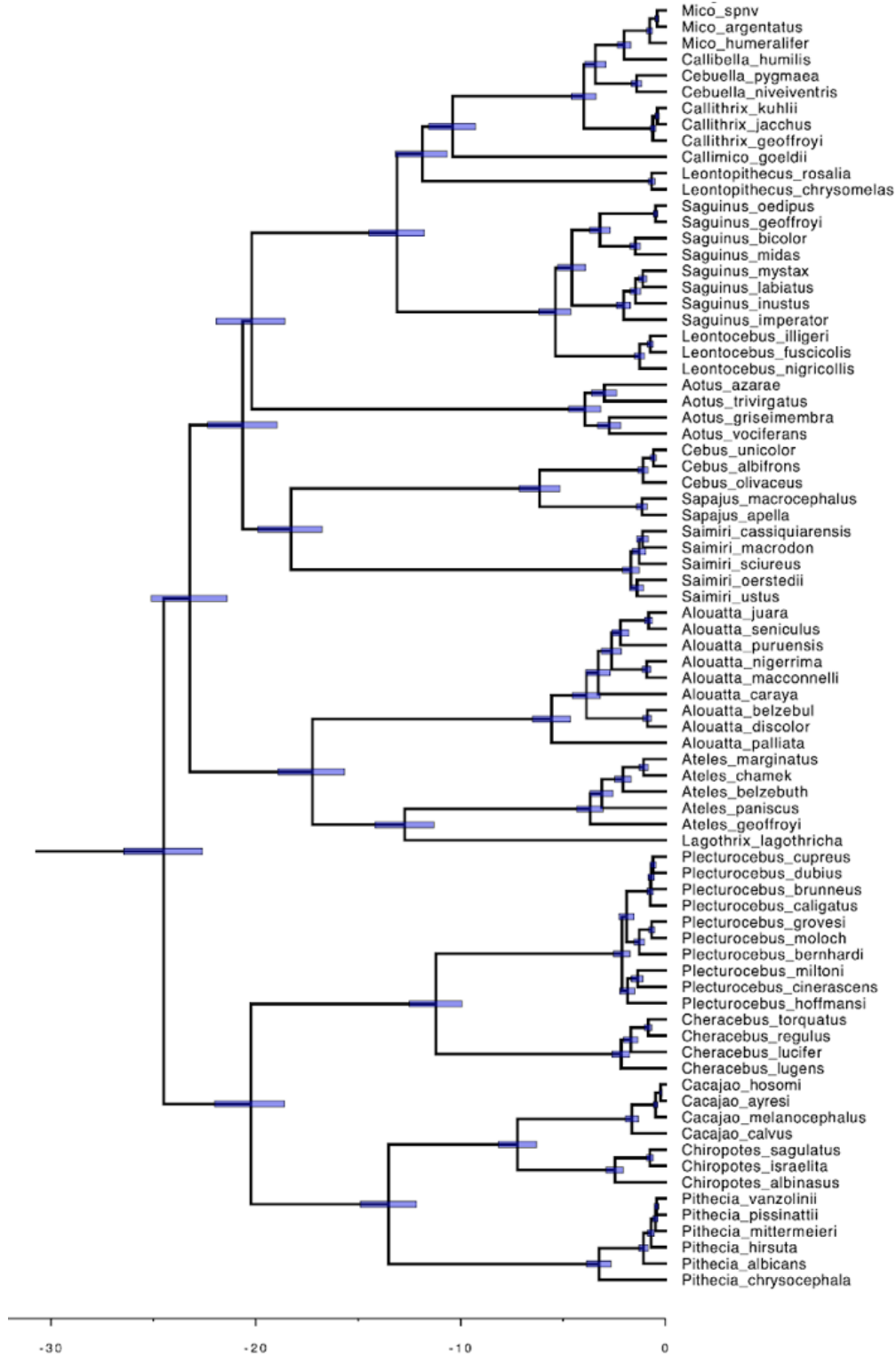


Fig. S 35.

Platyrrhine subtree of the fossil-calibrated time tree. Scaling units are million years, blue bars denote 95% highest posterior density.

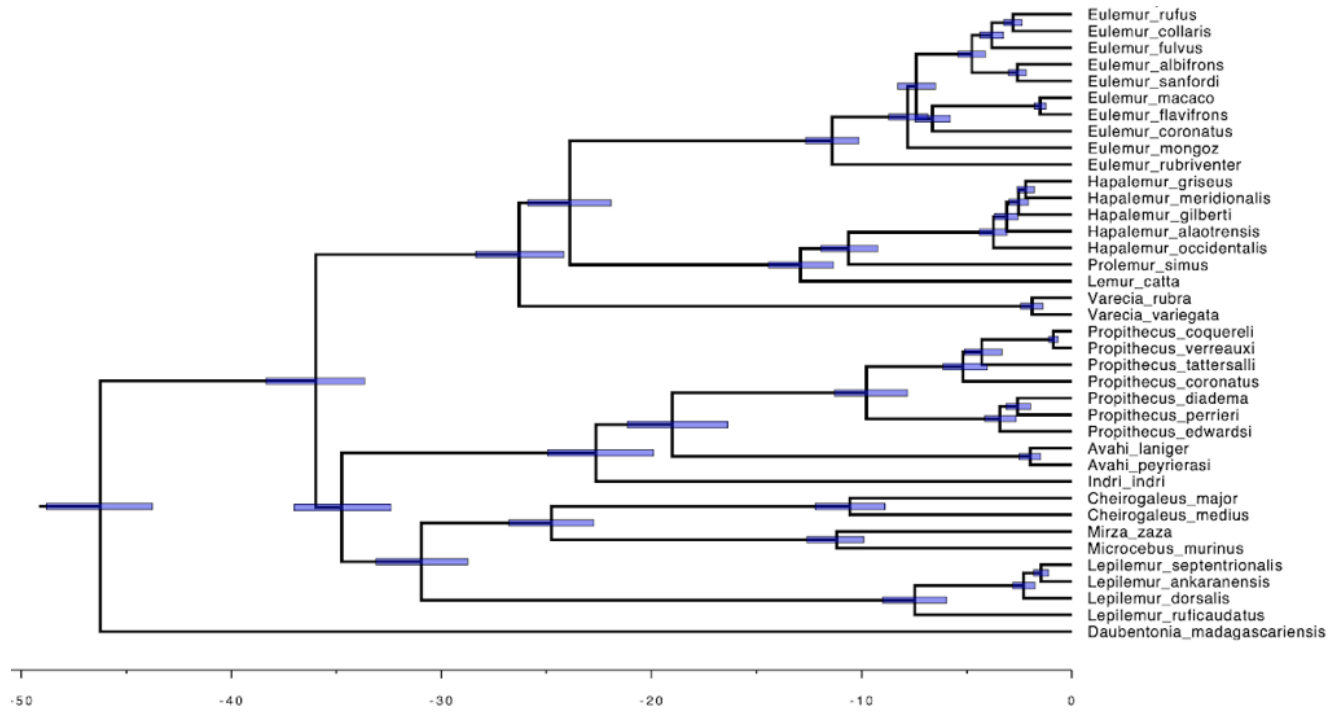


Fig. S 36

Lemur subtree of the fossil-calibrated time tree. Scaling units are million years, blue bars denote 95% highest posterior density.

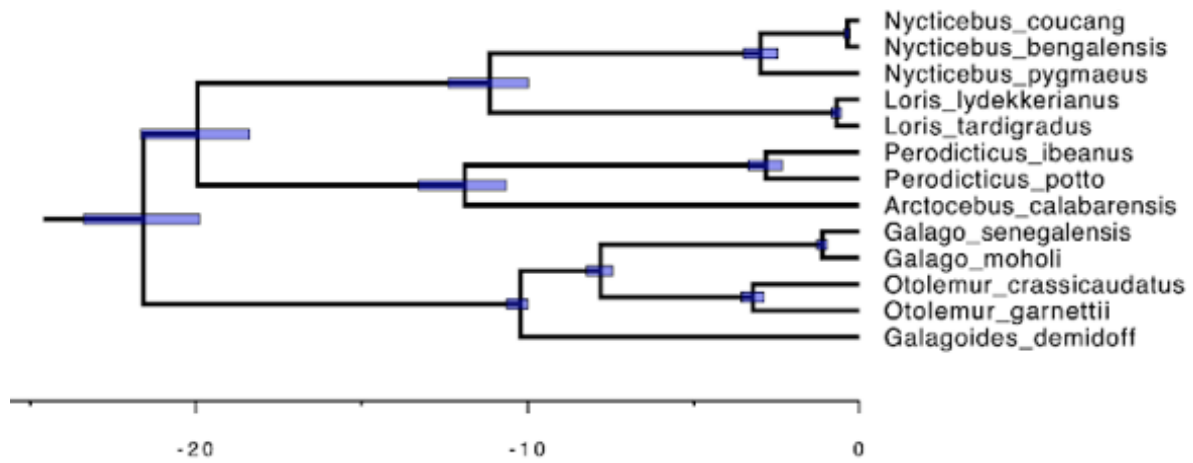


Fig. S 37

Loris and Galago subtree of the fossil-calibrated time tree. Scaling units are million years, blue bars denote 95% highest posterior density.

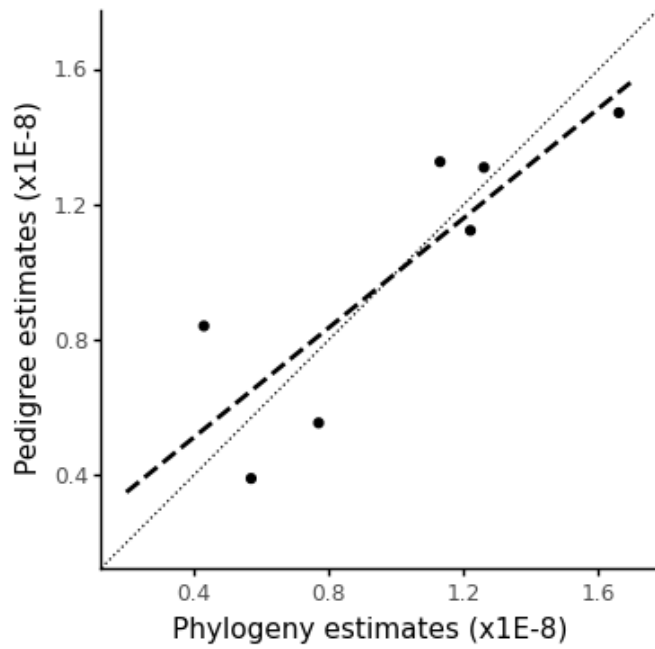


Fig. S38.

Comparison of trio-based and phylogeny-based estimates of mutation rates. The dashed line shows the linear regression of both estimates, the dotted line shows $x=y$. Estimates fall on either side of the $x=y$ slope.

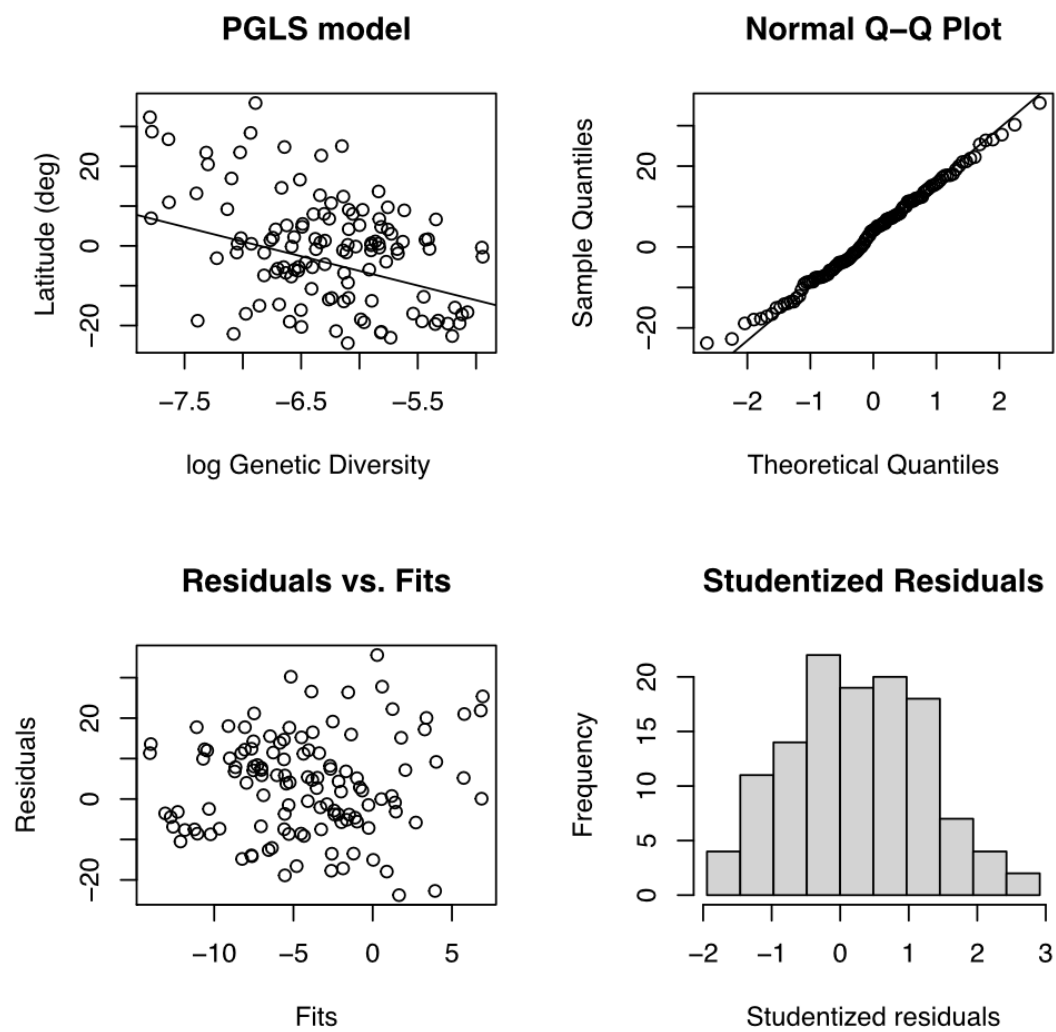


Fig. S39.
PGLS Diagnostic plots for log genetic diversity predicted by latitude.

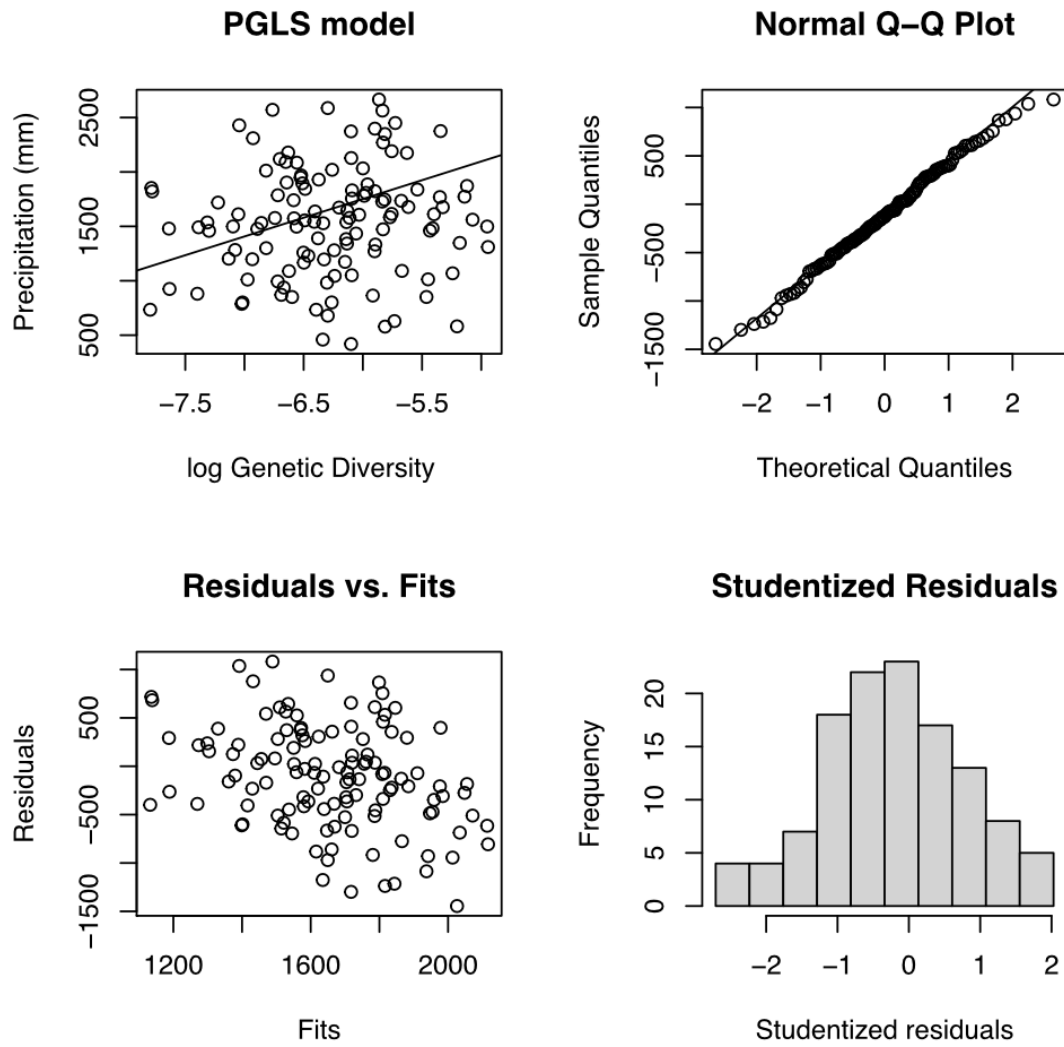


Fig. S40.
 PGLS Diagnostic plots for log genetic diversity predicted by precipitation.

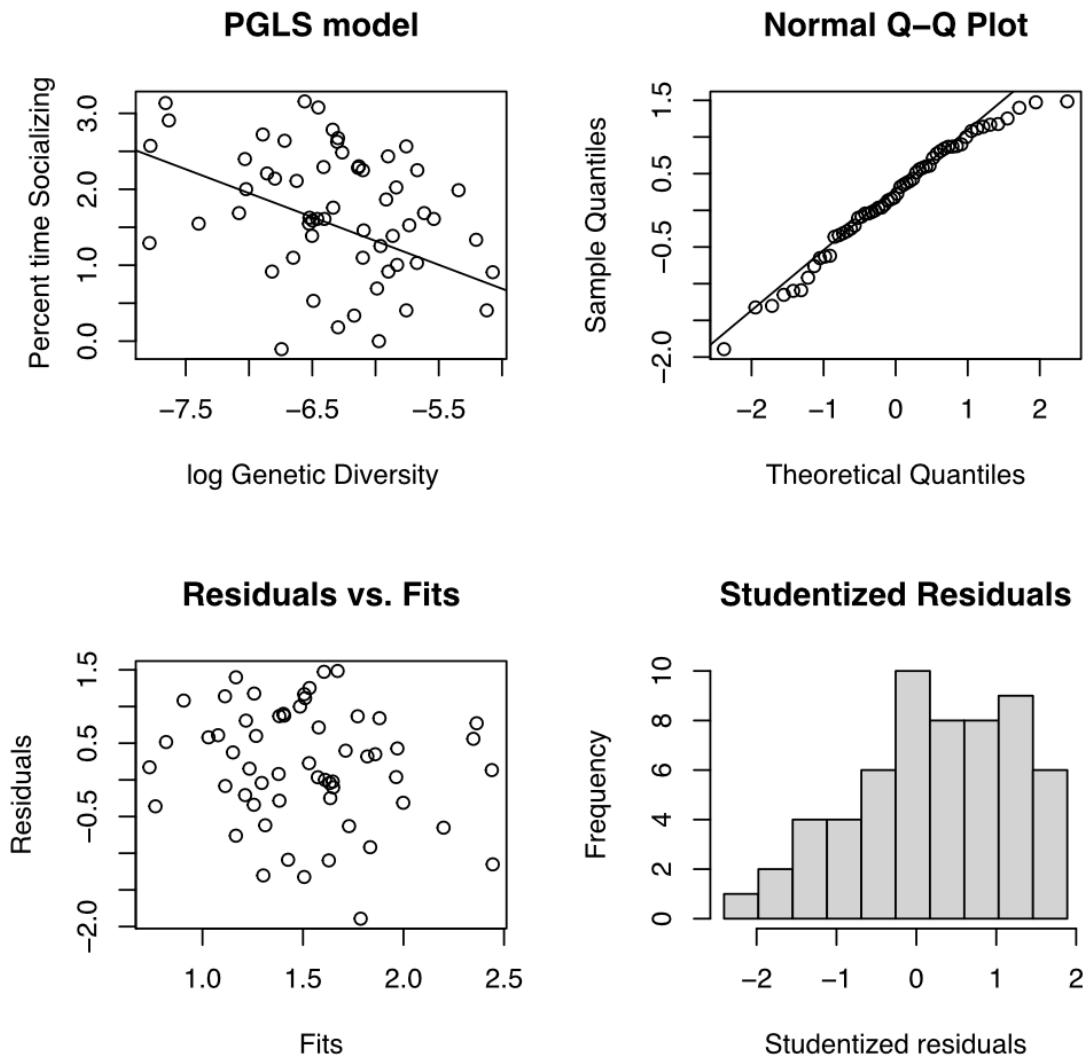


Fig. S41. PGLS Diagnostic plots for log genetic diversity predicted by time spent socializing.

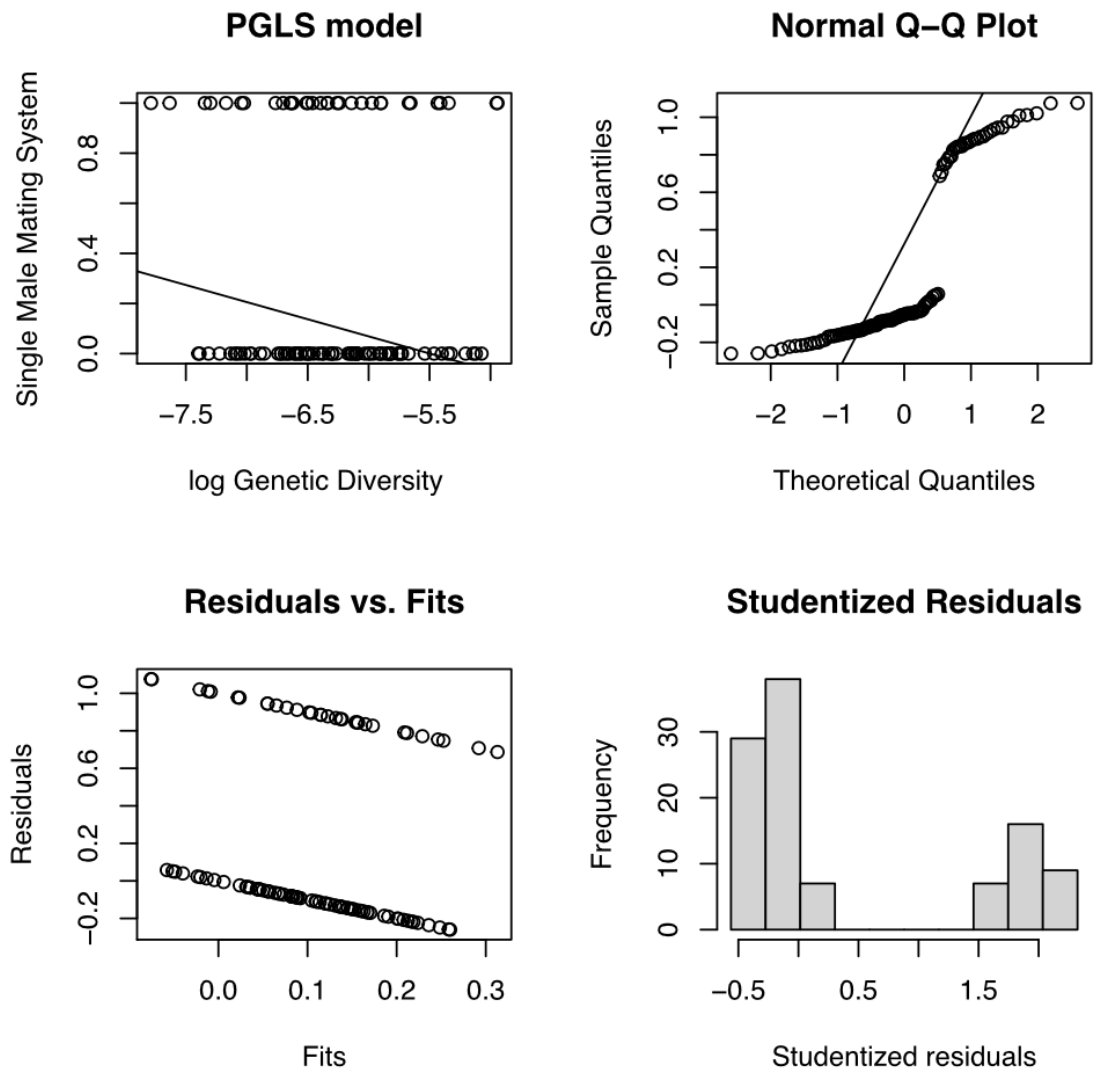


Fig. S42.
 PGLS Diagnostic plots for log genetic diversity predicted by single male mating system.

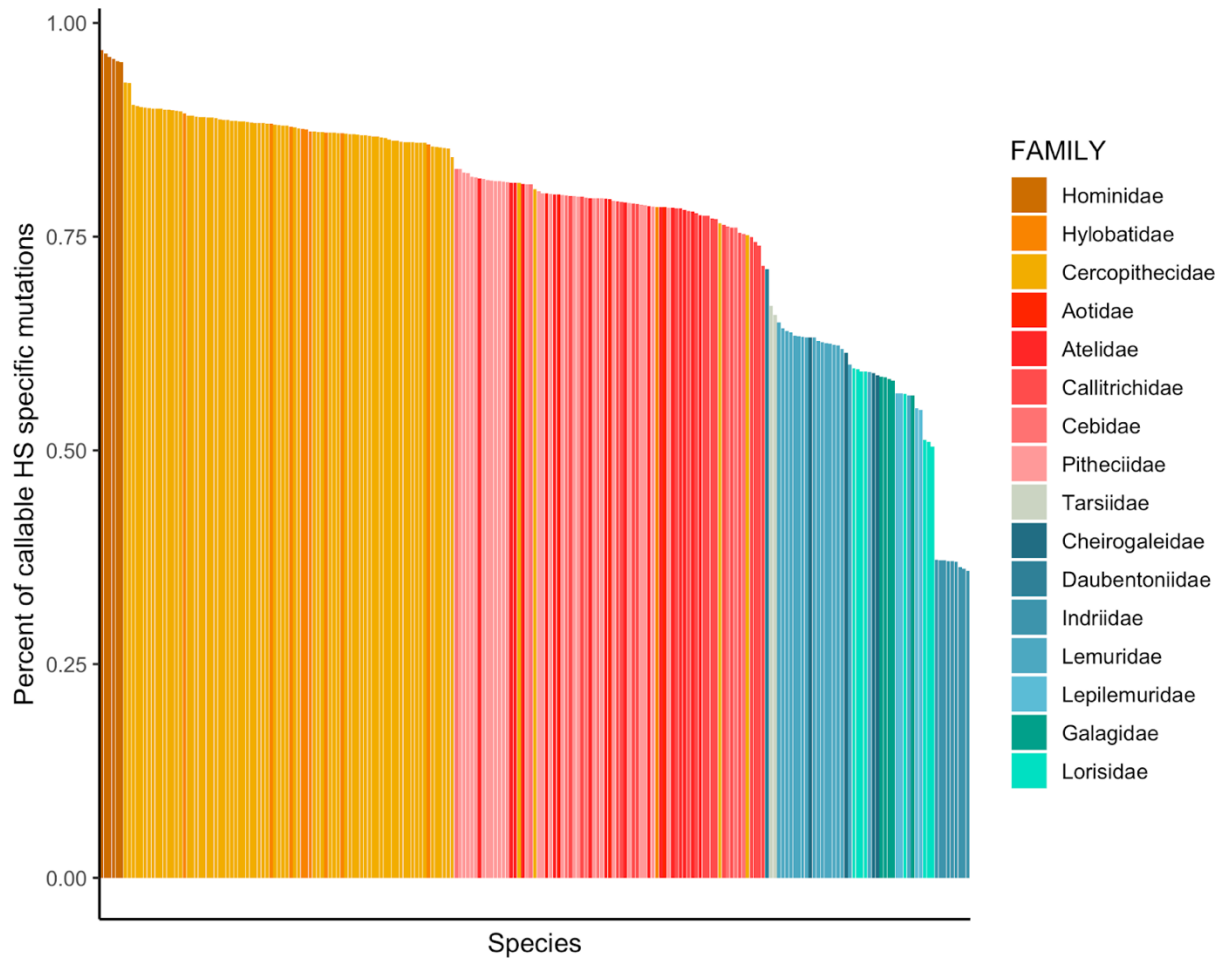


Fig. S43. Proportion of putatively human-specific candidate sites that are assessable across other species analyzed in this study

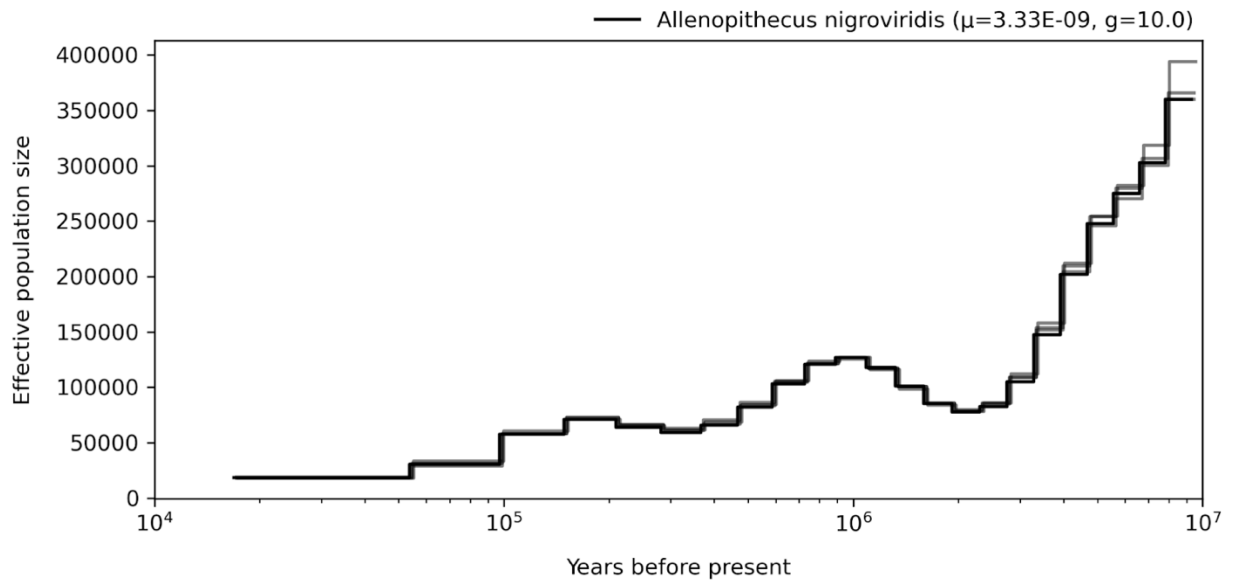


Fig. S44. Demographic history of the genus *Allenopithecus*, mutation rates for scaling are averaged per genus.

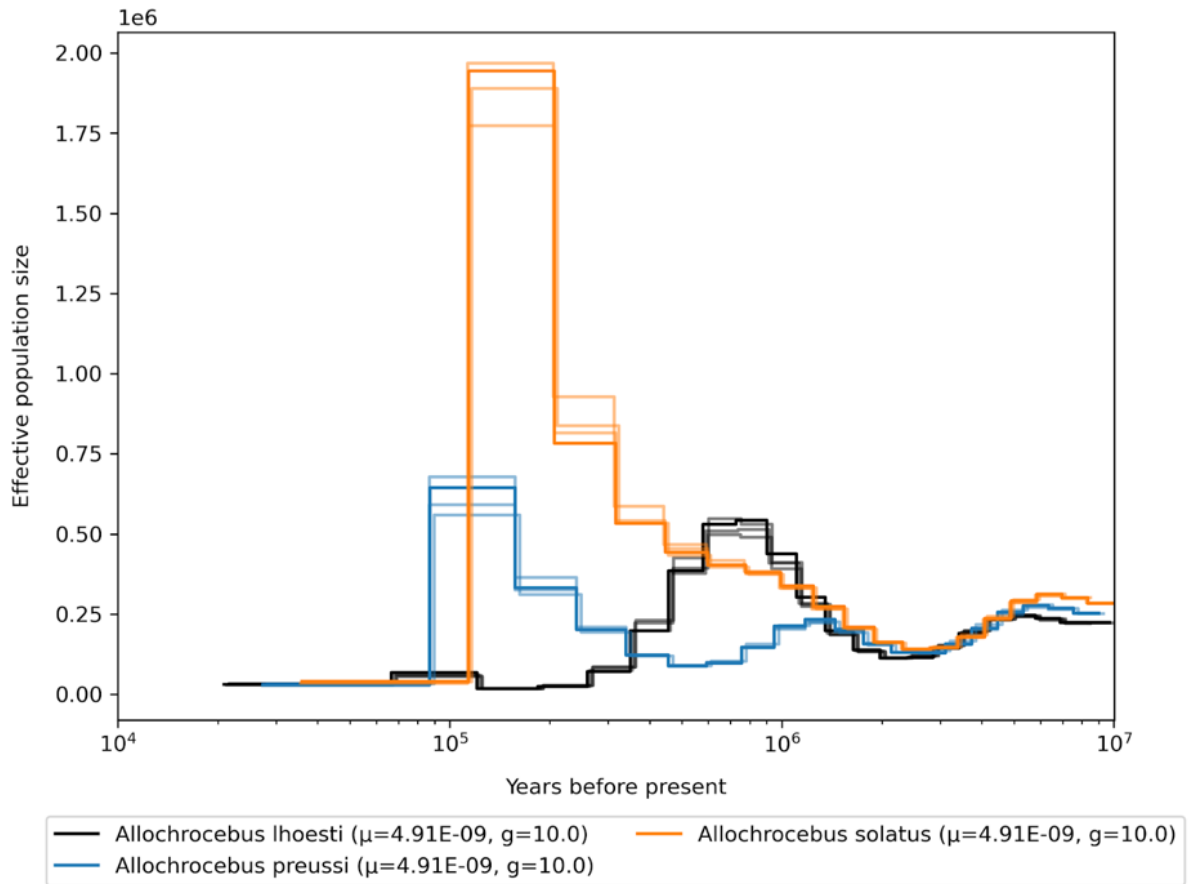


Fig. S45. Demographic history of the genus *Allochrocebus*, mutation rates for scaling are averaged per genus.

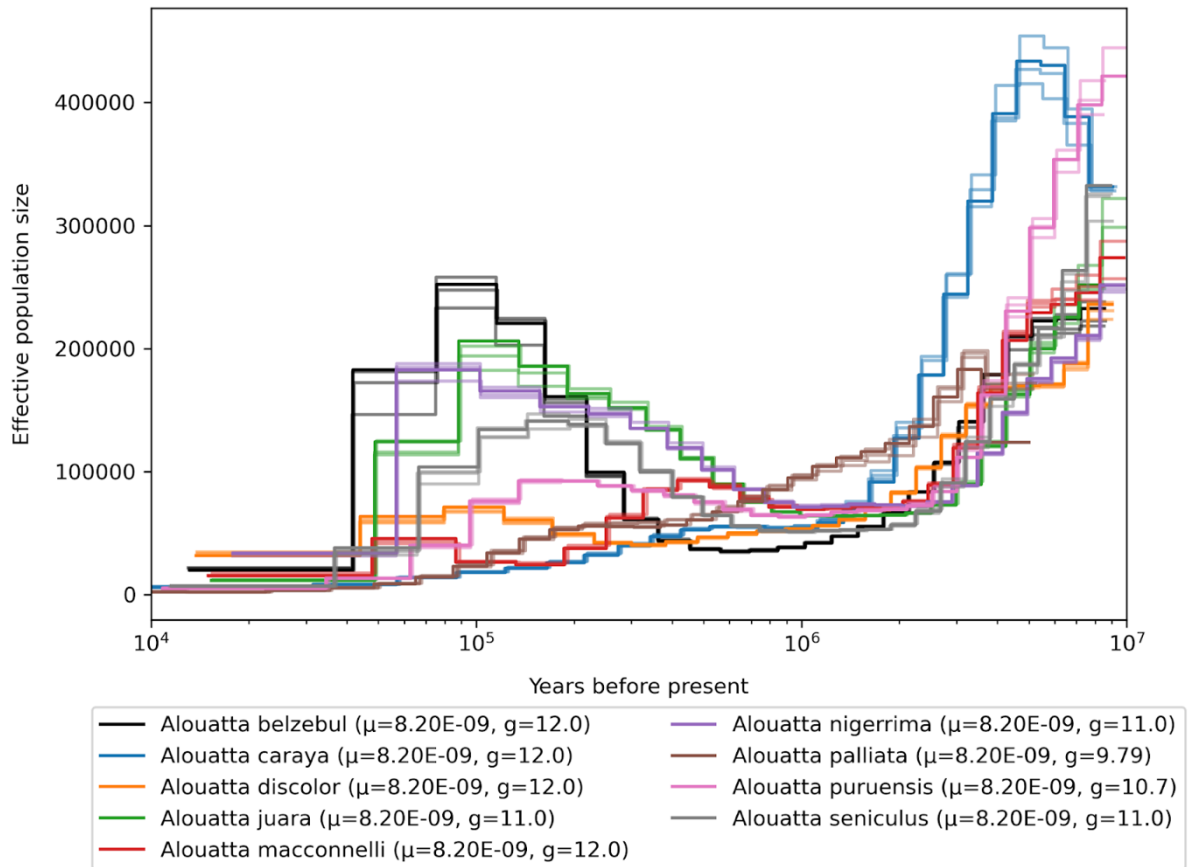


Fig. S46. Demographic history of the genus *Alouatta*, mutation rates for scaling are averaged per genus.

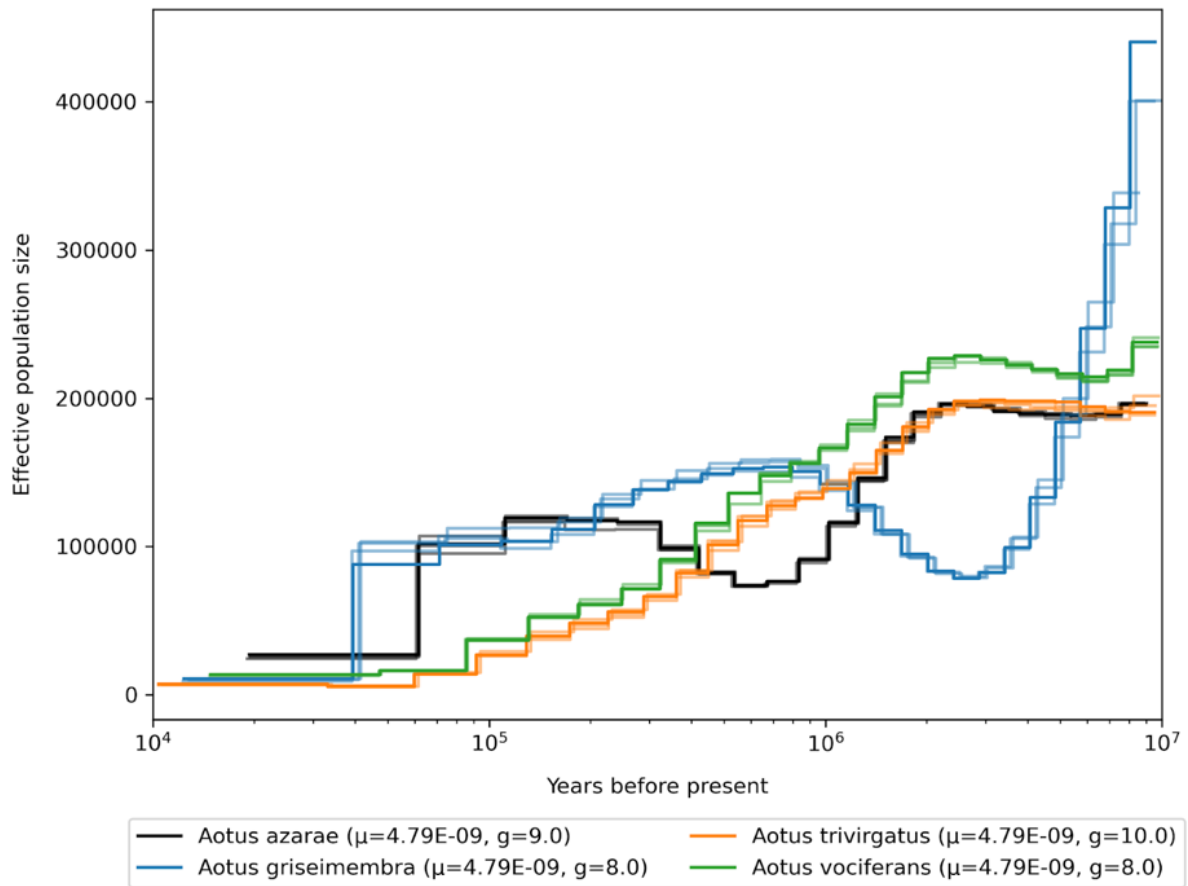


Fig. S47.
Demographic history of the genus *Aotus*, mutation rates for scaling are averaged per genus.

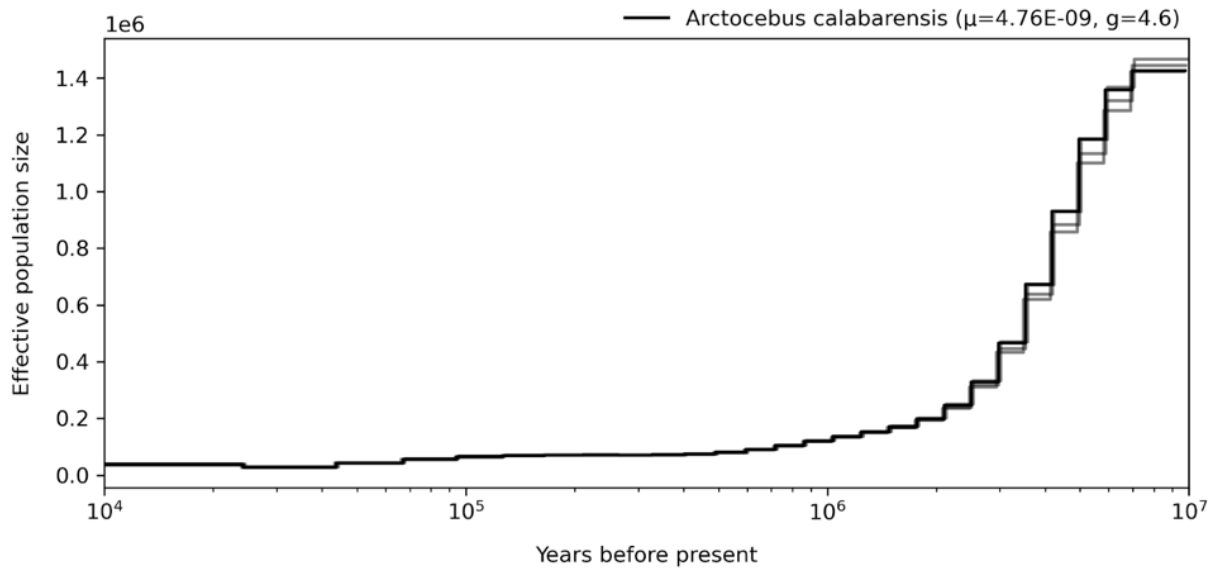


Fig. S48. Demographic history of the genus *Arctocebus*, mutation rates for scaling are averaged per genus.

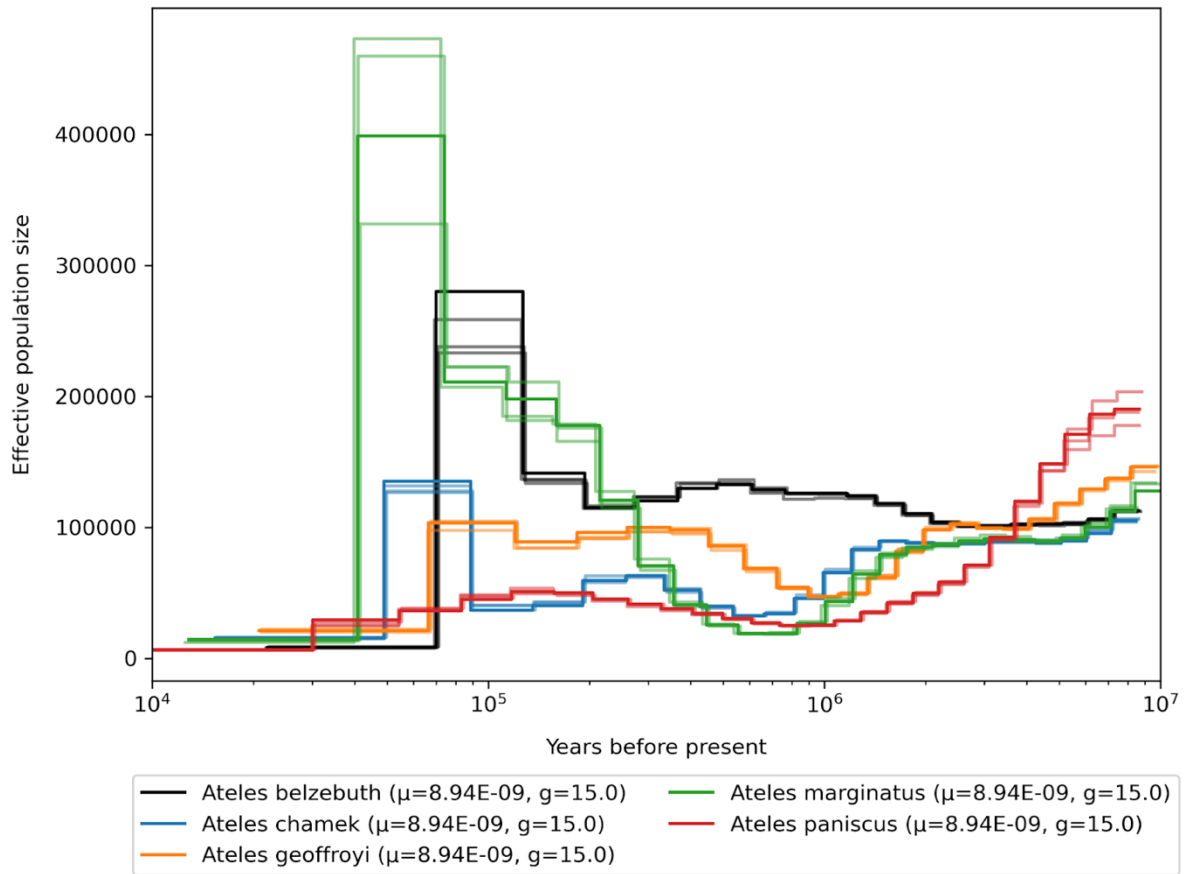


Fig. S49.
 Demographic history of the genus *Ateles*, mutation rates for scaling are averaged per genus.

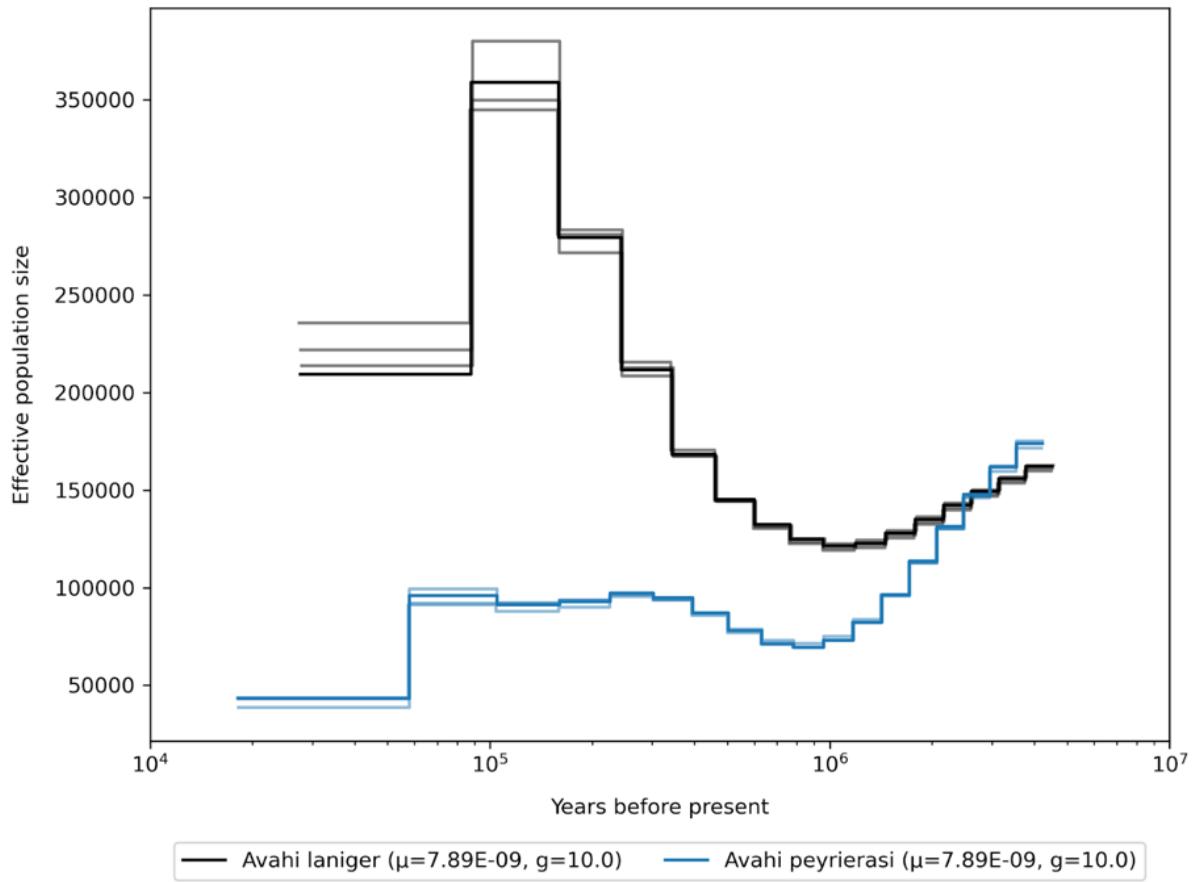


Fig. S50.
Demographic history of the genus *Avahi*, mutation rates for scaling are averaged per genus.

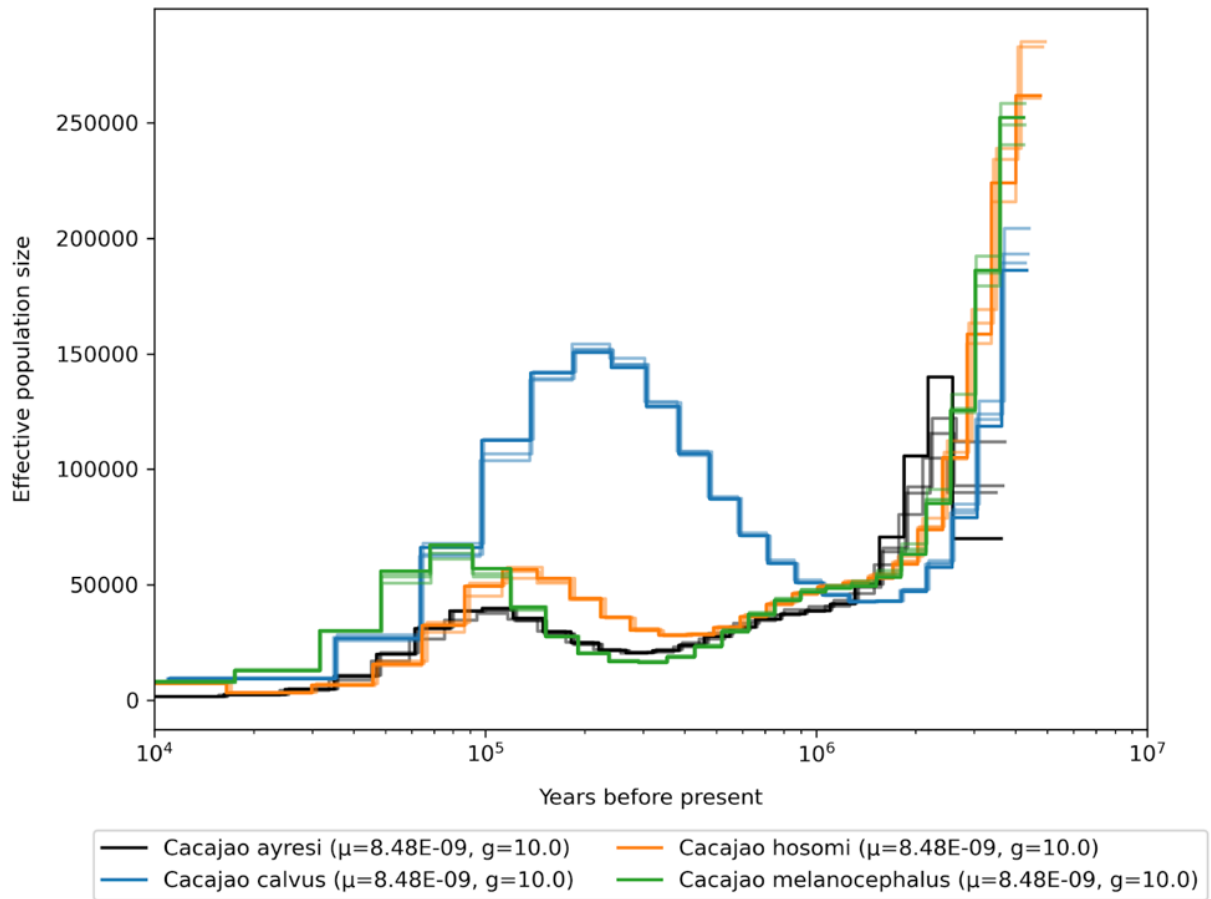


Fig. S51.
 Demographic history of the genus *Cacajao*, mutation rates for scaling are averaged per genus.

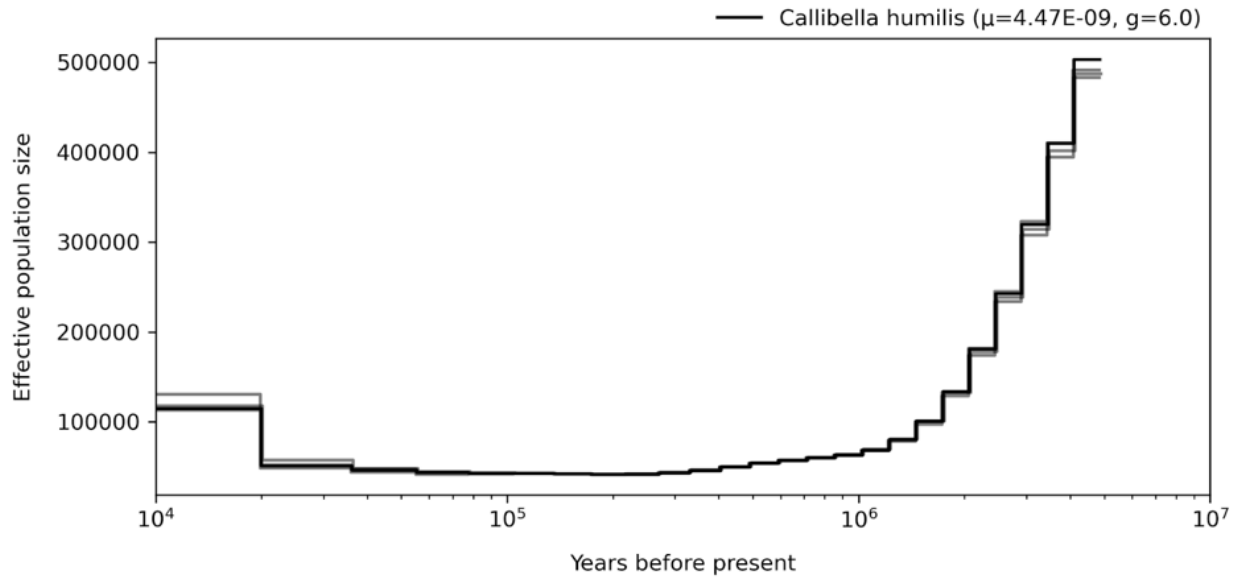


Fig. S52.
Demographic history of the genus *Calibella*, mutation rates for scaling are averaged per genus.

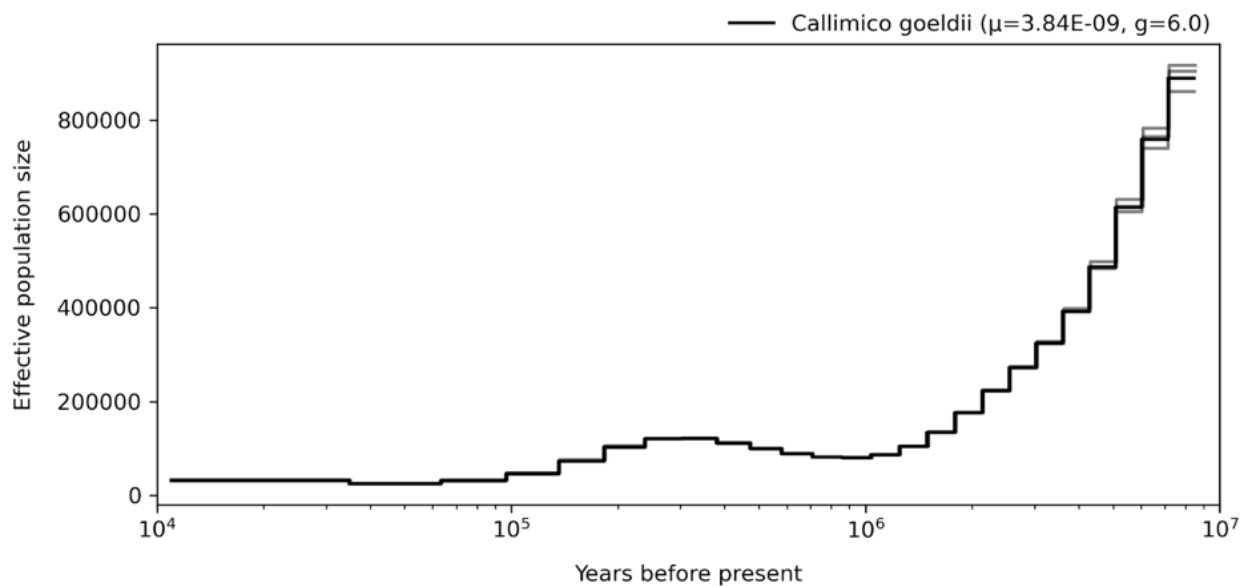


Fig. S53.

Demographic history of the genus *Callimico*, mutation rates for scaling are averaged per genus.

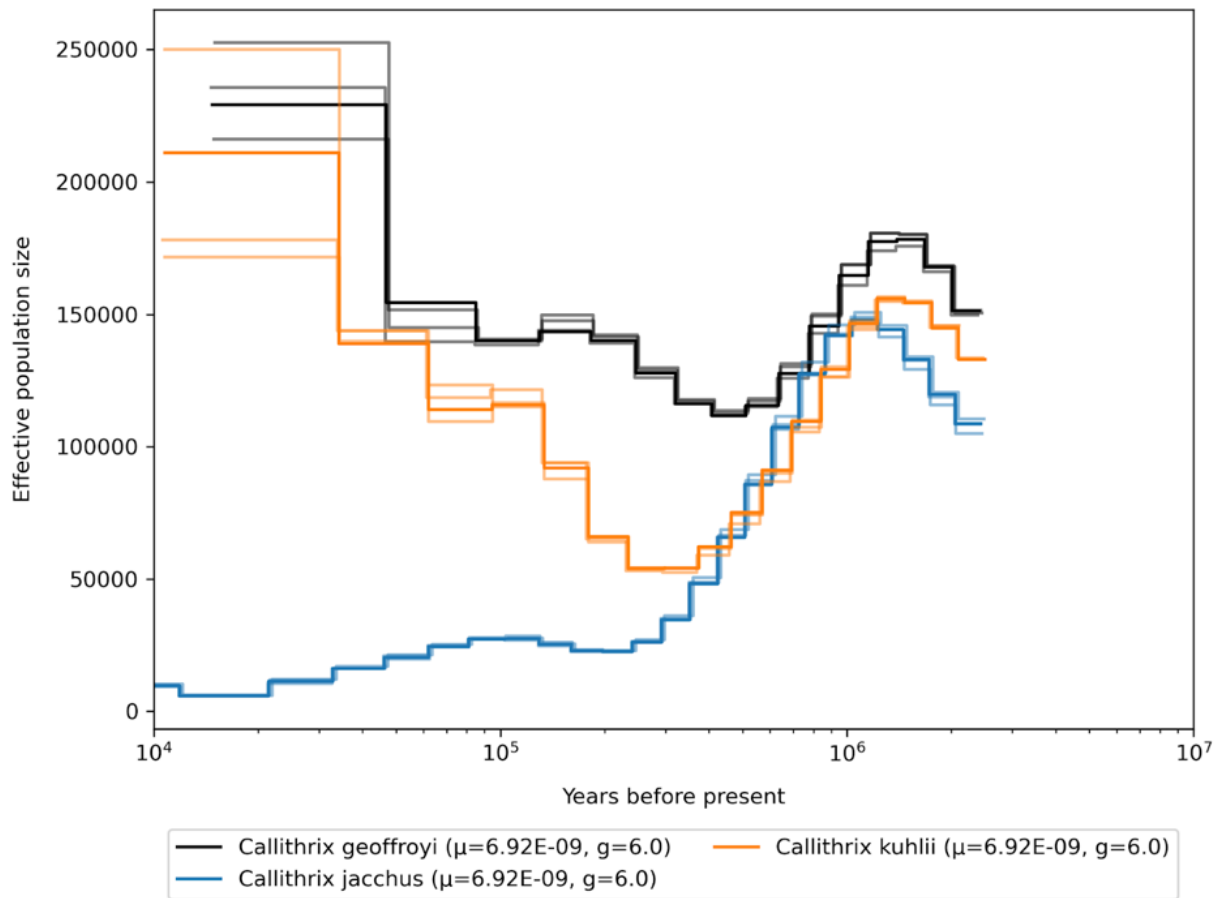


Fig. S54. Demographic history of the genus *Callithrix*, mutation rates for scaling are averaged per genus.

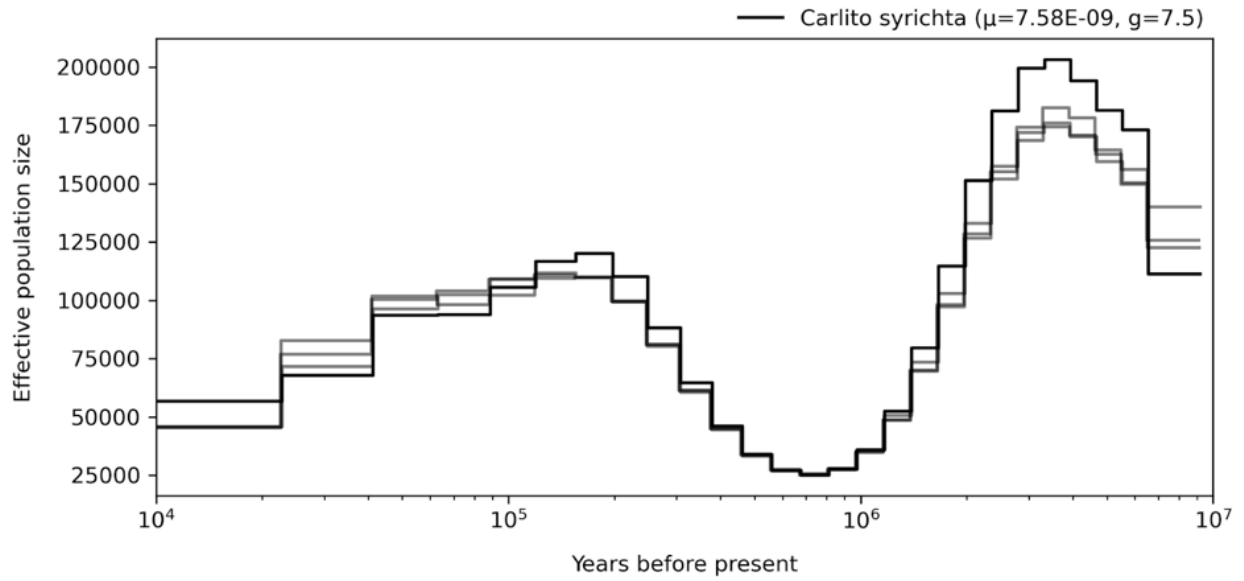


Fig. S55.
Demographic history of the genus *Carlito*, mutation rates for scaling are averaged per genus.

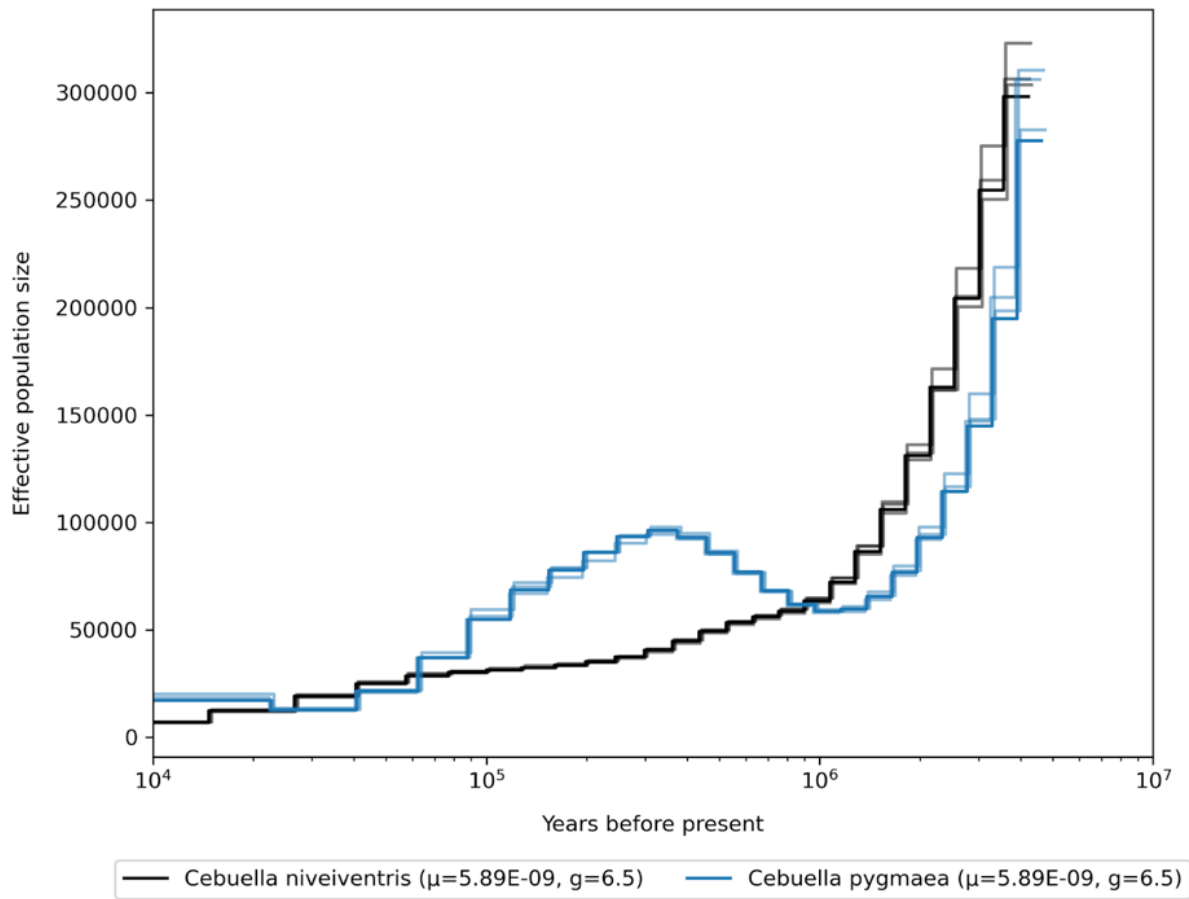


Fig. S56. Demographic history of the genus *Cebuella*, mutation rates for scaling are averaged per genus.

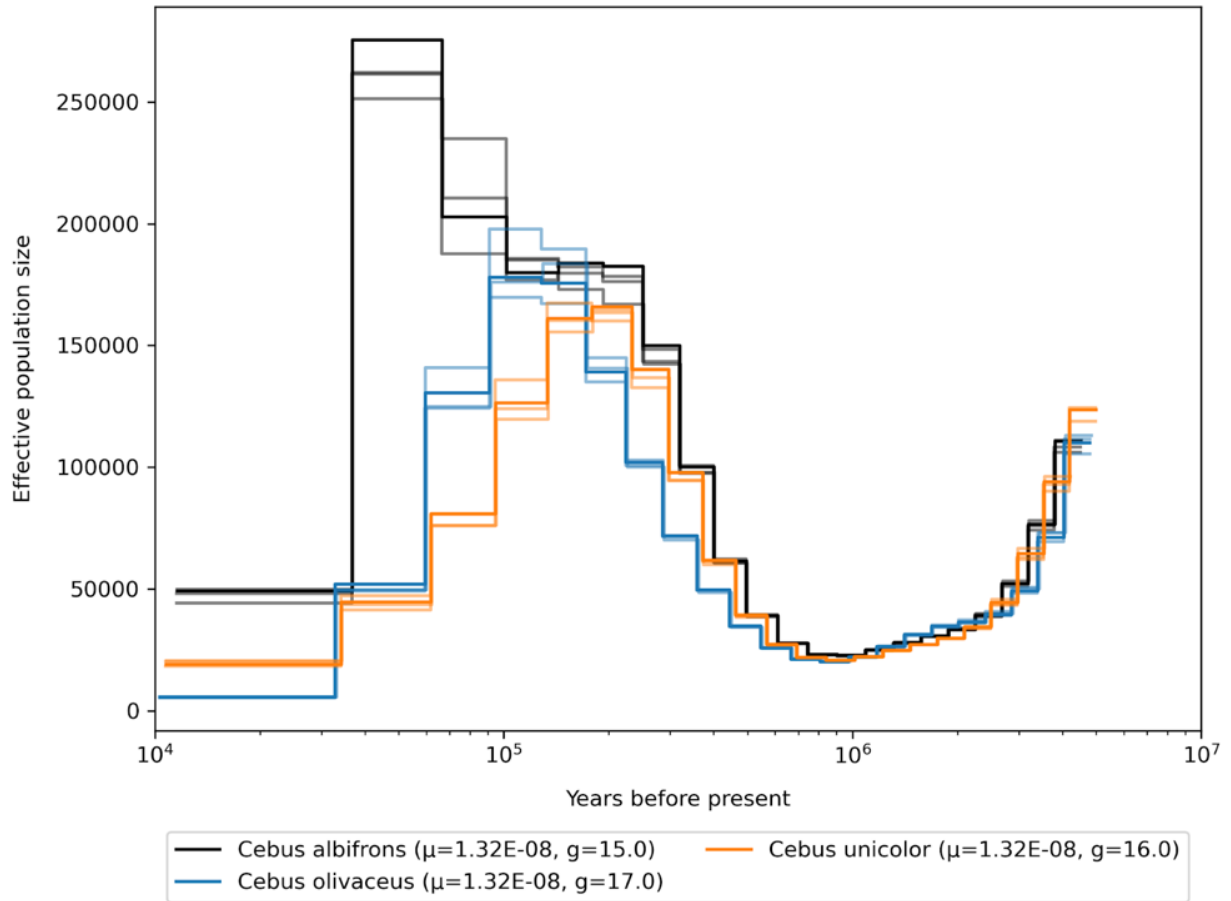


Fig. S57.
 Demographic history of the genus *Cebus*, mutation rates for scaling are averaged per genus.

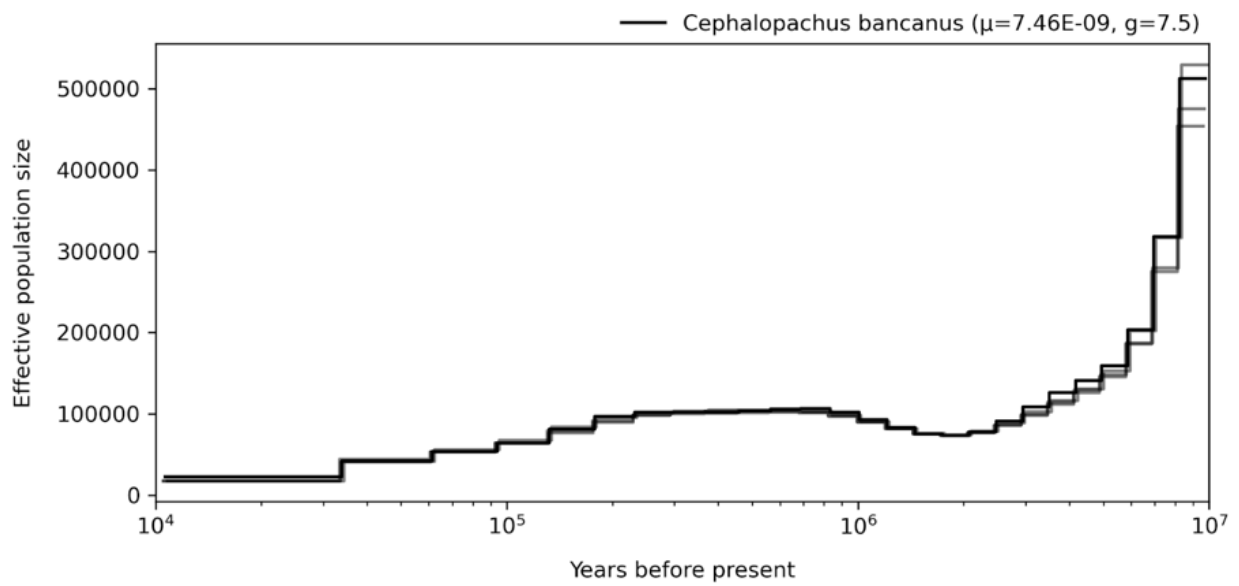


Fig. S58.

Demographic history of the genus *Cephalopachus*, mutation rates for scaling are averaged per genus.

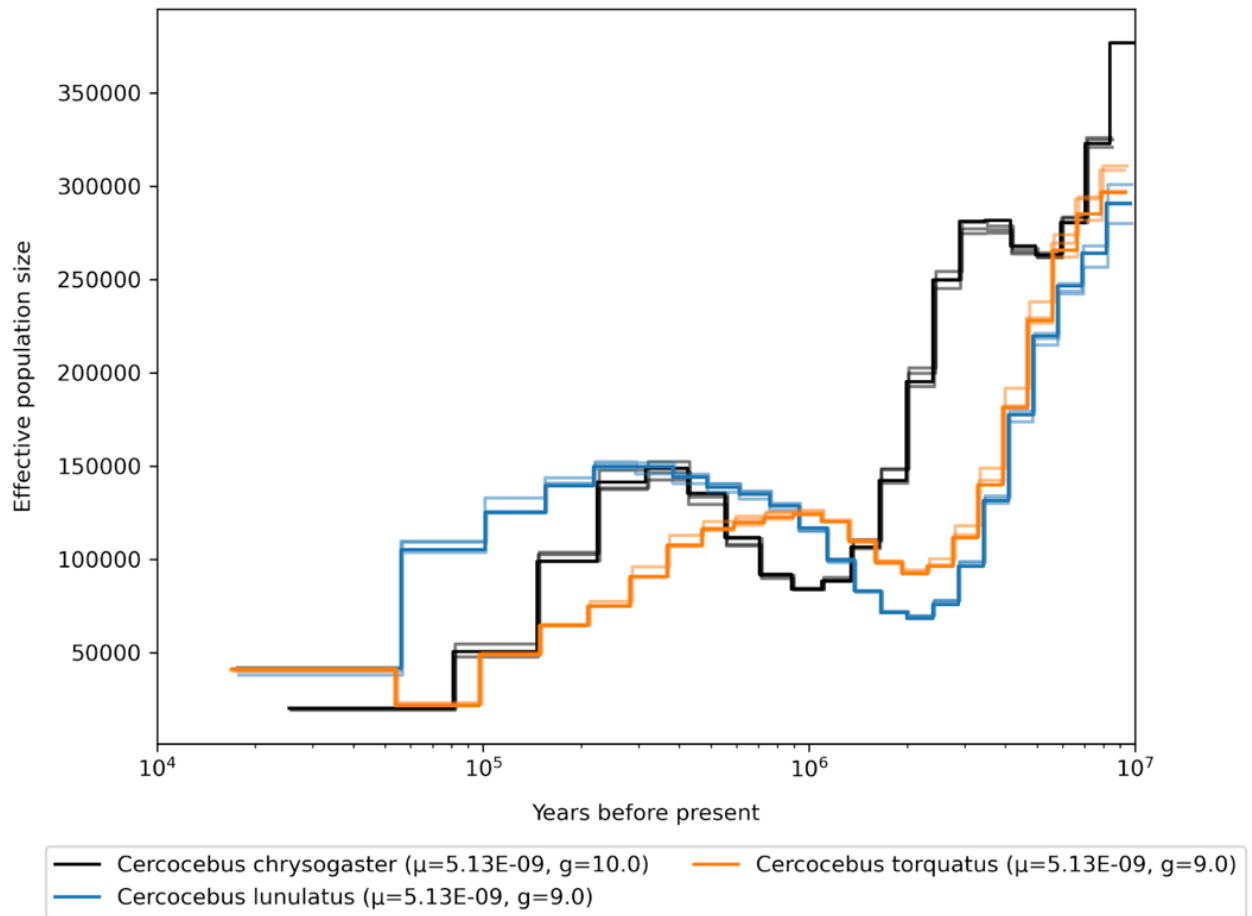


Fig. S59. Demographic history of the genus *Cercocebus*, mutation rates for scaling are averaged per genus.

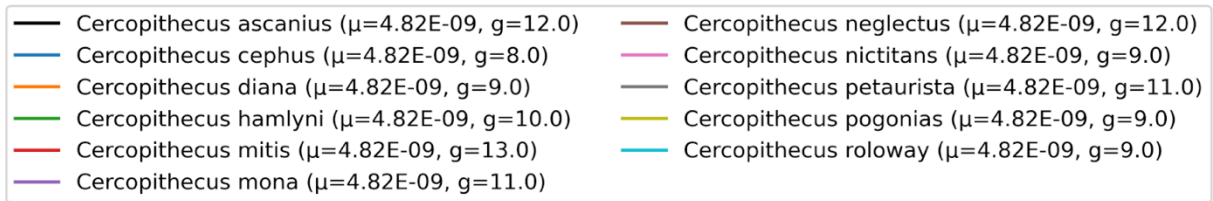
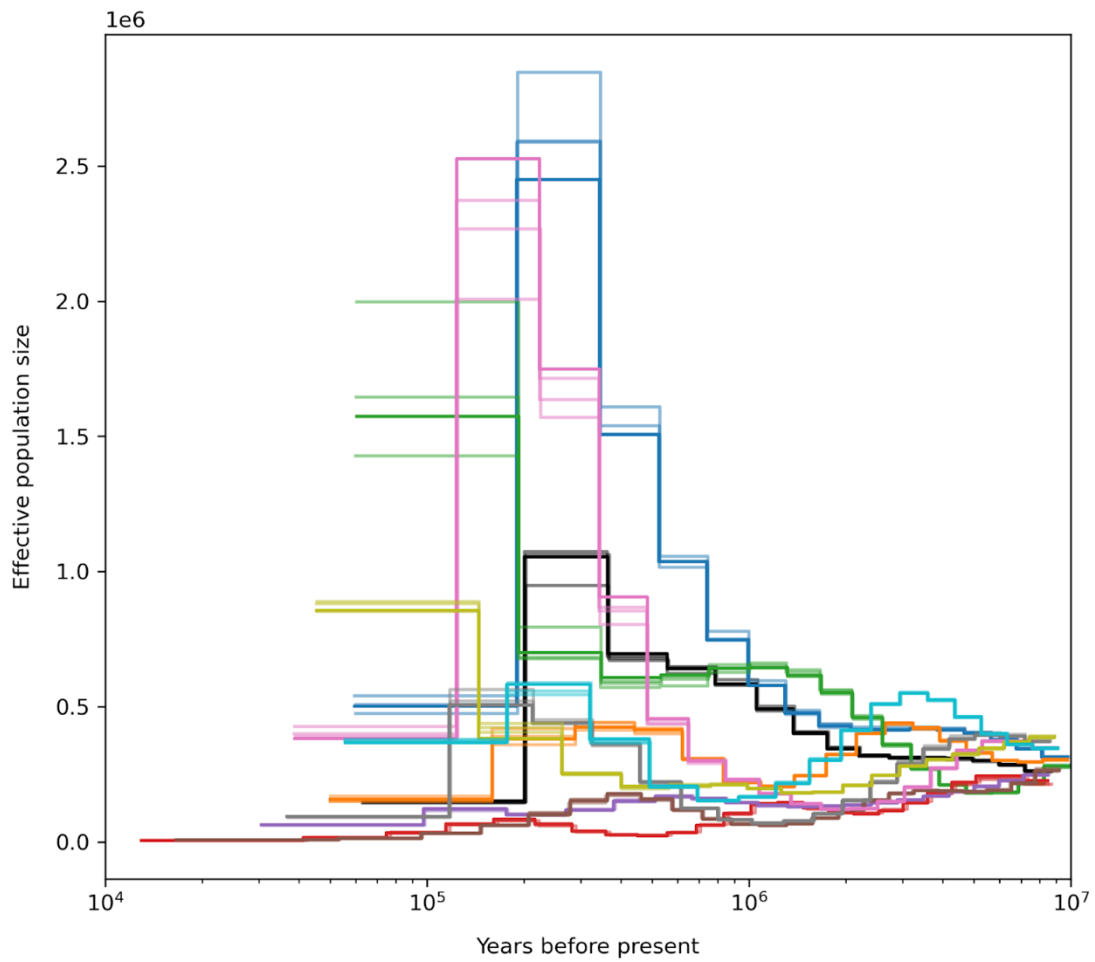


Fig. S60.

Demographic history of the genus *Cercopithecus*, mutation rates for scaling are averaged per genus.

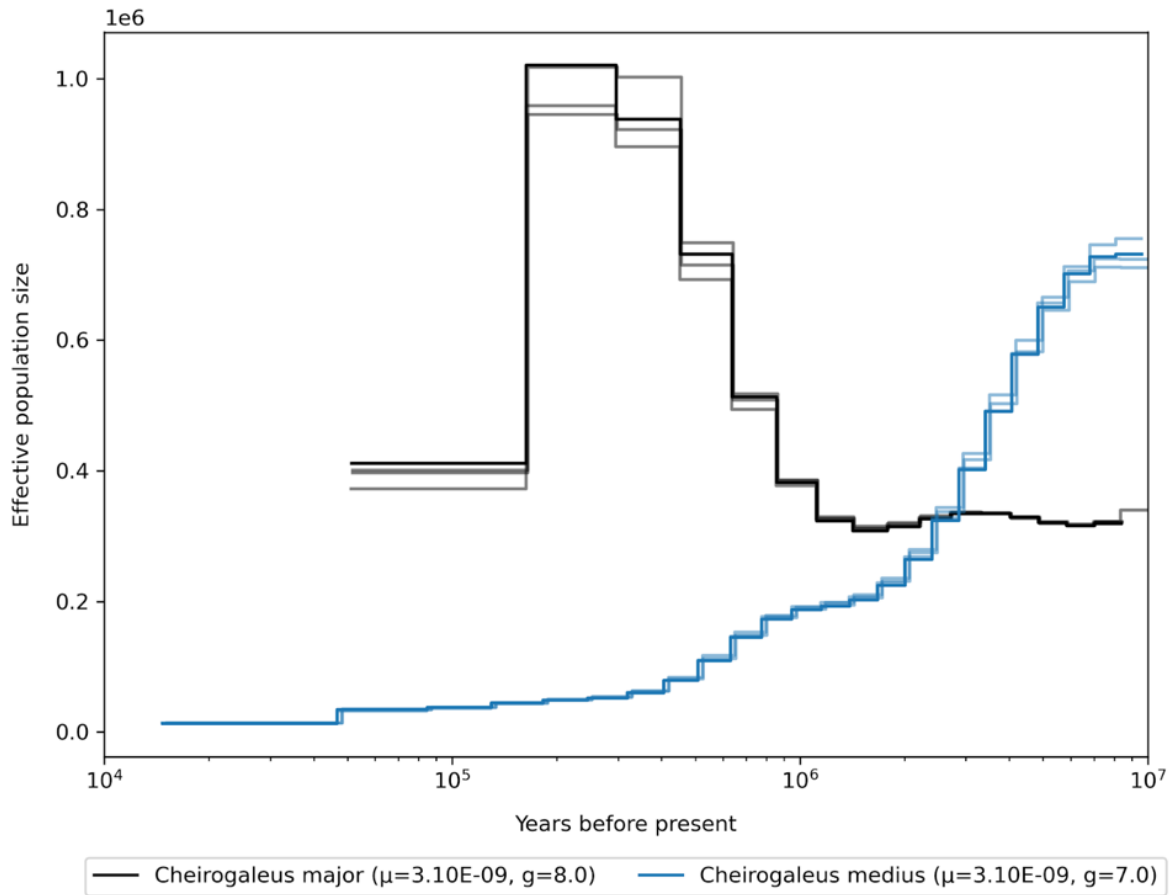


Fig. S61. Demographic history of the genus *Cheirogaleus*, mutation rates for scaling are averaged per genus.

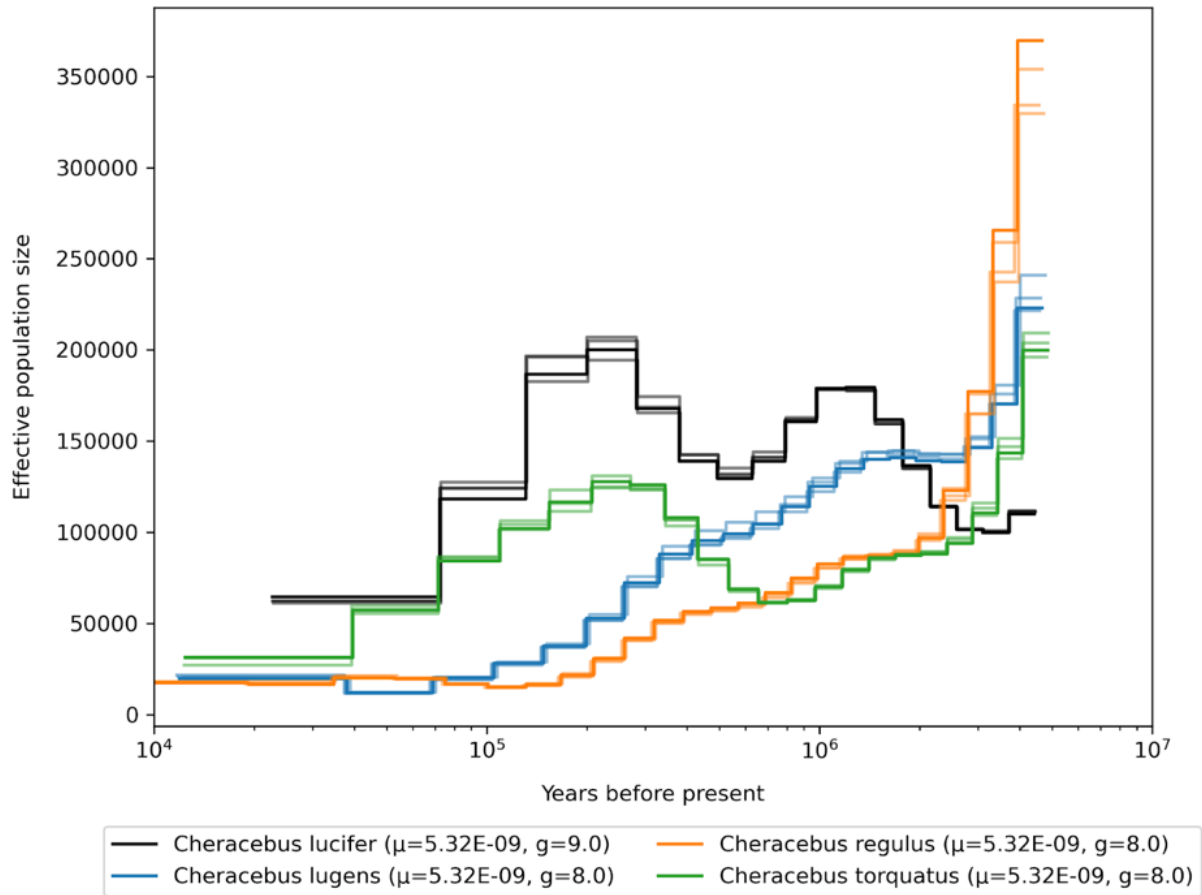


Fig. S62.
Demographic history of the genus *Cheracebus*, mutation rates for scaling are averaged per genus.

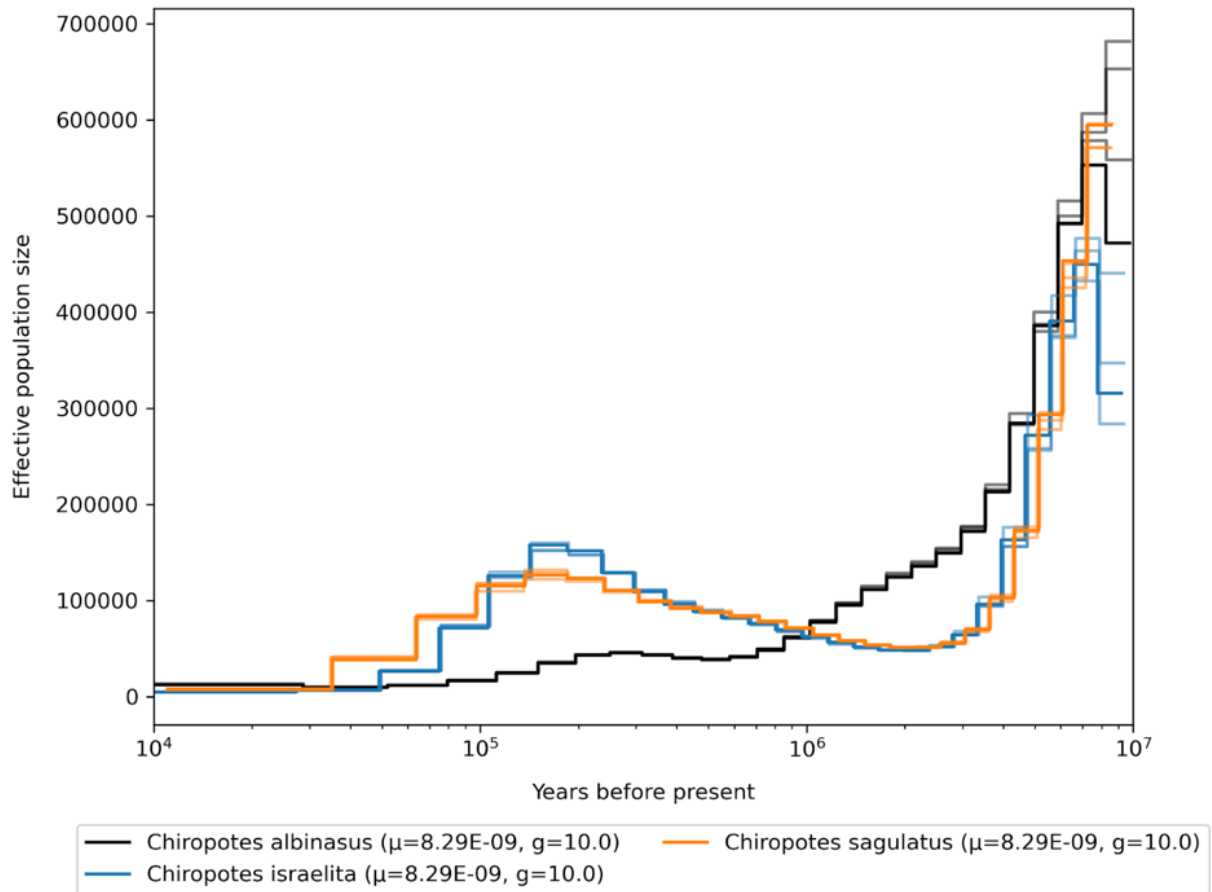


Fig. S63. Demographic history of the genus *Chiropotes*, mutation rates for scaling are averaged per genus.

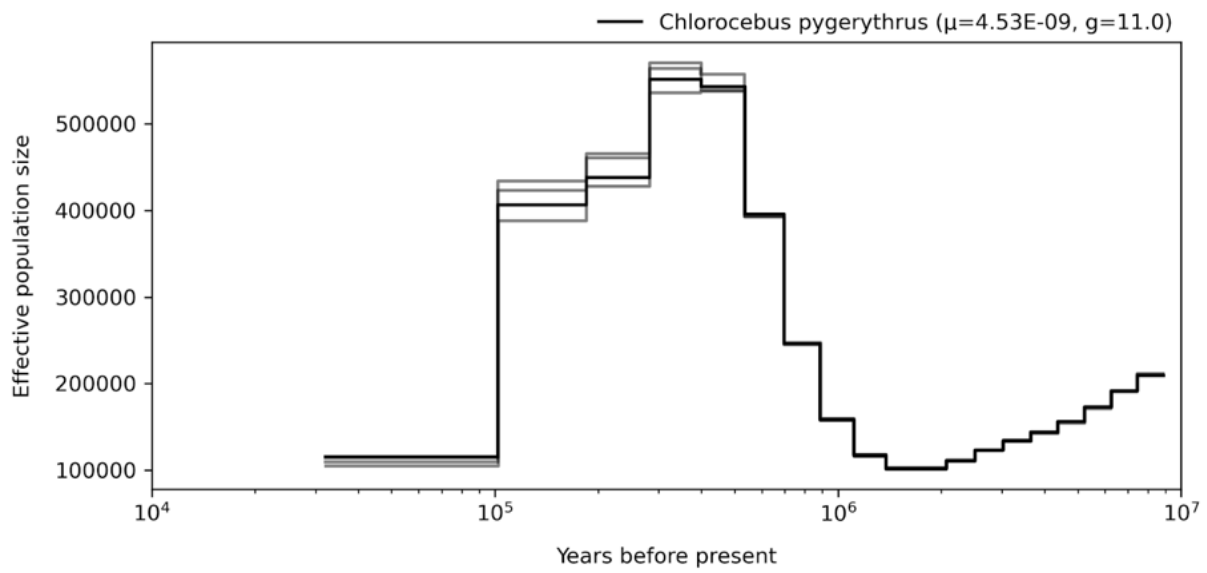


Fig. S64. Demographic history of the genus *Chlorocebus*, mutation rates for scaling are averaged per genus.

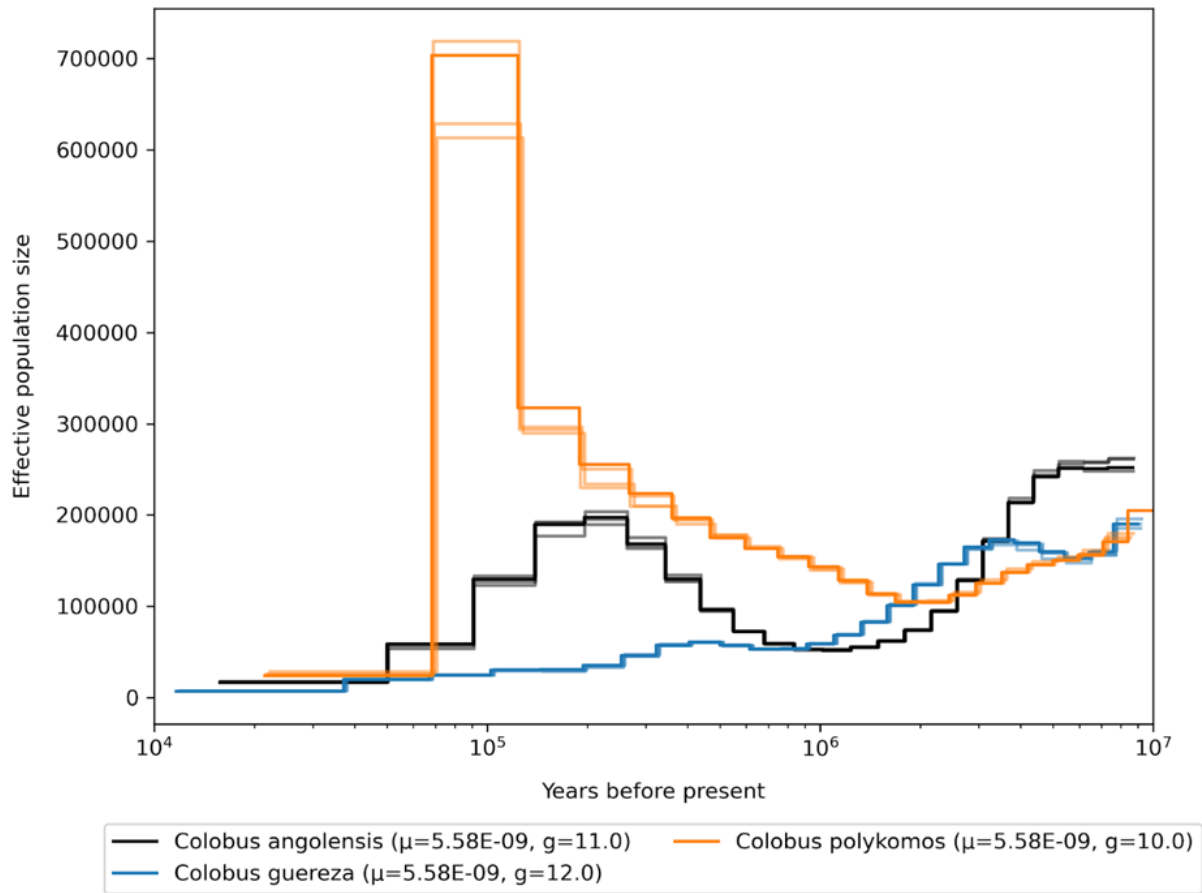


Fig. S65.
 Demographic history of the genus *Colobus*, mutation rates for scaling are averaged per genus.

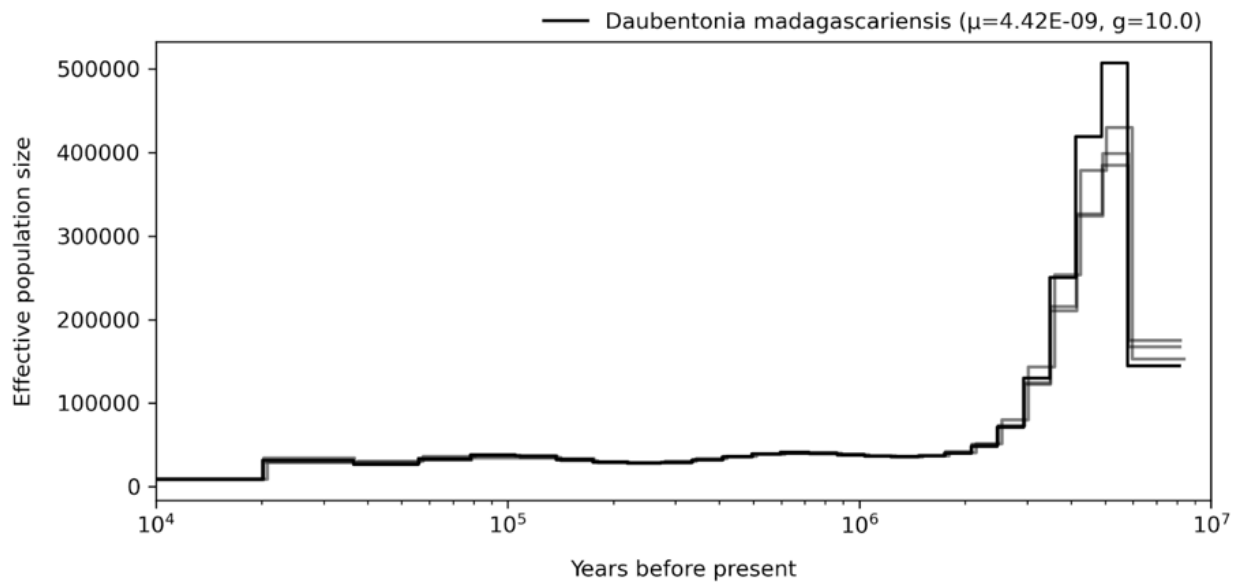


Fig. S66. Demographic history of the genus *Daubentonia*, mutation rates for scaling are averaged per genus.

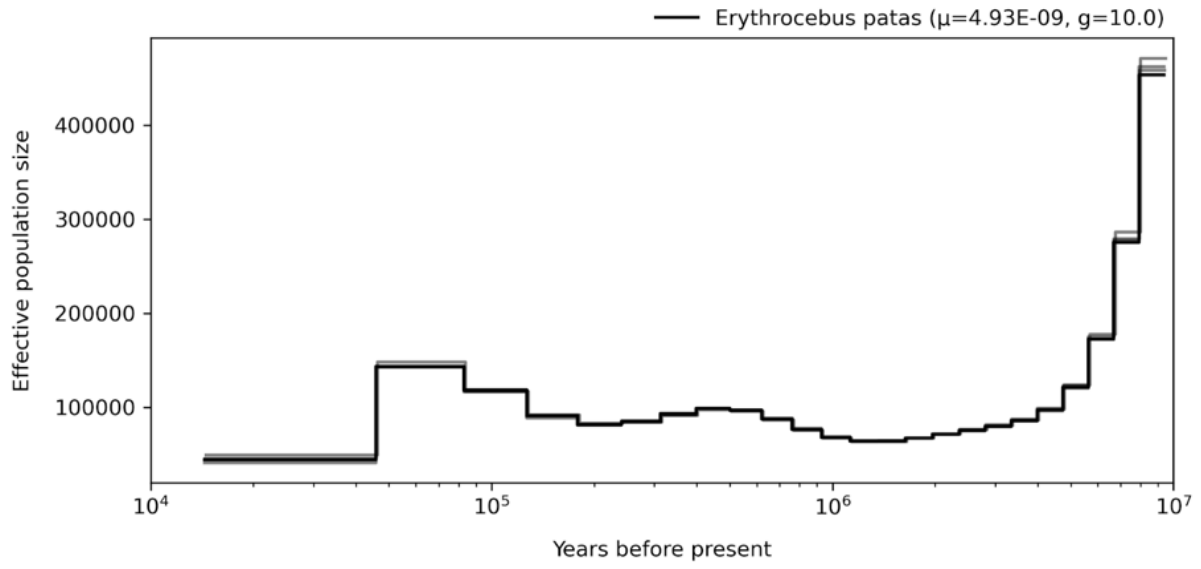


Fig. S67. Demographic history of the genus *Erythrocebus*, mutation rates for scaling are averaged per genus.

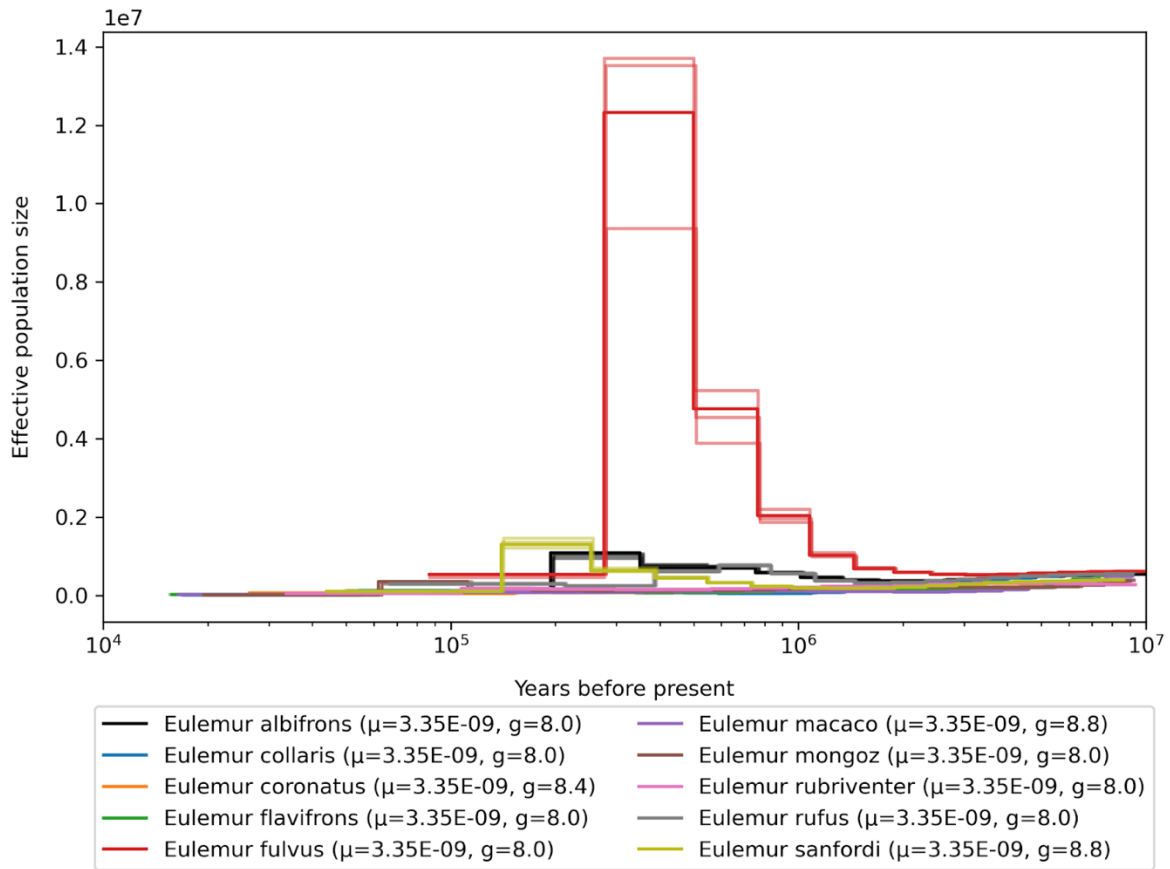


Fig. S68. Demographic history of the genus *Eulemur*, mutation rates for scaling are averaged per genus.

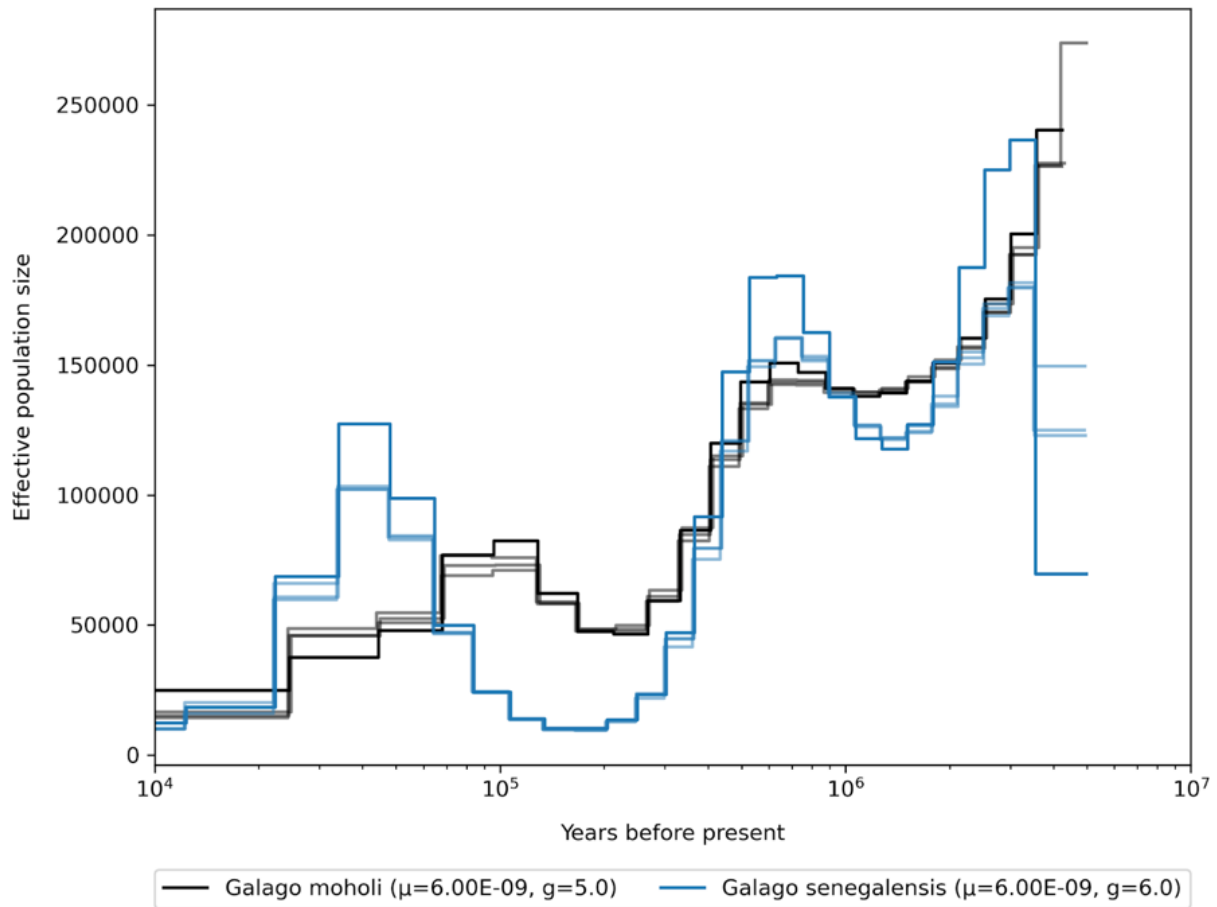


Fig. S69.
Demographic history of the genus *Galago*, mutation rates for scaling are averaged per genus.

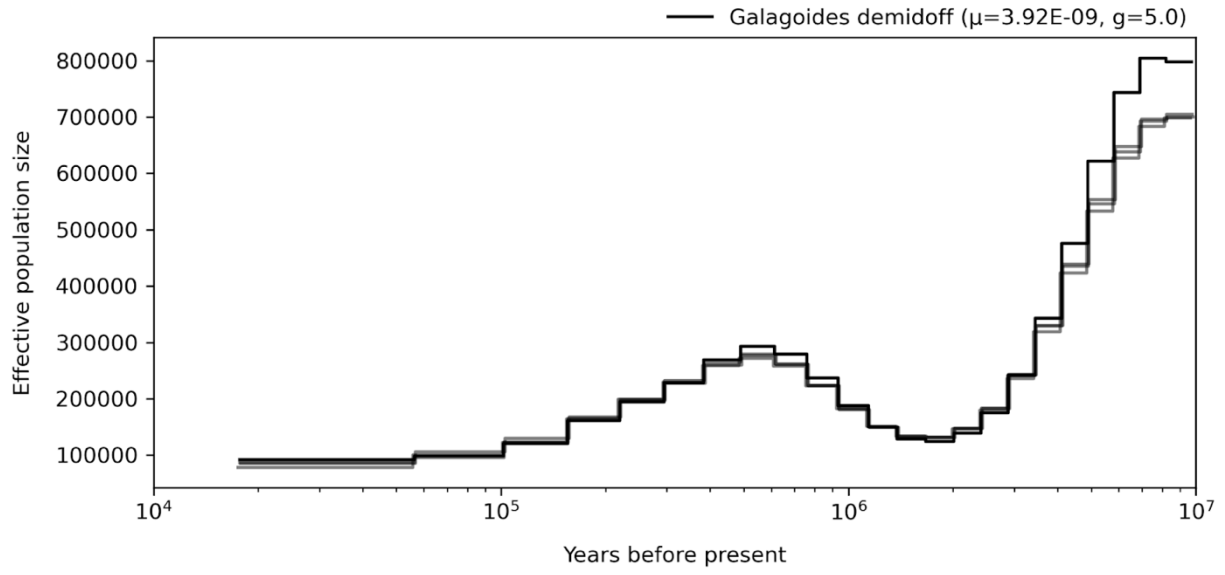


Fig. S70. Demographic history of the genus *Galagoides*, mutation rates for scaling are averaged per genus.

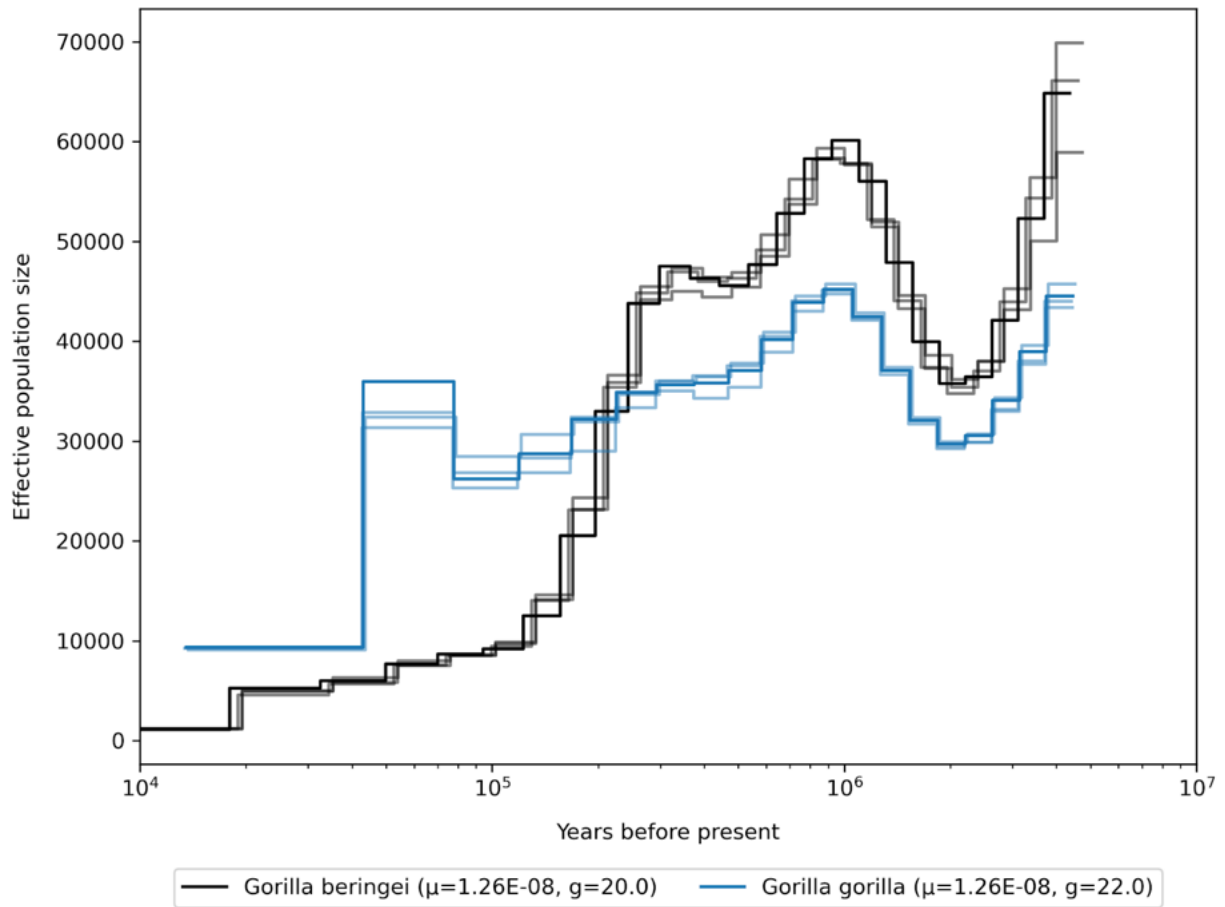


Fig. S71. Demographic history of the genus *Gorilla*, mutation rates for scaling are averaged per genus.

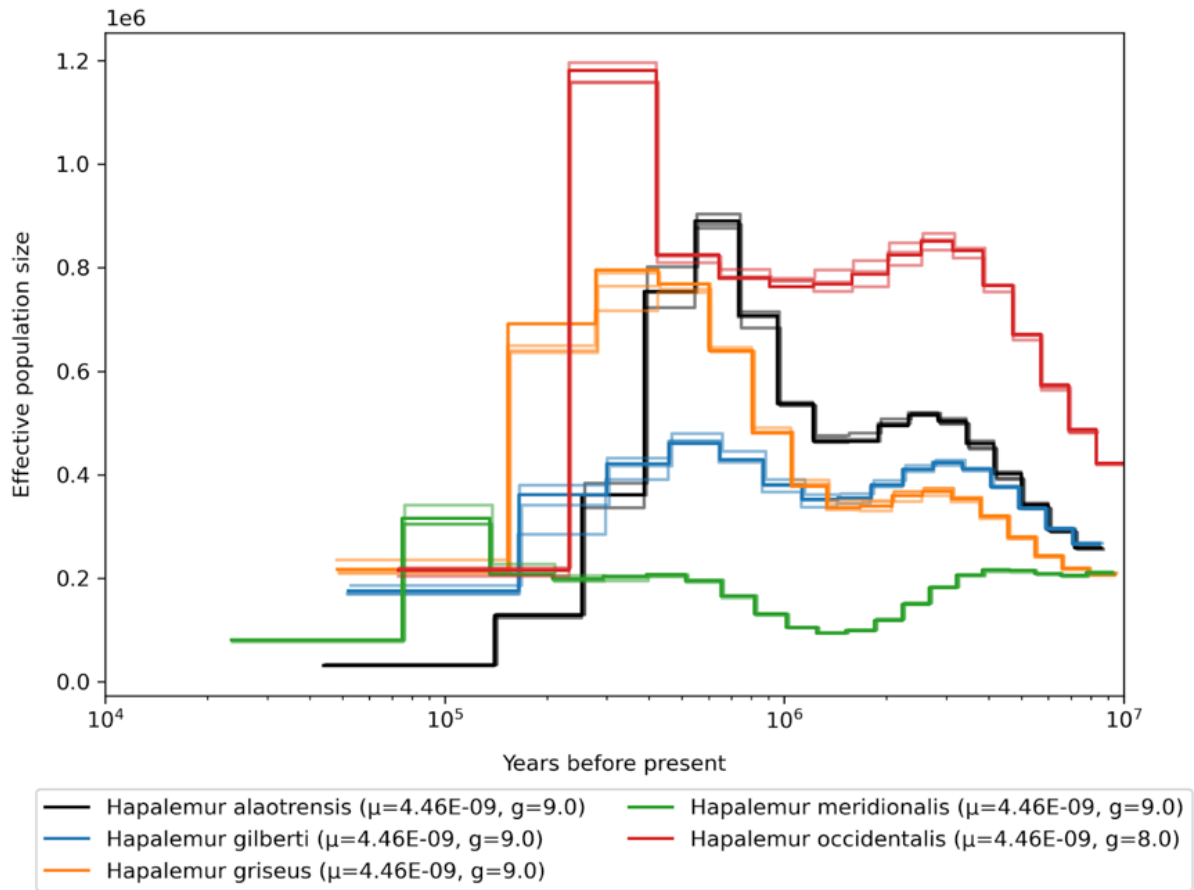


Fig. S72.
Demographic history of the genus *Hapalemur*, mutation rates for scaling are averaged per genus.

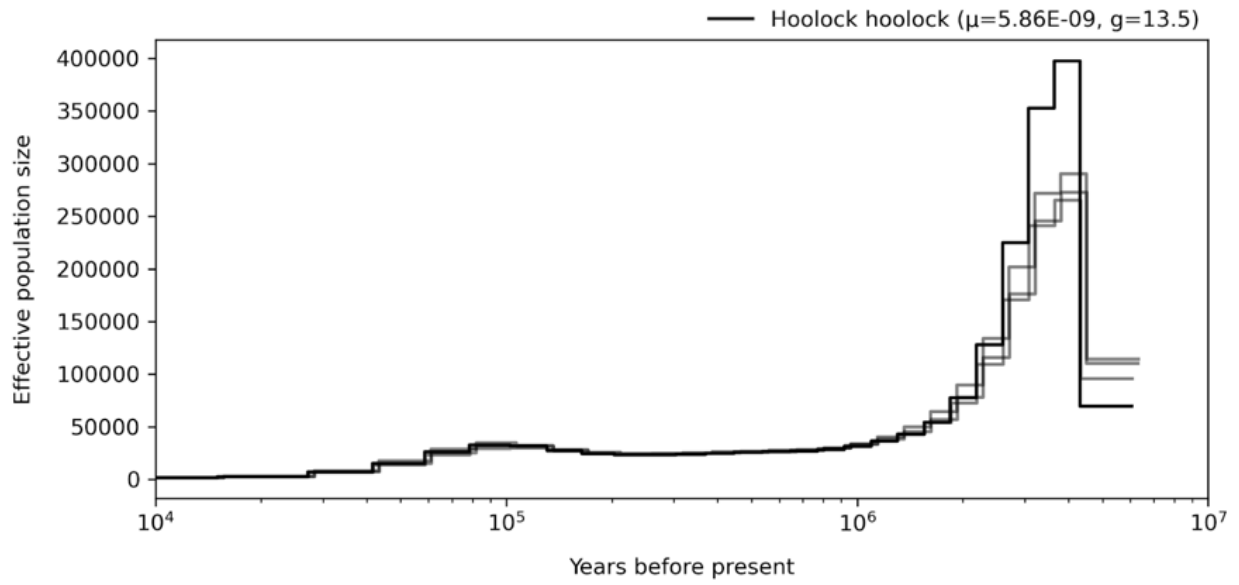


Fig. S73.
Demographic history of the genus *Hoolock*, mutation rates for scaling are averaged per genus.

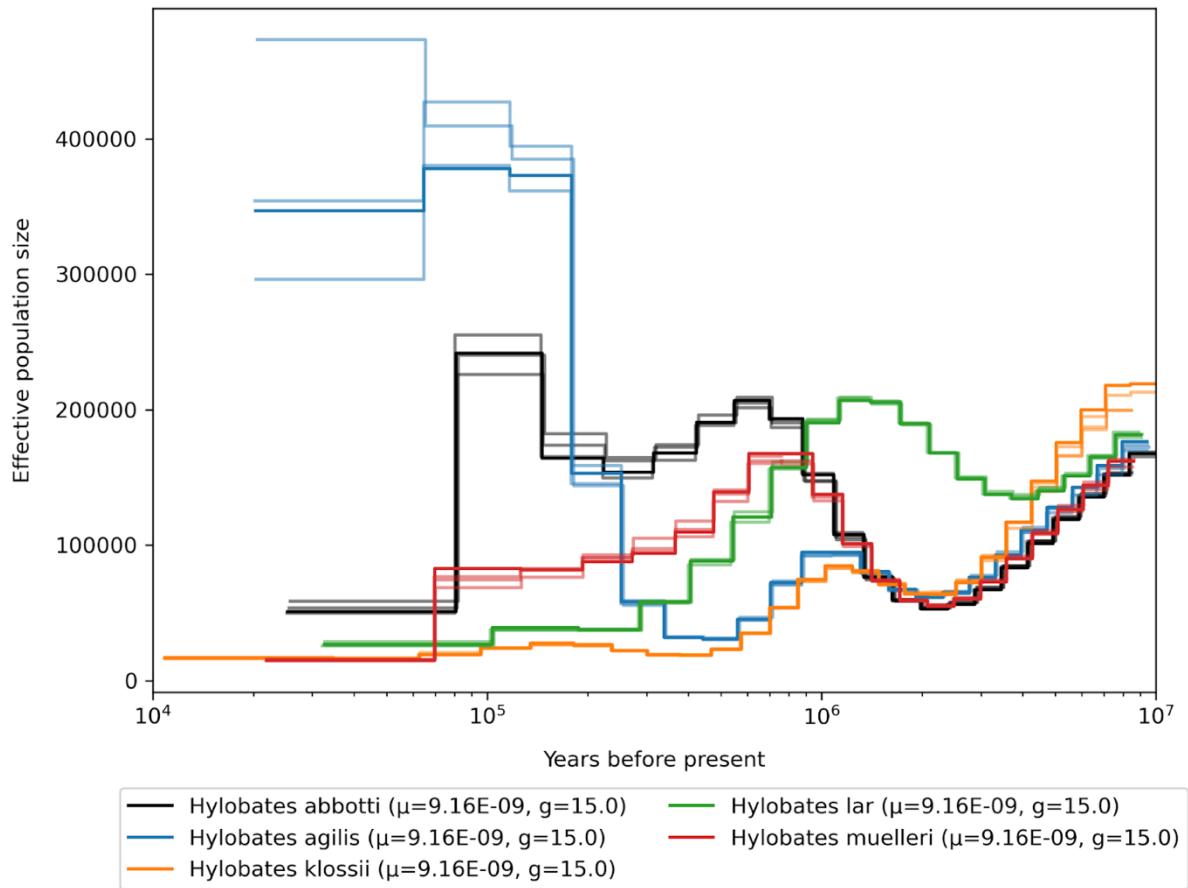


Fig. S74.
Demographic history of the genus *Hylobates*, mutation rates for scaling are averaged per genus.

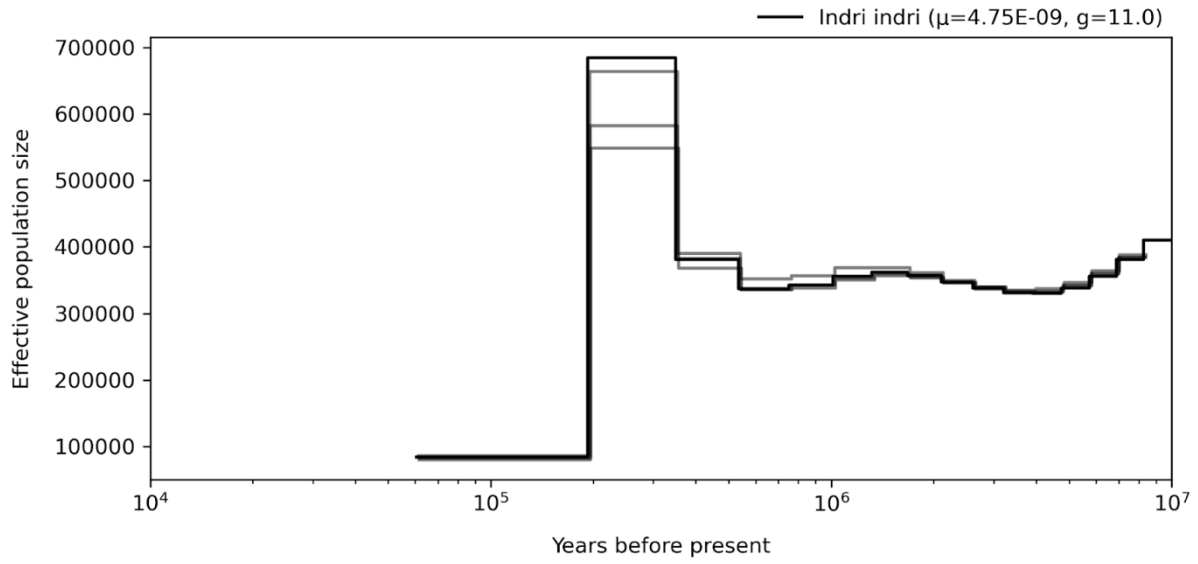


Fig. S75. Demographic history of the genus *Indri*, mutation rates for scaling are averaged per genus.

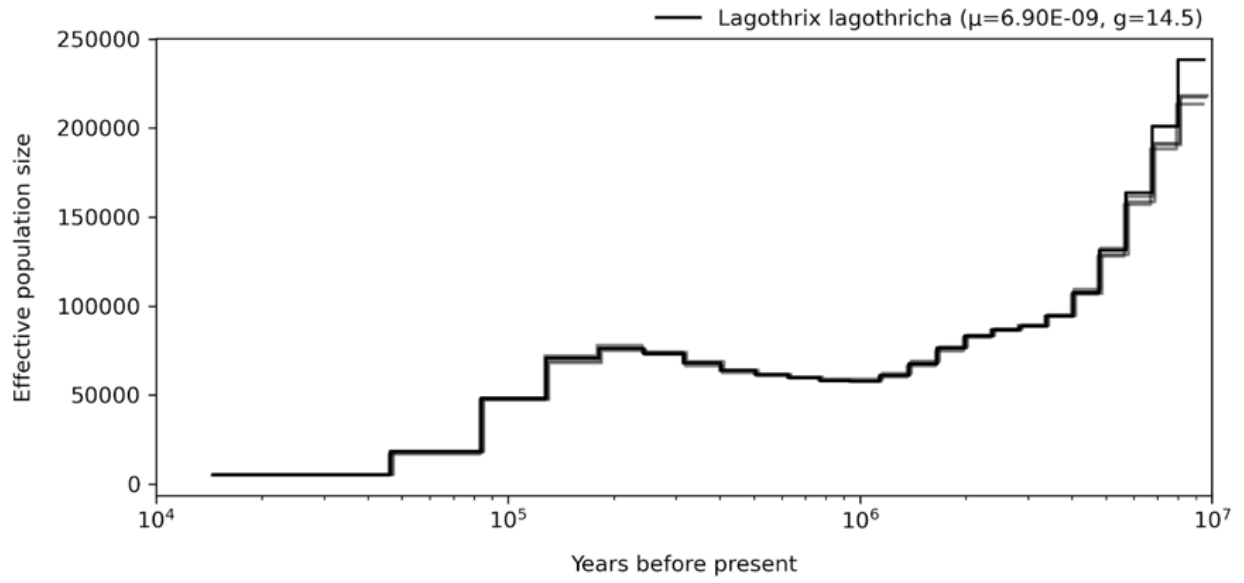


Fig. S76.
Demographic history of the genus *Lagothrix*, mutation rates for scaling are averaged per genus.

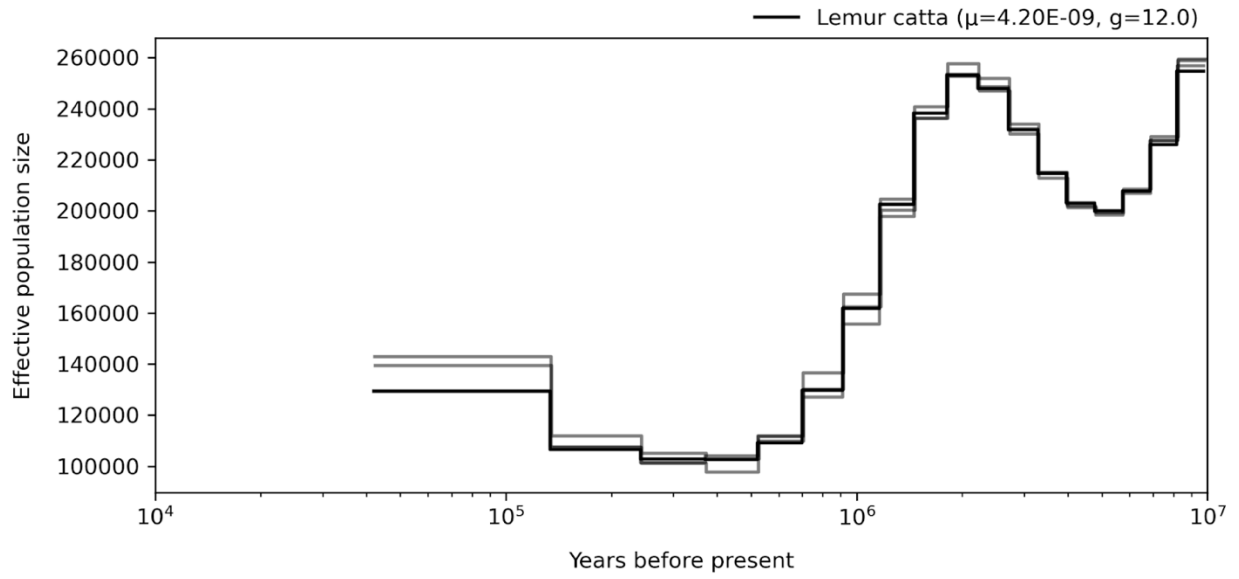


Fig. S77. Demographic history of the genus *Lemur*, mutation rates for scaling are averaged per genus.

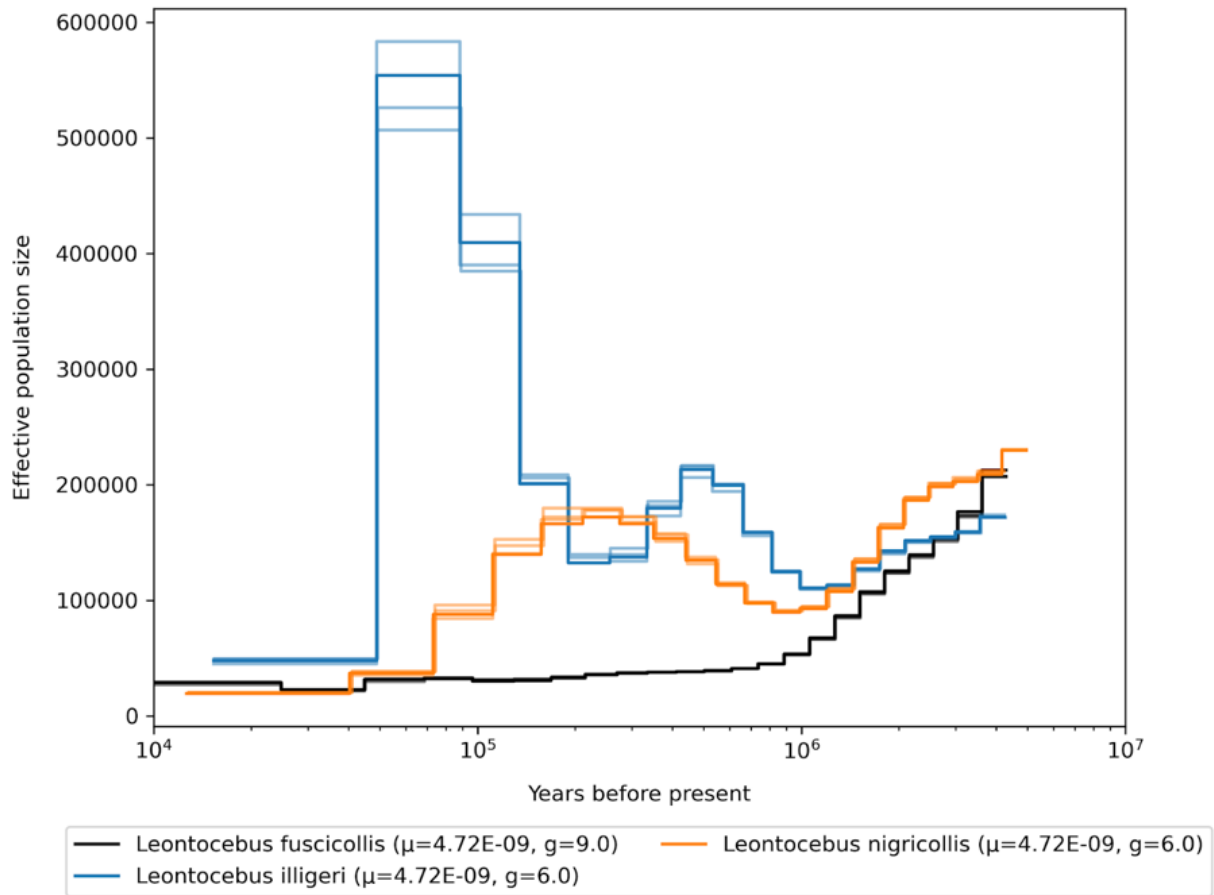


Fig. S78.
 Demographic history of the genus *Leontocebus*, mutation rates for scaling are averaged per genus.

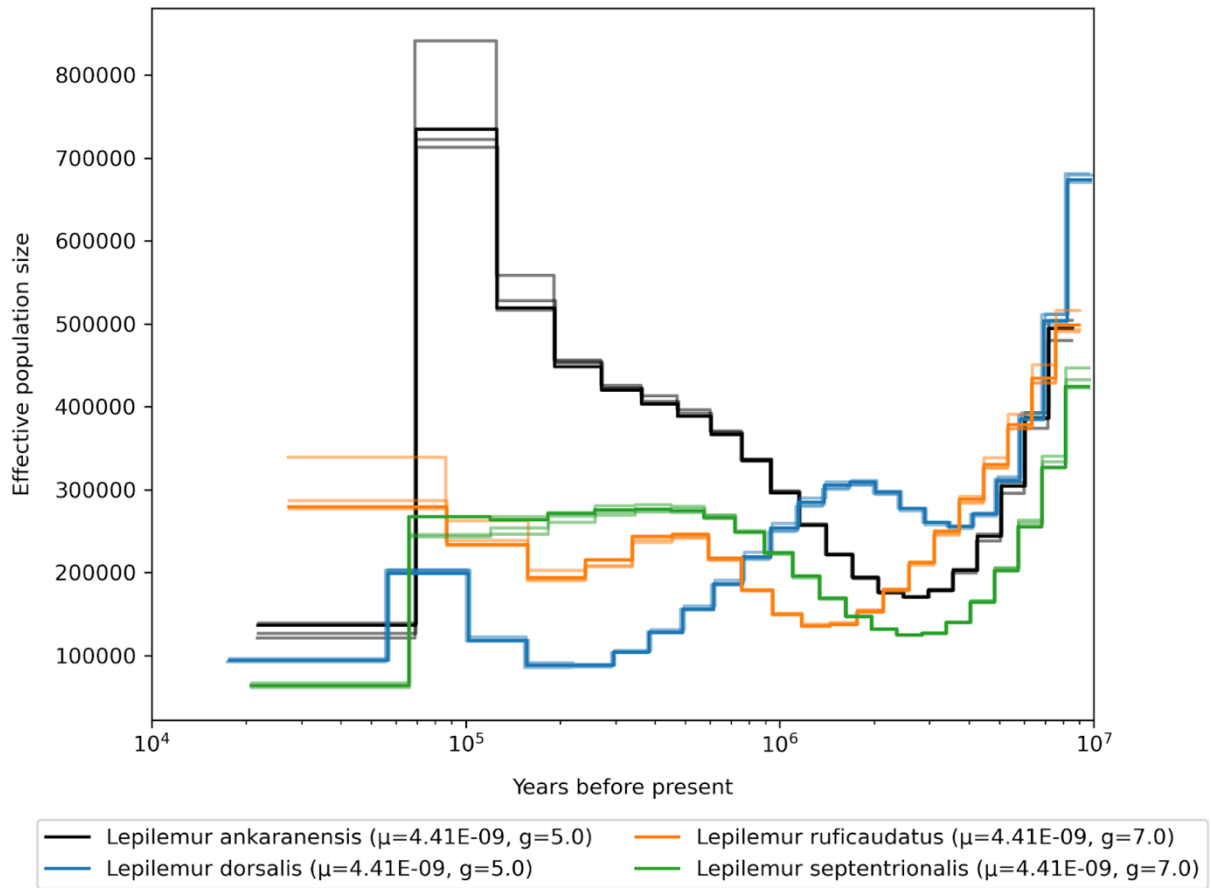


Fig. S79. Demographic history of the genus *Lepilemur*, mutation rates for scaling are averaged per genus.

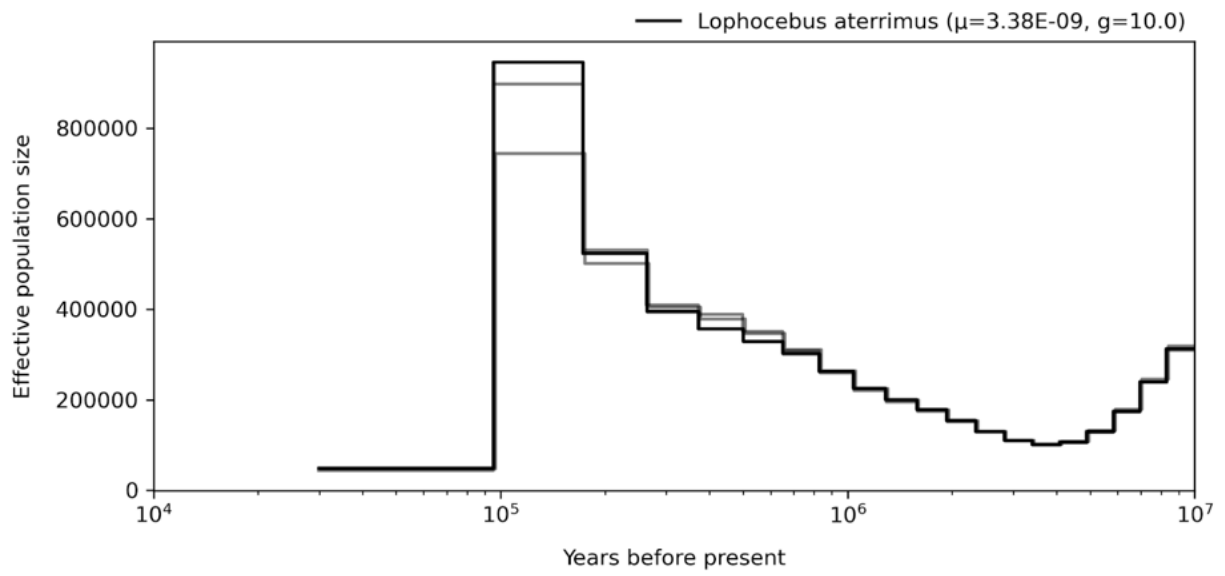


Fig. S80. Demographic history of the genus *Lophocebus*, mutation rates for scaling are averaged per genus.

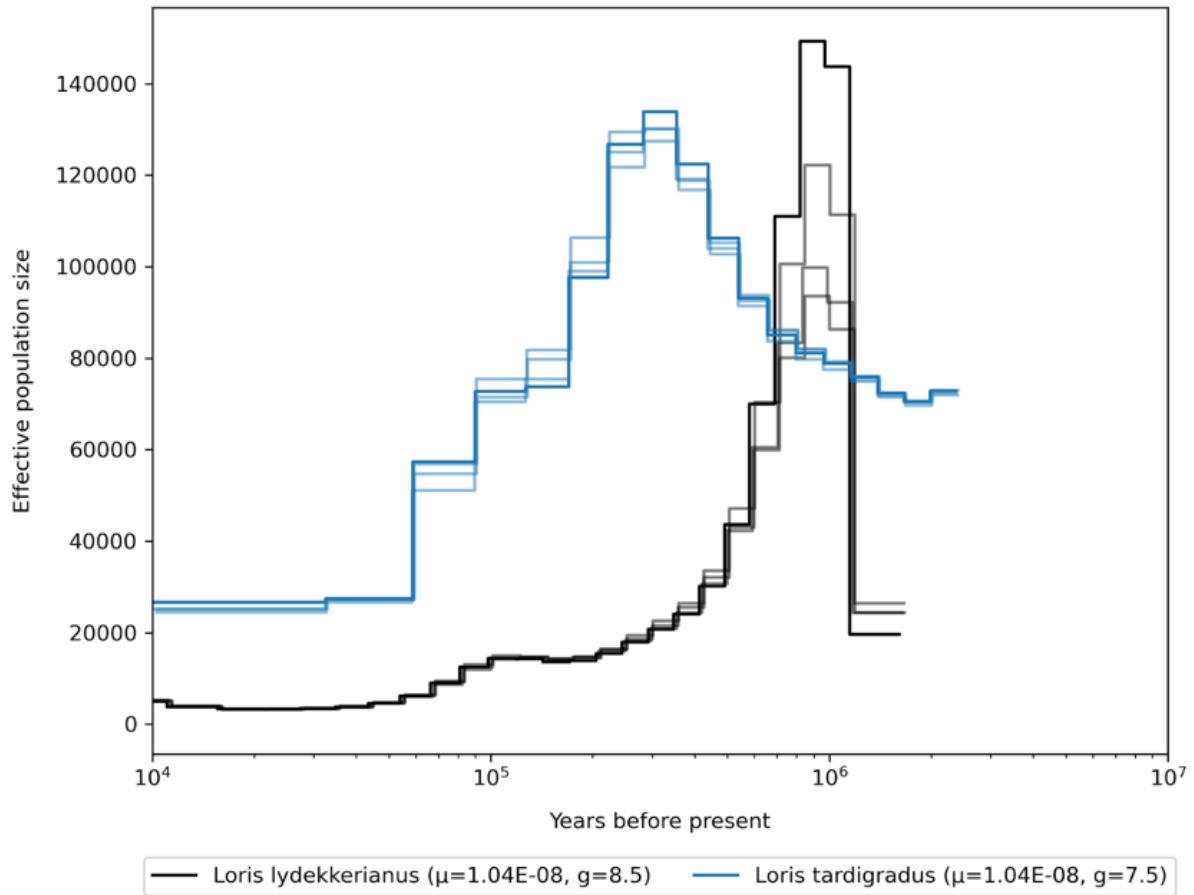


Fig. S81. Demographic history of the genus *Loris*, mutation rates for scaling are averaged per genus.

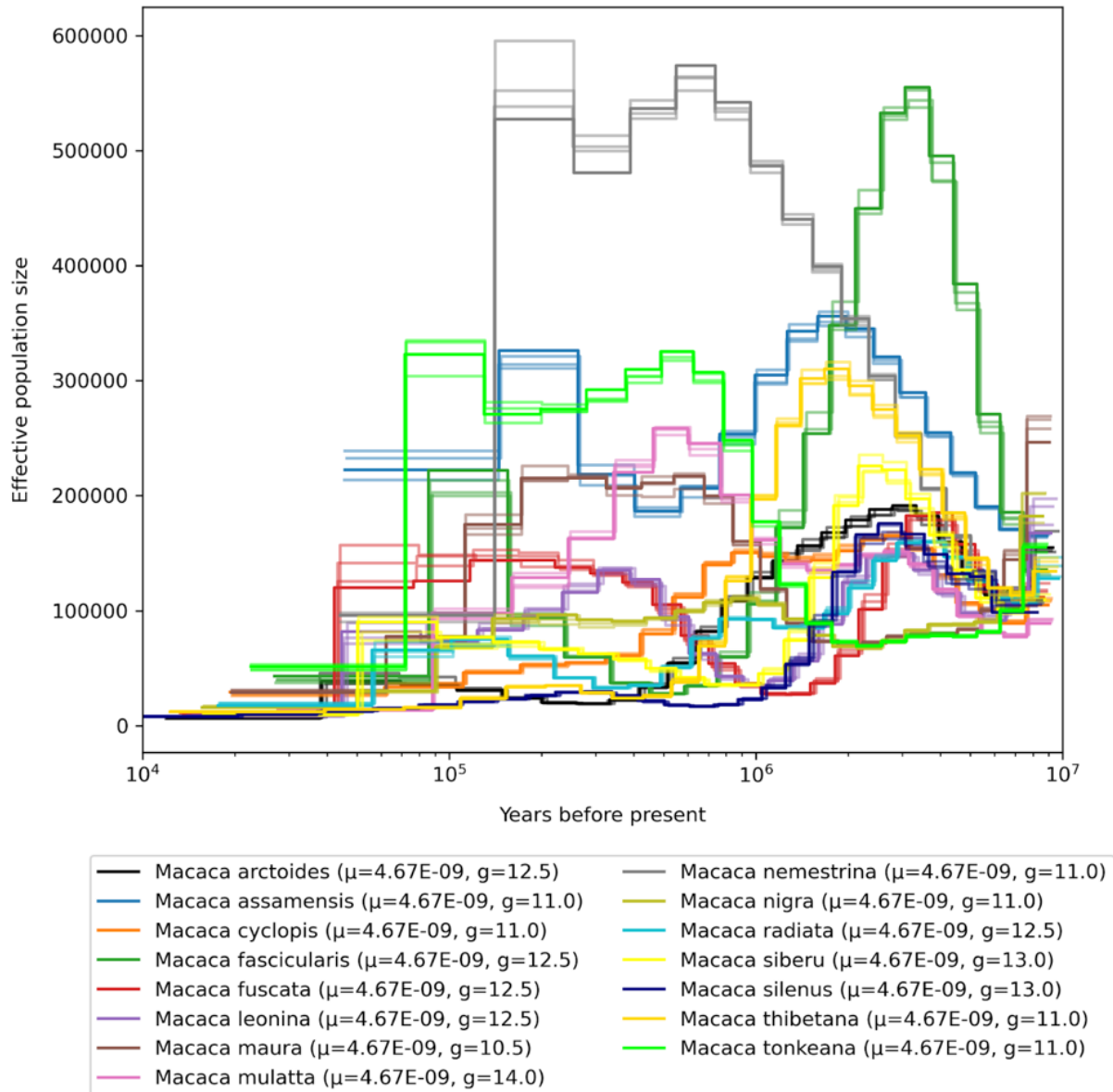


Fig. S82.
Demographic history of the genus *Macaca*, mutation rates for scaling are averaged per genus.

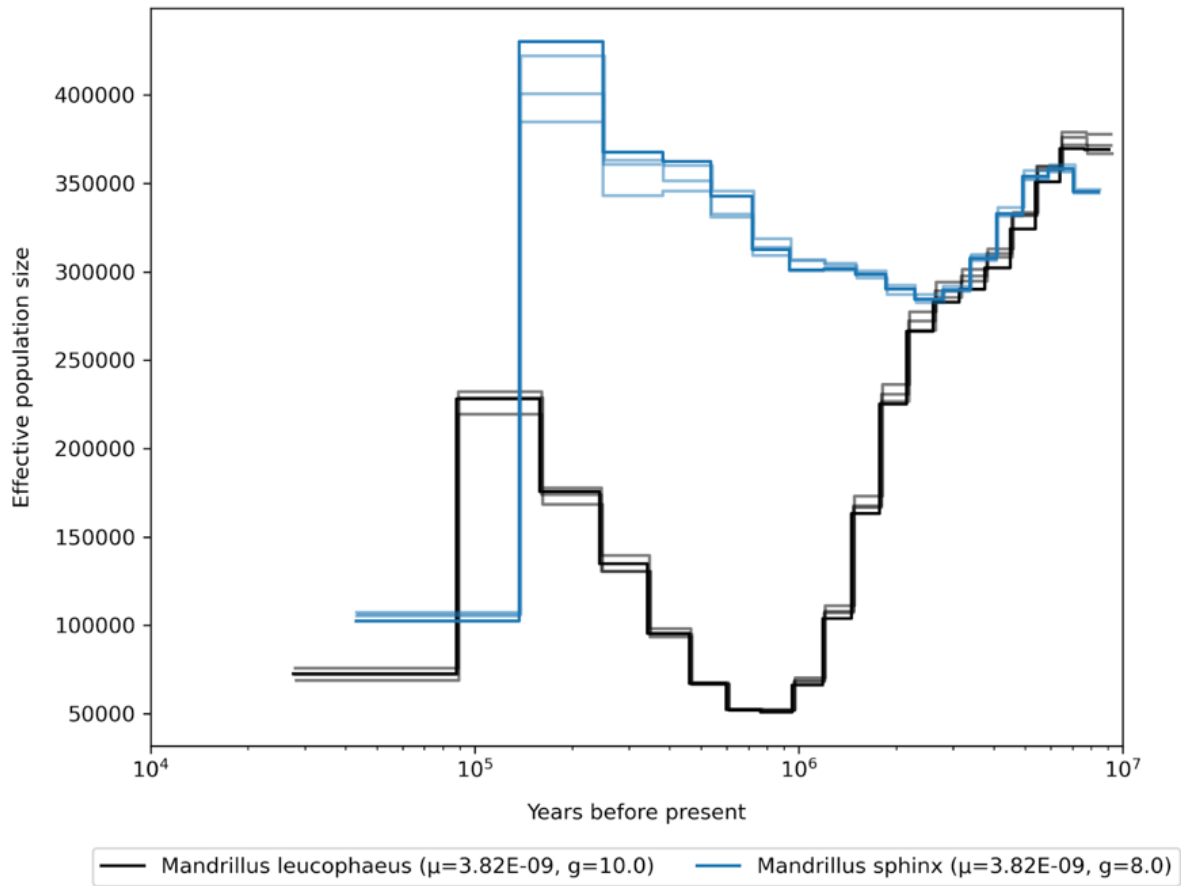


Fig. S83.
Demographic history of the genus *Mandrillus*, mutation rates for scaling are averaged per genus.

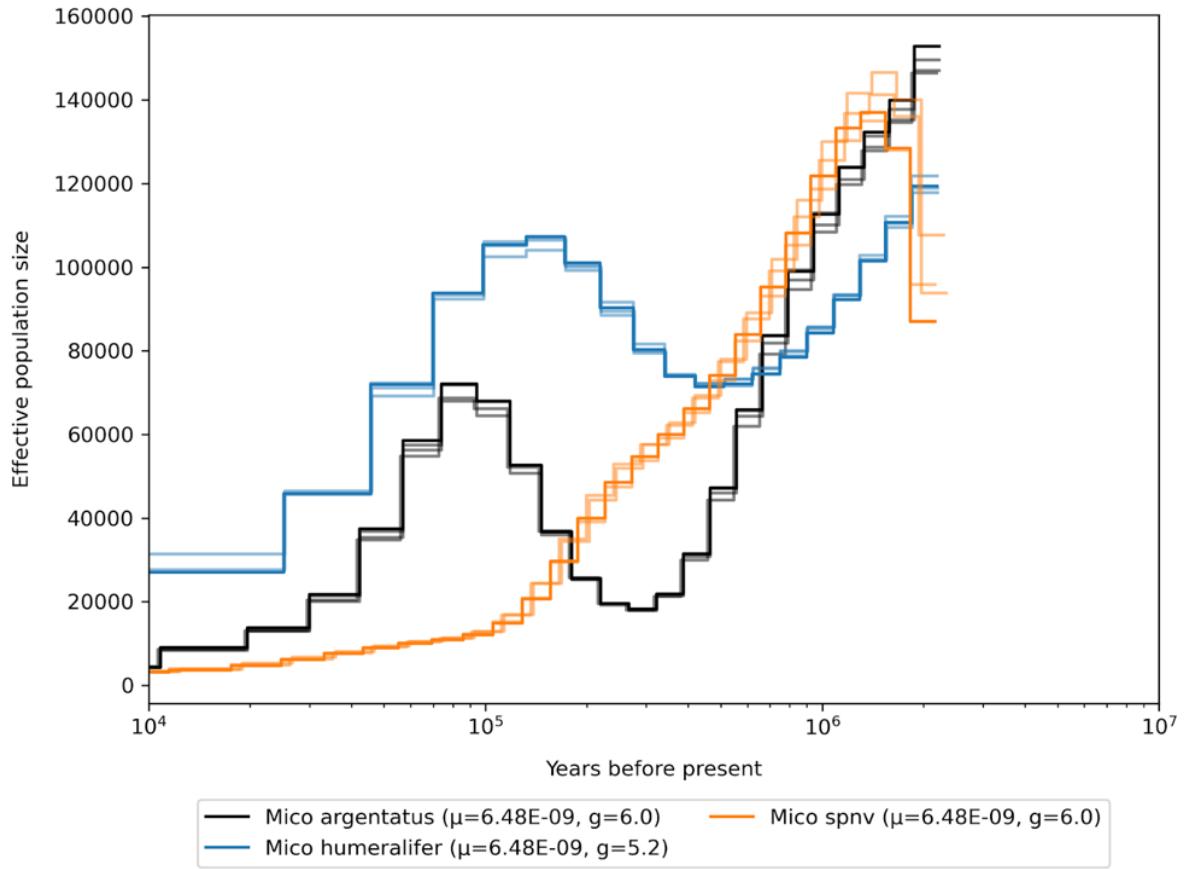


Fig. S84.
Demographic history of the genus *Mico*, mutation rates for scaling are averaged per genus.

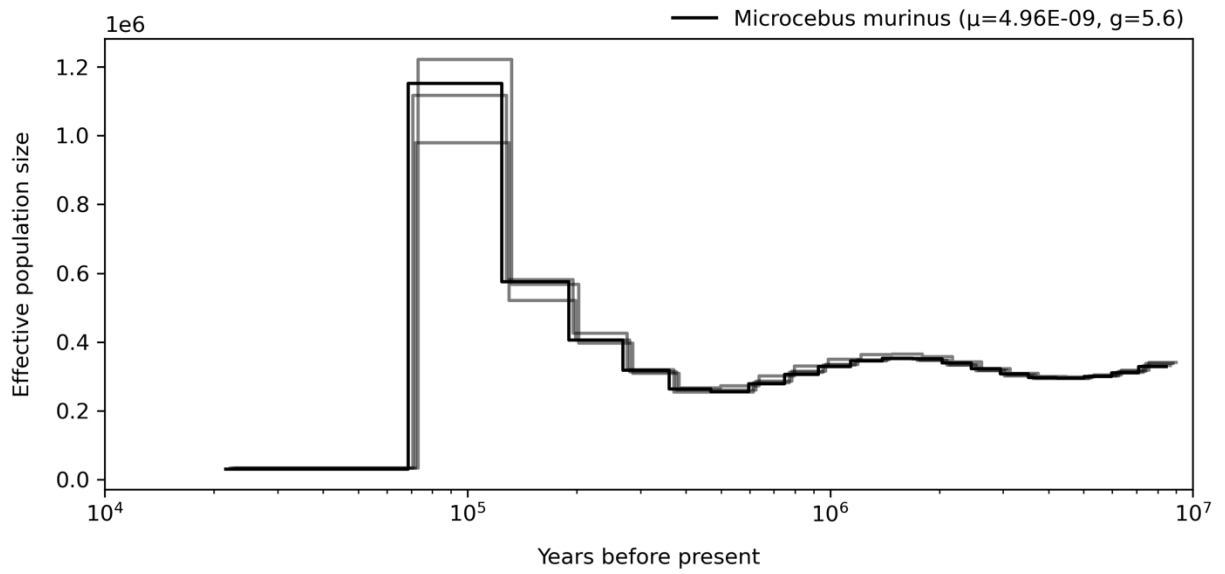


Fig. S85. Demographic history of the genus *Microcebus*, mutation rates for scaling are averaged per genus.

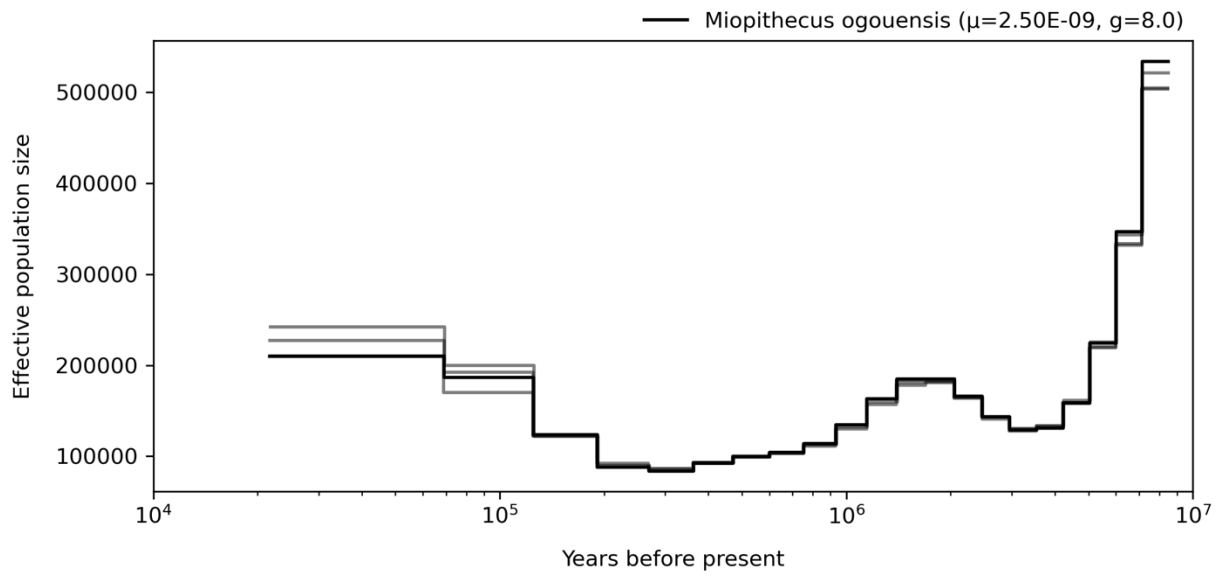


Fig. S86. Demographic history of the genus *Miopithecus*, mutation rates for scaling are averaged per genus.



Fig. S87.
Demographic history of the genus *Mirza*, mutation rates for scaling are averaged per genus.

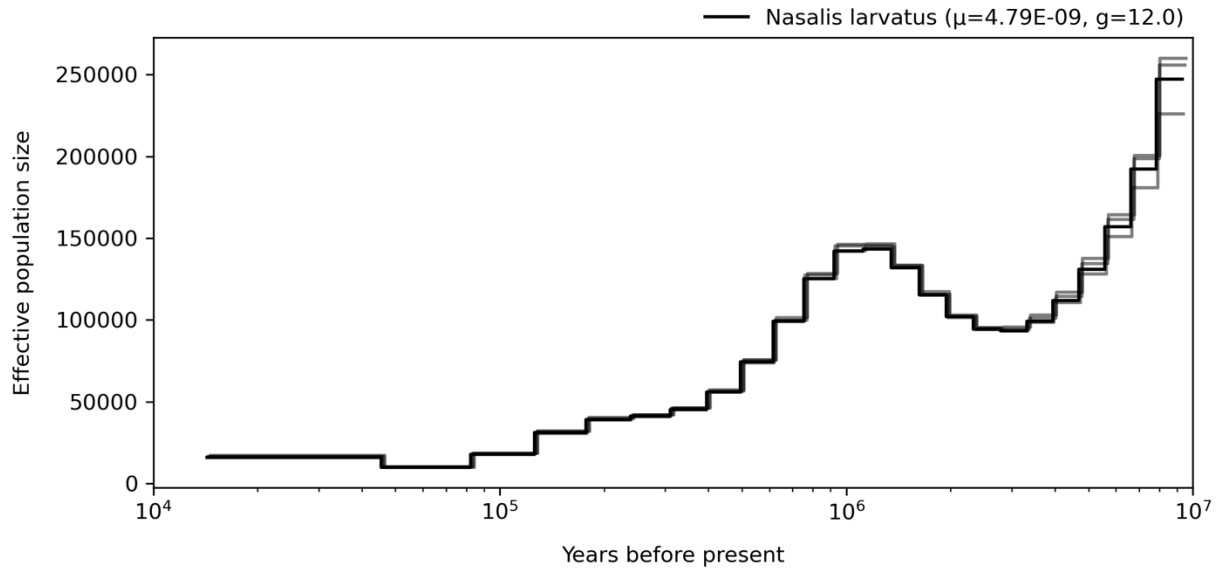


Fig. S88. Demographic history of the genus *Nasalis*, mutation rates for scaling are averaged per genus.

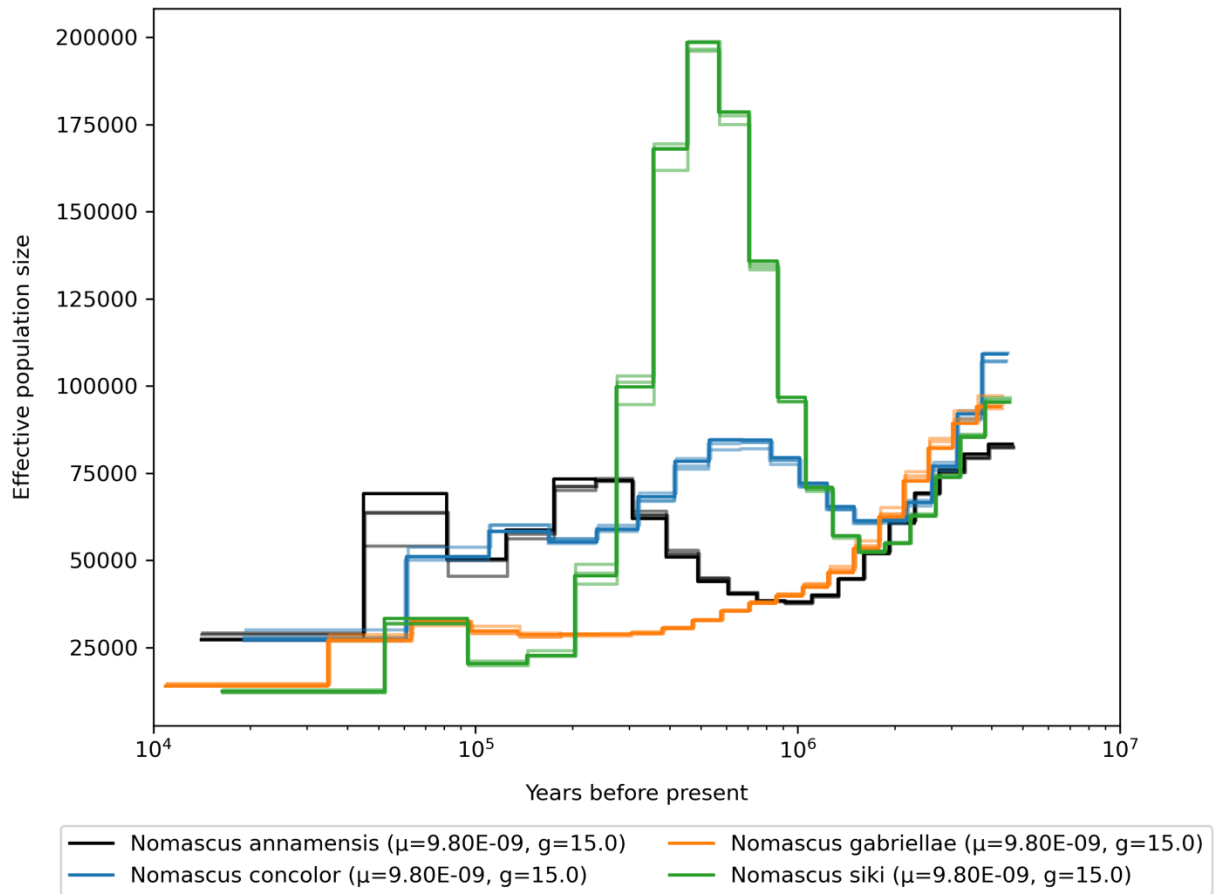


Fig. S89.
 Demographic history of the genus *Nomascus*, mutation rates for scaling are averaged per genus.

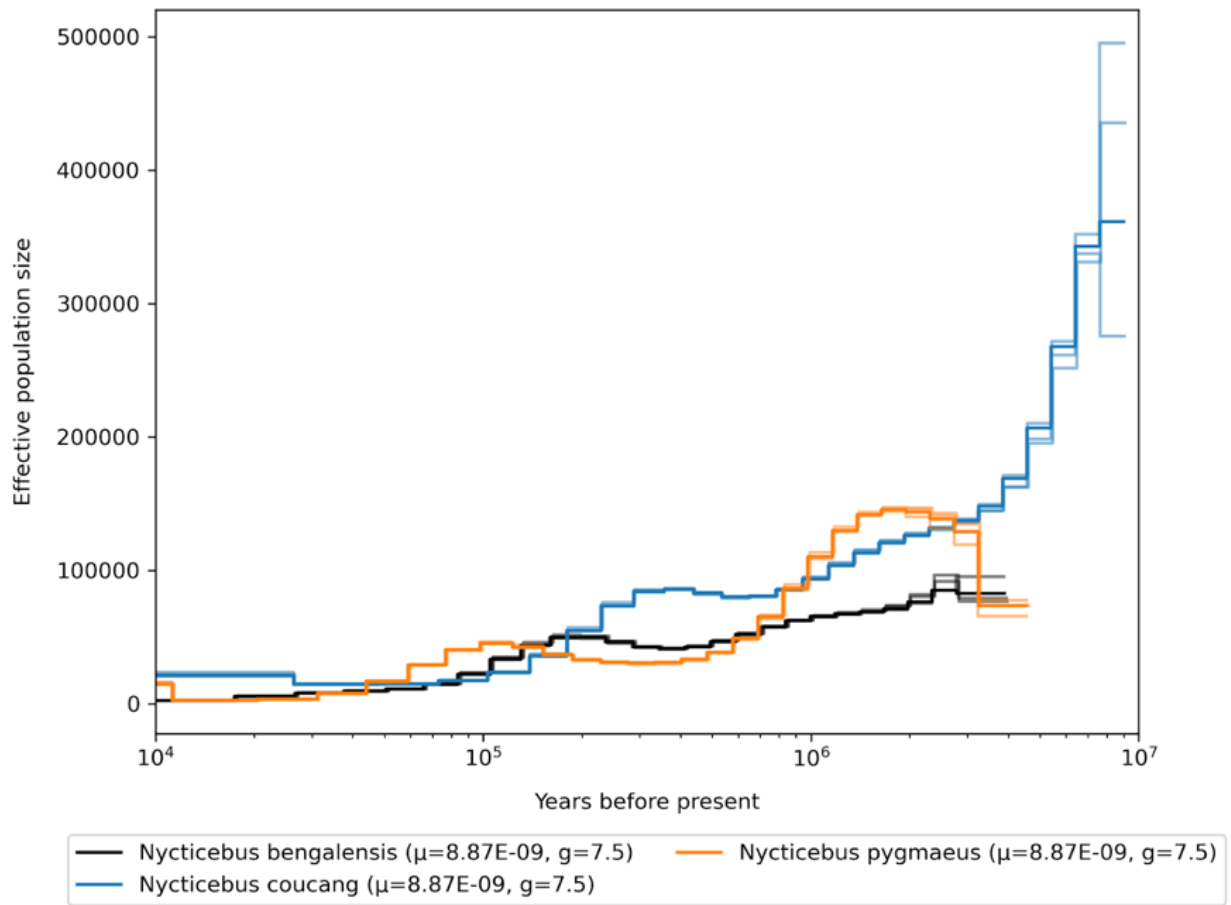


Fig. S90.
 Demographic history of the genus *Nycticebus*, mutation rates for scaling are averaged per genus.

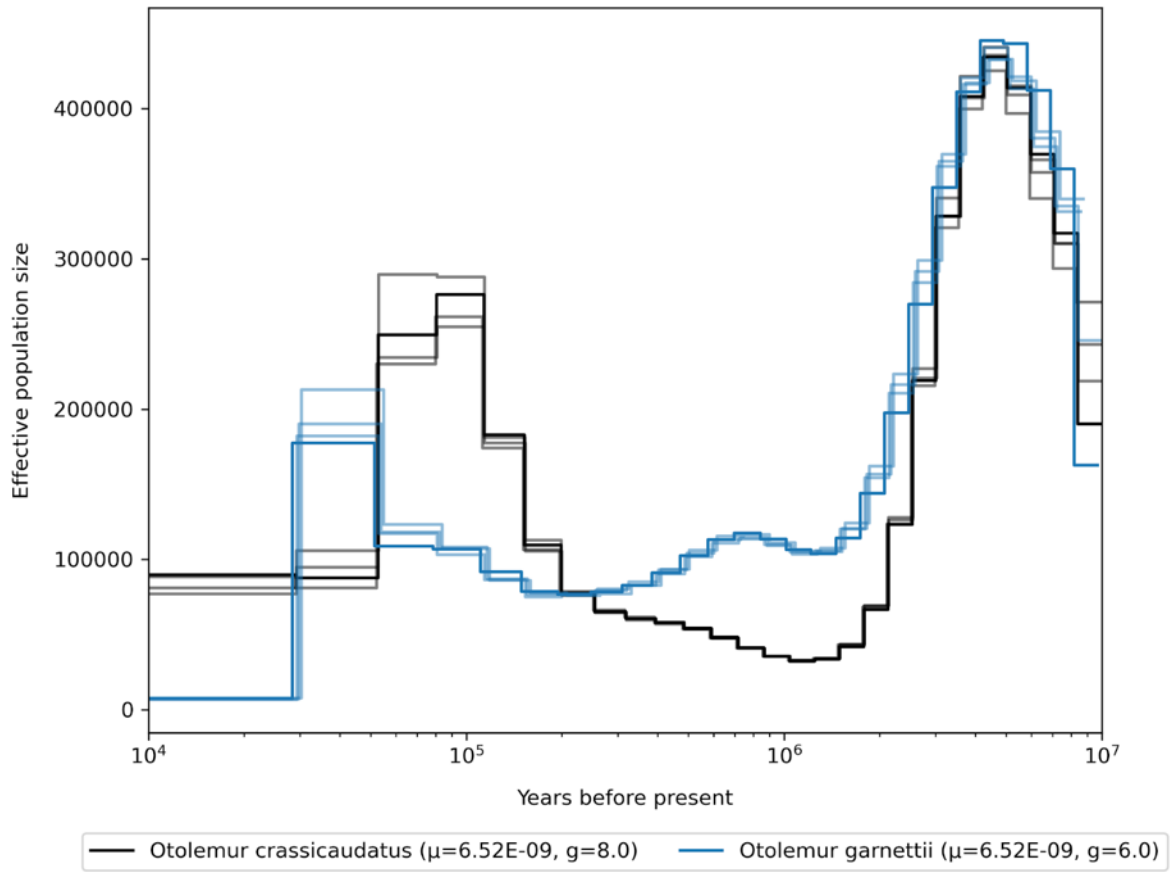


Fig. S91.
Demographic history of the genus *Otolemur*, mutation rates for scaling are averaged per genus.

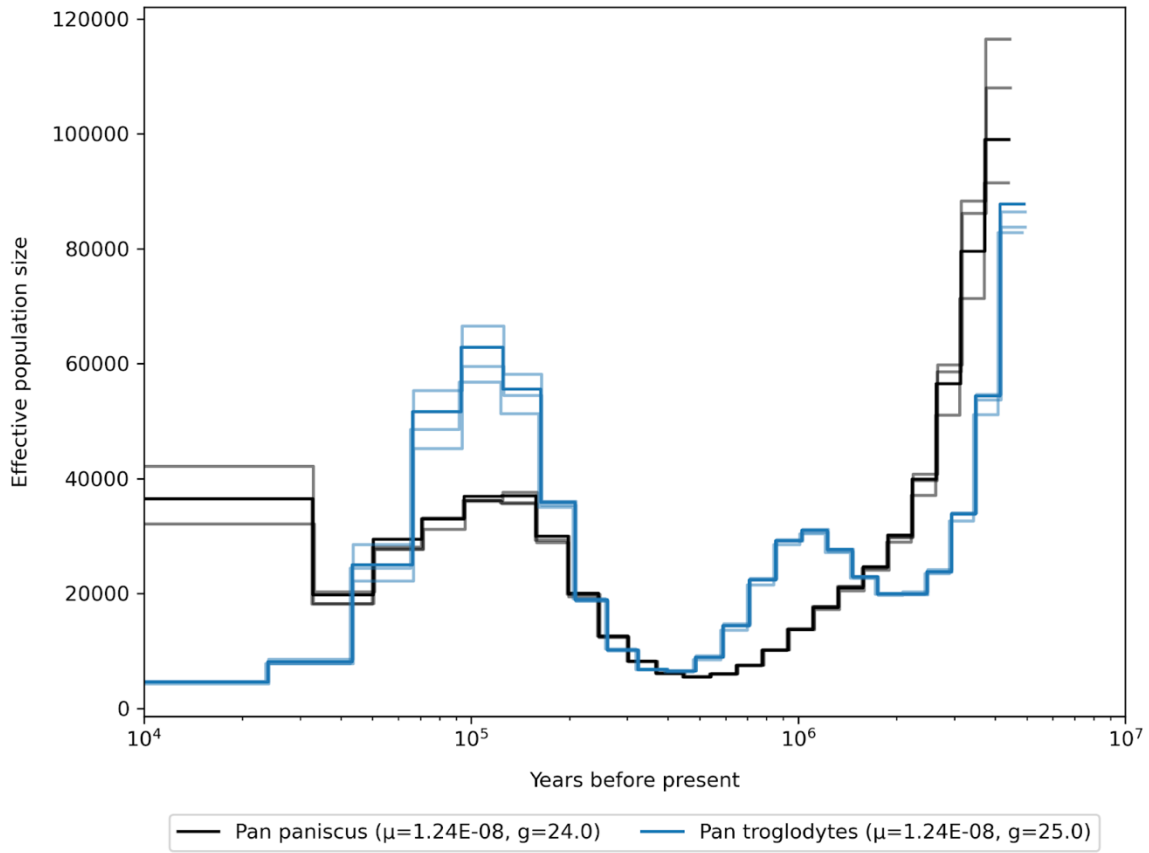


Fig. S92.
 Demographic history of the genus *Pan*, mutation rates for scaling are averaged per genus.

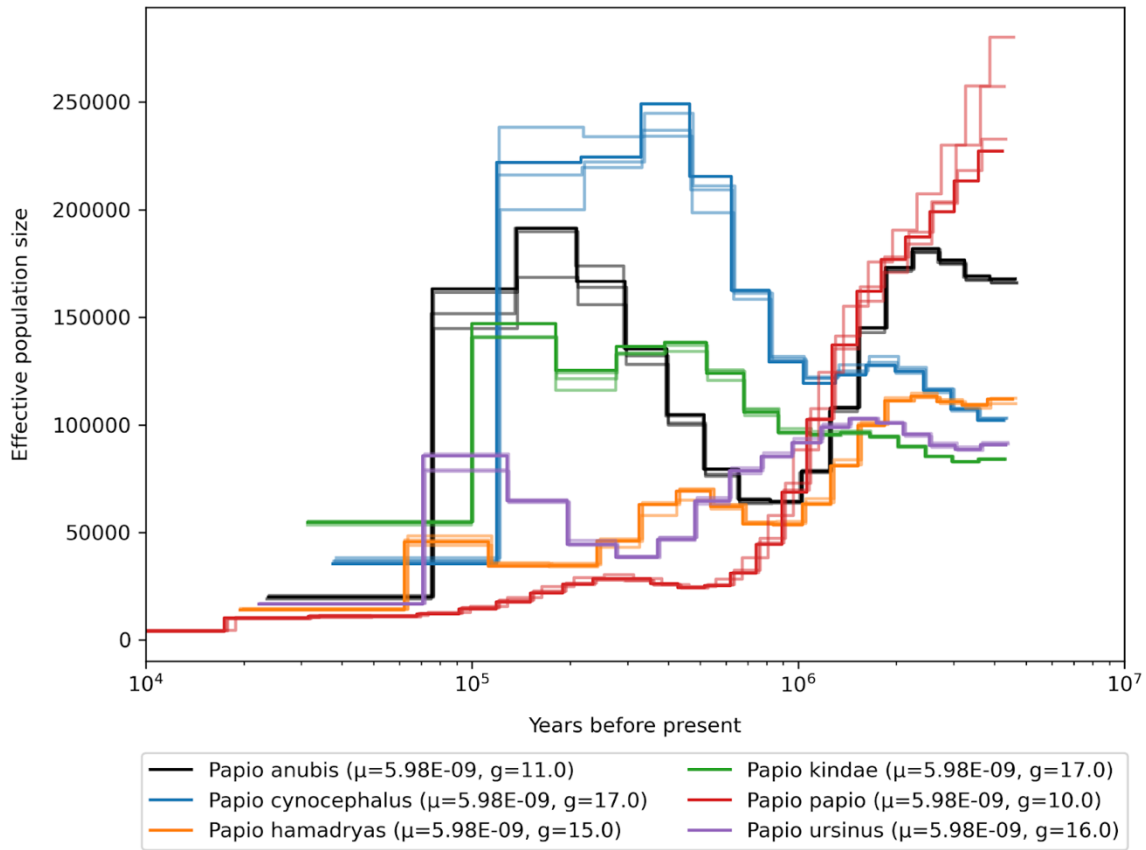


Fig. S93.
Demographic history of the genus *Papio*, mutation rates for scaling are averaged per genus.

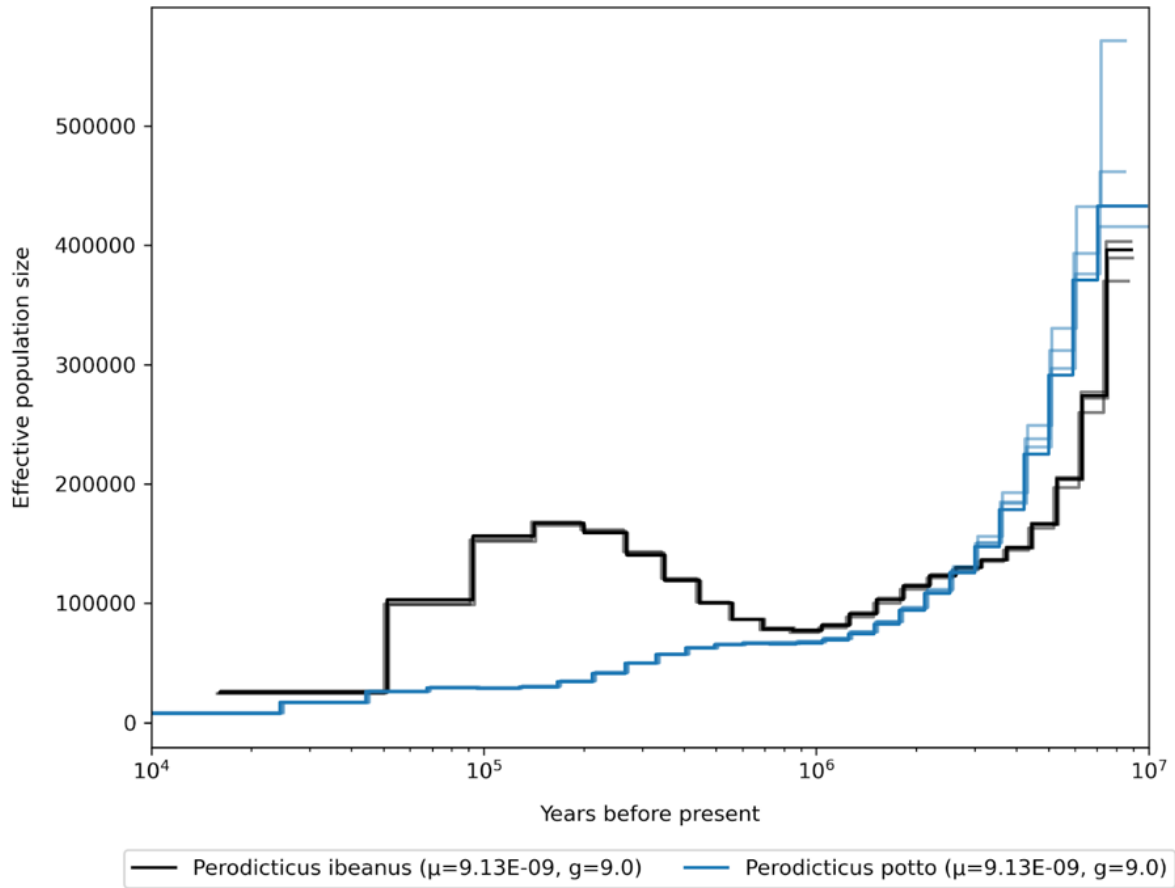


Fig. S94.
 Demographic history of the genus *Perodicticus*, mutation rates for scaling are averaged per genus.

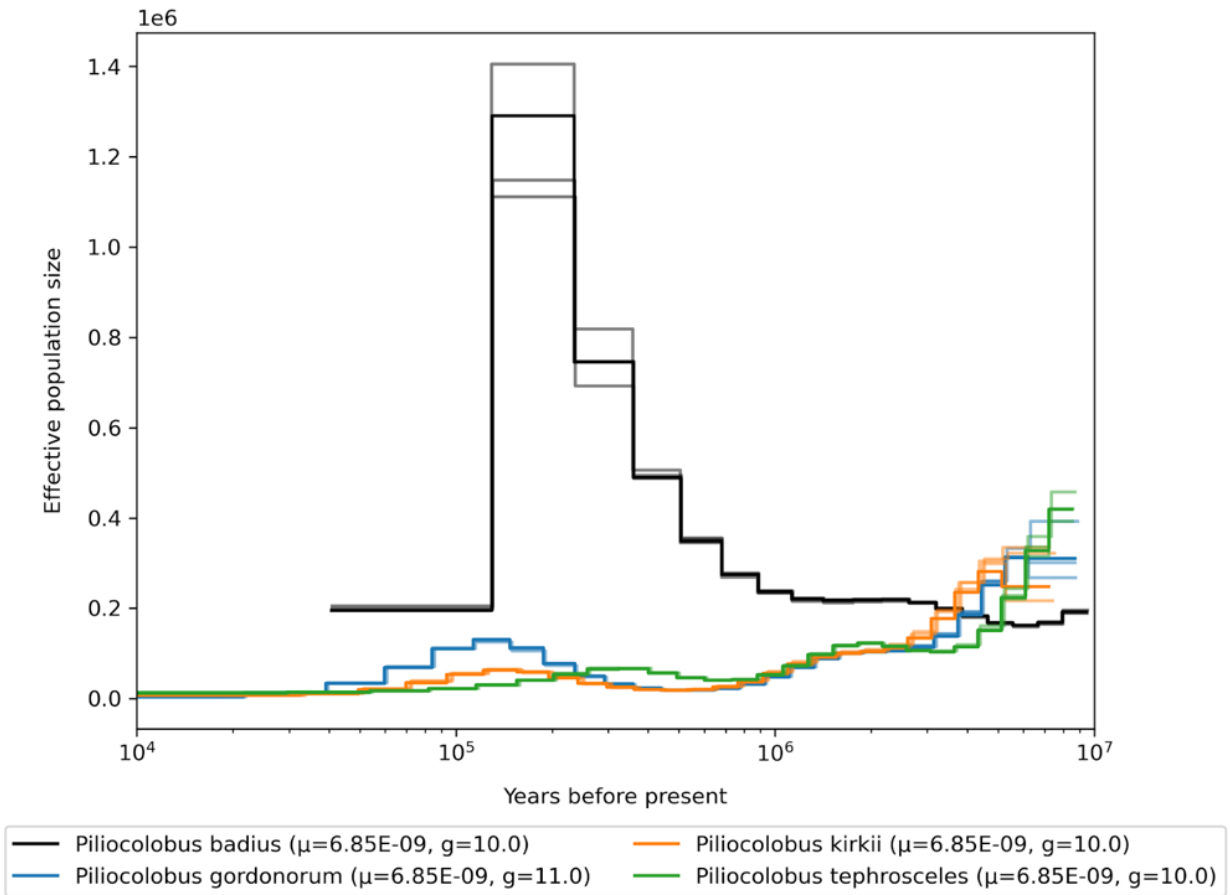


Fig. S95.
Demographic history of the genus *Piliocolobus*, mutation rates for scaling are averaged per genus.

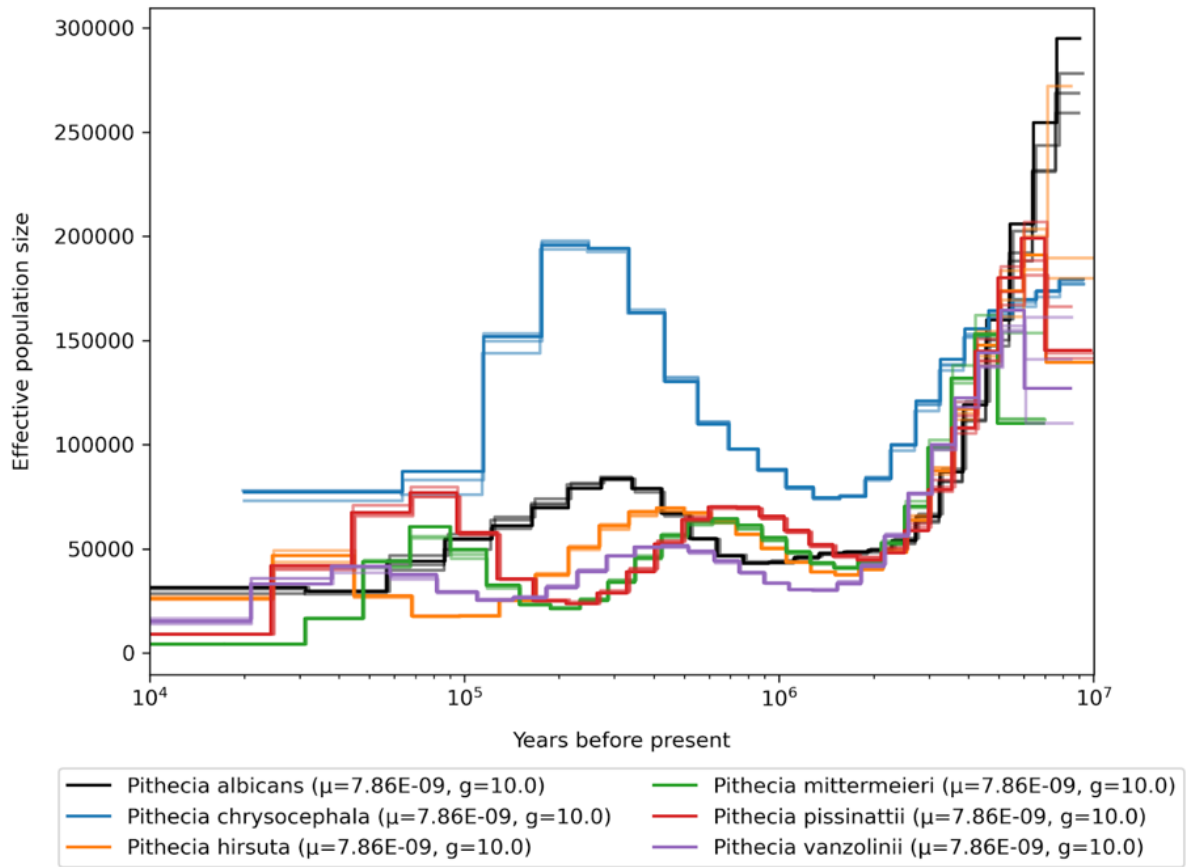


Fig. S96.
Demographic history of the genus *Pithecia*, mutation rates for scaling are averaged per genus.

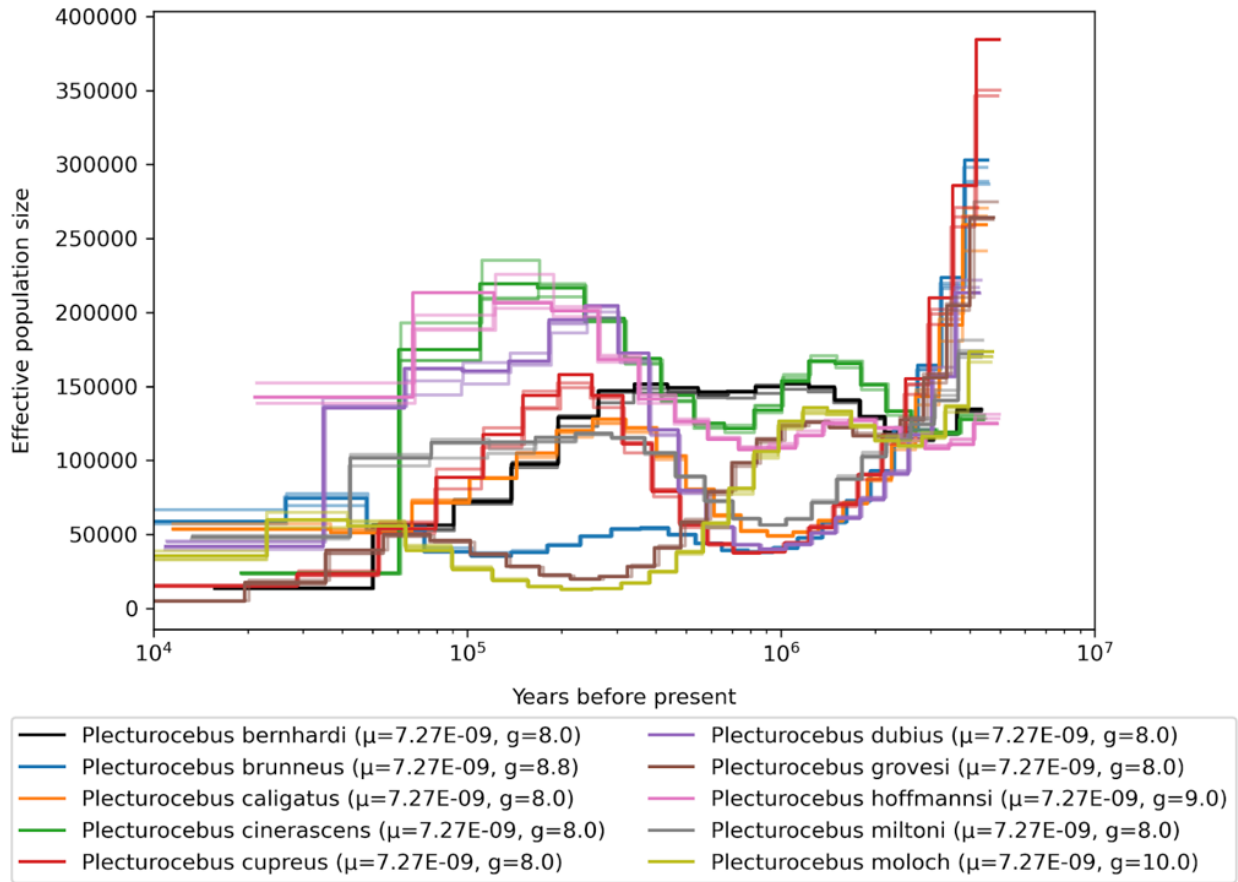


Fig. S97.

Demographic history of the genus *Plecturocebus*, mutation rates for scaling are averaged per genus.

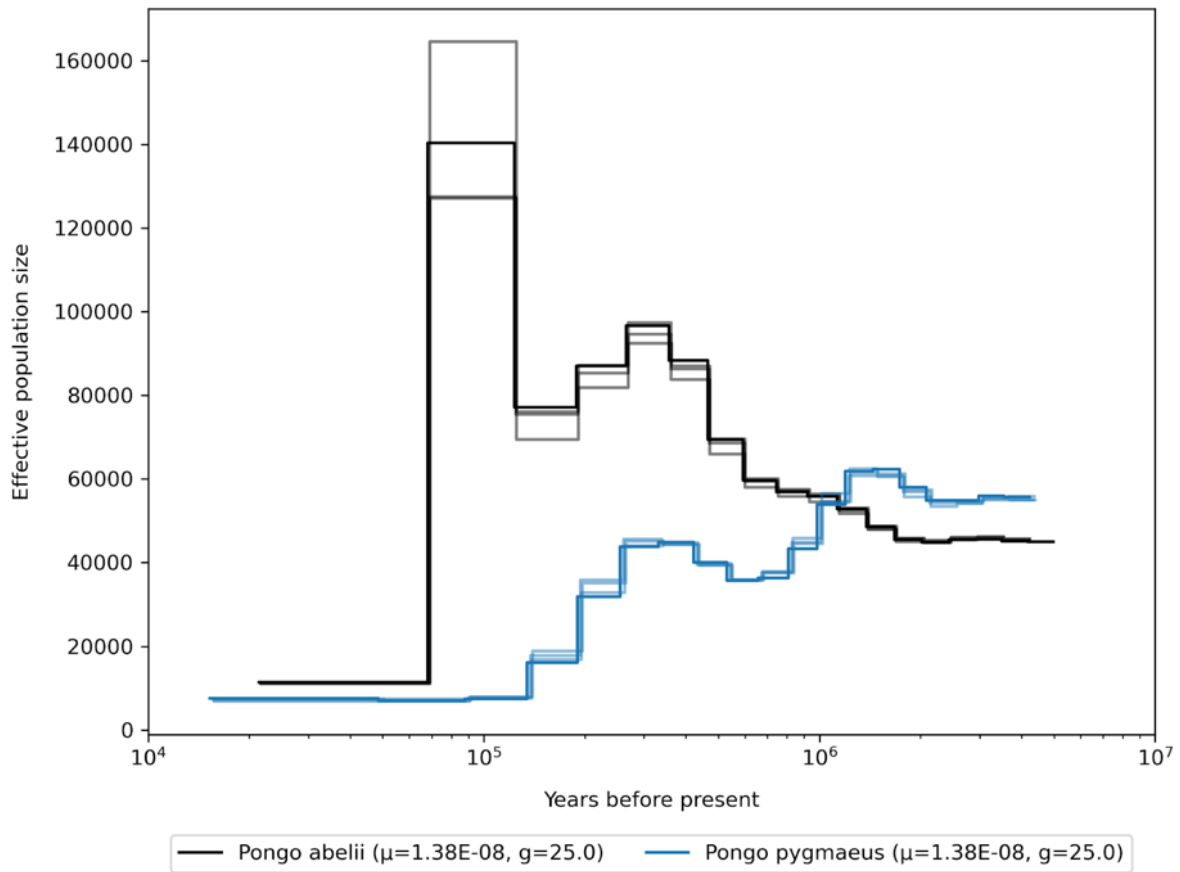


Fig. S98. Demographic history of the genus *Pongo*, mutation rates for scaling are averaged per genus.

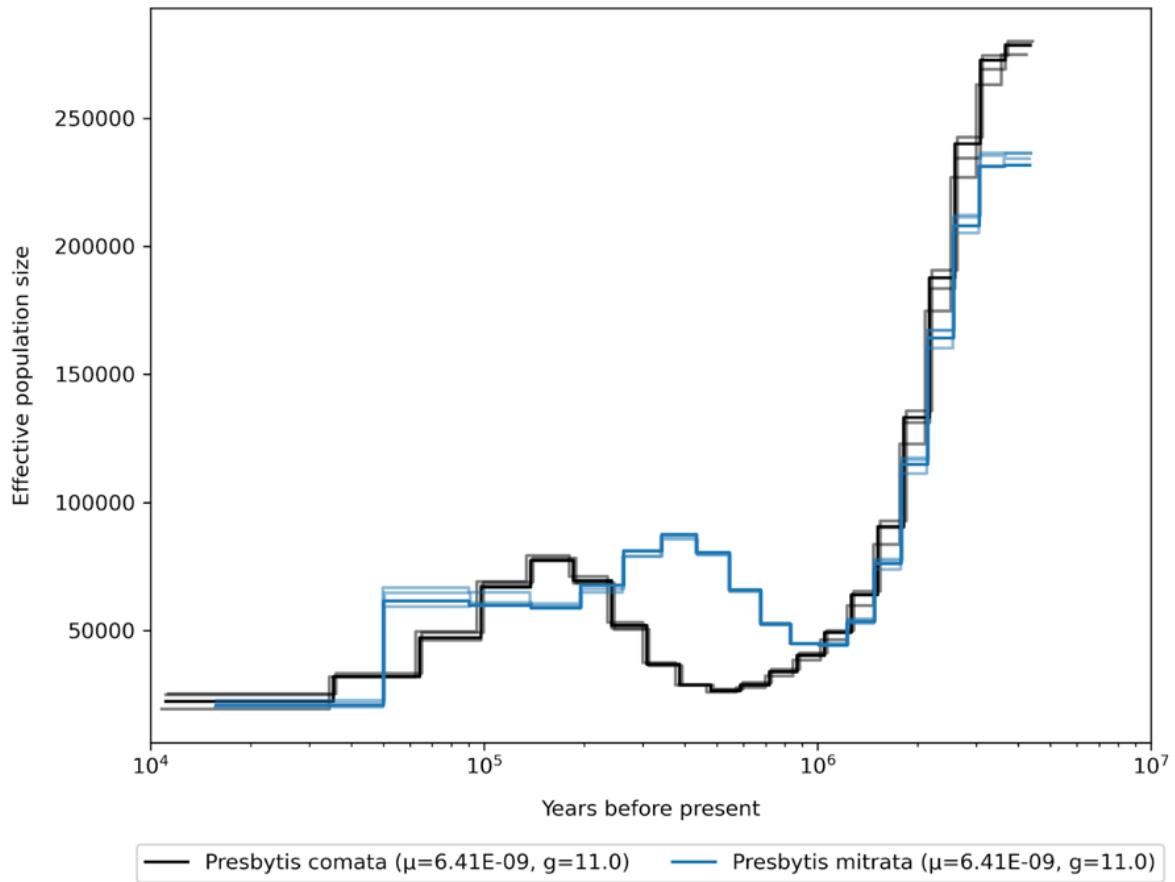


Fig. S99. Demographic history of the genus *Presbytis*, mutation rates for scaling are averaged per genus.

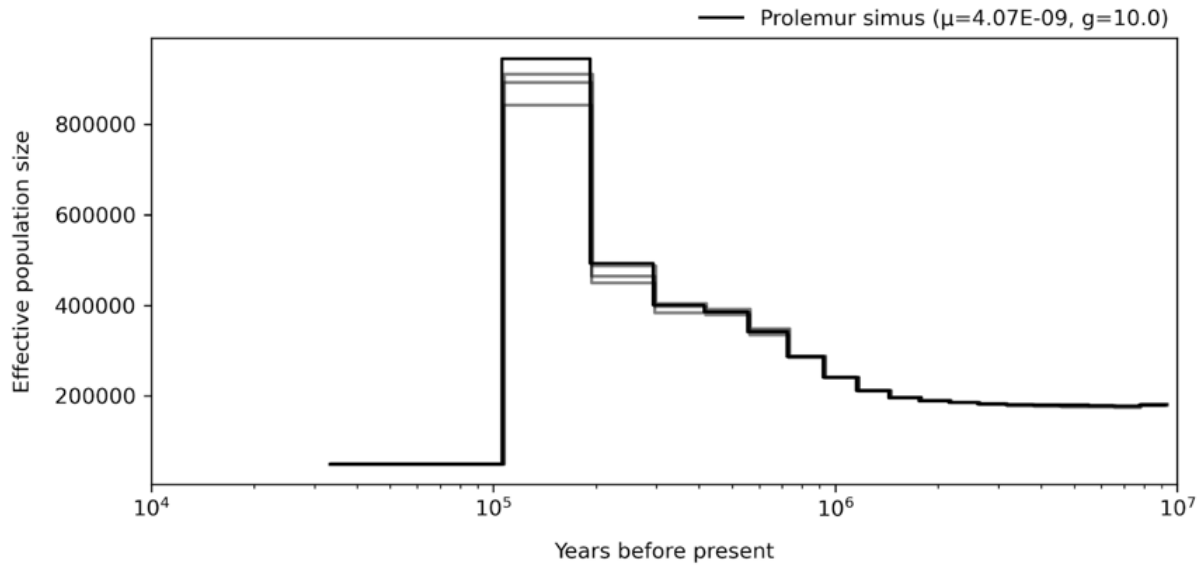


Fig. S100. Demographic history of the genus *Prolemur*, mutation rates for scaling are averaged per genus.

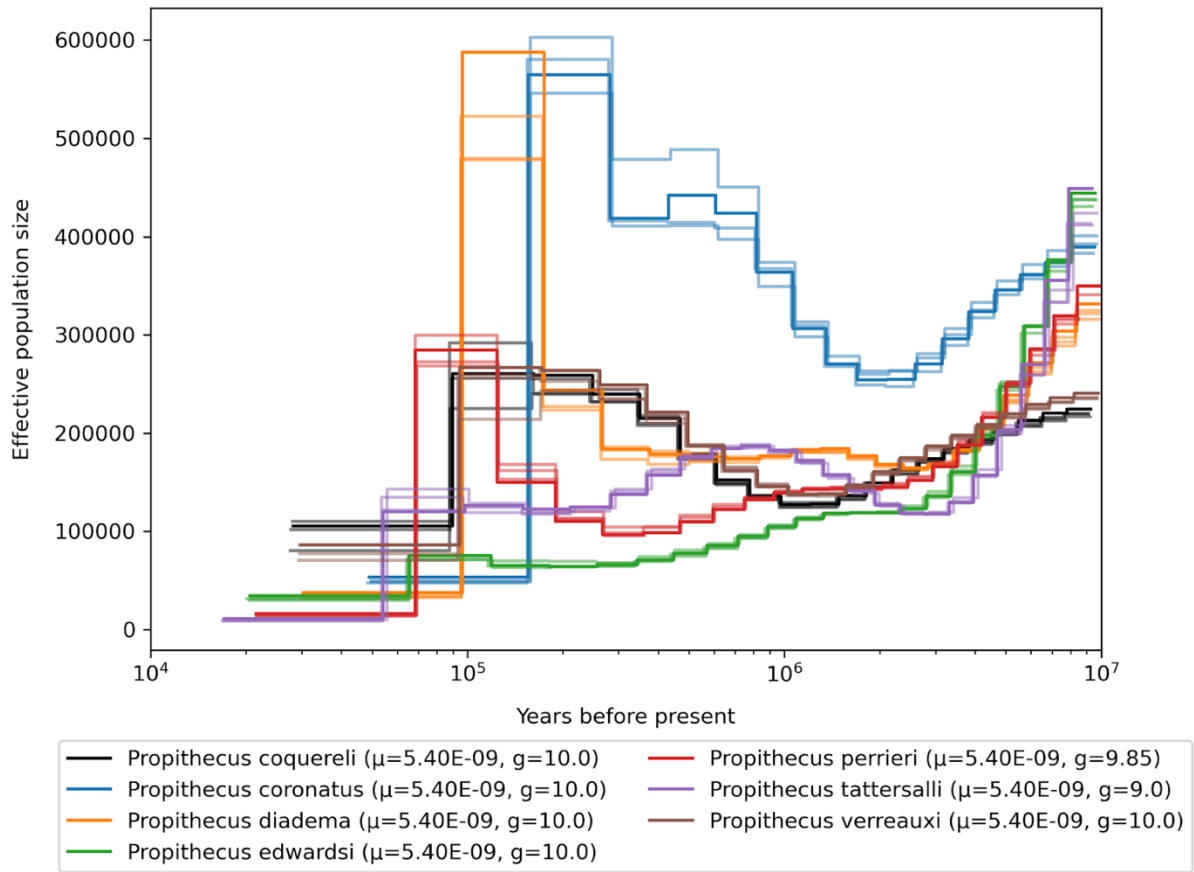


Fig. S101.
Demographic history of the genus *Propithecus*, mutation rates for scaling are averaged per genus.

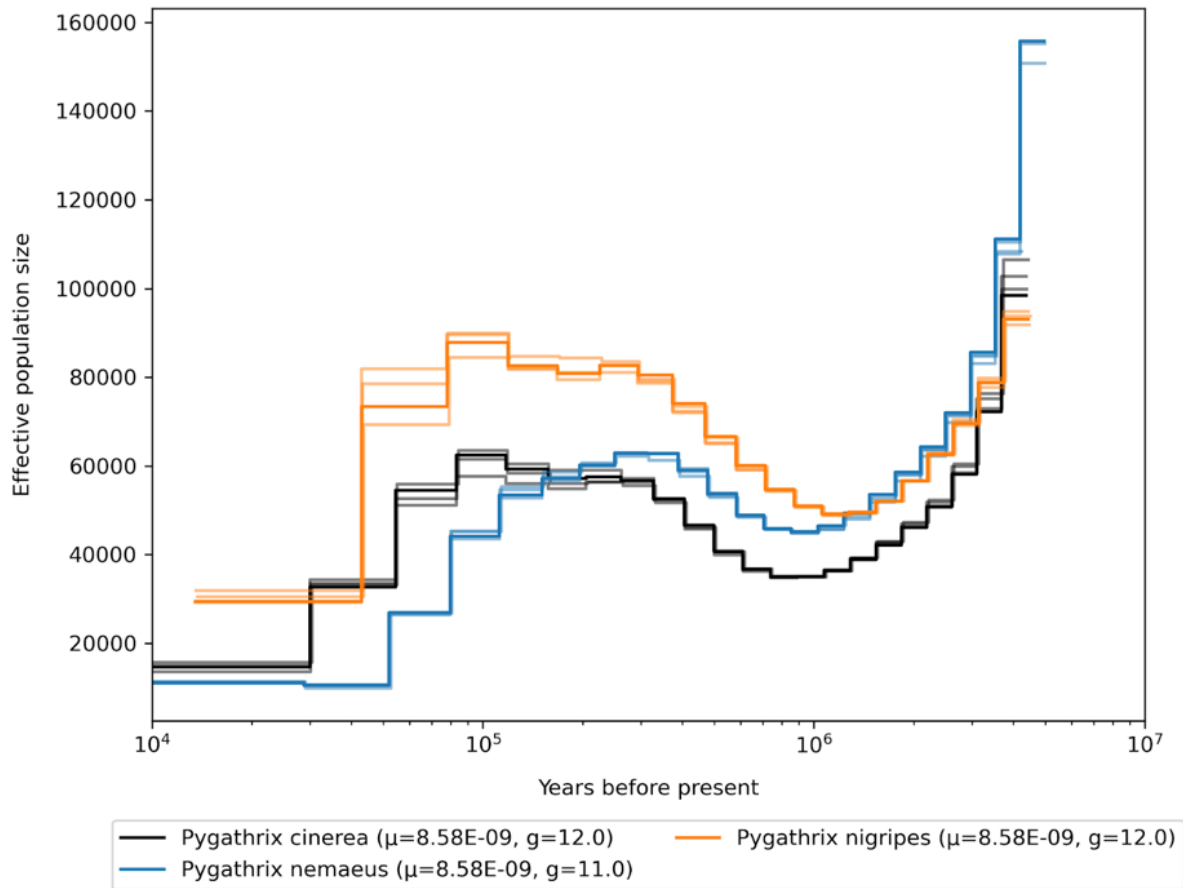


Fig. S102. Demographic history of the genus *Pygathrix*, mutation rates for scaling are averaged per genus.

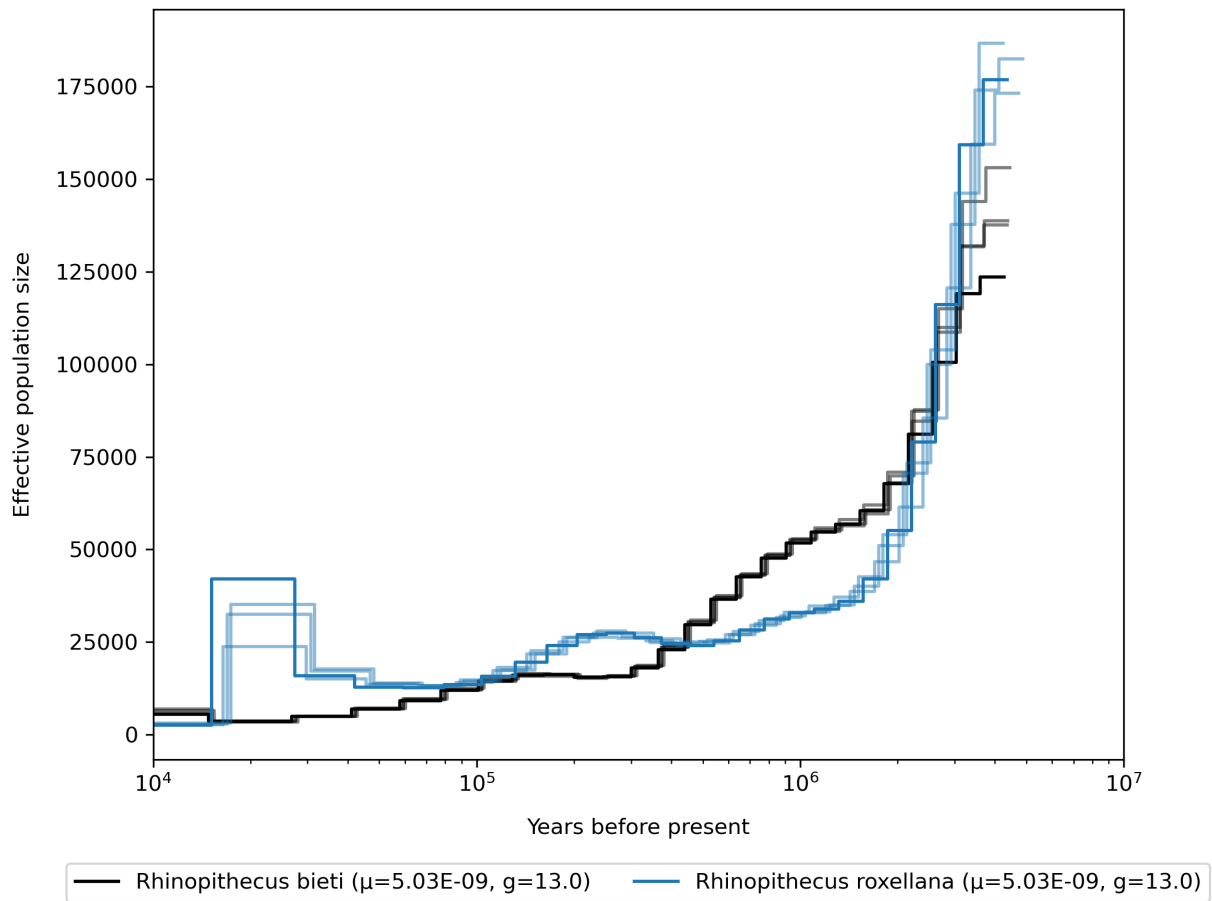


Fig. S103. Demographic history of the genus *Rhinopithecus*, mutation rates for scaling are averaged per genus.

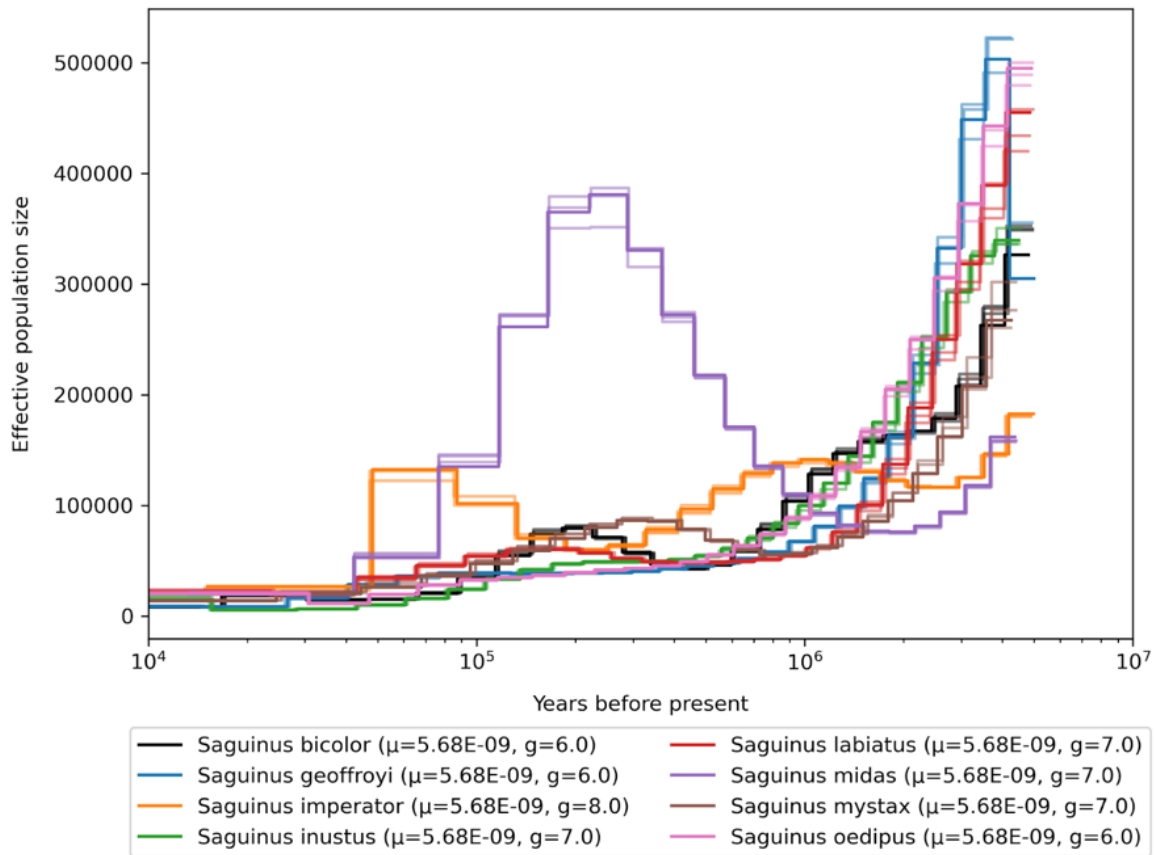


Fig. S104.
Demographic history of the genus *Saguinus*, mutation rates for scaling are averaged per genus.

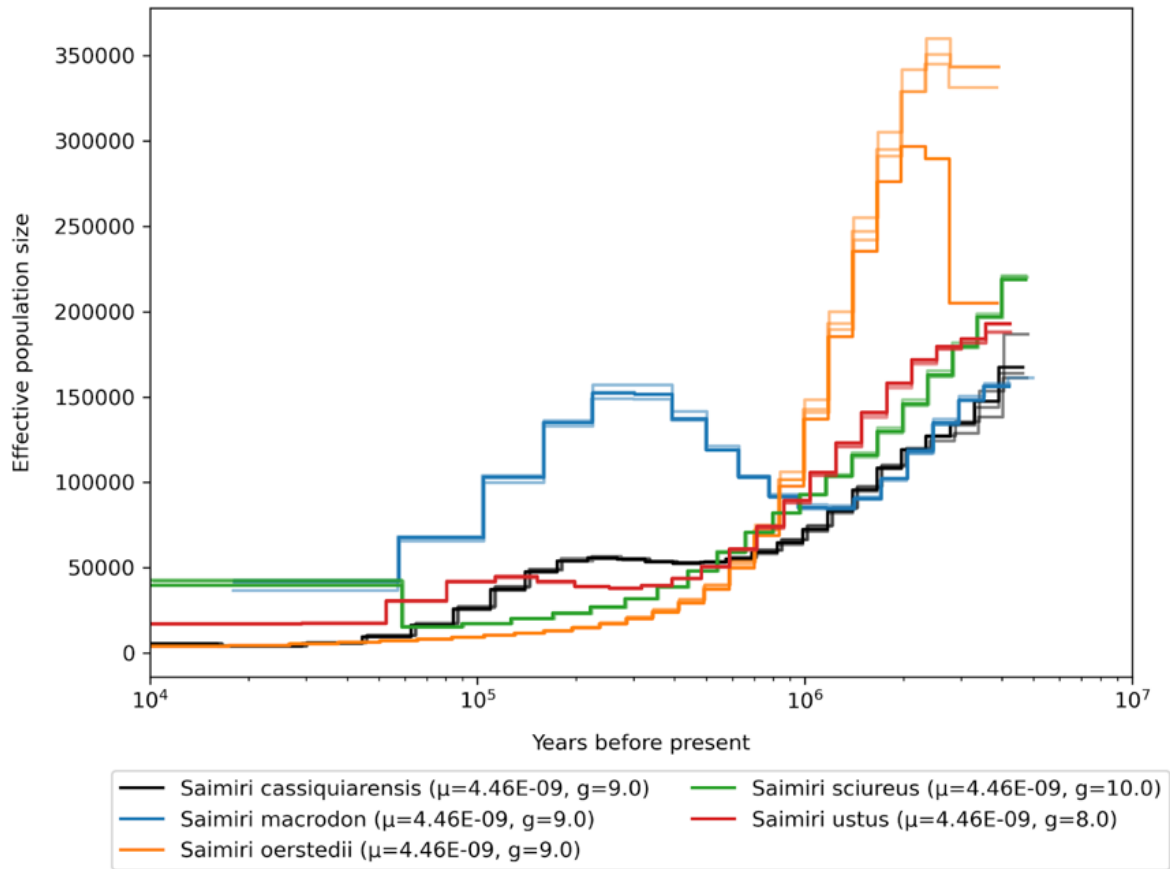


Fig. S105.

Demographic history of the genus *Saimiri*, mutation rates for scaling are averaged per genus.

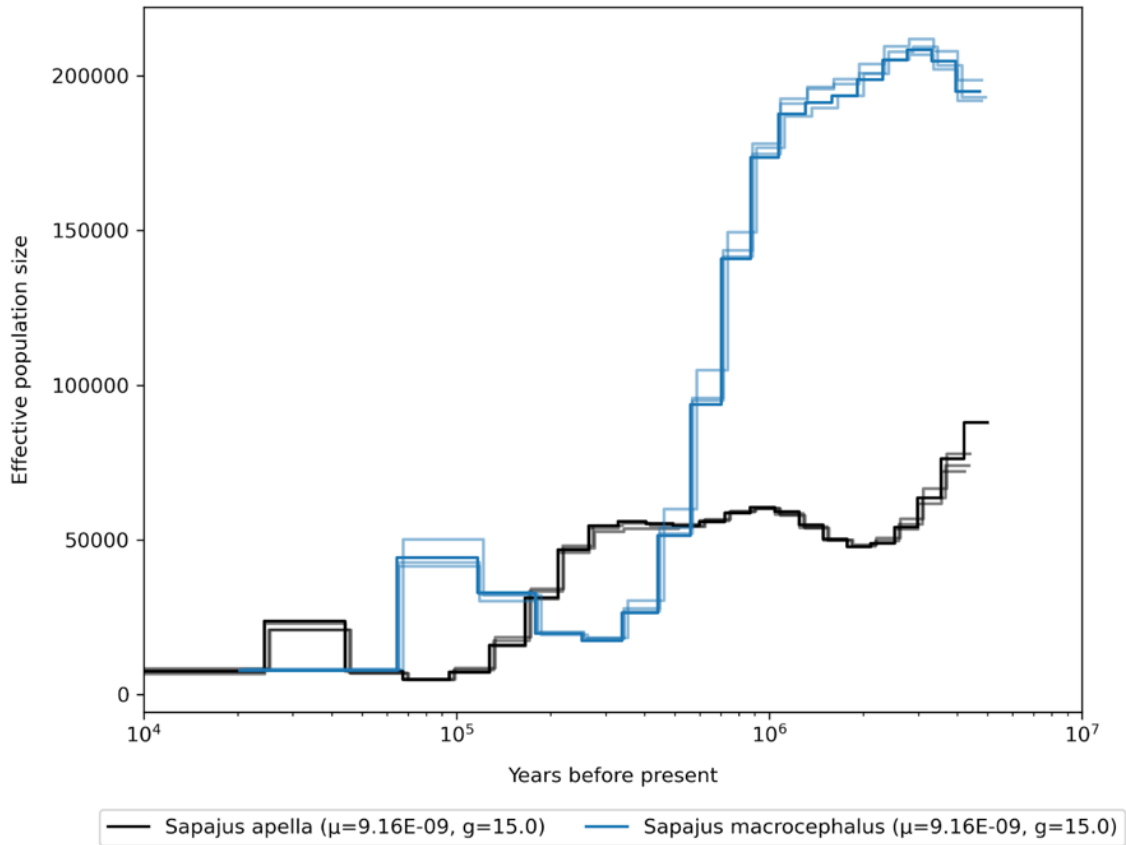


Fig. S106. Demographic history of the genus *Sapajus*, mutation rates for scaling are averaged per genus.

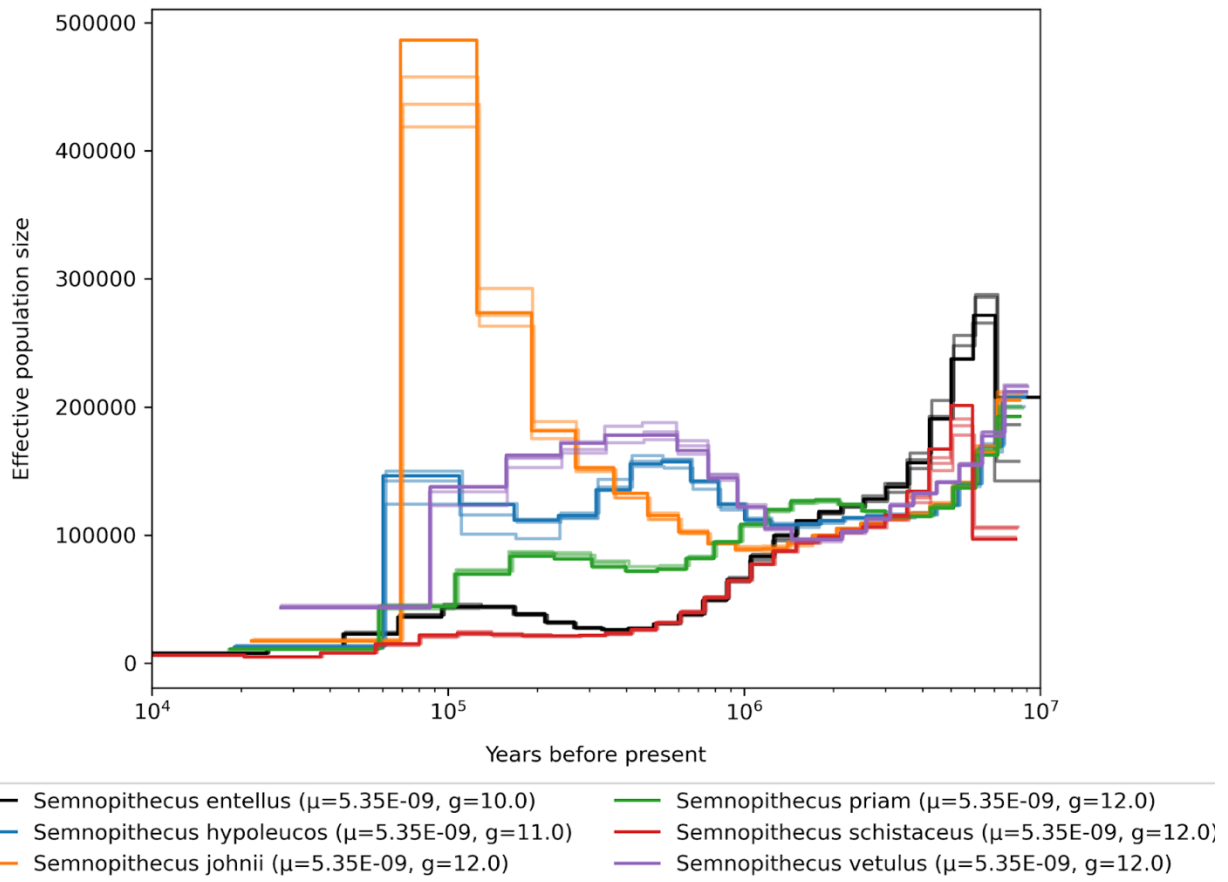


Fig. S107.

Demographic history of the genus *Semnopithecus*, mutation rates for scaling are averaged per genus.

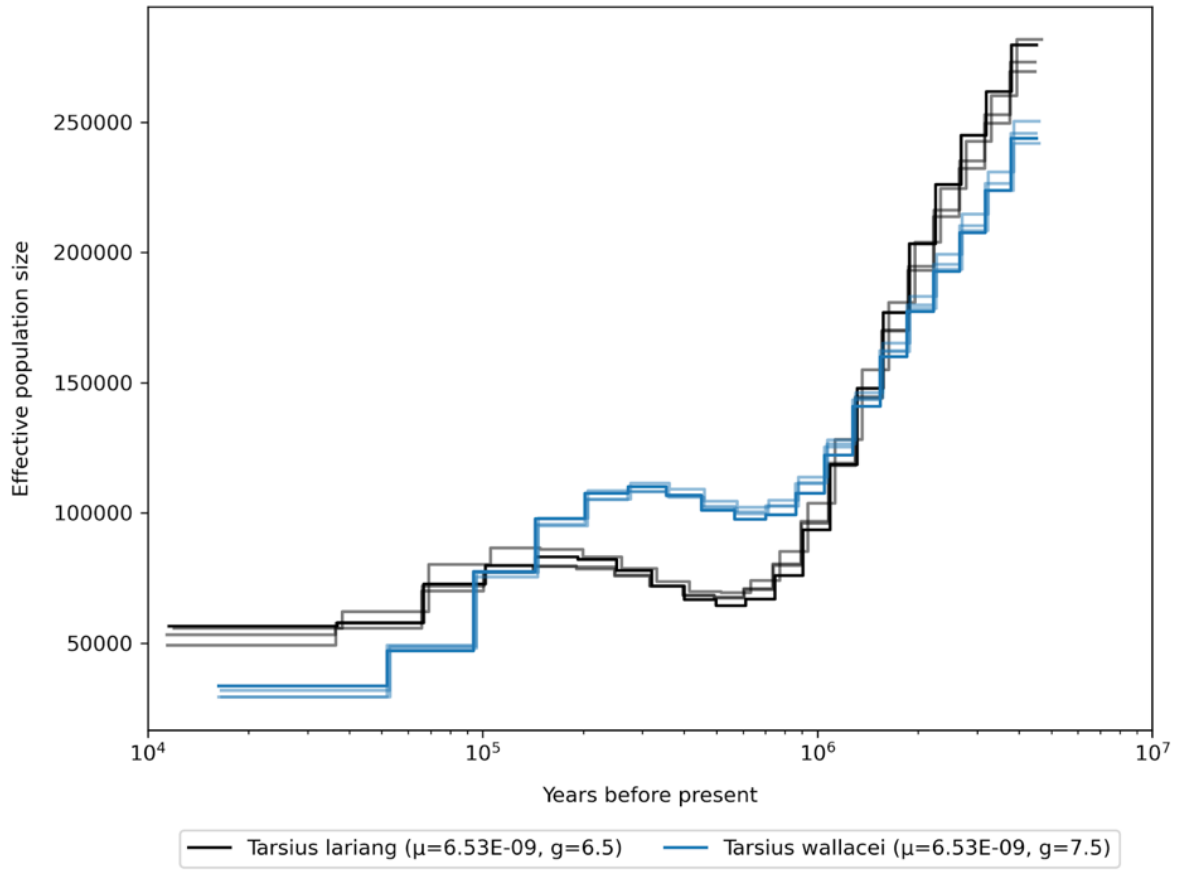


Fig. S108. Demographic history of the genus *Tarsius*, mutation rates for scaling are averaged per genus.

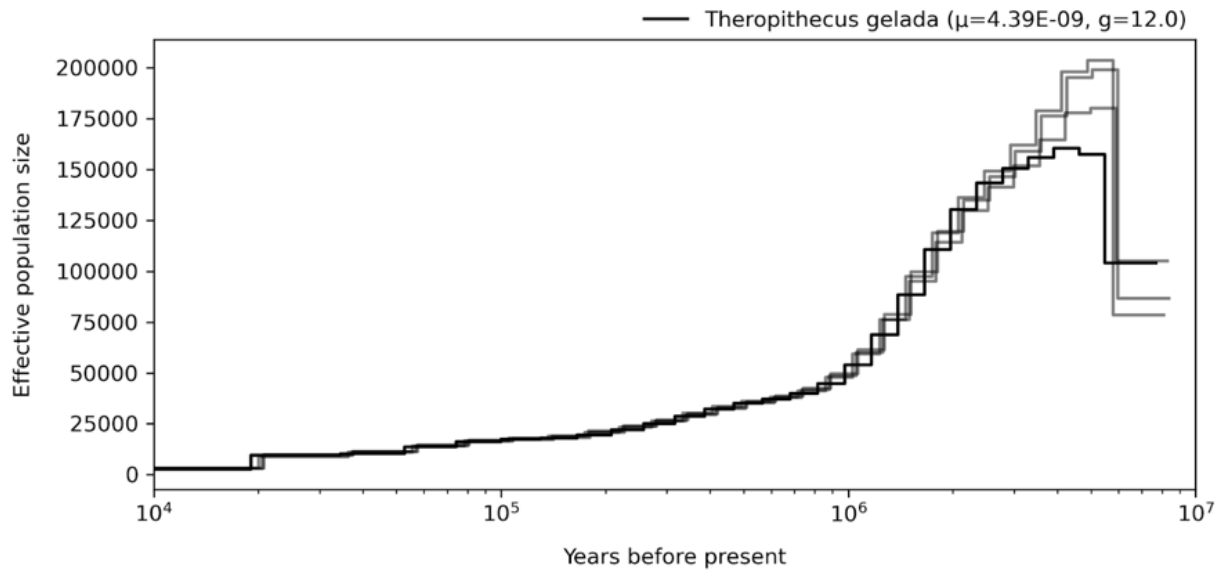


Fig. S109. Demographic history of the genus *Theropithecus*, mutation rates for scaling are averaged per genus.

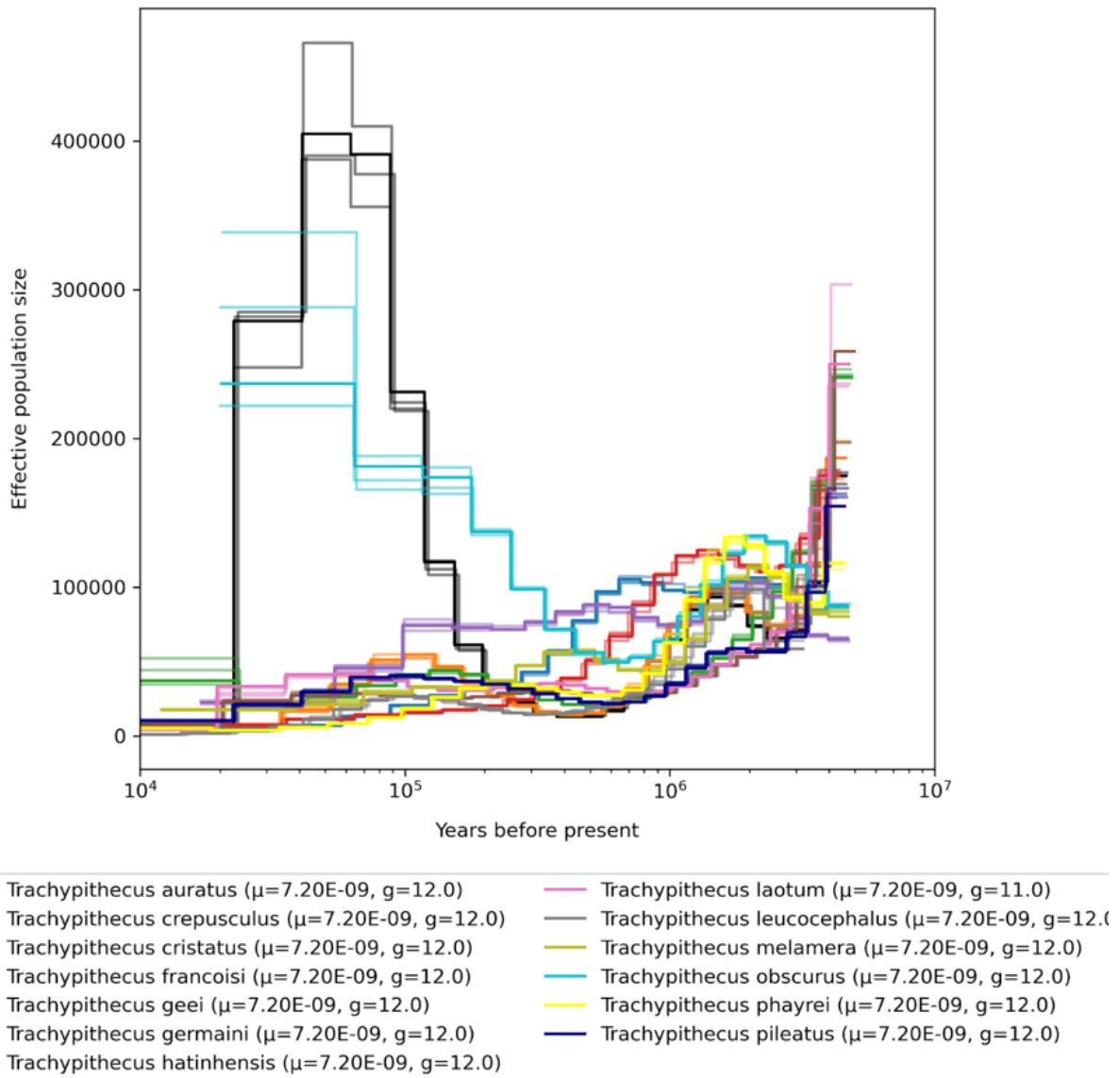


Fig. S110.

Demographic history of the genus *Trachypithecus*, mutation rates for scaling are averaged per genus.

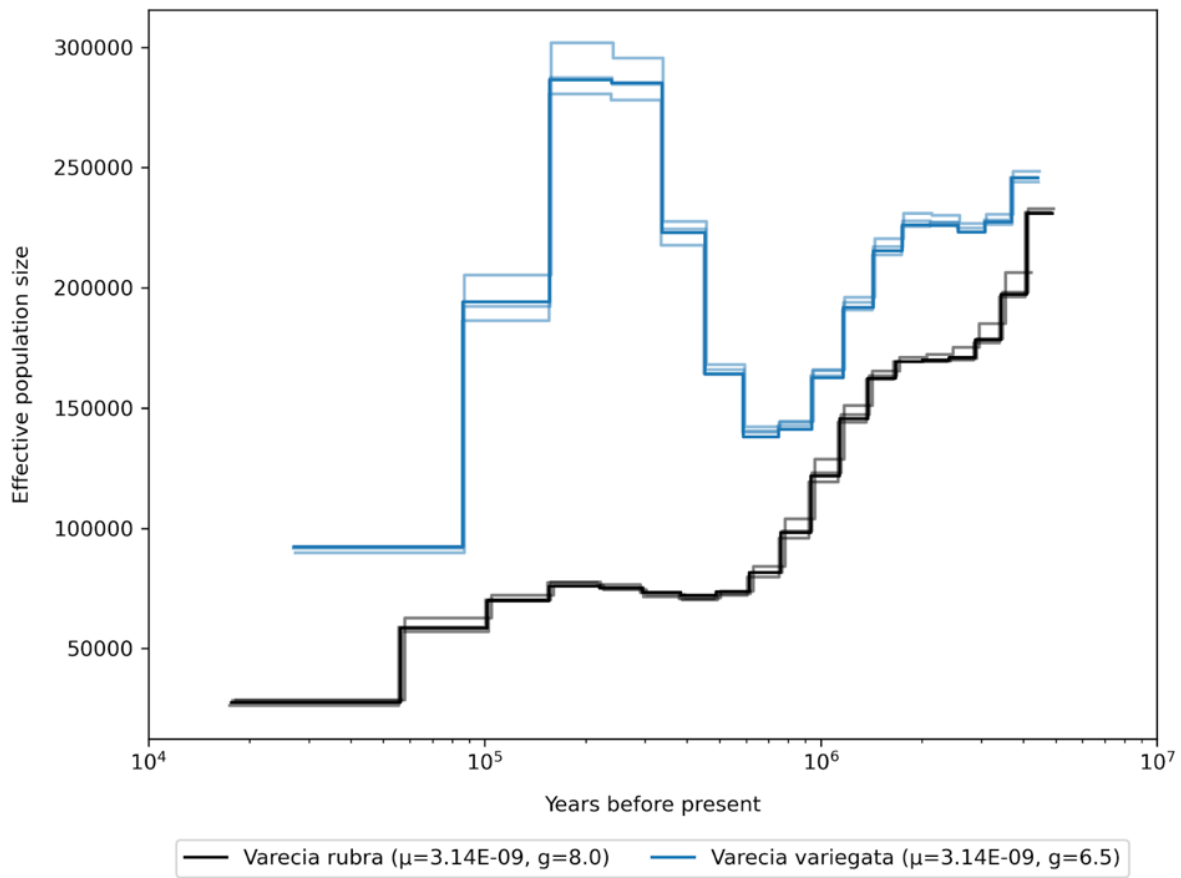


Fig. S111.
Demographic history of the genus *Varecia*, mutation rates for scaling are averaged per genus.

Supplementary Tables

Reference species	Name/Identifier	Source
<i>Aotus nancymaae</i>	GCA_000952055.2	NCBI
<i>Ateles fusciceps</i>	Ateles_fusciceps.fasta	Shao et al., this issue
<i>Callithrix jacchus</i>	GCA_000004665.1	NCBI
<i>Cebus albifrons</i>	Cebus_albifrons.fasta	Shao et al., this issue
<i>Cercocebus atys</i>	GCA_000955945.1	NCBI
<i>Cercopithecus mitis</i>	Cercopithecus_mitis.fasta	Shao et al., this issue
<i>Chlorocebus aethiops</i>	Chlorocebus_aethiops.fasta	Shao et al., this issue
<i>Colobus guereza</i>	Colobus_guereza.fasta	Shao et al., this issue
<i>Daubentonia madagascariensis</i>	Daubentonia_madagascariensis.fasta	Shao et al., this issue
<i>Erythrocebus patas</i>	Erythrocebus_patas.fasta	Shao et al., this issue
<i>Carlito syrichta</i>	GCF_000164805.1_Tarsius_syrichta-2.0.1_genomic.fna	NCBI
<i>Propithecus coquerelli</i>	GCA_000956105.1	NCBI
<i>Galago moholi</i>	Galago_moholi.fasta	Shao et al., this issue
<i>Gorilla gorilla</i>	GCA_008122165.1	NCBI
<i>Lemur catta</i>	GCA_020740605.1	NCBI
<i>Loris tardigradus</i>	Loris_tardigradus.fasta	Shao et al., this issue
<i>Macaca mulatta</i>	GCA_003339765.3	NCBI
<i>Mandrillus sphinx</i>	Mandrillus_sphinx.fasta	Shao et al., this issue
<i>Microcebus murinus</i>	GCA_000165445.3	NCBI
<i>Nomascus leucogenys</i>	Nomascus_leucogenys.fasta	Shao et al., this issue
<i>Nycticebus pygmaeus</i>	Nycticebus_pygmaeus.fasta	Shao et al., this issue
<i>Otolemur garnettii</i>	GCA_000181295.3	NCBI
<i>Pan troglodytes</i>	GCA_002880755.3	NCBI
<i>Papio anubis</i>	GCA_000264685.2	NCBI
<i>Pithecia pithecia</i>	Pithecia_pithecia.fasta	Shao et al., this issue
<i>Pongo abelii</i>	GCA_002880775.3	NCBI
<i>Pongo pygmaeus</i>	Pongo_pygmaeus.fasta	Shao et al., this issue
<i>Rhinopithecus roxellana</i>	GCA_007565055.1	NCBI
<i>Saguinus midas</i>	Saguinus_midas.fasta	Shao et al., this issue
<i>Sapajus apella</i>	Sapajus_apella.fasta	Shao et al., this issue
<i>Theropithecus gelada</i>	GCA_003255815.1	NCBI

Table S1.

Reference assemblies used in this study and their respective sources (see Supplementary Data S1 for sample level mappings).

Focal species	Comparison species	Pairwise distance (Substitutions per site)
<i>Callithrix jacchus</i>	<i>Saguinus midas</i>	1.91E-02
<i>Cebus albifrons</i>	<i>Sapajus apella</i>	6.69E-03
<i>Cercopithecus mitis</i>	<i>Cercopithecus mona</i>	6.01E-03
<i>Colobus guereza</i>	<i>Colobus angolensis</i>	3.05E-03
<i>Daubentonia madagascariensis</i>	<i>Lemur catta</i>	4.65E-02
<i>Erythrocebus patas</i>	<i>Chlorocebus aethiops</i>	6.27E-03
<i>Galago moholi</i>	<i>Otolemur garnettii</i>	1.50E-02
<i>Lemur catta</i>	<i>Prolemur simus</i>	1.00E-02
<i>Loris tardigradus</i>	<i>Nycticebus pygmaeus</i>	2.43E-02
<i>Macaca mulatta</i>	<i>Macaca assamensis</i>	3.13E-03
<i>Mandrillus sphinx</i>	<i>Mandrillus leucophaeus</i>	2.02E-03
<i>Microcebus murinus</i>	<i>Propithecus coquereli</i>	4.18E-02
<i>Nycticebus pygmaeus</i>	<i>Nycticebus bengalensis</i>	7.63E-03
<i>Otolemur garnettii</i>	<i>Galago moholi</i>	1.50E-02
<i>Papio anubis</i>	<i>Macaca mulatta</i>	7.05E-03
<i>Rhinopithecus roxellana</i>	<i>Rhinopithecus strykeri</i>	1.99E-03
<i>Saguinus midas</i>	<i>Callithrix jacchus</i>	1.91E-02
<i>Sapajus apella</i>	<i>Cebus albifrons</i>	6.69E-03
<i>Theropithecus gelada</i>	<i>Lophocebus aterrimus</i>	4.46E-03

Table S2.
Species pairs used to estimate effects of cross-reference mappings, and their pairwise distance.

Clade	Split	Distribution	MCMCTree prior age distribution
Crown Euarchontoglires	Euarchonta-Glires	skewT	ST(65.79,4.725,50,1)
Crown Glires	Rodentia-Lagomorpha	skewT	ST(62.22,0.298,50,1)
Crown Euarchonta	Primatesomorpha- Scandentia	skewT	ST(65.79,4.725,50,1)
Crown Primatomorpha	Primates-Dermoptera	lower	L(55.935)
Crown Primates	Haplorhini-Strepsirhini	skewT	ST(55.935,0.8,50,1)
Crown Strepsirrhini	Lorisiformes- (Lemuriformes+Chiromyiformes)	bound	B(37,55.8,1e-300,0.05)
Lemuriformes+ Chiromyiformes	Lemuriformes- Chiromyiformes	bound	B(33.4,55.8,1e-300,0.05)
Crown Lorisiformes	Lorisidae-Galagidae	bound	B(18.5,55.8,1e-300,0.05)
Crown Lorisidae	Lorisinae-Perodicticinae	bound	B(8.9,37.5,1e-300,0.05)
<i>Galagoides+Galago+ Paragalago+Otolemur+ Sciurocheirus</i>	<i>Galagoides-(Galago+Paragalago+ Otolemur+Sciurocheirus)</i>	bound	B(10,33.4,1e-300,0.05)
Crown Haplorhini	Anthropoidea-Tarsiiformes	lower	L(41)
Crown Anthropoidea	Catarrhini-Platyrrhini	bound	B(33.4,56.035,1e-300,0.05)
Crown Catarrhini	Cercopithecoidea-Hominoidea	bound	B(25.193,33.4,1e-300,0.05)
Crown Cercopithecidae	Cercopithecinae-Colobinae	skewT	ST(12.47,1.005,50,1)
Crown Colobinae	Colobini-Presbytini	bound	B(8.2,15,1e-300,0.05)
Crown Cercopithecinae	Cercopithecini-Papionini	bound	B(6.5,15,1e-300,0.05)
Crown Papionini	Macacina-Papionina	bound	B(5.33,12.51,1e-300,0.05)
Crown Cercopithecini	uncertain	bound	B(6.5,12.51,1e-300,0.05)
Crown Hominoidea	Hominidae-Hylobatidae	skewT	ST(13.4,0.932,50,1)
Crown Hominidae	Homininae-Ponginae	skewT	ST(12.3,1.018,50,1)
Crown Homininae	Gorillini-Homini	bound	B(7.5,15,1e-300,0.05)
Homo+Pan	Homo-Pan	bound	B(4.631,15,1e-300,0.05)
Crown Platyrrhini	Pitheciidae-(Aotidae+Atelidae+ Callitrichidae+Cebidae)	lower	L(13.363)
Crown Pitheciidae	Callicebinae-Pitheciinae	bound	B(13.032,34.5,1e-300,0.05)
Callitrichidae+Cebidae	Callitrichidae-Cebidae	bound	B(13.183,34.5,1e-300,0.05)
Crown Cebidae	Cebinae-Saimirinae	bound	B(13.032,34.5,1e-300,0.05)
Crown Atelidae	Alouattinae-Atelinae	bound	B(13.363,34.5,1e-300,0.05)

Table S3.
Calibrations used in MCMCTree

Genus	Mitochondrial Reference	GenBank Accessions
<i>Allenopithecus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Allochrocebus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Arctocebus</i>	<i>Nycticebus_cougang</i>	NC_040292
<i>Avahi</i>	<i>Propithecus_verreauxi</i>	NC_028210
<i>Carlito</i>	<i>Carlito_syrichta</i>	NC_012774
<i>Cephalopachus</i>	<i>Carlito_syrichta</i>	NC_012774
<i>Cercocebus</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Cercopithecus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Cheirogaleus</i>	<i>Microcebus_murinus</i>	NC_028718
<i>Chlorocebus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Colobus</i>	<i>Colobus_guereza</i>	NC_006901
<i>Daubentonia</i>	<i>Daubentonia_madagascariensis</i>	NC_010299
<i>Erythrocebus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Eulemur</i>	<i>Lemur_catta</i>	NC_059325
<i>Galago</i>	<i>Galago_moholi</i>	NC_021949
<i>Galagoides</i>	<i>Galago_moholi</i>	NC_021949
<i>Gorilla</i>	<i>Homo_sapiens</i>	NC_012920
<i>Hapalemur</i>	<i>Lemur_catta</i>	NC_059325
<i>Hoolock</i>	<i>Nomascus_leucogenys</i>	NC_021957
<i>Hylobates</i>	<i>Nomascus_leucogenys</i>	NC_021957
<i>Indri</i>	<i>Propithecus_verreauxi</i>	NC_028210
<i>Lemur</i>	<i>Lemur_catta</i>	NC_059325
<i>Lepilemur</i>	<i>Microcebus_murinus</i>	NC_028718
<i>Lophocebus</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Loris</i>	<i>Nycticebus_cougang</i>	NC_040292
<i>Macaca</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Mandrillus</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Microcebus</i>	<i>Microcebus_murinus</i>	NC_028718
<i>Miopithecus</i>	<i>Chlorocebus_sabaeus</i>	NC_008066
<i>Mirza</i>	<i>Microcebus_murinus</i>	NC_028718
<i>Nasalis</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Nomascus</i>	<i>Nomascus_leucogenys</i>	NC_021957
<i>Nycticebus</i>	<i>Nycticebus_cougang</i>	NC_040292
<i>Otolemur</i>	<i>Galago_moholi</i>	NC_021949
<i>Pan</i>	<i>Homo_sapiens</i>	NC_012920
<i>Papio</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Perodicticus</i>	<i>Nycticebus_cougang</i>	NC_040292
<i>Ptilocolobus</i>	<i>Colobus_guereza</i>	NC_006901
<i>Pongo</i>	<i>Homo_sapiens</i>	NC_012920

<i>Presbytis</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Prolemur</i>	<i>Lemur_catta</i>	NC_059325
<i>Propithecus</i>	<i>Propithecus_verreauxi</i>	NC_028210
<i>Pygathrix</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Rhinopithecus</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Semnopithecus</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Tarsius</i>	<i>Carlito_syrichta</i>	NC_012774
<i>Theropithecus</i>	<i>Macaca_mulatta</i>	NC_005943
<i>Trachypithecus</i>	<i>Rhinopithecus_bieti</i>	NC_015486
<i>Varecia</i>	<i>Lemur_catta</i>	NC_059325

Table S4.
Reference species and accessions used for mitochondrial assembly

Species	Type	Phylogeny
<i>Alouatta_juara</i>	polyphyletic	nuclear, mitochondrial
<i>Alouatta_seniculus</i>	polyphyletic	nuclear, mitochondrial
<i>Ateles_chamek</i>	polyphyletic	mitochondrial
<i>Cacajao_ayresi</i>	polyphyletic	nuclear
<i>Cacajao_hosomi</i>	polyphyletic	nuclear
<i>Cebus_albifrons</i>	polyphyletic	nuclear, mitochondrial
<i>Cebus_unicolor</i>	polyphyletic	nuclear
<i>Cercopithecus_cephus</i>	polyphyletic	mitochondrial
<i>Chiropotes_sagulatus</i>	paraphyletic	nuclear
<i>Macaca_mulatta</i>	polyphyletic	mitochondrial
<i>Macaca_nemestrina</i>	polyphyletic	mitochondrial
<i>Nycticebus_bengalensis</i>	polyphyletic	mitochondrial
<i>Pithecia_pissinattii</i>	polyphyletic	nuclear, mitochondrial
<i>Plecturocebus_caligatus</i>	polyphyletic	nuclear
<i>Plecturocebus_dubius</i>	polyphyletic	nuclear
<i>Pygathrix_cinerea</i>	polyphyletic	nuclear, mitochondrial
<i>Pygathrix_nemaeus</i>	polyphyletic	nuclear, mitochondrial
<i>Saimiri_sciureus</i>	polyphyletic	nuclear, mitochondrial
<i>Sapajus_apella</i>	polyphyletic	nuclear, mitochondrial
<i>Sapajus_macrocephalus</i>	polyphyletic	nuclear
<i>Semnopithecus_entellus</i>	polyphyletic	nuclear, mitochondrial
<i>Semnopithecus_hypoleucos</i>	polyphyletic	mitochondrial
<i>Semnopithecus_ohnii</i>	polyphyletic	mitochondrial
<i>Semnopithecus_vetulus</i>	polyphyletic	mitochondrial
<i>Trachypithecus_geei</i>	polyphyletic	nuclear

Table S5.

Species for which non-monophyletic relationships are observed in phylogenies including two individuals per species.

Species	Phylogeny-based estimates	Trio-based estimates	Source
<i>Pongo abelii</i>	1.66E-08	1.47E-08	Besenbacher et al., 2019
<i>Homo sapiens</i>	1.22E-08	1.12E-08	Kessler et al., 2020
<i>Pan troglodytes</i>	1.26E-08	1.31E-08	Besenbacher et al., 2019
<i>Gorilla gorilla</i>	1.13E-08	1.33E-08	Besenbacher et al., 2019
<i>Papio anubis</i>	5.70E-09	3.91E-09	Wu et al., 2020
<i>Macaca mulatta</i>	7.70E-09	5.55E-09	Bergeron et al., 202
<i>Callithrix jacchus</i>	4.30E-09	8.41E-09	Yang et al., 2021
<i>Microcebus murinus</i>	1.52E-08	4.96E-09	Campbell et al., 2021

Table S6.

Comparison of per generation mutation rates estimated in this study versus trio-based estimates available in the literature for overlapping species

Family	Median	Min	Max
Aotidae	4.79E-09	4.54E-09	5.37E-09
Atelidae	8.32E-09	5.58E-09	1.08E-08
Callitrichidae	5.75E-09	3.84E-09	8.41E-09
Cebidae	9.38E-09	4.10E-09	1.45E-08
Cercopithecidae	5.13E-09	2.50E-09	1.10E-08
Cheirogaleidae	4.00E-09	2.85E-09	4.96E-09
Daubentoniidae	4.42E-09	4.42E-09	4.42E-09
Galagidae	5.97E-09	3.92E-09	7.55E-09
Hominidae	1.30E-08	1.17E-08	1.47E-08
Hylobatidae	9.04E-09	5.86E-09	1.13E-08
Indriidae	5.40E-09	4.18E-09	8.64E-09
Lemuridae	3.81E-09	2.56E-09	5.60E-09
Lepilemuridae	4.41E-09	3.70E-09	6.99E-09
Lorisidae	9.10E-09	4.76E-09	1.62E-08
Pitheciidae	7.27E-09	3.91E-09	1.07E-08
Tarsiidae	7.52E-09	7.46E-09	7.58E-09

Table S7.

Summary of mutation rates per generation grouped by family. Estimates for individual species can be found in Supplementary Data S2.

	Median /1000	Min /1000	Max /1000		Median /1000	Min /1000	Max /1000
<i>Allenopithecus</i>	104	104	104	<i>Leontopithecus</i>	46	41	51
<i>Allochrocebus</i>	152	125	195	<i>Lepilemur</i>	193	156	232
<i>Alouatta</i>	64	19	94	<i>Lophocebus</i>	164	164	164
<i>Aotus</i>	74	49	198	<i>Loris</i>	44	10	79
<i>Arctocebus</i>	80	80	80	<i>Macaca</i>	72	28	225
<i>Ateles</i>	51	35	82	<i>Mandrillus</i>	229	162	297
<i>Avahi</i>	151	151	151	<i>Mico</i>	28	20	88
<i>Cacajao</i>	28	15	60	<i>Microcebus</i>	214	214	214
<i>Callimico</i>	84	84	84	<i>Miopithecus</i>	145	145	145
<i>Callithrix</i>	102	36	137	<i>Mirza</i>	87	87	87
<i>Carlito</i>	75	75	75	<i>Nasalis</i>	60	60	60
<i>Cebuella</i>	52	43	62	<i>Nomascus</i>	55	40	66
<i>Cebus</i>	32	30	37	<i>Nycticebus</i>	24	16	63
<i>Cephalopachus</i>	98	98	98	<i>Otolemur</i>	63	41	84
<i>Cercocebus</i>	110	110	144	<i>Pan</i>	23	19	27
<i>Cercopithecus</i>	288	108	398	<i>Papio</i>	91	32	117
<i>Cheirogaleus</i>	241	120	363	<i>Perodicticus</i>	71	48	93
<i>Cheracebus</i>	84	44	141	<i>Piliocolobus</i>	38	27	211
<i>Chiropotes</i>	50	47	61	<i>Pithecia</i>	43	29	111
<i>Chlorocebus</i>	164	164	164	<i>Plecturocebus</i>	81	43	138
<i>Colobus</i>	82	50	114	<i>Pongo</i>	41	36	47
<i>Daubentonia</i>	35	35	35	<i>Presbytis</i>	64	52	76
<i>Erythrocebus</i>	85	85	85	<i>Prolemur</i>	183	183	183
<i>Eulemur</i>	197	99	611	<i>Propithecus</i>	163	106	275
<i>Galago</i>	58	38	79	<i>Pygathrix</i>	43	41	59
<i>Galagoides</i>	207	207	207	<i>Rhinopithecus</i>	21	20	21
<i>Gorilla</i>	25	16	33	<i>Saguinus</i>	60	41	97
<i>Hapalemur</i>	289	143	497	<i>Saimiri</i>	60	49	117
<i>Hoolock</i>	28	28	28	<i>Sapajus</i>	61	29	93
<i>Hylobates</i>	77	38	114	<i>Semnopithecus</i>	88	23	125
<i>Indri</i>	314	314	314	<i>Theropithecus</i>	28	28	28
<i>Lagothrix</i>	53	53	53	<i>Trachypithecus</i>	30	10	86
<i>Lemur</i>	193	193	193	<i>Varecia</i>	184	125	244
<i>Leontocebus</i>	111	45	131				

Table S8.

Estimates of effective population sizes grouped by genus. Estimates for individual species can be found in Supplementary Data S2

Category	Trait	log-transform	Allometric
Brain Size	Species Mean Brain Size	yes	yes
Activity budget	Feed		
	Move		
	Social	yes	
	Rest		
Life History	Body Mass	yes	
	Interbirth Interval	yes	yes
	Gestation	yes	yes
	Weaning Age	yes	yes
	Maximum Longevity		yes
Sexual selection	Litter Size	yes	
	Mass Dimorphism	yes	yes
	Canine Dimorphism	yes	yes
	Testes Mass	yes	yes
Ranging patterns	Home range	yes	
	DayLength_km	yes	
	Territoriality Index	yes	
Climatic niche	Midrange Latitude		
	Mean Precipitation		
	Mean Temperature		
	Actual evapotranspiration		
	Potential evapotranspiration		
Social organization	Mean group size	yes	
	Adult males per group	yes	
	Adult females per group	yes	
	Adult sex ratio	yes	
Diet composition	Fruit		
	Leaves		
	Fauna		
Social/Mating Systems	Social System		
	Mating System		
Dispersal Pattern	Natal Dispersal		

Table S9.
Overview of traits and corresponding categories

Trait	P-value (raw)	P-value (corrected)	intercept	coefficient	R2-pred
GR_MidRangeLat_dd	4.68E-07	1.45E-05	-6.48	-0.03	0.28
Precip_Mean_mm	1.27E-05	1.97E-04	-7.09	0.00	0.33
Temp_Mean_degC	3.57E-04	3.69E-03	-7.66	0.05	0.26
Social_log	7.18E-04	5.56E-03	-6.04	-0.30	0.19
Mating.System_Harem_Polygyny	2.47E-03	1.53E-02	-6.55	-0.62	0.11
MaxLongevity_m_res	3.37E-02	1.74E-01	-6.48	0.00	0.04
Territoriality	4.11E-02	1.82E-01	-6.01	-0.16	0.08
DayLength_km_log	5.12E-02	1.98E-01	-6.56	-0.23	0.18
BodyMassMean_calc_log	6.77E-02	2.33E-01	-4.87	-0.23	0.16
HomeRange_km2_log	1.01E-01	3.14E-01	-6.70	-0.06	0.11
Gestation_res	1.37E-01	3.87E-01	-6.51	-0.96	0.11
Mating.System_Polyandry	2.55E-01	6.60E-01	-6.61	-0.57	0.05
Move	2.85E-01	6.80E-01	-6.80	0.01	0.21
MassDimorphism_res	3.70E-01	7.16E-01	-6.46	0.37	0.14
CanineDimorphism_res	3.35E-01	7.16E-01	-6.53	-0.39	0.07
Fauna	3.55E-01	7.16E-01	-6.50	-0.01	0.08
MeanGroupSize_log	4.52E-01	8.11E-01	-6.36	-0.08	0.14
AdultMales_log	4.71E-01	8.11E-01	-6.49	0.08	0.10
Feed	8.84E-01	9.31E-01	-6.51	0.00	0.17
InterbirthInterval_days_res	7.85E-01	9.31E-01	-6.61	-0.06	0.13
WeaningAge_d_res	9.31E-01	9.31E-01	-6.62	-0.02	0.03
LitterSz_log	9.13E-01	9.31E-01	-6.47	-0.05	0.12
CombinedTestesMass.in.g_res	7.98E-01	9.31E-01	-6.50	0.04	-0.24
AdultSexRatio_log	6.36E-01	9.31E-01	-6.43	-0.05	0.22
Fruit	5.92E-01	9.31E-01	-6.77	0.00	0.17
Leaves	9.17E-01	9.31E-01	-6.72	0.00	0.19
SocialSystem_Pair	6.67E-01	9.31E-01	-6.67	0.10	0.20
SocialSystem_Polygynandry	7.42E-01	9.31E-01	-6.60	-0.05	0.20
SocialSystem_Polygyny	9.09E-01	9.31E-01	-6.61	0.02	0.20
SocialSystem_Solitary	9.25E-01	9.31E-01	-6.60	-0.07	0.20
NatalDispersal_Sexbiased	6.17E-01	9.31E-01	-6.61	0.09	0.28

Table S10.

Results of genetic diversity trait PGLS models after removing traits that failed diagnostic or correlation tests. Traits that significantly predict log genetic diversity after multiple testing correction are listed in bold.

Regression Model	AIC	R ²	Intercept	Log10(Ne)	Gen. length	Log10(branch)
OLS	-310.12	0.512	-7.239	-0.257	0.014	-0.041
PGLS Brownian	-257.35	0.452	-7.135	-0.23	0.026	0.049
PGLS Grafen	-385.59	0.476	-7.855	-0.191	0.023	-0.113

Table S11.

Summary of Models. Effects reported are the slopes associated with each covariate (Ne, gen length, terminal branch length). Note that all three models have very similar R² and the slopes associated with Ne and Generation length are very similar. The log10(branch) effect (as measured by the slope coefficient) is more labile ranging from weakly positive to negative, with varying significance.

Regression Model	AIC	R ²	Intercept	log10(N _e)	Gen. length	log10(branch)
OLS	-99.15	0.458	-7.129	-0.275	0.012	-0.035
PGLS Brownian	-68.11	0.392	-8.076	-0.151	0.019	-0.145
PGLS Grafen	-111.88	0.411	-7.919	-0.176	0.02	-0.119

Table S12.

Summary of Models for a subset of 88 species that are separated by at least $4N_e$ generations, and thus unlikely to share polymorphisms. Effects reported are the slopes associated with each covariate (N_e, gen length, terminal branch length). While AIC varies considerably depending on the choice of models, the overall magnitude of R² and slopes estimates associated with log10(N_e), Generation length, and terminal branch length are qualitatively remarkably stable across different model choices and species subsets (see Table S11).

GO biological process complete	over/ under	Enrichment	P-value (raw)	P-value (corrected)
epithelial cilium movement involved in extracellular fluid movement (GO:0003351)	+	3.36	3.67E-05	2.40E-02
*extracellular transport (GO:0006858)	+	3.29	3.04E-05	2.07E-02
**microtubule-based movement (GO:0007018)	+	1.89	3.67E-08	8.22E-05
***microtubule-based process (GO:0007017)	+	1.75	4.69E-12	7.35E-08
***movement of cell or subcellular component (GO:0006928)	+	1.32	1.29E-05	1.06E-02
*cilium movement (GO:0003341)	+	2.26	1.07E-06	1.20E-03
non-motile cilium assembly (GO:1905515)	+	3.06	2.13E-05	1.67E-02
*cilium assembly (GO:0060271)	+	2.09	2.11E-09	6.61E-06
**organelle assembly (GO:0070925)	+	1.52	1.12E-06	1.17E-03
***organelle organization (GO:0006996)	+	1.2	5.84E-06	5.72E-03
**plasma membrane bounded cell projection assembly (GO:0120031)	+	1.84	8.68E-08	1.70E-04
***cell projection assembly (GO:0030031)	+	1.8	1.46E-07	2.29E-04
****cell projection organization (GO:0030030)	+	1.36	2.98E-05	2.12E-02
**cilium organization (GO:0044782)	+	2.05	1.65E-09	6.46E-06
axoneme assembly (GO:0035082)	+	3.06	1.03E-07	1.79E-04
*microtubule bundle formation (GO:0001578)	+	2.61	3.55E-07	4.64E-04
**microtubule cytoskeleton organization (GO:0000226)	+	1.83	1.01E-09	5.27E-06
***cytoskeleton organization (GO:0007010)	+	1.55	1.02E-10	8.02E-07
motile cilium assembly (GO:0044458)	+	2.87	4.38E-05	2.64E-02
cilium-dependent cell motility (GO:0060285)	+	2.3	6.74E-06	6.21E-03
*cilium or flagellum-dependent cell motility (GO:0001539)	+	2.3	6.74E-06	5.87E-03
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	+	2.26	8.15E-07	9.83E-04
*cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	+	1.78	5.77E-05	3.23E-02
**cell adhesion (GO:0007155)	+	1.56	6.08E-09	1.59E-05
multicellular organismal process (GO:0032501)	+	1.12	7.13E-05	3.85E-02
positive regulation of RNA metabolic process (GO:0051254)	-	0.75	5.00E-05	2.90E-02
*positive regulation of nucleobase-containing compound metabolic process (GO:0045935)	-	0.76	2.70E-05	2.01E-02
**positive regulation of nitrogen compound metabolic process (GO:0051173)	-	0.81	3.67E-05	2.30E-02
negative regulation of developmental process (GO:0051093)	-	0.66	8.72E-05	4.56E-02
adaptive immune response (GO:0002250)	-	0.49	1.68E-07	2.39E-04

Table S13.

Significant results in GO-term enrichment analysis of genes affected by non-recurring ape-specific missense changes. Terms without asterisk denote the most general term within a hierarchy, and subsequent terms within the same hierarchy are preceded by asterisks. Additional asterisks denote a further decrease in the hierarchy relating to the preceding term in the table.

Splits among outgroups	Mean	CI
Oryctolagus + Mus - Tupaia + Galeopterus + Primates	69.06	66.94 - 72.43
Tupaia - Galeopterus + Primates	67.82	65.67 - 70.64
Galeopterus - Primates	65.11	62.85 - 67.78
Oryctolagus - Mus	63.43	62.19 - 67.80

Table S 14

Split times in million years. CI denotes 95% HPD

Major splits among primates	Mean	CI
Strepsirrhini - Haplorhini	60.71	58.31 - 63.30
Lemuriformes - Lorisiformes	49.68	47.18 - 52.32
Daubentoniidae - Lemuridae + Indriidae + Cheirogaleidae + Lepilemuridae	46.25	43.76 - 48.81
Lemuridae - Indriidae + Cheirogaleidae + Lepilemuridae	35.99	33.67 - 38.37
Indriidae - Cheirogaleidae + Lepilemuridae	34.76	32.41 - 37.03
Cheirogaleidae - Lepilemuridae	30.97	28.76 - 33.13
Lorisidae - Galagidae	21.58	19.89 - 23.38
Tarsiidae - Anthropeida	57.87	55.41 - 60.41
Platyrrhini - Catarrhini	39.03	36.73 - 41.41
Pitheciidae - Atelidae + Cebidae + Aotidae + Callithrichidae	24.55	22.66 - 26.50
Atelidae - Cebidae + Aotidae + Callithrichidae	23.27	21.45 - 25.14
Cebidae - Aotidae + Callithrichidae	20.70	19.02 - 22.41
Aotidae - Callithrichidae	20.26	18.64 - 21.98
Cercopithecidae - Hylobatidae + Hominidae	31.08	29.16 - 33.08
Hylobatidae - Hominidae	22.89	21.02 - 24.80

Table S 15

Split times in million years. CI denotes 95% HPD

Splits within Lemuridae	Mean	CI
Varecia - Eulemur + Lemur + Prolemur + Hapalemur	26.32	24.20 - 28.39
Eulemur - Lemur + Prolemur + Hapalemur	23.90	21.93 - 25.90
Lemur - Prolemur + Hapalemur	12.92	11.35 - 14.45
Prolemur - Hapalemur	10.65	9.23 - 11.96
Varecia variegata - V. rubra	1.91	1.39 - 2.46
Eulemur rubriventer - other Eulemur sp.	11.42	10.14 - 12.69
Eulemur mongoz - other Eulemur sp.	7.82	6.86 - 8.74
Eulemur coronatus + E. flavifrons + E. macaco - other Eulemur sp.	7.41	6.50 - 8.30
Eulemur coronatus - E. flavifrons + E. macaco	6.65	5.81 - 7.46
Eulemur flavifrons - E. macaco	1.52	1.26 - 1.79
Eulemur albifrons + E. sanfordi - E. fulvus + E. collaris + E. rufus	4.77	4.11 - 5.43
Eulemur albifrons - E. sanfordi	2.61	2.18 - 3.03
Eulemur fulvus - E. collaris + E. rufus	3.83	3.28 - 4.40
Eulemur collaris - E. rufus	2.82	2.38 - 3.25
Hapalemur occidentalis - other Hapalemur sp.	3.74	3.09 - 4.42
Hapalemur alaotrensis - other Hapalemur sp.	3.11	2.56 - 3.69
Hapalemur gilberti - H. meridionalis + H. griseus	2.55	2.08 - 3.01
Hapalemur meridionalis - H. griseus	2.22	1.81 - 2.63

Table S 16
Split times in million years. CI denotes 95% HPD

Splits within Indriidae	Mean	CI
Indri - Avahi + Propithecus	22.67	19.92 - 24.96
Avahi - Propithecus	19.03	16.40 - 21.17
Avahi laniger - A. peyrierasi	2.00	1.50 - 2.52
Propithecus edwardsi + P. diadema + P. perrieri – P. coronatus + P. tattersalli + P. verreauxi + P. coquereli	9.79	7.84 - 11.32
Propithecus edwardsi - P. diadema + P. perrieri	3.44	2.65 - 4.15
Propithecus diadema - P. perrieri	2.59	1.97 - 3.14
Propithecus coronatus - P. tattersalli + P. verreauxi + P. coquereli	5.20	4.03 - 6.15
Propithecus tattersalli - P. verreauxi + P. coquereli	4.30	3.33 - 5.12
Propithecus verreauxi - P. coquereli	0.90	0.67 - 1.12

Table S 17

Split times in million years. CI denotes 95% HPD

Splits within Cheirogaleidae	Mean	CI
Cheirogaleus - Mirza + Microcebus	24.78	22.76 - 26.80
Mirza - Microcebus	11.20	9.91 - 12.63
Cheirogaleus major - C. medius	10.58	8.89 - 12.23

Table S 18

Split times in million years. CI denotes 95% HPD

Splits within Lepilemuridae	Mean	CI
Lepilemur ruficaudatus - other Lepilemur sp.	7.49	5.99 - 9.01
Lepilemur dorsalis - L. septentrionalis + L. ankaranensis	2.31	1.77 - 2.84
Lepilemur septentrionalis - L. ankaranensis	1.48	1.12 - 1.84

Table S 19

Split times in million years. CI denotes 95% HPD

Splits within Lorisidae	Mean	CI
Loris + Nycticebus - Arctocebus + Perodicticus	19.96	18.40 - 21.65
Loris - Nycticebus	11.14	9.99 - 12.39
Arctocebus - Perodicticus	11.89	10.64 - 13.30
Loris tardigradus - L. lydekkerianus	0.68	0.54 - 0.84
Nycticebus pygmaeus - N. bengalensis + N. coucang	2.98	2.47 - 3.50
Nycticebus bengalensis - N. coucang	0.37	0.29 - 0.45
Perodicticus potto - P. ibeanus	2.82	2.32 - 3.36

Table S 20
Split times in million years. CI denotes 95% HPD

Splits within Galagidae	Mean	CI
Galagoides - Galago + Otolemur	10.22	10.00 - 10.64
Galago - Otolemur	7.80	7.44 - 8.22
Galago senegalensis - G. moholi	1.12	0.96 - 1.29
Otolemur crassicaudatus - O. garnettii	3.22	2.89 - 3.36

Table S 21

Split times in million years. CI denotes 95% HPD

Splits within Tarsiidae	Mean	CI
Tarsius - Carlito + Cephalopachus	12.14	9.51 - 15.25
Carlito - Cephalopachus	5.45	4.13 - 7.01
Tarsius lariang - T. wallacei	2.46	1.78 - 3.24

Table S 22

Split times in million years. CI denotes 95% HPD

Splits within Pitheciidae	Mean	CI
Plecturocebus + Cheracebus - Pithecia + Cacajao + Chiropotes	20.29	18.65 - 22.05
Plecturocebus - Cheracebus	11.26	9.99 - 12.56
Pithecia - Cacajao + Chiropotes	13.57	12.21 - 14.95
Cacajao - Chiropotes	7.28	6.35 - 8.20
Plecturocebus hoffmannsi + P. miltoni + P. cinerascens - other Plecturocebus sp.	2.18	1.79 - 2.60
Plecturocebus hoffmannsi - P. miltoni + P. cinerascens	1.90	1.55 - 2.27
Plecturocebus miltoni - P. cinerascens	1.42	1.15 - 1.72
Plecturocebus bernhardi + P. moloch + P. grovesi – P. calligatus + P. brunneus + P. cupreus + P. dubius	1.95	1.59 - 2.33
Plecturocebus bernhardi - P. moloch + P. grovesi	1.33	1.08 - 1.60
Plecturocebus moloch - P. grovesi	0.73	0.58 - 0.88
Plecturocebus calligatus - P. brunneus + P. cupreus + P. dubius	0.80	0.64 - 0.96
Plecturocebus brunneus - P. cupreus + P. dubius	0.75	0.61 - 0.91
Plecturocebus cupreus - P. dubius	0.66	0.53 - 0.80
Cheracebus lugebns - C. lucifer + C. regulus + C. torquatus	2.22	1.88 - 2.66
Cheracebus lucifer - C. regulus + C. torquatus	1.74	1.41 - 2.09
Cheracebus regulus - C. torquatus	0.90	0.72 - 1.09
Pithecia chrysocephala - other Pithecia sp.	3.30	2.71 - 3.91
Pithecia albicans - P. hirsuta + P. mittermeieri + P. pissinattii + P. vanzolinii	1.12	0.90 - 1.36
Pithecia hirsuta - P. mittermeieri + P. pissinattii + P. vanzolinii	0.77	0.62 - 0.93
Pithecia mittermeieri - P. pissinattii + P. vanzolinii	0.55	0.44 - 0.67
Pithecia pissinattii - P. vanzolinii	0.49	0.38 - 0.60
Cacajao calvus - C. melanocephalus + C. ayresi + C. hosomi	1.69	1.39 - 2.01
Cacajao melanocephalus - C. ayresi + C. hosomi	0.54	0.44 - 0.66
Cacajao ayresi - C. hosomi	0.29	0.23 - 0.36
Chiropotes albinasus - C. israelita + C. sagulatus	2.53	2.12 - 2.96
Chiropotes israelita - C. sagulatus	0.82	0.66 - 0.98

Table S 23

Split times in million years. CI denotes 95% HPD

Splits within Atelidae	Mean	CI
Alouatta - Lagothrix + Ateles	17.29	15.72 - 18.97
Lagothrix - Ateles	12.78	11.36 - 14.23
Alouatta palliata - other Alouatta sp.	5.62	4.70 - 6.56
Alouatta discolor + A. belzebul - other Alouatta sp.	3.91	3.24 - 4.61
Alouatta discolor - A. belzebul	0.94	0.75 - 1.15
Alouatta caraya - other Alouatta sp.	3.35	2.76 - 3.94
Alouatta macconnelli + A. nigerrima - A. puruensis + A. juara + A. seniculus	2.68	2.21 - 3.19
Alouatta macconnelli - A. nigerrima	0.98	0.78 - 1.19
Alouatta puruensis - A. juara + A. seniculus	2.25	1.84 - 2.67
Alouatta juara - A. seniculus	0.88	0.70 - 1.06
Ateles geoffroyi - other Ateles sp.	3.74	3.09 - 4.40
Ateles paniscus - other Ateles sp.	3.16	2.61 - 3.74
Ateles belzebuth - A. chamek + A. marginatus	2.13	1.74 - 2.54
Ateles chamek - A. marginatus	1.12	0.90 - 1.35

Table S 24

Split times in million years. CI denotes 95% HPD

Splits within Cebidae	Mean	CI
Saimiri - Cebus + Sapajus	18.33	16.81 - 19.95
Cebus - Sapajus	6.19	5.20 - 7.21
Saimiri oerstedii + S. ustus - S. sciureus + S. macrodon + S. cassiquiarensis	1.75	1.35 - 2.15
Saimiri oerstedii - S. ustus	1.45	1.12 - 1.79
Saimiri sciureus - S. macrodon + S. cassiquiarensis	1.34	1.02 - 1.66
Saimiri macrodon - S. cassiquiarensis	1.17	0.89 - 1.46
Cebus olivaceus - C. albifrons + C. unicolor	1.15	0.92 - 1.40
Cebus albifrons - C. unicolor	0.65	0.51 - 0.80
Sapajus apella - S. macrocephalus	1.20	0.94 - 1.48

Table S 25
Split times in million years. CI denotes 95% HPD

Splits within Aotidae	Mean	CI
Aotus vociferans + A. griseimembra - A. trivirgatus + A. azarae	3.99	3.21 - 4.79
Aotus vociferans - A. griseimembra	2.80	2.23 - 3.38
Aotus trivirgatus - A. azarae	3.04	2.44 - 3.67

Table S 26

Split times in million years. CI denotes 95% HPD

Splits within Callithrichidae	Mean	CI
Leontocebus + Saguinus – Leontopithecus + Callimico + Callithrix + Cebuella + Callibella + Mico	13.16	11.84 - 14.54
Leontocebus - Saguinus	5.43	4.67 - 6.23
Leontopithecus - Callimico + Callithrix + Cebuella + Callibella + Mico	11.92	10.71 - 13.23
Callimico - Callithrix + Cebuella + Callibella + Mico	10.45	9.32 - 11.60
Callithrix - Cebuella + Callibella + Mico	4.04	3.46 - 4.64
Cebuella - Callibella + Mico	3.48	2.97 - 4.01
Callibella - Mico	2.08	1.75 - 2.41
Leontocebus nigricollis - L. fuscicollis + L. illigeri	1.32	1.08 - 1.57
Leontocebus fuscicollis - L. illigeri	0.79	0.64 - 0.95
Saguinus imperator + S. inustus + S. labiatus + S. mystax – S. midas + S. bicolor + S. oedipus + S. geoffroyi	4.63	3.96 - 5.32
Saguinus imperator - S. inustus + S. labiatus + S. mystax	2.10	1.77 - 2.46
Saguinus inustus - S. labiatus + S. mystax	1.51	1.27 - 1.78
Saguinus labiatus - S. mystax	1.17	0.97 - 1.38
Saguinus midas + S. bicolor - S. oedipus + S. geoffroyi	3.25	2.77 - 3.76
Saguinus midas - S. bicolor	1.53	1.28 - 1.79
Saguinus oedipus - S. geoffroyi	0.54	0.44 - 0.65
Lentopithecus rosalia - L. chrysomelas	0.73	0.57 - 0.89
Callithrix geoffroyi - C. jacchus + C. kuhlii	0.66	0.54 - 0.80
Callithrix jacchus - C. kuhlii	0.46	0.37 - 0.56
Cebuella pygmaea - C. niveiventris	1.47	1.23 - 1.71
Mico humeralifer - M. argentatus + Mico sp. nov.	0.83	0.69 - 0.98
Mico argentatus - Mico schneideri	0.47	0.38 - 0.56

Table S 27

Split times in million years. CI denotes 95% HPD

Splits within Cercopithecidae	Mean	CI
Cercopithecinae - Colobinae	19.30	18.04 - 20.58
Cercopithecini - Papionini	14.92	13.96 - 15.92
Allenopithecus + Erythrocebus + Chlorocebus + Allochocebus – Miopithecus + Cercopithecus	11.70	10.96 - 12.54
Allenopithecus - Erythrocebus + Chlorocebus + Allochocebus	10.66	9.95 - 11.44
Erythrocebus - Chlorocebus + Allochocebus	6.35	5.75 - 6.97
Chlorocebus - Allochocebus	5.84	5.27 - 6.43
Miopithecus - Cercopithecus	10.41	9.72 - 11.19
Macaca - Mandrillus + Cercocebus + Theropithecus + Lophocebus + Papio	11.20	10.39 - 12.08
Mandrillus + Cercocebus - Theropithecus + Lophocebus + Papio	9.29	8.56 - 10.11
Mandrillus - Cercocebus	8.07	7.39 - 8.81
Theropithecus - Lophocebus + Papio	6.29	5.69 - 6.94
Lophocebus - Papio	5.42	4.86 - 6.00
Colobini - Presbytini	13.51	12.36 - 14.57
Colobus - Piliocolobus	11.11	10.07 - 12.04
Presbytis - Nasalis + Rhinopithecus + Pygathrix + Semnopithecus + Trachypithecus	8.98	7.93 - 9.89
Nasalis + Rhinopithecus + Pygathrix - Semnopithecus + Trachypithecus	8.86	7.82 - 9.76
Nasalis - Rhinopithecus + Pygathrix	7.50	6.57 - 8.28
Rhinopithecus - Pygathrix	7.25	6.37 - 8.03
Semnopithecus - Trachypithecus	4.51	3.89 - 5.10
Allochocebus solatus - A. lhoesti + A. preussi	2.53	2.20 - 2.89
Allochocebus lhoesti - A. preussi	1.83	1.57 - 2.11
Cercopithecus hamlyni - other Cercopithecus sp.	7.71	7.11 - 8.36
Cercopithecus diana + C. roloway - other Cercopithecus sp.	6.80	6.23 - 7.39
Cercopithecus diana - C. roloway	2.46	2.12 - 2.82
Cercopithecus neglectus + C. pogonias + C. mona + C. lowei – C. mitis + C. nictitans + C. petaurista + C. ascanius + C. cephus	6.47	5.92 - 7.04
Cercopithecus neglectus - C. pogonias + C. mona + C. lowei	5.99	5.47 - 6.53
Cercopithecus pogonias - C. mona + C. lowei	3.39	3.01 - 3.79
Cercopithecus mona - C. lowei	1.93	1.66 - 2.21
Cercopithecus mitis + C. nictitans - C. petaurista + C. ascanius + C. cephus	4.79	4.33 - 5.26
Cercopithecus mitis - C. nictitans	2.60	2.28 - 2.91
Cercopithecus petaurista - C. ascanius + C. cephus	3.88	3.48 - 4.30
Cercopithecus ascanius - C. cephus	2.69	2.34 - 3.04
Macaca nigra + M. maura + M. tonkeana + M. siberu + M. nemestrina + M. leonina + M. silenus – other Macaca sp.	5.49	4.89 - 6.11
Macaca nigra + M. maura + M. tonkeana – M. siberu + M. nemestrina + M. leonina + M. silenus	3.83	3.37 - 4.31
Macaca nigra - M. maura + M. tonkeana	2.08	1.79 - 2.40
Macaca maura - M. tonkeana	1.63	1.37 - 1.89
Macaca siberu + M. nemestrina - M. leonina + M. silenus	2.66	2.31 - 3.02
Macaca siberu - M. nemestrina	2.49	2.15 - 2.84

Macaca leonina - M. silenus	1.21	1.01 - 1.41
Macaca fascicularis + M. fuscata + M. mulatta + M. cyclopis – M. assamensis + M. arctoides + M. thibetana + M. radiata	5.12	4.56 - 5.72
Macaca fascicularis - M. fuscata + M. mulatta + M. cyclopis	3.66	3.21 - 4.13
Macaca fuscata - M. mulatta + M. cyclopis	2.33	2.00 - 2.66
Macaca mulatta - M. cyclopis	1.80	1.53 - 2.08
Macaca assamensis - M. arctoides + M. thibetana + M. radiata	3.31	2.90 - 3.74
Macaca arctoides - M. thibetana + M. radiata	3.21	2.80 - 3.62
Macaca thibetana - M. radiata	2.83	2.46 - 3.20
Mandrillus sphinx - M. leucophaeus	2.72	2.33 - 3.14
Cercocebus chrysogaster - C. torquatus + C. lunulatus	3.70	3.25 - 4.18
Cercocebus torquatus - C. lunulatus	1.72	1.46 - 1.99
Papio cynocephalus + P. kindae + P. ursinus – P. hamadryas + P. anubis + P. papio	2.44	2.10 - 2.78
Papio cynocephalus - P. kindae + P. ursinus	1.90	1.61 - 2.19
Papio kindae - P. ursinus	1.68	1.42 - 1.95
Papio hamadryas - P. anubis + P. papio	1.74	1.48 - 2.00
Papio anubis - P. papio	1.27	1.06 - 1.47
Colobus angolensis - C. guereza + C. polykomos	3.39	2.88 - 3.93
Colobus guereza - C. polykomos	1.92	1.60 - 2.26
Piliocolobus badius - P. tephrosceles + P. kirkii + P. gordonorum	3.44	2.92 - 3.98
Piliocolobus tephrosceles - P. kirkii + P. gordonorum	1.04	0.85 - 1.23
Piliocolobus kirkii - P. gordonorum	0.55	0.44 - 0.66
Presbytis comata - P. mitrata	1.00	0.80 - 1.22
Rhinopithecus roxellana - R. bieti	2.13	1.78 - 2.49
Pygathrix nigripes - P. nemaesus + P. cinerea	1.71	1.42 - 2.01
Pygathrix nemaesus - P. cinerea	0.59	0.47 - 0.72
Semnopithecus vetulus - S. johnii + S. hypoleucos + S. priam + S. entellus + S. schistaceus	2.90	2.46 - 3.32
Semnopithecus johnii - S. hypoleucos + S. priam + S. entellus + S. schistaceus	2.24	1.89 - 2.58
Semnopithecus hypoleucos + S. priam - S. entellus + S. schistaceus	1.57	1.32 - 1.83
Semnopithecus hypoleucos - S. priam	1.37	1.14 - 1.59
Semnopithecus entellus - S. schistaceus	0.58	0.47 - 0.69
Trachypithecus pileatus + T. geei - other Trachypithecus sp.	3.56	3.05 - 4.05
Trachypithecus pileatus - T. geei	0.65	0.52 - 0.78
Trachypithecus obscurus + T. crepusculus + T. melamerus + T. phayrei – other Trachypithecus sp.	2.59	2.20 - 2.96
Trachypithecus obscurus - T. crepusculus + T. melamerus + T. phayrei	1.97	1.66 - 2.27
Trachypithecus crepusculus - T. melamerus + T. phayrei	1.61	1.35 - 1.87
Trachypithecus melamerus - T. phayrei	1.18	0.98 - 1.38
Trachypithecus germaini + T. cristatus + T. auratus – T. francoisi + T. leucocephalus + T. laotum + T. hatinhensis	2.41	2.05 - 2.76

Trachypithecus germaini - T. cristatus + T. auratus	1.29	1.08 - 1.51
Trachypithecus cristatus - T. auratus	0.33	0.26 - 0.40
Trachypithecus francoisi + T. leucocephalus - T. laotum + T. hatinhensis	0.76	0.63 - 0.89
Trachypithecus francoisi - T. leucocephalus	0.35	0.28 - 0.43
Trachypithecus laotum - T. hatinhensis	0.40	0.32 - 0.48

Table S 28

Split times in million years. CI denotes 95% HPD

Splits within Hylobatidae	Mean	CI
Nomascus - Hylobates + Hoolock	8.62	7.19 - 10.62
Hylobates - Hoolock	8.04	6.65 - 9.90
Nomascus concolor - N. siki + N. annamensis + N. gabriellae	1.84	1.42 - 2.36
Nomascus siki - N. annamensis + N. gabriellae	1.55	1.19 - 1.99
Nomascus annamensis - N. gabriellae	0.81	0.61 - 1.04
Hylobates lar - H. klossii + H. agilis + H. abbotti + H. muelleri	3.23	2.58 - 4.10
Hylobates klossii - H. agilis + H. abbotti + H. muelleri	2.24	1.77 - 2.84
Hylobates agilis - H. abbotti + H. muelleri	1.68	1.31 - 2.15
Hylobates abbotti - H. muelleri	1.39	1.08 - 1.78

Table S 29

Split times in million years. CI denotes 95% HPD

Splits within Hominidae	Mean	CI
Pongo - Gorilla + Pan + Homo	20.32	18.58 - 22.19
Gorilla - Pan + Homo	10.12	8.79 - 11.24
Pan - Homo	8.01	6.91 - 8.96
Pongo pygmaeus - P. abelii	1.55	1.18 - 1.95
Gorilla gorilla - G. beringei	1.03	0.81 - 1.26
Pan troglodytes - P. paniscus	2.39	1.98 - 2.81

Table S 30

Split times in million years. CI denotes 95% HPD

Data S1. (separate file)

Tabular sample level metadata for all samples included in this study.

Data S2. (separate file)

Tabular species level metadata for all samples that did not fail QC as annotated in Data S1.

Data S3. (separate file)

Newick tree: UCE-based species tree with branch lengths as substitutions per site

Data S4. (separate file)

Newick tree: Fossil calibrated time tree

Data S5. (separate file)

Newick tree: Mitochondrial phylogeny

Data S6. (separate file)

Spreadsheet with list of variants specific to the human lineage, and list of non-recurring variants specific to the ape lineage