PRIMATE GENOMES

Phylogenomic analyses provide insights into primate evolution

Yong Shao[1]†, Long Zhou[2]†, Fang Li[3,4], Lan Zhao[5], Bao-Lin Zhang[1], Feng Shao[6], Jia-Wei Chen[7], Chun-Yan Chen[8], Xupeng Bi[2], Xiao-Lin Zhuang[1,9], Hong-Liang Zhu[7], Jiang Hu[10], Zongyi Sun[10], Xin Li[10], Depeng Wang[10], Iker Rivas-González[11], Sheng Wang[1], Yun-Mei Wang[1], Wu Chen[12], Gang Li[13], Hui-Meng Lu[14], Yang Liu[13], Lukas F. K. Kuderna[15,16], Kyle Kai-How Farh[16], Peng-Fei Fan[17], Li Yu[18], Ming Li[19], Zhi-Jin Liu[20], George P. Tiley[21], Anne D. Yoder[21], Christian Roos[22], Takashi Hayakawa[23,24], Tomas Marques-Bonet[15,25,33,34], Jeffrey Rogers[26], Peter D. Stenson[27], David N. Cooper[27], Mikkel Heide Schierup[11], Yong-Gang Yao[9,28,29,30], Ya-Ping Zhang[1,29,30], Wen Wang[1,8,29], Xiao-Guang Qi[5]*, Guojie Zhang[1,2,3,31]*, Dong-Dong Wu[1,29,30,32]*

Comparative analysis of primate genomes within a phylogenetic context is essential for understanding the evolution of human genetic architecture and primate diversity. We present such a study of 50 primate species spanning 38 genera and 14 families, including 27 genomes first reported here, with many from previously less well represented groups, the New World monkeys and the Strepsirrhini. Our analyses reveal heterogeneous rates of genomic rearrangement and gene evolution across primate lineages. Thousands of genes under positive selection in different lineages play roles in the nervous, skeletal, and digestive systems and may have contributed to primate innovations and adaptations. Our study reveals that many key genomic innovations occurred in the Simiiformes ancestral node and may have had an impact on the adaptive radiation of the Simiiformes and human evolution.

The order Primate contains >500 species from 79 genera and 16 families (*1*), with new species continuing to be discovered (*2–5*), making primates the third most speciose order of living mammals after bats (Chiroptera) and rodents (Rodentia). As our closest living relatives, nonhuman primates play important roles in the cultures and religions of human societies (*1*). Many nonhuman primate species have been widely used as animal models because of their genetic, physiological, and anatomical similarities to humans, allowing the efficacy and safety of newly developed drugs and vaccines to be tested (*6*). For example, since the emergence of COVID-19, macaques have served as important models in the research and development of vaccines (*7–16*). Primates display considerable morphological, behavioral, and physiological diversity and hold the key to understanding the evolution of our own species, particularly the evolution of human phenotypes such as high-level cognition (*17, 18*).

Nonhuman primates occupy a wide range of diverse habitats in the tropical forest, savanna, semidesert, and subtropical regions of Asia, Central and South America, and Africa, and humans have spread across much of the earth's surface. Nevertheless, according to the International Union for Conservation of Nature (IUCN) Red Lists, >33% of primate species are critically endangered or vulnerable, ~60% are threatened with extinction, and ~75% are experiencing population decline (*1*). With global climate change and increasing anthropogenic interference, the conservation status of primates has attracted global scientific and public awareness.

Despite the importance of nonhuman primates, reference genomes have been sequenced in <10% of species (*19–27*), which both impedes research and hampers conservation efforts. Here, we present high-quality reference genomes for 27 primate species with long-read sequencing generated from our first-phase program of the Primate Genome Project.

Assembly and annotation of 27 new primate reference genomes

We applied long-read genome-sequencing technologies, including Pacbio and Nanopore, to sequence the genomes of 27 nonhuman primate species from 26 genera of 11 families (table S1). Long reads were self-polished and assembled, and the genome assemblies were further corrected and polished by paired-end short reads sequenced from the same individuals (tables S2 to S4). We also used sequencing data generated by high-throughput chromosome conformation capture technology (*28*) to anchor assembled contigs into chromosomes for four species (fig. S1 and table S4). The sizes of the new genome assemblies of the primate species under study ranged from ~2.4×10^9 base pairs (Gbp) (*Daubentonia madagascariensis*) to ~3.1 Gbp (*Erythrocebus patas*), which were mostly consistent with the k-mer–based esimations (fig. S2 and table S5), with a high average contig N50 length of ~15.9×10^6 base pairs (Mbp) (table S6). All of the genome assemblies yielded BUSCO complete scores >92% (table S6). A method that integrates de novo and homology-based strategies was applied to annotate all genomes with protein sequences from human, chimpanzee, gorilla, orangutan, and mouse as references for homology-based gene model prediction. Between 20,066 and 21,468 protein-coding genes were predicted in these genome assemblies (table S7). Further, we also identified ~24.2 Mbp of primate-specific highly conserved elements by using whole-genome alignments between all primates and nine other mammals (fig. S3).

[1]State Key Laboratory of Genetic Resources and Evolution, Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China. [2]Center of Evolutionary & Organismal Biology, and Women's Hospital at Zhejiang University School of Medicine, Hangzhou 310058, China. [3]Section for Ecology and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. [4]Institute of Animal Sex and Development, ZhejiangWanli University, Ningbo 315100, China. [5]Shaanxi Key Laboratory for Animal Conservation, College of Life Sciences, Northwest University, Xi'an 710069, China. [6]Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education), Southwest University School of Life Sciences, Chongqing 400715, China. [7]BGI-Shenzhen, Shenzhen 518083, China. [8]School of Ecology and Environment, Northwestern Polytechnical University, Xi'an 710072, China. [9]Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming 650204, China. [10]Grandomics Biosciences, Beijing 102206, China. [11]Bioinformatics Research Centre, Aarhus University, DK-8000 Aarhus, Denmark. [12]Guangzhou Zoo & Guangzhou Wildlife Research Center, Guangzhou 510070, China. [13]College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China. [14]School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, China. [15]Institute of Evolutionary Biology (UPF-CSIC), PRBB, 08003 Barcelona, Spain. [16]Illumina Artificial Intelligence Laboratory, Illumina Inc, San Diego, CA 92122, USA. [17]School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275, China. [18]State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, School of Life Sciences, Yunnan University, Kunming 650091, China. [19]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. [20]College of Life Sciences, Capital Normal University, Beijing 100048, China. [21]Department of Biology, Duke University, Durham, NC 27708, USA. [22]Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany. [23]Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan. [24]Japan Monkey Centre, Inuyama, Aichi 484-0081, Japan. [25]Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain. [26]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [27]Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK. [28]Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China. [29]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650201, China. [30]National Resource Center for Non-Human Primates, Kunming Primate Research Center, and National Research Facility for Phenotypic & Genetic Analysis of Model Animals (Primate Facility), Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650107, China. [31]Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou 311121, China. [32]KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650204, China. [33]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain. [34]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain.
*Corresponding author. Email: wudongdong@mail.kiz.ac.cn (D.-D.W.); guojiezhang@zju.edu.cn (G.Z.); qixg@nwu.edu.cn (X.-G.Q.)   †These authors contributed equally to this work.

The Primate Genome Project also generated high-quality reference genomes for another 16 primate species that were used in the accompanying papers to reveal hybrid speciation during the rapid radiation of the macaques (*29*), the homoploid hybrid speciation in the snub-nosed monkey *Rhinopithecus* genus (*30*), social evolution in the Asian colobines driven by cold adaptation (*31*), and the evolutionary adaptations of slow lorises (*32*). All genomic data have been published openly and can be freely accessed in the National Center for Biotechnology Information (NCBI) Assembly Database under the accession information described in this study.

A genomic phylogeny of living primates

We next performed phylogenomic analyses comprising the 27 newly generated genomes, another 22 published primate genomes, one long-read genome from *Nycticebus pygmaeus* reported in an accompanying paper (*32*), and two close relatives of primates, the Sunda flying lemur (*Galeopterus variegatus*) and the Chinese tree shrew (*Tupaia belangeri chinensis*) (*33*), as outgroups (table S8). We constructed whole-genome–wide phylogenetic trees using ExaML under a GTR+GAMMA model (*34*). Altogether, ~433.5 Mbp of gap-free data for syntenic orthologous sequences were retrieved
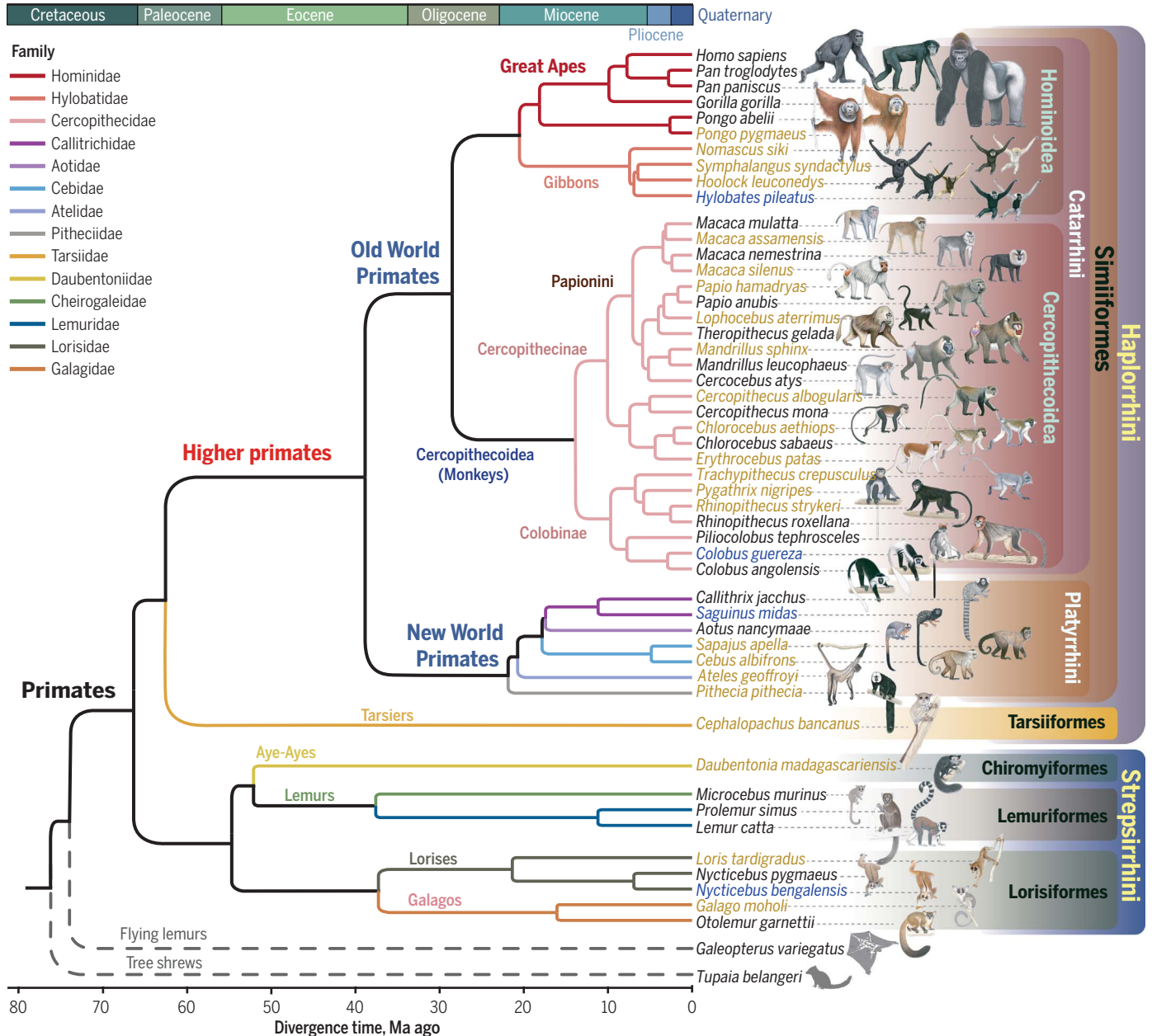

Fig. 1. Genomic phylogeny of primates. The maximum likelihood method was used to infer the primate species tree from whole-genome sequences across 52 species, including 50 primate species and two outgroup species (the Sunda flying lemur and the Chinese tree shrew) with 100 bootstraps under a GTR+GAMMA model. The divergence time was estimated using fossil calibrations (fig. S11) and the MCMCtree algorithm. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level. The genomes of the species marked in black were downloaded from the NCBI and Ensembl databases (table S8). Monkey pictures are copyrighted by Stephen D. Nash/IUCN/SSC Primate Specialist Group and are used in this study with their permission.
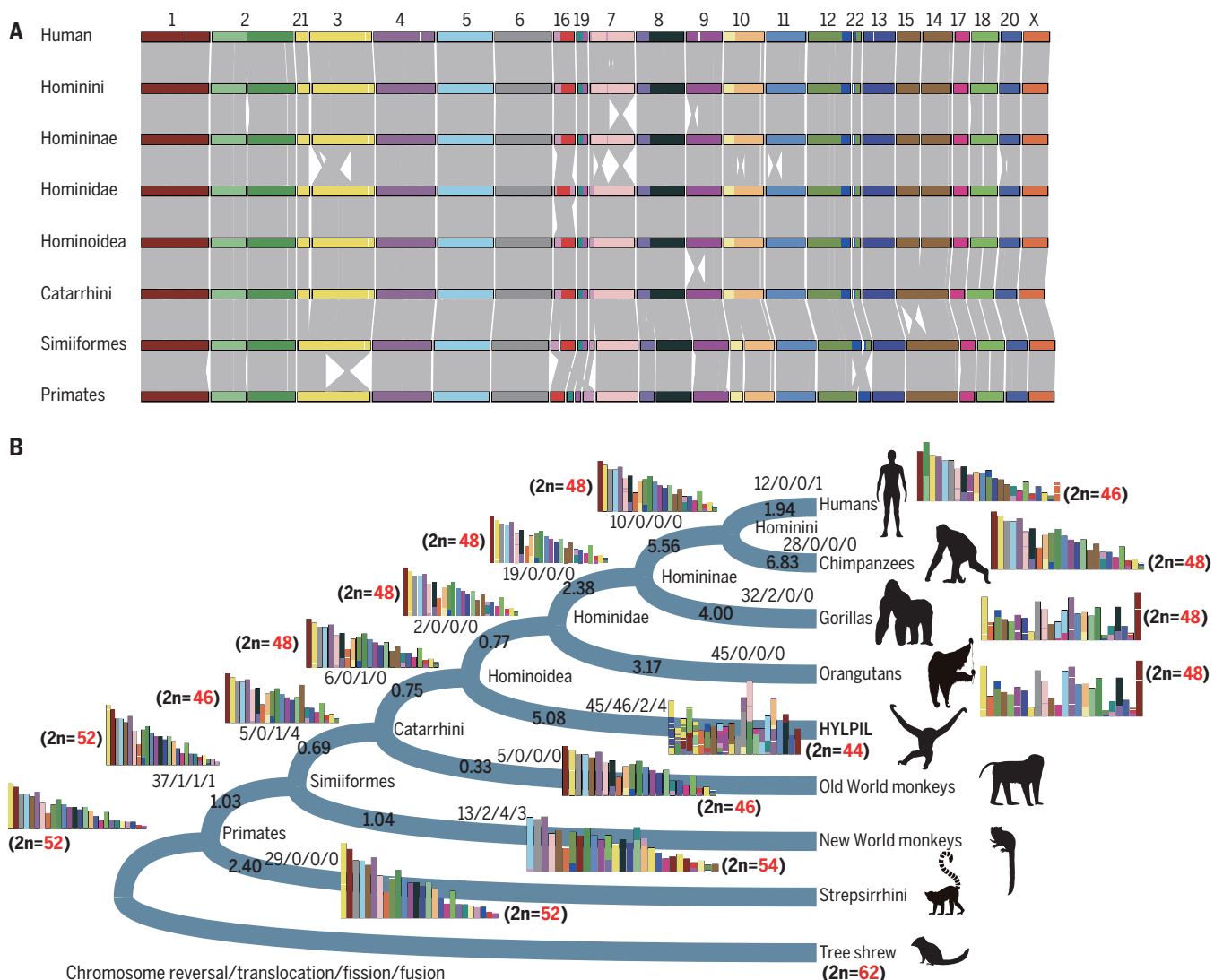
Fig. 2. Reconstruction of primate ancestral chromosomes. (**A**) Chromosome evolution patterns from the primate common ancestral lineage leading to the human lineage. Chromosomes are colored on the basis of human homologies. (**B**) Karyotype evolution and genome rearrangement. The rates of genomic rearrangement are highlighted in black bold font. Chromosome variations from ancestral nodes to derived branches are shown by pathways including chromosome reversal, translocation, and fission and fusion events, which are shown by number, e.g., reversal, translocation, fission, and fusion. "HYLPIL" represents the gibbon *Hylobates pileatus*, the genome of which was assembled at the chromosome level.

from the whole-genome alignments (table S9) and used to infer the primate phylogeny, yielding a high-resolution whole-genome nucleotide evidence tree with identical topology to a previous tree derived from 54 nuclear gene regions from 186 living primates (*35*). This tree has 100% bootstrap support for all evolutionary nodes, with the exception of the node ((*Symphalangus syndactylus*, *Hoolock leuconedys*), *Hylobates pileatus*) among gibbon genera with 90% bootstrap support (Fig. 1 and figs. S4 and S5). The evolution of gibbons has been characterized by their rapid karyotypic changes and remains controversial in primate phylogeny at the genus level (*24*, *35*, *36*). To confirm the phylogeny of this node, we also

generated partitioned trees with orthologous protein-coding genes, exon codons with first and second positions, fourfold degenerate sites, and conserved nonexonic elements (figs. S6 to S9). The tree from conserved nonexonic elements yielded the identical topologies for the gibbon lineages with the whole-genome nucleotide evidence trees (fig. S9). However, the trees from orthologous protein-coding genes and exon codons with first and second positions and fourfold degenerate sites, respectively, supported the alternative topologies, ((*Nomascus*, *Hylobates*), (*Symphalangus*, *Hoolock*)) and ((*Nomascus*, (*Symphalangus*, *Hoolock*)), *Hylobates*) (figs. S6 to S8). The two topologies were shown in previous studies based on var-

iants called by mapping short reads to the reference genome of *Nomascus leucogenys* (*24*, *36*).

Our analyses again confirmed the phylogenetic challenge within the gibbon lineage, which has experienced pronounced adaptive radiation within an extremely short evolutionary time period (*24*, *35*). Consistently, we observed extremely short internal branches in this lineage on the phylogeny. A comparative analysis using CoalHMM (*37*) across primate lineages showed that the gibbon lineage represents one of the lineages with the highest frequency of incomplete lineage sorting (*38*), supporting a previous study based on population data (*24*). Specifically, the two gibbon branches showed incomplete lineage-sorting

proportions of 57 and 61%, respectively, but the species topology inferred from incomplete lineage-sorting analyses was identical to those presented herein (figs. S4 and S10).

Using the whole-genome nucleotide evidence tree and fossil calibration data (*35, 39*) (Fig. 1 and fig. S11), the divergence dating of living primates was estimated by means of the MCMCtree algorithm (*40*) (Fig. 1 and fig. S12). We estimated that the most recent common ancestor of all primates evolved between 64.95 and 68.29 million years (Ma) ago, which is close to the estimate given in the latest phylogenetic study across mammals (*41*), suggesting that the origin of the primate group was near the Cretaceous–Tertiary boundary at 66 Ma ago. We also estimated that the most recent common ancestor of Strepsirrhini appeared between 52.57 and 56.56 Ma ago, and that of the Simiiformes emerged between 35.65 and 42.55 Ma ago (Fig. 1 and fig. S12).

Genomic structure and evolution of primates
Karyotype evolution and genome rearrangement

The speciation process is often accompanied by karyotypic evolution, which also affects genome evolution and gene function (*42–44*). We reconstructed the ancestral karyotype evolutionary process across primate lineages (table S10) and observed an overall conserved pattern of chromosome-level synteny (Fig. 2A). The numbers of ancestral karyotypes of Catarrhini ($2n = 46$) and Hominoidea ($2n = 48$) were consistent with previous inferences derived from the fluorescence in situ hybridization data of bacterial artificial chromosomes (*45*) (Fig. 2A). However, we deduced that both of the ancestral karyotypes of primates and Simiiformes had a diploid number of $2n = 52$ (Fig. 2A) rather than $2n = 50$ as previously suggested (*45*), recovering a fission event in chromosome 8 that was observed in the common ancestor of primates (Fig. 2A and fig. S13). Fusion and fission are the most common mechanisms of karyotype evolution in primates, as exemplified by the fusion of chromosome 2, which occurred specifically in the human lineage (*45*). Our analyses further identified at least one fission and one fusion during the emergence of the Simiiformes, as well as one fission and four fusions associated with the Catarrhini node (Fig. 2B and fig. S13), resulting in the contemporary karyotype structure of our own. The rapid change of karyotypes in the Simiiformes also led to an increased chromosome number in New World monkeys, which have the largest number of chromosomes across primates. We further estimated the rate of genome rearrangement by taking into account all large-scale genomic rearrangement events, including reversions, translocations, fusions, and fissions, in key evolutionary nodes from the primate common ancestral lineage leading to the human lineage. We observed an increasing rate of

rearrangement in the Homininae (*Gorilla-Homo-Pan*) (~2.38/Ma) and particularly in the Hominini (*Homo-Pan*) (~5.56/Ma) (Fig. 2B), which contradicts the Hominini slowdown hypothesis on the nucleotide substitution rates (*35*).

Lineage-specific segmental duplication

We next compiled segmental duplication maps (segmental duplication length ≥5 kbp) for primates and five outgroup species (fig. S14 and table S11). Compared with other primate lineages, we observed a marked increase in the number of lineage-specific segmental duplications ($n = 221$) in the great ape genomes (Fig. 3A and table S12), consistent with previous findings describing a burst of segmental duplications in the great ape ancestor (*46*). These specific segmental duplications in great apes overlapped with 57 protein-coding genes (table S13), 20 of which were highly expressed in the human brain (fig. S15). We also observed lineage-specific segmental duplications in other primate groups producing lineage-specific new genes that might have contributed to the evolution of these lineages (table S13). We further explored the functions of all genes overlapping segmental duplications in primate genomes (table S13) against the Human Gene Mutation Database (*47*), and found that a high proportion of these genes (52.8%) have been reported to be associated with inherited conditions including autism, intellectual disabilities, and other developmental disorders (Fig. 3B and table S14).

Evolution of genome size and transposable elements

Compared with other mammalian groups, the primates on average have a relatively large genome size (*48, 49*). Among primates, the lemurs (Lemuriformes and Chiromyiformes) were found to be characterized by a significantly smaller genome size (~2.36 Gbp) than other groups such as the lorisoids (Lorisiformes: Lorisdae and Galagidae, ~2.70 Gbp), New World monkeys (~2.82 Gbp), Old World monkeys (~2.91 Gbp), and Hominoidea (~2.96 Gbp) ($P < 0.05$, Mann-Whitney U test) (fig. S16). The increase of genome size in the Simiiformes can be attributed to the expansion of transposable elements (figs. S16 to S18 and table S15), especially *Alu* elements, ~300 nucleotide short interspersed sequence elements (SINEs) that make up ~11% of the human genome (*50–54*). We observed that the genomes of lemurs exhibited a relative paucity of SINEs, especially *Alu* (~3.87%), which is less than one-third of the proportion noted in other lineages (figs. S16 to S18). By contrast, the *Alu* elements in both Simiiformes and Lorisiformes experienced major bursts of retrotranspositional activity at ~40 to 45 and ~34 to 39 Ma ago independently (fig. S19). Specifically, we noticed a

substantial expansion of the *AluS*-related subclasses, especially *AluSx* in the Simiiformes, whereas the *AluJ*-related subclasses (especially *AluJb*) were the dominant subclasses of *Alu* in the Lorisiformes (fig. S20).

Variation in the nucleotide substitution rate

We estimated the overall nucleotide substitution rate in primates to be $\sim 1.1 \times 10^{-3}$ substitutions per site per million years (Fig. 3C, fig. S21, and table S16), which is much lower than the average rate for mammals ($\sim 2.7 \times 10^{-3}$) and birds ($\sim 1.9 \times 10^{-3}$) (*55*). However, the nucleotide substitution rate exhibited a high degree of heterogeneity between primate lineages, potentially caused by differences with respect to life history traits (*56–58*). The New World monkeys evolved the fastest at $\sim 1.4 \times 10^{-3}$ substitutions per site per million years (Fig. 3C and fig. S21). We confirmed the hominoid "slowdown" (*35, 59–61*) hypothesis by detecting a reduced substitution rate in hominoids ($\sim 0.8 \times 10^{-3}$ substitutions per site per million years) (fig. S21). Our analysis and a previous study (*62*) suggested that tarsiers, as the most basal haplorrhines, potentially evolved with a rapid substitution rate compared with other primates (fig. S21).

Evolution of protein-coding genes

We obtained a high-confidence orthologous gene set comprising 10,185 orthologs across 50 primate species, along with the Sunda flying lemur and the Chinese tree shrew. On the basis of the whole-genome nucleotide evidence tree topology of primates, we calculated the ratio of the rates of nonsynonymous (d_N) to synonymous (d_S) substitutions for each ortholog to explore the evolutionary constraints operating on coding regions. We estimated the evolutionary rate of tissue-specific expressed genes for different tissues across evolutionary clades in primates based on the observation that tissue-specific expressed genes are generally conserved across diverse species (*63, 64*), and observed that testis- and spleen-specific expressed genes generally displayed higher values of d_N/d_S (Fig. 3D and figs. S22 and S23) than other tissue-specific expressed genes, corroborating the rapid evolution of the reproductive and immune systems in primates (*65, 66*). By contrast, brain-specific expressed genes generally showed a high degree of conservation with lower d_N/d_S values, as previously reported, despite the rapid evolution of primate cognitive functions (*67*).

Next, we detected 82 positively selected genes in the common ancestral lineage of primates by comparison with other mammalian species (table S17) using the codeml algorithm under the branch-site model with a likelihood rate test in PAML4 (*40, 68*). We found that these positively selected genes were significantly enriched in genes exhibiting high-level

expression in brain, bone marrow, and testis (table S18). In particular, close to 37% (30 genes) of positively selected genes exhibited biased expression in the brain (tables S18 and S19), and we found that some of them (e.g., *SPTAN1*, *MYT1L*, and *SHMT1*) could have important roles in brain function, because deleterious mutations of these genes have been reported to cause brain disorders (*69–71*) such as epilepsy and schizophrenia. These genes may be important candidates for involvement in the evolution of the primate brain because of their functional importance. Our results suggest that some positively selected genes in the primate

ancestral lineage may have been involved in the rapid evolution of their brain functions despite the general conservation of brain-specific expressed genes. In addition, several immune-related genes (e.g., *XRCC6* and *CD2*) (table S17) also experienced positive selection in the primate ancestor, suggesting that the adaptive immune system might also have contributed to primate evolution.

An increased level of genomic change in the ancestor of the Simiiformes

To provide new insights into the genetic underpinnings of primate phenotypic evolu-

tion, we performed various comparative genomic analyses, including the identification of positively selected genes, genes having conserved noncoding regions that have been subject to lineage-specific accelerated evolution (*72*), and expanded gene families in different primate lineages (*68*). An increased level of genomic evolutionary changes, as reflected by the high numbers of positively selected genes, lineage-specific accelerated regions, and expanded gene families, was observed in the Simiiformes ancestor (Fig. 4A). Consistently, the Simiiformes have also experienced rapid evolution of a series of complex traits, unlike
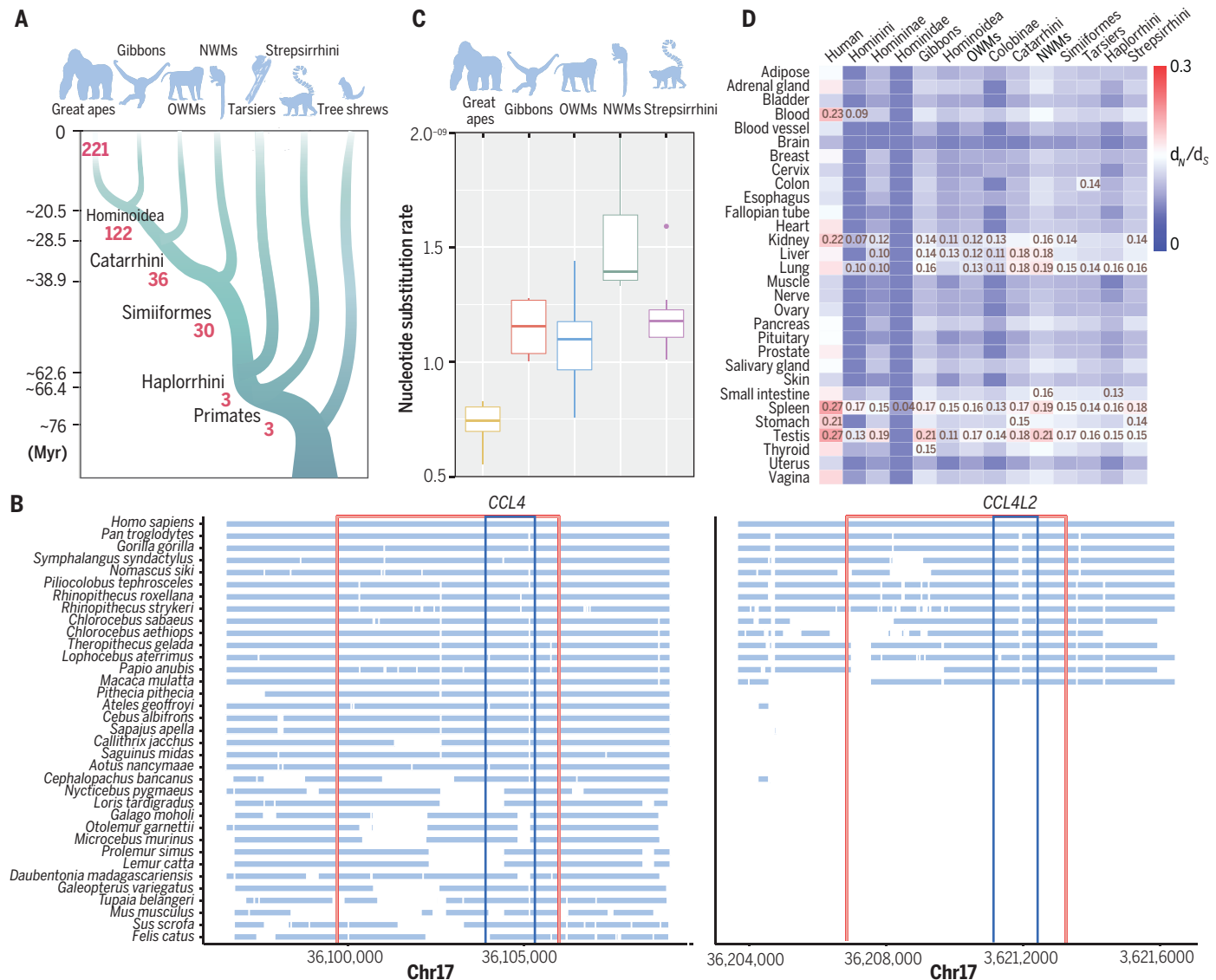

Fig. 3. Structural evolution in primate genomes. (**A**) Evolutionary pattern of lineage-specific segmental duplications in primates. The numbers of lineage-specific segmental duplications are shown in red. The largest number of segmental duplications was found in the great ape lineage. OWMs, Old World monkeys; NWMs, New World monkeys. (**B**) Example of specific segmental duplications during evolution of the genome in Catarrhini. A gene pair overlapping the segmental duplication (left, *CCL4*; right, *CCL4L2*) is associated

with HIV susceptibility. The red and green boxes represent the segmental duplication region and the overlapping gene pair, respectively. (**C**) Substitution rates across five evolutionary branches in primates. (**D**) Evolutionary constraints of tissues across diverse lineages in primates. The evolutionary constraints of tissues are shown by the d_N/d_S median of tissue-specific expressed genes in different evolutionary nodes among primates.

the Strepsirrhini and Tarsiiformes. For example, the Simiiformes generally exhibit a larger brain volume and body mass than the Strepsirrhini and Tarsiiformes (Fig. 4B) (73, 74). Functional enrichment analyses showed that the associated genes relevant to these rapid genomic changes in the Simiiformes ancestor (tables S20 to S22) were overrepresented in functions related to the nervous system and development, such as postsynaptic density, synapses, and the negative regulation of the canonical Wnt signaling pathway (table S23).

Additional analyses indicated that various candidate genes in the Simiiformes ancestral lineage, comprising 168 positively selected genes, 273 genes associated with lineage-specific ac-

celerated regions, and 14 expanded gene families, were enriched in central nervous system terms, i.e., brain, cerebrum, cerebellum, hippocampus, and cerebral cortex (table S24). More specifically, five genes participated in pathway axon guidance (Fig. 4C), being expressed in the human brain at a high level (table S25). Axon guidance represents a key stage in the formation of a neural network (75, 76) and may have been an important influence on brain volume. In this pathway, two semaphorin genes, *SEMA3B* and *SEMA3D*, which are critical for central nervous system patterning (77, 78), experienced positive selection and served as a gene associated with the lineage-specific accelerated region, respectively. These

two genes, together with another three genes associated with the lineage-specific accelerated regions, *EPHA3*, *RAC1*, and *NTNG2*, are known to be important for brain development (79–81). Furthermore, eight genes were assigned under the term "Hippo signaling pathway" (Fig. 4D), an evolutionarily conserved signaling pathway that controls organ or body size by regulating cell growth, proliferation, and apoptosis in a range of animals from flies to humans (82–84). Genes involved in neuronal network formation and the control of organ size appear to have undergone adaptive evolution in the Simiiformes ancestral lineage and may have been responsible for specific phenotypic changes, particularly the progressive
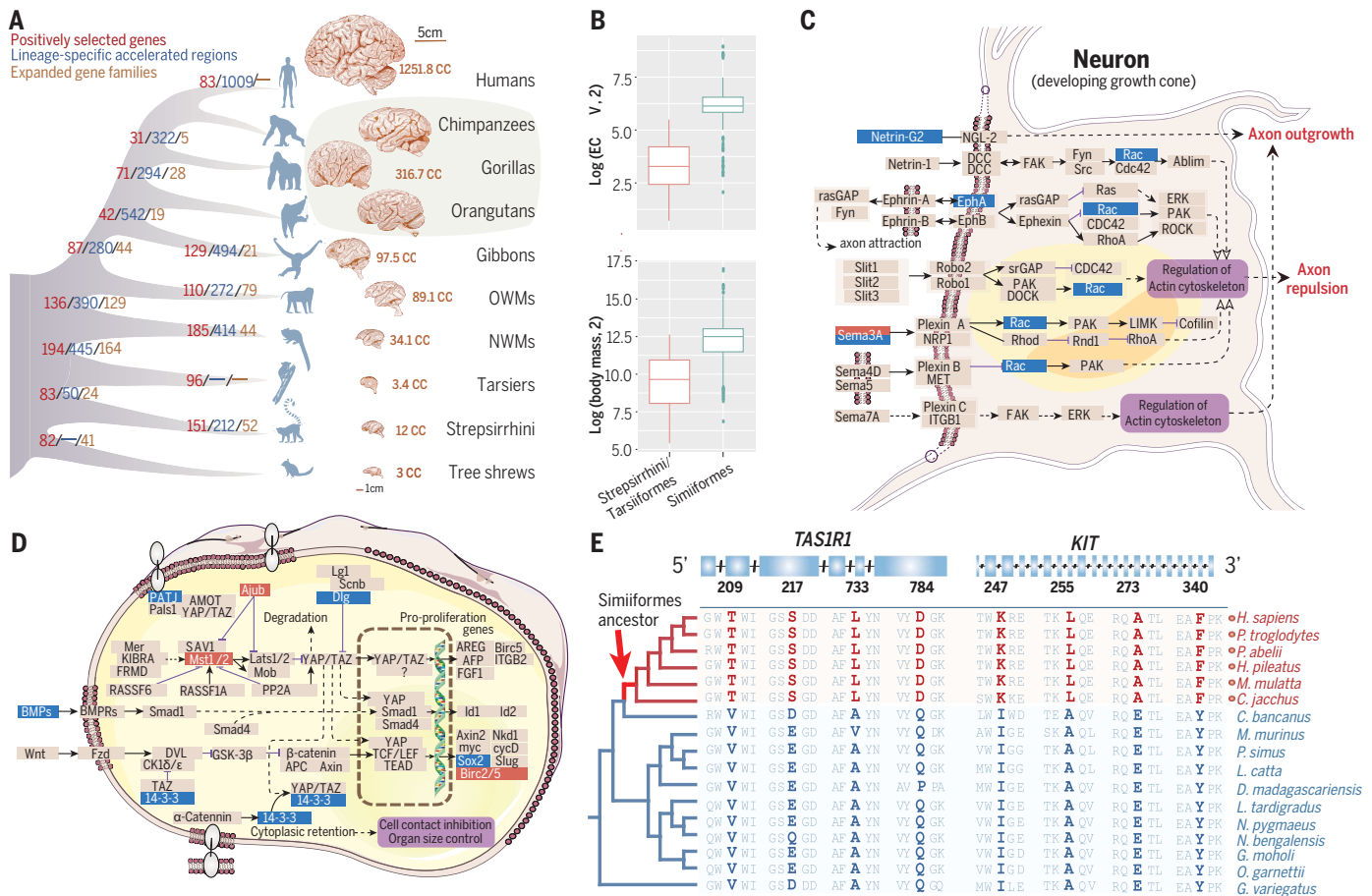

Fig. 4. Genomic changes and phenotype evolution in the ancestor of the Simiiformes. (**A**) Increased level of genomic evolutionary change, including positively selected genes, lineage-specific accelerated regions, and significantly expanded gene families, seen in the Simiiformes ancestral lineage. The brain sizes and brain structures are shown in representative evolutionary groups of primates. The brain sizes across primate and outgroup species are derived from previous studies (156, 157). Brain images are from the Michigan State University Comparative Mammalian Brain Collections (www.brainmuseum.org). (**B**) Representative phenotype variations, including brain size and body mass, between the Strepsirrhini and Tarsiiformes and the Simiiformes. Statistical significance was assessed by the Mann-Whitney *U* test as *P* < 0.05. (**C**) Candidate genes involved in the axon guidance KEGG pathway (hsa04360). Genes relating to genomic changes in the Simiiformes ancestral lineage are shown in this pathway. The

protein product of the positively selected gene in the Simiiformes ancestral lineage, *SEMA3B*, is shown in red. The protein products of genes associated with lineage-specific accelerated regions, *EPHA3*, *RAC1*, *NTNG2*, and *SEMA3D*, are shown in blue. (**D**) The Hippo signaling pathway (hsa04390), which is involved in organ size and body size, with candidates including positively selected genes and genes associated with lineage-specific accelerated regions. The gene products for positively selected genes (*LIMD1*, *BIRC3*, and *STK3*) in the Simiiformes ancestral lineage are shown in red, and the products of genes associated with lineage-specific accelerated regions (*PATJ*, *SOX2*, *BMP2*, *DLG2*, and *YWHAQ*) in the Simiiformes ancestral lineage are shown in blue. (**E**) Multiple sequence alignments of two positively selected genes, *TAS1R1* and *KIT*, along the Simiiformes ancestral lineage. The phylogenetic position of the Simiiformes ancestor is indicated by a red arrow.

increase in brain volumes and body sizes compared with the Tarsiiformes and Strepsirrhini.

A major phenotypic difference between the Strepsirrhini and Tarsiiformes and the Simiiformes is nocturnal versus diurnal life history. The visual system has diverged substantially between the Strepsirrhini and Tarsiiformes and the Simiiformes such that the diurnal Simiiformes have much smaller corneal sizes (relative to their eyes) and higher visual acuity than the Strepsirrhini and Tarsiiformes (85). Consistent with this phenotypic difference, we detected positive selection signals in three genes, *NPHP4*, *GRHL2*, and *SLC39A5*, which are associated with eye development (Gene Ontology identifier: 0001654) in the Simiiformes ancestral lineage. An intragenic deletion in *NPHP4* causes recessive cone-rod dystrophy with a predominant loss of cone function in the dachshund (86). *GRHL2* encodes a transcription factor that suppresses epithelial-to-mesenchymal transition; ectopic *GRHL2* expression caused by mutation accelerates cell state transition and leads to posterior polymorphous corneal dystrophy and vision function disruption (87). The *GRHL2* gene has the highest number of positively selected sites in the Simiiformes ancestor compared with the other genes involved in eye development (fig. S24). *TAS1R1* encodes a taste receptor that can form a heterodimer with TAS1R3 to elicit the umami taste (88). We found that *TAS1R1* also experienced positive selection with four positively selected sites in the Simiiformes ancestor (Fig. 4E). The rapid and concerted evolution of taste receptors and vision could have helped the diurnal Simiiformes to locate and identify food. The detailed functional consequences of these amino acid changes might be worthy of further study.

Compared with the Strepsirrhini and Tarsiiformes, the Simiiformes generally exhibit darker skin pigmentation and a less bright coat color (fig. S25) (89). We identified two pigmentation-related genes, *KIT* and *CREB3L4*, that participate in the melanogenesis pathway that evolved under positive selection (detected by the branch-site model) in the Simiiformes ancestor (Fig. 4E). Melanocytes play an important role during the formation of skin and coat colors in mammals by regulating melanin-related genes (90). *KIT*, a proto-oncogene, encodes a receptor tyrosine kinase that regulates cell migration, proliferation, and differentiation in melanocytes and plays a key role in melanin deposition (91, 92). *KIT* also communicates with *MITF*, a key gene in the formation of melanin that regulates the development of melanocytes (93–95).

Genetic mechanisms underlying primate phenotype evolution

Primates have evolved diverse phenotypic traits to adapt to their challenging environ-

ments. Here, we sought to investigate the evolution of complex phenotypes in the brain, skeletal system, digestive system, and sense organs, as well as body size, in primates.

Brain evolution

In primates, brain volumes range from $<\sim2\ cm^3$ in the mouse lemur to $\sim1300\ cm^3$ in human (73). To reveal the genetic changes that might underlie brain evolution in primates, we detected signals of positive selection in brain development genes using a branch-site model in PAML in key evolutionary nodes in the primate phylogeny. A total of 34 brain genes were found to be under positive selection in one of the primate evolutionary nodes (table S26) (68). Four of them, *SLC6A4*, *NR2E1*, *NIPBL*, and *XRCC6*, were under positive selection in the common ancestor of all primates, whereas 30 were under positive selection in other primate ancestral nodes leading to the evolution of humans (table S26). These results appear to suggest that primates underwent continuous brain evolution over an extended period of evolutionary time. Knockout experiments in mice on many of these positively selected genes have shown brain function impairment. For instance, the *NIPBL* gene interacts with *ZFP609* to regulate the migration of cortical neurons, and its mutations are frequently involved in brain neurological defects encompassing intellectual disability and seizures (96). We identified two amino acid residues in the NIPBL protein that experienced adaptive change in the common ancestor of all primate lineages (fig. S26).

Microcephaly is characterized by severe neurological defects, the small brain size being caused by a disturbance of the proliferation of nerve cells (97). Some genes involved in microcephaly have been proposed as candidates for involvement in the evolution of brain size (98–100). We also searched for positive selection signals in the 1113 coding genes involved in microcephaly (g:Profiler identifier HP:0000252). In total, 65 positively selected genes with functional roles in microcephaly were identified, along with the primate ancestor leading to the human lineage (table S27), suggesting that microcephaly genes may have been involved in the marked evolutionary expansion of brain size that characterizes primates, especially in those crucial evolutionary nodes characterized by a sharp increase in the degree of cortical folding (gyrification) and brain volume (101).

We next sought to investigate the roles of regulatory elements in the evolution of primate brain size. We first identified noncoding regions that were highly conserved and under strong purifying selection across all primates and detected signals of accelerated evolution in four lineages: the Simiiformes ancestor (table S21), the Catarrhini ancestor (table S28), the ancestor of great apes (table S29), and the

human lineage (table S30), representing crucial evolutionary nodes for the enlargement of primate brain size (101) (fig. S27). These lineage-specific accelerated regions should be under strong positive selection specifically in the targeted lineages and might contribute to the adaptation or innovation of these lineages (72). We found 15 genes associated with lineage-specific accelerated regions in the common ancestor of the great apes that showed particularly high expression in the human fetal brain (fig. S27 and table S31) ($P = 0.023$, modified Fisher's exact test). More than half of these genes have been reported to have roles in brain development and function (102–109). For example, knockout of the transcription factor–encoding *MEF2C* in a mouse model resulted in impaired neuronal differentiation and smaller somal size among neural progenitor cells (108). Coincidentally, the lineage-specific accelerated region of this gene was detected in the great ape ancestral lineage. The *DLG5* gene, which is required for the polarization of citron kinase in mitotic neural precursors, also contains a lineage-specific accelerated region in the great ape lineage, and $DLG5^{-/-}$ mice have smaller brains and thinner neocortices (109, 110).

We further investigated the evolution of neurotransmitters, which mediate the neurogenesis process (111, 112) and also play a role in the regulation of brain size (111). We detected 12 positively selected genes and 39 genes associated with lineage-specific accelerated regions in the ancestral nodes leading to the human lineage that were found to be involved in the release, transportation, and reception of neurotransmitter signals (Fig. 5A and fig. S28). These genes participate in diverse neurotransmitter systems: glutamatergic, dopaminergic, cholinergic, and GABAergic synapses and the synaptic vesicle cycle. Among these, five positively selected genes and 33 genes associated with lineage-specific accelerated regions are highly expressed in the human brain (table S32). It is likely that at least some of these genomic changes affecting the neurotransmitter signaling pathway might have played a role in primate brain evolution.

Evolution of the skeletal system and limbs

The arboreal lifestyle coevolved with adaptive changes of the skeletal system and limb development. Genes functioning in bone development are likely to have been especially important for the adaptive radiation of the primates. We identified four positively selected genes, *PIEZO1*, *EGFR*, *BMPER*, and *NOTCH2*, that were involved in bone development (113–116) in the ancestral lineage of primates (table S17). Bone development requires the recruitment of osteoclast precursors from the surrounding mesenchyme, thereby actuating the key events of bone growth, such as marrow cavity formation, capillary
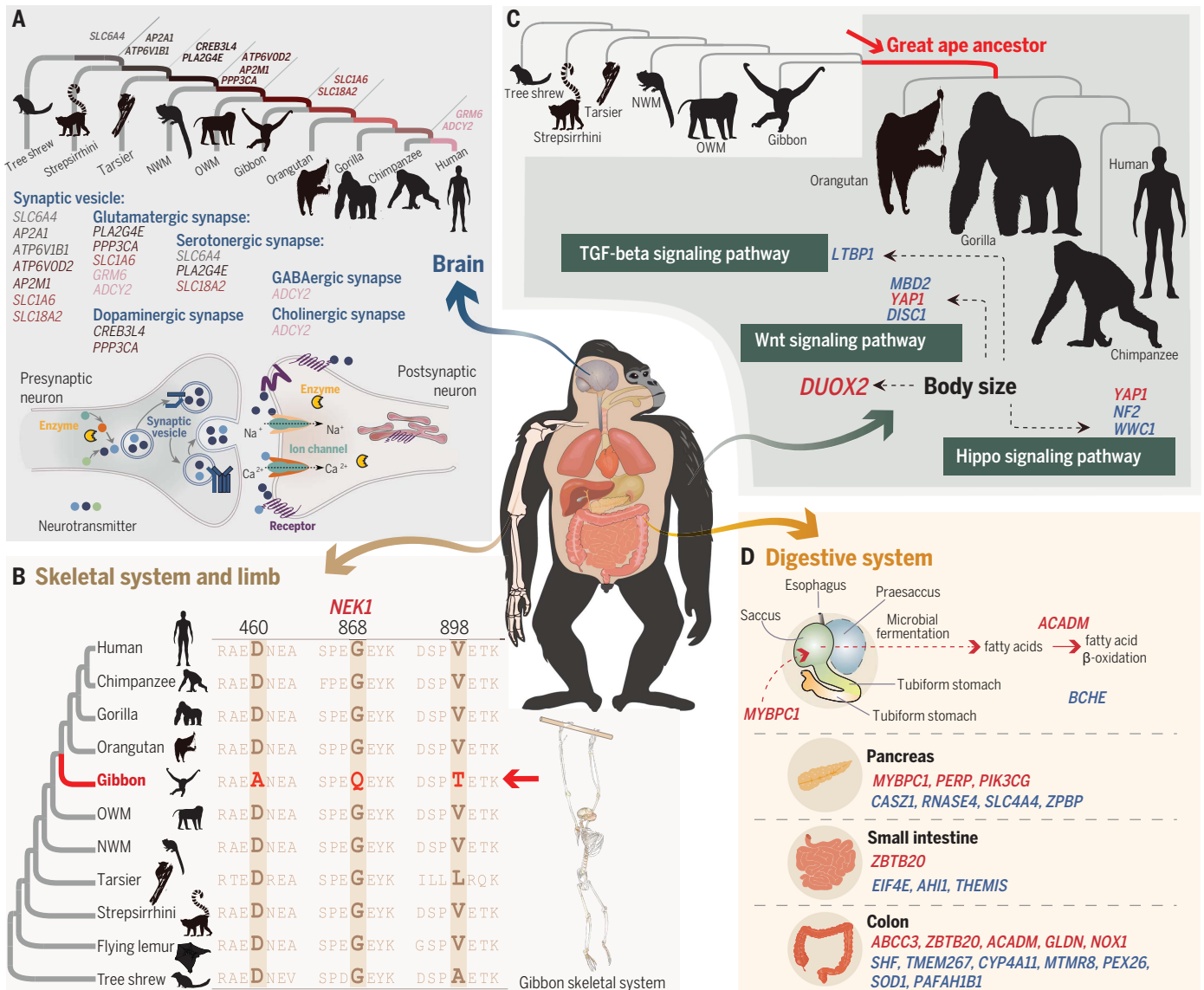
Fig. 5. Associations between genomic evolutionary characteristics and phenotypic traits in primates. (**A**) Positively selected genes and genes associated with lineage-specific accelerated regions from the primate ancestral lineage leading to the human lineage that are involved in transport, release, and receptors in neurotransmitter signaling. (**B**) The *NEK1* gene, which is involved in upper limb bone development, was under positive selection with three positively selected sites in the gibbon ancestral lineage. The gibbon ancestor is shown in red. (**C**) Eight positively selected genes and genes associated with lineage-specific accelerated regions from the great ape ancestral lineage involved in the TGF-β, Wnt, and Hippo signaling pathways. (**D**) Positively selected genes and genes associated with lineage-specific accelerated regions involved in the evolution of the digestive system in the Colobinae ancestral lineage. Genes marked in red and blue represent positively selected genes and genes associated with lineage-specific accelerated regions, respectively, in this lineage.

invasion, and matrix remodelling. The mechanical sensing protein PIEZO1 accommodates bone homeostasis through osteoclast-osteoblast cross-talk (*113*). Osteoclasts then influence osteoblast formation and differentiation through the secretion of some soluble factors (*117*). *EGFR* negatively regulates mTOR signaling during osteoblast differentiation to control bone development (*114*). The *NOTCH2* gene regulates cancellous bone volume and microarchitecture in osteoblast precursors (*116, 118*).

Although tails vary in length and shape across the primates, they generally play key roles in relation to locomotion (*119*). This notwithstanding, the tail was lost in some primate lineages, including the common ancestor of the apes (*120, 121*). We retrieved 151 genes associated with lineage-specific accelerated regions in the common ancestral lineage of the apes (table S33), including *KIAA1217* (sickle tail protein homolog) (figs. S29 and S30). Mutations in *KIAA1217* are associated with malformations of the notochord and caudal vertebrae in humans, and in mice they affect the development of the vertebral column, leading to a characteristic short tail due to a reduced number of caudal vertebrae (*122, 123*). Thus, the lineage-specific accelerated region may serve as a regulator of the expression of *KIAA1217*, because this lineage-specific accelerated region, residing in the vicinity of *KIAA1217* in the ape lineage, overlaps with the enhancer EH38E1455433 (pELS) (fig. S31). High-throughput chromosome conformation capture data (fig. S32) also showed that this lineage-specific accelerated region is located in the same topologically associated domain as *KIAA1217*, suggesting that they may physically interact with each other. Furthermore, the

lesser apes (gibbons) are of particular interest because of their dominant locomotor style, brachiation (*124, 125*). This locomotor adaptation was accompanied by the acquisition of distinct morphological characteristics, particularly the elongated forelimb, representing one of the most intriguing phenotypic traits in gibbons that enables them to travel through the canopy at high speed (*126*). We found that positive selection has operated on four genes related to upper limb bone morphology in the gibbon ancestral lineage (table S34). Of these, *NEK1*, which encodes a serine or threonine kinase, contains the most positively selected sites (Fig. 5B). Functional studies have shown that genetic variants in this gene can influence bone length and shorten the humerus and femur in humans (*127, 128*). Therefore,
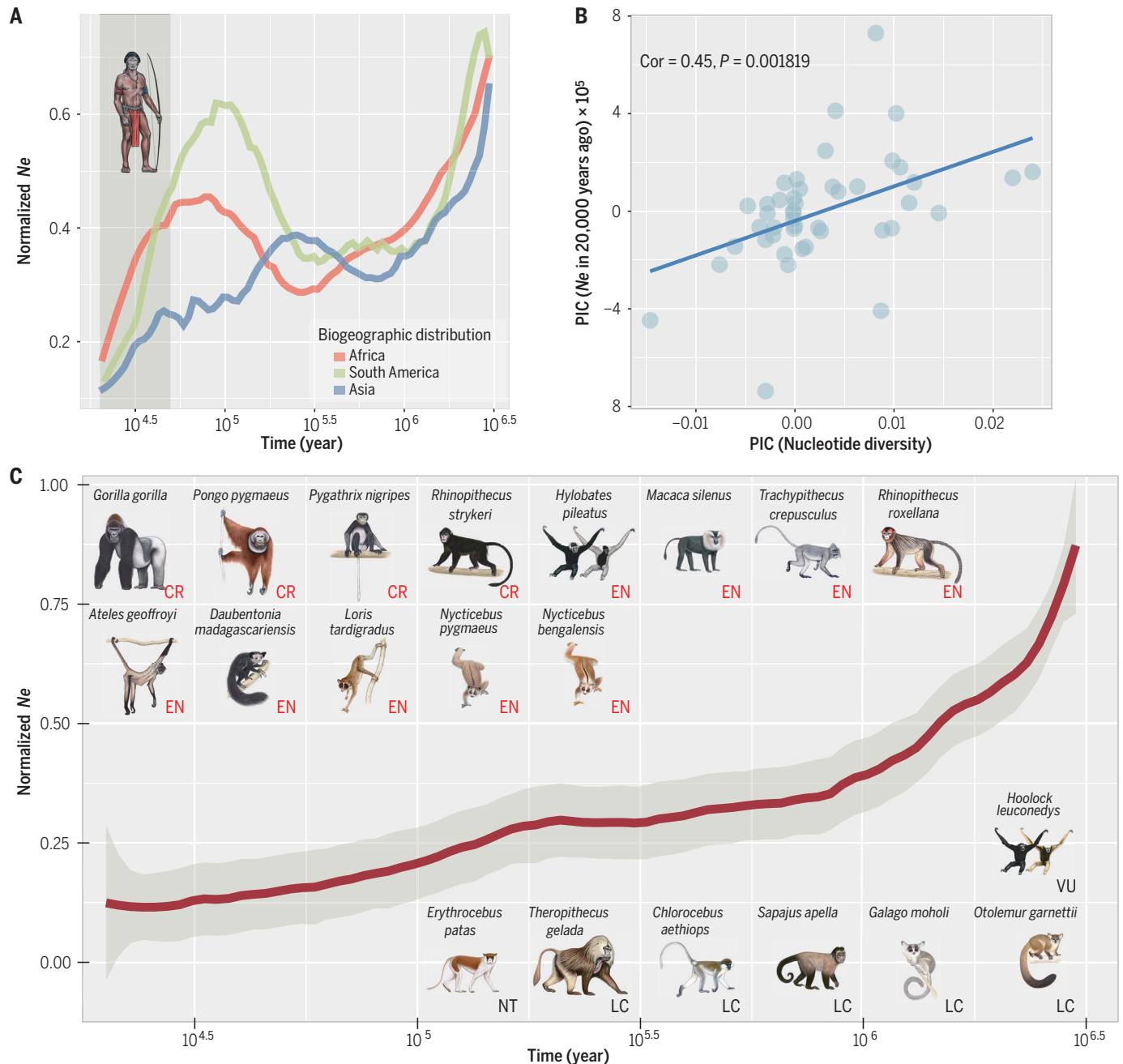
Fig. 6. Demographic history of nonhuman primates. (**A**) Primate species grouped according to their biogeographic distribution (Africa, Asia, or South America). The plot shows the normalized demographic history of all species within each biogeographic region. The normalized N_e was inferred by dividing the estimated value of N_e for each species at each time point by its maximum value. *Callithrix jacchus* was removed from this analysis because the genome was derived from an inbred individual. The time period from 50,000 to 20,000 years ago (late Pleistocene) is indicated by a gray background. (**B**) Correlation analysis between nucleotide diversity and N_e after phylogenetic corarection using the Ape library in R (http://ape-package.ird.fr/). N_e represents the median value of effective population size for each species 20,000 years ago. (**C**) Nearly half (*n* = 20) of all nonhuman primate species experienced a continual decline in N_e over the past 3 million years. These include the 13 critically endangered or endangered species shown in red. The IUCN Red List status is shown for each species in the inserted plot: CR, critically endangered; EN, endangered; VU, vulnerable; NT, near threatened; and LC, least concern.

positive selection acting on genes related to upper limb bone morphology may have been important in the acquisition of the elongated forelimb, a key adaptive trait for the unique brachiating locomotion style of gibbons.

Evolution of body size in primates

Like other mammalian groups (*129, 130*), extant primate species exhibit a large range of body sizes, from dwarf galagos and mouse lemurs (~60 to 70 g) at one end of the spectrum to male gorillas (>200 kg in some individuals) at the other (*131*). Thus, primate body size has experienced significant divergence, particularly for the great apes with their substantial enlargement in body size. We detected several positively selected genes in the common ancestors of the great apes that might have contributed to the evolution of this trait. *DUOX2* encodes a protein involved in a critical step of thyroid hormone synthesis, and mutations in *DUOX2* are known to cause decreased body size in mouse and panda (*132, 133*). This gene experienced strong positive selection in the great ape ancestral lineage ($P = 0.018$, χ^2 test) (Fig. 5C and table S35). Additionally, we found several genes involved in the transforming growth factor-β (TGF-β) signaling pathway (e.g., *LTBP1*) or the Wnt signaling pathway (e.g., *MBD2, YAP1*, and *DISC1*), two of the best known pathways participating in bone development and body size (*48*), that were either under strong positive selection in the great apes or had lineage-specific accelerated regions in this lineage (Fig. 5C and tables S29 and S35).

Several positively selected genes and genes associated with lineage-specific accelerated regions in the great ape ancestor were also significantly overrepresented in the Hippo signaling pathway ($P = 0.045$, modified Fisher's exact test) (table S36), which has been implicated in the determination of organ and body size (*82*). When combining all positively selected genes, genes associated with lineage-specific accelerated regions, and expanded gene families in the Simiiformes ancestral lineage, which markedly increased their body size compared with non-Simiiformes lineages (Fig. 4B), we also detected diverse candidate genes with adaptive changes in the Hippo signaling pathway. These results indicate potentially important roles for the Hippo pathway in body size changes in these two nodes during primate evolution.

Evolution of the digestive system

Primate lineages have evolved diverse dietary habits and specialized digestive functions (*134*). In particular, leaf-eating colobines, an African and Asian subfamily (Colobinae) of Old World monkeys, have evolved a uniquely specialized and compartmentalized foregut in which there are discrete alkaline and acidic sections to cope with their folivorous diet

and microbial fermentation can take place (*135, 136*). Although colobines eat leaves, fruits, flowers, and seeds, they typically focus much of their feeding time on leaves (estimated range: ~34 to 81% of their annual diet) (*135*). Accordingly, these leaf-eaters are well adapted in terms of meeting their energy metabolism requirements and balancing micronutrients and protein intake while also dealing with the toxins contained in their food plants (*137*).

In the ancestor of the Colobinae, we identified a number of pivotal digestive genes that underwent positive selection (table S37). Acyl-CoA dehydrogenase, encoded by the *ACADM* gene, is an important lipolytic enzyme that catalyzes the initial step in each cycle of mitochondrial fatty acid β-oxidation and plays a key role in metabolizing fatty acids derived from ingested foods (*138*). Energy-rich short-chain volatile fatty acids are produced by the microbial fermentation process and absorbed by the host, thus making an important contribution to the energy budget of colobines (*135*). Therefore, rapid evolution of this gene, with two positively selected sites (V75M and A138C), may have been important for the absorption of fatty acids by colobines (Fig. 5D and fig. S33). *NOX1*, which is highly expressed in the colon, was identified as being under positive selection in the ancestor of the Colobinae (Fig. 5D and tables S37 and S38). *NOX1*-dependent reactive oxygen species production can further regulate microorganism homeostasis in the ileum of mice (*139*). The rumens of ruminants and the saccus stomachs of colobines have developed a similar adaptive strategy to allow the microbial fermentation of high-fiber foods, and therefore are an example of convergent evolution. We found that *MYBPC1*, which has been shown to contribute to morphological and functional differences in the bovine rumen (*140*), also underwent positive selection in the ancestor of the Colobinae (Fig. 5D and table S37). In addition, 100 genes associated with lineage-specific accelerated regions were identified in the ancestral lineage of the Colobinae (table S39). Several of these genes were also highly expressed in the stomach, colon, pancreas, and small intestine (Fig. 5D and table S38). Of these, *RNASE4* encodes a vital digestive enzyme, pancreatic ribonuclease 4, and is a paralog of *RNASE1*, which is known to have undergone adaptive evolution by gene duplication in leaf-eating colobines and howler monkeys (*26, 141*). Colobines may therefore have acquired adaptations to allow them to digest fatty acids and ribonucleic acids, and their unique foregut and intestinal microbiota enabled them to cope with their folivorous diet.

Evolution of sensory organs

In many mammals, olfaction is the dominant sense and provides much of the sensory infor-

mation upon which animals rely to navigate, forage, and avoid predators or for social behavior and courtship (*134*). Most Strepsirrhini species are nocturnal, whereas most Simiiformes are diurnal with well-developed color vision systems attuned to their priorities in diurnal activity (*142–145*). By contrast, olfactory sensitivity appears to have decreased in the Simiiformes compared with the Strepsirrhini (*134, 146, 147*). Consistent with these findings, we found that the copy number of several specific olfactory receptor gene families was significantly reduced in the Simiiformes. For example, the olfactory receptor gene family *OR52A* underwent a significant contraction in the Simiiformes (40 species), with only ~0.7 copies on average, in contrast to the ~3.4 average copies in the Strepsirrhini (nine species) (figs. S34 and S35) ($P = 4.072 \times 10^{-5}$, Mann-Whitney U test). Anatomically, Strepsirrhini are characterized by the presence of a rhinarium, a moist and naked surface around the tip of the nose that is present in most mammals, including dogs and cats, but has been lost in the Simiiformes (*134, 147*). Olfactory bulb volume, which correlates with olfactory receptor neuron population size, is also larger in the Strepsirrhini than in the Simiiformes (*146, 148*). The *LHX2* gene, which participates in olfactory bulb development (*149, 150*), experienced positive selection in the ancestor of the Strepsirrhini ($P = 0.03$, χ^2 test; table S40).

Demographic history of nonhuman primates

The IUCN lists more than one-third of primates as critically endangered or vulnerable (*1*). To evaluate the effects of climate change and human activity on the recent population declines in these primates, we inferred their demographic histories over the past million years by using the pairwise sequentially Markovian coalescent model (*151*) for each species in this study (fig. S36 and tables S16 and S41). Our data showed that most nonhuman primate species experienced rapid population declines during the late Pleistocene (Fig. 6A and fig. S37), consistent with the record of a large mass extinction of mammals during this period (*48, 152*). Although we did not observe a significant difference between endangered species and other species in terms of nucleotide diversity (fig. S38 and table S42), we did detect a significant positive correlation between the median effective population size (N_e) over the past ~20,000 years and nucleotide diversity ($P = 0.002$, Pearson's product-moment correlation after phylogenetic correction) (Fig. 6B and table S42), indicating a long-term effect of N_e decline on the loss of genetic diversity. According to the historical demographic patterns, we further clustered all nonhuman primate species with similar trends of historical N_e, and found that 20 species experienced a

continual N_e decline over the past 3 million years (Fig. 6C). Sixty-five percent of these species are now listed as endangered or critically endangered (Fig. 6C and fig. S39). This ratio is twice that of the remaining species, suggesting that the prehistoric environmental effects (e.g., habitat fragmentation) (*26*) may also have driven population decline and contributed to the current endangered status of these species well before human interference in the modern era.

Conclusions

Understanding the evolution and genetic basis of human-specific traits requires a systematic comparison of genomes along the primate lineages. Previous studies of primate genomes have focused on genomic changes in the human lineage that influenced brain functions and other traits (*120*, *153–155*). Our comparative phylogenomic analyses across primate lineages have revealed some of the accumulated genomic changes at different primate ancestral nodes that may have contributed to the evolution of unique human traits. Of particular interest, we report a hitherto unreported increase in the rate of genomic change in the Simiiformes common ancestor that may have played a role in the later diversification of Simiiformes and the evolution of humans. Our comparative genomic analyses also yielded insights into the genetic basis of phenotypic diversity across primate lineages. With the rich diversity of morphology and physiology among nonhuman primates, further genomic analyses covering all primate species will provide an indispensable resource for comparative studies allowing expansion of the scope of biomedical research programs using primates as model systems. Further, increased knowledge of the genomic makeup and variations of nonhuman primates should help to identify risk factors for genetic disorders and enhance wildlife health management in both wild and captive members of these species.

REFERENCES AND NOTES

1. A. Estrada *et al.*, *Sci. Adv.* **3**, e1600946 (2017).
2. C. Roos *et al.*, *Zool. Res.* **41**, 656–669 (2020).
3. A. Nater *et al.*, *Curr. Biol.* **27**, 3487–3498.e10 (2017).
4. P. F. Fan *et al.*, *Am. J. Primatol.* **79**, e22631 (2017).
5. C. Li, C. Zhao, P. F. Fan, *Am. J. Primatol.* **77**, 753–766 (2015).
6. J. Rogers, R. A. Gibbs, *Nat. Rev. Genet.* **15**, 347–359 (2014).
7. B. Rockx *et al.*, *Science* **368**, 1012–1015 (2020).
8. A. Chandrashekar *et al.*, *Science* **369**, 812–817 (2020).
9. Q. Gao *et al.*, *Science* **369**, 77–81 (2020).
10. J. Yu *et al.*, *Science* **369**, 806–811 (2020).
11. V. J. Munster *et al.*, *Nature* **585**, 268–272 (2020).
12. N. B. Mercado *et al.*, *Nature* **586**, 583–588 (2020).
13. K. S. Corbett *et al.*, *N. Engl. J. Med.* **383**, 1544–1555 (2020).
14. N. van Doremalen *et al.*, *Nature* **586**, 578–582 (2020).
15. B. N. Williamson *et al.*, *Nature* **585**, 273–276 (2020).
16. T. Z. Song *et al.*, *Zool. Res.* **41**, 503–516 (2020).
17. W. Enard, S. Pääbo, *Annu. Rev. Genomics Hum. Genet.* **5**, 351–378 (2004).
18. Z. N. Kronenberg *et al.*, *Science* **360**, eaar6343 (2018).
19. Chimpanzee Sequencing and Analysis Consortium, *Nature* **437**, 69–87 (2005).
20. R. A. Gibbs *et al.*, *Science* **316**, 222–234 (2007).
21. A. Scally *et al.*, *Nature* **483**, 169–175 (2012).
22. Marmoset Genome Sequencing and Analysis Consortium, *Nat. Genet.* **46**, 850–857 (2014).
23. D. P. Locke *et al.*, *Nature* **469**, 529–533 (2011).
24. L. Carbone *et al.*, *Nature* **513**, 195–201 (2014).
25. L. Yu *et al.*, *Nat. Genet.* **48**, 947–952 (2016).
26. X. Zhou *et al.*, *Nat. Genet.* **46**, 1303–1310 (2014).
27. A. O. Ayoola *et al.*, *Mol. Biol. Evol.* **38**, 876–890 (2021).
28. D. M. Bickhart *et al.*, *Nat. Genet.* **49**, 643–650 (2017).
29. B.-L. Zhang *et al.*, *Sci. Adv.* **9**, eadd3580 (2023).
30. H. Wu *et al.*, *Science* **380**, eabl4997 (2023) .
31. X.-G. Qi *et al.*, *Science* **380**, eabl8621 (2023).
32. M.-L. Li *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2123030119 (2022).
33. M. S. Ye *et al.*, *Zool. Res.* **42**, 692–709 (2021).
34. A. M. Kozlov, A. J. Aberer, A. Stamatakis, *Bioinformatics* **31**, 2577–2579 (2015).
35. P. Perelman *et al.*, *PLOS Genet.* **7**, e1001342 (2011).
36. C. M. Shi, Z. Yang, *Mol. Biol. Evol.* **35**, 159–179 (2018).
37. A. Hobolth, O. F. Christensen, T. Mailund, M. H. Schierup, *PLOS Genet.* **3**, e7 (2007).
38. I. Rivas-González *et al.*, *Science* **380**, eabn4409 (2022).
39. D. Vanderpool *et al.*, *PLOS Biol.* **18**, e3000954 (2020).
40. Z. Yang, *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
41. S. Álvarez-Carretero *et al.*, *Nature* **602**, 263–267 (2022).
42. C. Liu *et al.*, *Sci. Adv.* **9**, eabe9459 (2021).
43. E. E. Eichler, D. Sankoff, *Science* **301**, 793–797 (2003).
44. Y. Yin *et al.*, *Nat. Commun.* **12**, 6858 (2021).
45. R. Stanyon *et al.*, *Chromosome Res.* **16**, 17–39 (2008).
46. T. Marques-Bonet *et al.*, *Nature* **457**, 877–881 (2009).
47. P. D. Stenson *et al.*, *Hum. Genet.* **139**, 1197–1207 (2020).
48. L. Chen *et al.*, *Science* **364**, eaav6202 (2019).
49. J. D. Smith, J. W. Bickham, T. R. Gregory, *Genome* **56**, 457–472 (2013).
50. S. Shen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2837–2842 (2011).
51. G. E. Liu, C. Alkan, L. Jiang, S. Zhao, E. E. Eichler, *Genome Res.* **19**, 876–885 (2009).
52. T. Hayakawa, Y. Satta, P. Gagneux, A. Varki, N. Takahata, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11399–11404 (2001).
53. P. Kuehnen *et al.*, *PLOS Genet.* **8**, e1002543 (2012).
54. J. Jurka, *Curr. Opin. Genet. Dev.* **14**, 603–608 (2004).
55. G. Zhang *et al.*, *Science* **346**, 1311–1320 (2014).
56. P. Moorjani, C. E. Amorim, P. F. Arndt, M. Przeworski, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10607–10612 (2016).
57. E. Fontanillas, J. J. Welch, J. A. Thomas, L. Bromham, *BMC Evol. Biol.* **7**, 95 (2007).
58. A. Wong, *Mol. Biol. Evol.* **31**, 1432–1436 (2014).
59. W. H. Li, M. Tanimura, *Nature* **326**, 93–96 (1987).
60. M. E. Steiper, N. M. Young, *Mol. Phylogenet. Evol.* **41**, 384–394 (2006).
61. S. H. Kim, N. Elango, C. Warden, E. Vigoda, S. V. Yi, *PLOS Genet.* **2**, e163 (2006).
62. J. Schmitz *et al.*, *Nat. Commun.* **7**, 12997 (2016).
63. L. Fang *et al.*, *Genome Res.* **30**, 790–801 (2020).
64. B. Y. Liao, J. Zhang, *Mol. Biol. Evol.* **23**, 1119–1128 (2006).
65. G. J. Wyckoff, W. Wang, C. I. Wu, *Nature* **403**, 304–309 (2000).
66. T. Boehm, *Curr. Biol.* **22**, R722–R732 (2012).
67. H. Y. Wang *et al.*, *PLOS Biol.* **5**, e13 (2007).
68. Materials and methods are available as supplementary materials.
69. J. Tohyama *et al.*, *J. Hum. Genet.* **60**, 167–173 (2015).
70. P. Mansfield, J. N. Constantino, D. Baldridge, *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **183**, 227–233 (2020).
71. M. Maekawa *et al.*, *J. Neurochem.* **115**, 1374–1385 (2010).
72. X. Bi *et al.*, *Sci. Adv.* 10.1126/sciadv.adc9507 (2023).
73. J. K. Rilling, T. R. Insel, *Neuroreport* **10**, 1453–1459 (1999).
74. K. Isler *et al.*, *J. Hum. Evol.* **55**, 967–978 (2008).
75. C. Plachez, L. J. Richards, *Curr. Top. Dev. Biol.* **69**, 267–346 (2005).
76. M. A. Robichaux, C. W. Cowan, *Curr. Top. Behav. Neurosci.* **16**, 19–48 (2014).
77. J. Falk *et al.*, *Neuron* **48**, 63–75 (2005).
78. M. A. Wolman, Y. Liu, H. Tawarayama, W. Shoji, M. C. Halloran, *J. Neurosci.* **24**, 8428–8435 (2004).
79. C. Kudo, I. Ajioka, Y. Hirata, K. Nakajima, *J. Comp. Neurol.* **487**, 255–269 (2005).
80. M. V. Tejada-Simon, *J. Neurochem.* **133**, 767–779 (2015).
81. S. L. Eastwood, P. J. Harrison, *Neuropsychopharmacology* **33**, 933–945 (2008).
82. D. Pan, *Genes Dev.* **21**, 886–897 (2007).
83. S. H. Patel, F. D. Camargo, D. Yimlamai, *Gastroenterology* **152**, 533–545 (2017).
84. R. H. Gokhale, A. W. Shingleton, *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 335–356 (2015).
85. E. C. Kirk, *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **281**, 1095–1103 (2004).
86. A. C. Wiik *et al.*, *Genome Res.* **18**, 1415–1421 (2008).
87. P. Liskova *et al.*, *Am. J. Hum. Genet.* **102**, 447–459 (2018).
88. Y. Toda *et al.*, *Curr. Biol.* **31**, 4641–4649.e5 (2021).
89. J. M. Kamilar, B. J. Bradley, *J. Biogeogr.* **38**, 2270–2277 (2011).
90. S. Hu *et al.*, *PeerJ* **8**, e9402 (2020).
91. M. C. Garrido, B. C. Bastian, *J. Invest. Dermatol.* **130**, 20–27 (2010).
92. J. M. Grichnik, *J. Invest. Dermatol.* **126**, 945–947 (2006).
93. Y. Mizutani, N. Hayashi, M. Kawashima, G. Imokawa, *Arch. Dermatol. Res.* **302**, 283–294 (2010).
94. R. Kitamura *et al.*, *J. Pathol.* **202**, 463–475 (2004).
95. B. Wen *et al.*, *Pigment Cell Melanoma Res.* **23**, 441–447 (2010).
96. D. L. C. van den Berg *et al.*, *Neuron* **93**, 348–361 (2017).
97. G. H. Mochida, C. A. Walsh, *Curr. Opin. Neurol.* **14**, 151–156 (2001).
98. S. H. Montgomery, I. Capellini, C. Venditti, R. A. Barton, N. I. Mundy, *Mol. Biol. Evol.* **28**, 625–638 (2011).
99. L. Shi, M. Li, Q. Lin, X. Qi, B. Su, *BMC Biol.* **11**, 62 (2013).
100. L. Shi, B. Su, *Zool. Res.* **40**, 236–238 (2019).
101. J. Rogers *et al.*, *Neuroimage* **53**, 1103–1108 (2010).
102. S. V. Puram *et al.*, *Genes Dev.* **25**, 2659–2673 (2011).
103. A. Yamada *et al.*, *Mol. Cell. Neurosci.* **56**, 234–243 (2013).
104. R. Kusano *et al.*, *FEBS Lett.* **590**, 3606–3615 (2016).
105. M. Talarowska, J. Szemraj, M. Kowalczyk, P. Gałecki, *Med. Sci. Monit.* **22**, 152–160 (2016).
106. A. K. Pandey, L. Lu, X. Wang, R. Homayouni, R. W. Williams, *PLOS ONE* **9**, e88889 (2014).
107. A. Graziano, G. Foffani, E. B. Knudsen, J. Shumsky, K. A. Moxon, *PLOS ONE* **8**, e54350 (2013).
108. H. Li *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9397–9402 (2008).
109. Y. Chang, O. Klezovitch, R. S. Walikonis, V. Vasioukhin, J. J. LoTurco, *Cell Cycle* **9**, 1990–1997 (2010).
110. M. R. Sarkisian, *Cell Cycle* **9**, 1876 (2010).
111. D. A. Berg, L. Belnoue, H. Song, A. Simon, *Development* **140**, 2548–2561 (2013).
112. P. Levitt, J. A. Harvey, E. Friedman, K. Simansky, E. H. Murphy, *Trends Neurosci.* **20**, 269–274 (1997).
113. L. Wang *et al.*, *Nat. Commun.* **11**, 282 (2020).
114. M. Linder *et al.*, *Cell Death Differ.* **25**, 1094–1106 (2018).
115. F. Xiao *et al.*, *Cell. Physiol. Biochem.* **45**, 1927–1939 (2018).
116. S. Zanotti, E. Canalis, *Bone* **62**, 22–28 (2014).
117. J. M. Kim, C. Lin, Z. Stavre, M. B. Greenblatt, J. H. Shim, *Cells* **9**, 2073 (2020).
118. S. Zanotti, E. Canalis, *Endocr. Rev.* **37**, 223–253 (2016).
119. M. Schmidt, *Adv. Sci. Res.* **5**, 23–39 (2011).
120. Y. He *et al.*, *Nat. Commun.* **10**, 4233 (2019).
121. S. A. Williams, G. A. Russo, *Evol. Anthropol.* **24**, 15–32 (2015).
122. K. Semba *et al.*, *Genetics* **172**, 445–456 (2006).
123. N. Al Dhaheri *et al.*, *Am. J. Med. Genet. A.* **182**, 1664–1672 (2023).
124. J. R. Usherwood, J. E. Bertram, *J. Exp. Biol.* **206**, 1631–1642 (2003).
125. J. R. Usherwood, S. G. Larson, J. E. Bertram, *Am. J. Phys. Anthropol.* **120**, 364–372 (2003).
126. S. M. Cheyne, in *Primate Locomotion: Linking Field and Laboratory Research*, K. D'Août, E. E. Vereecke, Eds. (Springer, NY, 2011), pp. 201–213.
127. C. Thiel *et al.*, *Am. J. Hum. Genet.* **88**, 106–114 (2011).
128. J. El Hokayem *et al.*, *J. Med. Genet.* **49**, 227–233 (2012).
129. J. M. Vazquez, V. J. Lynch, *eLife* **10**, e65041 (2021).
130. J. G. M. Thewissen, L. N. Cooper, J. C. George, S. Bajpai, *Evolution (N. Y.)* **2**, 272–288 (2009).
131. W. L. Jungers, in *Size and Scaling in Primate Biology*, W. L. Jungers, Ed. (Springer, 1985), pp. 345–381.
132. A. M. Rudolf *et al.*, *Natl. Sci. Rev.* **9**, nwab125 (2021).
133. K. R. Johnson *et al.*, *Mol. Endocrinol.* **21**, 1593–1602 (2007).
134. J. G. Fleagle, *Primate Adaptation and Evolution* (Academic, 2013).
135. K. Milton, *Int. J. Primatol.* **19**, 513–548 (1998).
136. I. Matsuda, C. A. Chapman, M. Clauss, *J. Morphol.* **280**, 1608–1616 (2019).
137. M. C. Janiak, *Evol. Anthropol.* **25**, 253–266 (2016).
138. J. J. Kim, R. Miura, *Eur. J. Biochem.* **271**, 483–493 (2004).

139. C. Matziouridou et al., Mucosal Immunol. 11, 774–784 (2018).
140. C.-J. Li, R. W. Li, R. L. Baldwin Vi, Agric. Sci. 9, 619–638 (2018).
141. M. C. Janiak, A. S. Burrell, J. D. Orkin, T. R. Disotell, Sci. Rep. 9, 20366 (2019).
142. P. Pontarotti, Evolutionary Biology: Mechanisms and Trends (Springer, 2012).
143. N. J. Dominy, P. W. Lucas, Nature 410, 363–366 (2001).
144. N. G. Caine, N. I. Mundy, Proc. Biol. Sci. 267, 439–444 (2000).
145. A. C. Smith, H. M. Buchanan-Smith, A. K. Surridge, D. Osorio, N. I. Mundy, J. Exp. Biol. 206, 3159–3165 (2003).
146. S. Heritage, PLOS ONE 9, e113904 (2014).
147. A. Matsui, Y. Go, Y. Niimura, Mol. Biol. Evol. 27, 1192–1200 (2010).
148. T. D. Smith, K. P. Bhatnagar, Anat. Rec. B New Anat. 279, 24–31 (2004).
149. A. Berghard, A. C. Hägglund, S. Bohm, L. Carlsson, FASEB J. 26, 3464–3472 (2012).
150. J. Hirota, P. Mombaerts, Proc. Natl. Acad. Sci. U.S.A. 101, 8751–8755 (2004).
151. H. Li, R. Durbin, Nature 475, 493–496 (2011).
152. A. D. Barnosky, P. L. Koch, R. S. Feranec, S. L. Wing, A. B. Shabel, Science 306, 70–75 (2004).
153. X. Luo et al., Cell 184, 723–740.e21 (2021).
154. C. Yang et al., Nature 594, 227–233 (2021).
155. G. Dumas, S. Malesys, T. Bourgeron, Genome Res. 31, 484–496 (2021).
156. J. K. Rilling, Evol. Anthropol. 15, 65–77 (2006).
157. H. Stephan, H. Frahm, G. Baron, Folia Primatol. (Basel) 35, 1–29 (1981).
158. Genome annotation GFF files at Mendeley Data for: Y. Shao et al., Phylogenomic analyses provide insights into primate evolution, Mendeley (2023).
159. Genome annotation GFF files at Figshare for: Y. Shao et al., Phylogenomic analyses provide insights into primate evolution, Figshare (2023); https://doi.org/10.5061/dryad.8w9ghx3qj.
160. Gene sequences for: Y. Shao et al., Phylogenomic analyses provide insights into primate evolution, Dryad (2023).

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn6919
Materials and Methods
Figs. S1 to S39
Tables S1 to S42
References (161–237)
MDAR Reproducibility Checklist

Phylogenomic analyses provide insights into primate evolution

Yong Shao, Long Zhou, Fang Li, Lan Zhao, Bao-Lin Zhang, Feng Shao, Jia-Wei Chen, Chun-Yan Chen, Xupeng Bi, Xiao-Lin Zhuang, Hong-Liang Zhu, Jiang Hu, Zongyi Sun, Xin Li, Depeng Wang, Iker Rivas-Gonzlez, Sheng Wang, Yun-Mei Wang, Wu Chen, Gang Li, Hui-Meng Lu, Yang Liu, Lukas F. K. Kuderna, Kyle Kai-How Farh, Peng-Fei Fan, Li Yu, Ming Li, Zhi-Jin Liu, George P. Tiley, Anne D. Yoder, Christian Roos, Takashi Hayakawa, Tomas Marques-Bonet, Jeffrey Rogers, Peter D. Stenson, David N. Cooper, Mikkel Heide Schierup, Yong-Gang Yao, Ya-Ping Zhang, Wen Wang, Xiao-Guang Qi, Guojie Zhang, and Dong-Dong Wu

View the article online
https://www.science.org/doi/10.1126/science.abn6919
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service

Science

Supplementary Materials for

Phylogenomic analyses provide insights into primate evolution

Yong Shao *et al*.

Corresponding authors: Xiao-Guang Qi, qixg@nwu.edu.cn; Guojie Zhang, guojiezhang@zju.edu.cn; Dong-Dong Wu, wudongdong@mail.kiz.ac.cn

The PDF file includes:

Materials and Methods
Figs. S1 to S39
Tables S1 to S20, S23 to S27, S31, S32, S34 to S38, S40 to S42
References

Other Supplementary Material for this manuscript includes the following:

Tables S21, S22, S28 to S30, S33, and S39
MDAR Reproducibility Checklist

Materials and Methods

SM Text 1 Species and sample information

The complete set of primate genomes used in our project encompassed a total of 50 species, including 27 newly assembled genomes (tables S1–S4), one genome (for *Nycticebus pygmaeus,* the *pygmy slow loris*) from our accompanying paper (*32*), and 22 previously published genomes (table S8).

Sample collection strictly followed the rules of CITES (the Convention on International Trade in Endangered Species of Wild Fauna and Flora). The blood samples from each species for 25 primate species were obtained from zoos in China, during physical examination. The frozen liver tissue for the species including *Daubentonia madagascariensis* and *Galago moholi* were obtained from the Duke Lemur Center (https://lemur.duke.edu/), and sequenced by RTL Genomics, Marsha Sharp Fwy, USA. All animal specimens were obtained legally and in accordance with the code of ethics for the Care and Use of Animals of the Kunming Institute of Zoology, Chinese Academy of Sciences, China (Approval ID: SMKX-20180701-01), which conforms to the regulatory standards for the human care and treatment of animals in research.

SM Text 2 Genome sequencing, assembly and annotation
<u>DNA sample collection</u>

Genomic DNA from 27 primate species (table S1) was extracted from fresh blood or frozen tissues according to the protocol of the DNAeasy Blood & Tissue kit (Qiagen, USA). Briefly, 20 μl proteinase K and 220 μl anticoagulated blood were added to 200 μl Buffer AL in a 2 ml microcentrifuge tube. The solution was then mixed thoroughly by vortexing before being incubated at 56°C for 10 min. Next, 200 μl ethanol was added to the solution, and the sample mixed fully by vortexing. The mixture was further pipetted into a DNeasy Mini spin column and centrifuged at 8,000 rpm for 1 min. The DNeasy Mini spin column was placed in a new 2 ml collection tube with 500 μl Buffer AW1, and the mixture was centrifuged for 1 min at 8,000 rpm. Around 500 μl Buffer AW2 was added to the spin column and was centrifuged for 3 min at 14,000 rpm to dry the DNeasy membrane. Finally, 50 μl Buffer AE was placed directly onto the DNeasy membrane and centrifuged for 2 min at 8,000 rpm prior to elution after incubation at room temperature for 5 min.

<u>DNA library construction and sequencing</u>

A NanoDrop spectrophotometer, a Qubit fluorometer and gel electrophoresis were used to evaluate the DNA purity, concentration and integrity.

13 DNA libraries (representing 12 primate species, namely *Pongo pygmaeus*, *Nomascus siki*, *Symphalangus syndactylus*, *Lophocebus aterrimus*, *Cercopithecus albogularis*, *Chlorocebus aethiops*, *Sapajus apella*, *Ateles geoffroyi*, *Pithecia pithecia*, *Cephalopachus bancanus*, *Daubentonia madagascariensis*, *Loris tardigradus* and

Galago moholi) were constructed for PacBio SMRT sequencing following the Pacific Biosciences recommended protocols (*28*). Subsequently, 20 kbp libraries were sequenced on a PacBio RS II platform (Pacific Biosciences, USA). On average, ~147.50 Gbp (~53.81X) of clean subread bases was generated for each of 13 primate species for downstream analyses (table S1).

14 DNA libraries (representing 15 primate species, specifically *Hoolock leuconedys*, *Hylobates pileatus*, *Macaca assamensis*, *Macaca silenus*, *Papio hamadryas*, *Mandrillus sphinx*, *Erythrocebus patas*, *Trachypithecus crepusculus*, *Pygathrix nigripes*, *Rhinopithecus strykeri*, *Colobus guereza*, *Saguinus midas*, *Cebus albifrons* and *Nycticebus bengalensis*) were constructed for Oxford Nanopore Technology sequencing following the manufacturer's instructions. Subsequently, libraries were sequenced on a Nanopore PromethI ON platform. On average, ~148.53 Gbp (~51.32X) of clean subread bases were generated for each of 14 primate species for downstream analyses (table S1).

Four Hi-C libraries for 4 species including *Hylobates pileatus* (in Hominoidea), *Colobus guereza* (in Old World monkeys [OWMs]), *Saguinus midas* (in New World monkeys [NWMs]) and *Nycticebus bengalensis* (in Strepsirrhini) were constructed and sequenced according to the manufacturer's instructions to link contigs into scaffolds. On average, ~332.61 Gb (~114.15X) clean Hi-C data were generated for each of the 4 species (table S4).

For 4 newly sequenced primate species, namely *Nomascus siki*, *Symphalangus syndactylus*, *Ateles geoffroyi* and *Pithecia pithecia*, we generated 10X Genomics Chromium sequencing data to link contigs into scaffolds and to further polish sequencing errors from the long-read sequencing platforms (PacBio RS II and Nanopore PromethI ON) (table S3). For the remaining 23 newly sequenced primate species, we additionally produced Illumina clean short-read sequencing data to correct sequencing errors from the long-read platforms (table S3). A total of ~4.91 tera base-pairs (Tbp) sequencing data [on average ~189.04 Gbp (~67.11X) per species] were generated during this study (table S3).

Genome assembly

In order to obtain an optimal assembly for each species, we adopted diverse assembly strategies and ultimately selected the best assembly version (table S2). We generated optimal assemblies using wtdbg2 (*161*) for 15 species (*Pongo pygmaeus*, *Nomascus siki*, *Hoolock leuconedys*, *Lophocebus aterrimus*, *Cercopithecus albogularis*, *Chlorocebus aethiops*, *Trachypithecus crepusculus*, *Pygathrix nigripes*, *Rhinopithecus strykeri*, *Ateles geoffroyi*, *Cephalopachus bancanus*, *Daubentonia madagascariensis*, *Loris tardigradus*, *Nycticebus bengalensis* and *Galago moholi*), NextDenovo (https://github.com/Nextomics/NextDenovo) for 10 species (*Hylobates pileatus*, *Macaca assamensis*, *Macaca silenus*, *Papio hamadryas*, *Mandrillus sphinx*, *Erythrocebus patas*, *Colobus guereza*, *Saguinus midas*, *Sapajus paella*, and *Cebus albifrons*), and FALCON (*162*) for two species (*Symphalangus syndactylus* and

Pithecia pithecia) (table S2).

For the 15 genome assemblies derived by using wtdbg2 (*161*), the short-read sequencing data for each species were aligned back to their assembled contigs using BWA (v0.7.12) (*163*) in order to correct sequencing errors from long-read sequencing platforms; an additional polish was performed by means of wtdbgcns from wtdbg2 (*161*). For the 10 genome assemblies derived by using NextDenovo (https://github.com/Nextomics/NextDenovo), we first performed self-error correction for subreads using NextDenovo with parameters "read_cutoff = 1k; seed_cutoff = 25k". Then the self-error corrected reads were assembled into contigs using wtdbg (v1.2.8) (https://github.com/ruanjue/wtdbg) with parameters "wtdbg -k 0 -p 17 -S 2; wtdbg-cns -c 0 -k 15; kbm-1.2.8 -k 0 -p 15 -S 2 -O 0". We utilized the short-read sequencing data for each species aligned to their assembled contigs using BWA (v0.7.12) (*163*); nextpolish (https://github.com/Nextomics/NextPolish) was utilized to polish each assembly three times. For the two genome assemblies derived by using FALCON (*162*), the ARROW program in SMRTLink (v5.1.0.26412) (resequencing pipeline) was employed to polish assemblies using the default parameters.

For 4 newly sequenced primate species (specifically *Nomascus siki*, *Symphalangus syndactylus*, *Ateles geoffroyi* and *Pithecia pithecia*), we generated scaffolds by 10X Chromium barcoded linked reads through two rounds of scaffolding with Scaff10x (v2.1) (https://github.com/wtsi-hpag/Scaff10X) (table S3). The 10X Chromium barcoded linked reads were mapped to the contigs using BWA (v0.7.12) (*163*). The output of the first round of Scaff10x (v2.1) was used as input for the second round. For 4 new sequenced species (specifically *Hylobates pileatus*, *Colobus guereza*, *Saguinus midas* and *Nycticebus bengalensis*), the Hi-C read pair data were aligned to the genome (table S4); both duplicates and near-duplicates were removed, and read pairs that aligned to three or more locations were set aside. The Hi-C contacts were listed as input for 3d-dna. Then 3d-dna were used to correct misassemblies, serving to anchor, order and orient fragments of DNA based on Hi-C contacts to generate high quality scaffolds.

Genome completeness evaluation

The completeness of the newly assembled primate genomes was evaluated by Benchmarking Universal Single-Copy Orthologs (BUSCO, v3.0.2) (*164*) based on the mammalian single-copy orthologs (mammalia_odb9). In this study, >92% of BUSCO genes were completely annotated in each newly assembled genome, indicating that the genome sequence assemblies were of the highest quality (table S6).

Genome annotation

We performed repeat sequence and gene annotation for each *de novo* assembled genome. To further standardize the annotation results, we utilized the same method to re-annotate the downloaded primate assemblies except for human (*Homo sapiens*).

Repeat annotation

Transposable elements (TEs) were identified in each genome by means of an integration of homology-based and *de novo* approaches. We used RepeatMasker (v4.0.6) (http://www.repeatmasker.org/), TRF (v4.07) (*165*), and RepeatModeler (v1.0.8) (http://www.repeatmasker.org/RepeatModeler/) to identify repetitive sequences for each genome. Genome sequences were aligned to RepBase (v21.11) (https://www.girinst.org/repbase/) through RepeatMasker (v4.0.6), and then each hit was further subclassified into a specific category. Tandem repeats, which are defined as DNA sequences containing >2 adjacent copies, were identified using TRF (v4.07) (*165*) using the default parameters. At the protein level, RepeatProteinMask, an updated program in RepeatMasker (v4.0.6), was used to perform RMBlast against the TE protein database. In addition, RepeatModeler (v1.0.8) was employed to build a *de novo* repeat library based on each genome. Using this library as a database, RepeatMasker (v4.0.6) was run to identify repeats in each genome. All identified repeat elements were classified into diverse categories (DNA, LINE, SINE, LTR and Unknown) according to the classification generally employed in repeat databases. Repeat annotations were combined into a non-redundant repeat annotation for each genome. In total, we predicted that ~40.55% of bases in each primate genome are derived from TEs (table S15).

Protein-coding gene annotation

In order to predict a gene set for each genome, we employed a number of different methods and integrated the diverse results obtained. In total, we obtained ~20,000 genes for each species (table S7). Homologous proteins from well-annotated mammalian genomes including mouse (*Mus musculus*, GRCm38), human (*Homo sapiens*, GRCh38), chimpanzee (*Pan troglodytes*, GCF_002880755.1), gorilla (*Gorilla gorilla*, GCA_900006655.3) and orangutan (*Pongo abelii*, GCA_002880775.3) were mapped to each primate genome using TBLASTN (v.2.2.26) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (*166*) with an E-value cutoff of 1e-5. In addition to human (*Homo sapiens*, GRCh38) and mouse (*Mus musculus*, GRCm38), these great ape species with high genome assembly and annotation quality were often used as reference annotation resources applied for the annotation of genomes of other different primate groups [e.g., the Chinese rhesus macaque (*Macaca mulatta*) (*120*) and African mona monkey (*Cercopithecus mona*) (*27*) in Old World monkeys]. Together with a relatively small divergence time for Crown Primates (appearring at ~66.36 million years ago), gene annotations should be largely conserved across primate species. Therefore, we concluded that these species could be used as references for homologous annotations in this study. Multiple adjacent hits from the same protein were then linked together using genBlastA (v1.0.4) (*167*) in order to obtain the candidate gene boundaries. The aligned sequences as well as their query proteins were filtered and passed to GeneWise (v.2.4.1) (*168*) to search for accurately spliced alignments. We randomly selected 1,000 high score homology-based genes to train Augustus (v3.0.3) (*169*) for *de novo* prediction on the repeat N-masked genome with default parameters.

Single CDS genes < 300bp or multiple-CDS genes < 150bp were filtered out. Homology annotation results were first merged to generate a homology gene set. One best gene model was retained for each locus according to the GeneWise score. Finally, results obtained by means of these methods were merged into a non-redundant gene set. In brief, the homologous annotation was first used as a backbone and the *de novo* annotation was merged into it, thereby providing the additional CDS information. Genes from the *de novo* annotation that failed to overlap with genes from the homology annotation were aligned to Swiss-Prot (v2019_03) (https://www.uniprot.org/) and InterPro (v68.0) (http://www.ebi.ac.uk/interpro/) (*170*). Only genes that aligned to these databases and had multiple CDS were retained to be merged into the final gene set.

SM Text 3. Primate phylogeny analysis and divergence time

<u>Primate phylogeny analysis</u>

Whole genome alignments

We performed whole genome alignments across all 52 species representing 27 newly assembled primate species (table S1), 22 downloaded primate species (table S8), one long-read genome from *Nycticebus pygmaeus* reported in an accompanying paper (*32*), and two outgroup species, Sunda flying lemur (*Galeopterus variegatus*) and Chinese tree shrew (*Tupaia belangeri chinensis*) (table S8). First, the pairwise whole genome alignments were obtained between human (*Homo sapiens*) and each of the other species using LASTZ (v1.04.00) (https://www.geneious.com/plugins/lastz-plugin/) with the parameter settings '--step=19 --hspthresh=2200 --inner=2000 --ydrop=3400 --gappedthresh=10000 --scores=birdMatrix --format=axt'. The Chain/Net was performed by axtChain with parameters '–minScore=5000, -linearGap=loose'. Multi-way alignments across all species were generated using MULTIZ (v11.2) (*171*) by merging MAF files sequentially according to phylogenetic distance. Closely related species in one group were respectively merged, and then MAF files from all groups were merged into one. Segments in the final merged MAF file that were lacking more than five species were discarded; then, all segments were concatenated by means of an in-house script. In total, we generated final alignments containing ~ 433.51 Mbp for all 52 species (table S9).

Whole genome tree

The whole genome alignments of 50 primate genomes and two outgroup genomes (Sunda flying lemur and Chinese tree shrew) were utilized to construct the whole genome tree. Exascale Maximum Likelihood (ExaML) (v3.0.14) is relatively new code for large-scale phylogenetic analyses on supercomputers, and both addresses and provides generally applicable solutions to several performance bottlenecks in parallel phylogenetic likelihood calculations (*34*). In total, ~433.51 Mbp of syntenic

orthologous sequences was obtained from whole genome alignments across 52 species. These data were concatenated and used to construct the complete genome trees using ExaML (v3.0.14) (*34*) (https://github.com/stamatak/ExaML) under the GAMMA model run 100 times. The tree with the highest likelihood was selected as the final tree (Figs. 1 and S4).

Coalescent-based genome tree inference

Standard concatenation approaches may do not completely model the discordance among gene trees beyond differences in sequence evolution rates (*172*). Previous studies have also shown that the incomplete lineage sorting (ILS) could lead to incorrect topology, possibly due to estimation bias in concatenated analyses where the mixture of gene trees represents a model violation (*173*). These possible limitations can theoretically be overcome with multispecies coalescent methods using e.g., ASTRAL (*174*).

Therefore, we used IQ-TREE (v1.6.12) (http://www.iqtree.org/) to generate 28,034 trees for each 5 kbp window with the window interval of 50 kbp across the whole genome alignments of 52 species. Then all trees were parsed by ASTRAL (v5.5.4) to obtain the coalescent-based consensus species tree (Fig. S5).

Identification of pair-wise orthologous genes

Pair-wise orthologous genes were identified between human (*Homo sapiens*, GRCh38), each of the other primate species and the outgroup species, the Chinese tree shrew (*175*), based on criteria including reciprocal best blastp hit (RBH), gene synteny and genome synteny.

We first selected syntenic gene pairs and genomic syntenic regions using pair-wise genome alignments between human and other species. For each gene pair, we further calculated the synteny ratio of the two genes respectively (the length of the coding region overlapping with the syntenic region/total length of the coding region). The synteny ratio for each gene should be ≥ 30%. The alignment rate for a pair of genes is the length of aligned sequence of the genes/the shortest gene length. The alignment rate should also be ≥ 30%. Then, we filtered out those gene pairs without gene synteny. Two gene pairs on an identical chromosome/scaffold should meet the requirement that the number of genes between the two gene pairs is less than 5. Where there was only one gene pair on the chromosome/scaffold, the gene pair was retained for further analyses. Finally, we obtained those pair-wise orthologous genes meeting these criteria including reciprocal best blastp hit (RBH), gene synteny and genome synteny. Orthologous pairs from different comparisons were merged across 52 species including 50 primate species and two outgroup species, the Sunda flying lemur and Chinese tree shrew (*175*).

In total, 10,185 orthologous genes were identified across 52 species including 50 primate species and two outgroup species, the Sunda flying lemur and Chinese tree shrew (*175*), when permitting a species-missing threshold of ≤ 5.

Orthologous protein-coding gene tree

10,185 one-to-one orthologous genes were co-shared among 52 species including 50 primate species, Sunda flying lemur, and the Chinese tree shrew (*175*) when we permitted a species-missing threshold of ≤5. For each species, a "–" symbol was inserted if the locus could not be aligned to the human genome.

Then, 10,185 one-to-one orthologous genes across 52 species were concatenated to generate a supergene sequence, which was used for the construction of a phylogenetic tree. Orthologous gene trees were constructed using RAxML (v8.2.12) (*176*) with 100 bootstrap replicates under the GTRGAMMA model. The resulting tree with the highest likelihood score was selected as the best tree (Fig. S6).

Exon codon tree with 1st and 2nd positions

Considering the composition heterogeneity of different codon positions, we further partitioned the orthologous gene sequences into 1st, 2nd and 3rd codon positions. Then, the positions with codon 1st and codon 2nd for 10,185 one-to-one orthologous genes were further concatenated into a super-sequence to generate the exon codon tree with 1st and 2nd positions. The concatenated super-sequence from exon codons with 1st and 2nd positions was employed to generate the exon codon ML tree with 1st and 2nd positions using RAxML (v8.2.12) (*176*) with 100 bootstrap replicates under the GTRGAMMA model (Fig. S7).

Fourfold degenerate site tree

The 4d-sites of 10,185 one-to-one orthologous genes were extracted using our in-house scripts. All 4d-sites of orthologous genes were obtained and concatenated to construct ML trees based on RAxML (v8.2.12) (*176*) with 100 replicates for nodal supports under the GTRGAMMA model (Fig. S8).

Conserved non-exonic element tree

We first filtered the 52-way primate genome alignments and at least 90% of the species were presented in the alignments. The 4d-sites were employed to extract MAFs from the 52-way primate genome alignments. We employed the phyloFit program in PHAST (v1.5) (*177*) to estimate the non-conserved model. Then, we ran the phastCons program in PHAST (v1.5) (*177*) with the primates' non-conserved model to create the primates' conserved models with the option '--estimate-rho'. The highly conserved elements were predicted in primates using the phastCons program in PHAST (v1.5) with options '--most-conserved –score options'. Eventually, we identified 1,309,699 HCEs with an average length of 147 bp and a total length of 192 Mbp.

We generated the conserved non-exonic elements by excluding exon regions from highly conserved elements and finally obtained 1,118,099 conserved non-exonic elements with a total length of 156.9 Mbp and an average length of 140 bp.

These conserved non-exonic element data, generated across 52 species, were then concatenated and used to construct the conserved non-exonic element tree using

ExaML (v3.0.14) (*34*) (https://github.com/stamatak/ExaML), under the GAMMA model and run 100 times. The tree with the highest likelihood was selected as the final tree (Fig. S9).

<u>Primate divergence time evaluation</u>

To evaluate precisely the divergence times in primates, we added a further two outgroup species genomes, specifically the domestic cat (*Felis catus*; v9.0.97) and the pig (*Sus scrofa*; v11.1.97) from Ensembl 97 (http://jul2019.archive.ensembl.org/index.html) to our previous whole genome alignments across 52 species based on the identical pipeline above. Taken together, we obtained whole genome alignments of 54 species including 50 primate species and 4 outgroup species.

Then, our whole genome alignments across 54 species and utilized the MCMCtree program in PAML (v4.9) (http://abacus.gene.ucl.ac.uk/software/paml.html) (*40*) to infer the primate divergence times. Eleven fossils from Vanderpool *et al*. (*39*) and Perelman *et al*. (*35*) were used to calibrate the nodal divergence time, and the divergence time (~76-88 Mya) between primates and *Tupaia* was obtained from TimeTree (http://www.timetree.org/). The dating of eleven fossils (*35, 39*) was listed as follows: "(5.7, 10) Mya for the most recent common ancestor (TMRCA) of *Homo-Pan*; (14, 34) Mya for TMRCA of Hominidae; (3.5, 4.5) Mya for TMRCA of the Theropithecus clade; (6, 8) Mya for TMRCA of Papionini; (25, 34) Mya for TMRCA of Catarrhini; (36, 50) Mya for TMRCA of Simiiformes; (20.5, 26.5) Mya for TMRCA of Platyrrhini; (37, 43) Mya for TMRCA of galagids and lorisids; (38, 56) Mya for TMRCA of Strepsirrhini; (55.8, 65.8) Mya for TMRCA of Primate; (61, 165) Mya for TMRCA of Primatomorpha" (Fig. S11).

SM Text 4. Structural evolution of primate genomes

<u>Karyotype evolution</u>

Karyotype evolution is involved in many aspects of biology e.g. speciation and genome evolution (*42, 43*). In this primate genome project, a mass of species-sampling genomes of high quality (from Pacific Biosciences and Oxford Nanopore Technology platforms) allowed us to accurately explore the evolution of primate karyotypes.

Here, we utilized the human genome (*Homo sapiens*; GRCh38) as a reference to reconstruct the ancestral nodal karyotype in primates. First, we selected 37 primate genomes with scaffold N50 >=13 Mbp (table S10) as inputs. Then we used the DESCHRAMBLER algorithm (https://github.com/jkimlab/DESCHRAMBLER), Chains and Nets generated by LASTZ (v1.04.00) (*178*) (https://www.geneious.com/plugins/lastz-plugin/) to generate conserved syntenic fragments across 37 species (table S10). Syntenic fragments of length ≥ 500 kbp were used to build the conserved syntenic fragments. Finally, the structures of ancestral

genomes were constructed by using ANGES (v1.01) (*179*) across ancestral nodes in phylogenetic tree: Primates, Strepsirrhini, Simiiformes, Platyrrhini, Catarrhini, Cercopithecoidea, Hominoidea, Hominidae, Hylobatidae, Homininae, *Pongo*, Hominini and *Pan* (Fig. 2). GRIMM (v2.01) (*180*) was utilized to analyze genome rearrangements (e.g. chromosome reversals, fusions, fissions, translocations) from ancestral genomes to next nodal genomes.

We further investigated the karyotype changes from primate common ancestor to human (*Homo sapiens*; GRCh38). The different colors represent different chromosomes in human (*Homo sapiens*; GRCh38), whilst the gray lines denote the movement of genome markers from ancestral nodes to next nodes (Fig. 2).

Evolution of segmental duplications

Segmental duplications are DNA fragments longer than 1 kbp, distributed within and between chromosomes and sharing more than 90% identity at the DNA sequence level (*181-183*).

In this study, we considered segmental duplications of length ≥5 kbp and identity >90%. We selected 39 well-assembled genomes with a high scaffold N50 and a genome of flying lemur and used ASGART (v2.3) (*184*) to predict segmental duplications for each species (table S11). Pairwise whole genome alignments with reciprocal best hit using human (*Homo sapiens*; GRCh38) as reference were produced. Then lineage-specific segmental duplications were further grouped according to genome synteny criteria. All segmental duplications were grouped based on their co-shared sequence homology. For lineage-specific SDs, they were required to meet the following conditions: firstly, these segmental duplications had to be found only in species from the identical group although they were missing from other groups; secondly, the number of species in the group had to be greater than a certain threshold (table S12).

Comparative genomics analyses of genome sizes and transposable elements in primates

Genome size analyses in different evolutionary clades of primates

In this study, we integrated 27 newly assembled primate genomes and 23 downloaded genomes. The primate genome sizes ranged from ~2.12 Gbp (*Lemur catta*) to ~3.29 Gbp (*Pan paniscus*, GCA_000258655.2). Thus, we found an average primate genome size of ~2.85 Gbp, which is larger than that of Ruminants (~2.7 Gbp), Chiroptera (~2.35 Gbp) and Carnivora (~2.3 Gbp) (*48, 49*).

Then, we further compared the heterogeneity of genome sizes for different evolutionary clades of primate including Hominoidea, Old World monkeys, New World monkeys, lemurs [*Microcebus murinus* (GCA_000165445.3), *Prolemur simus* (GCA_003258685.1), *Lemur catta*), and *Daubentonia madagascariensis* (newly assembled genome in this study)], and Lorisiformes [*Loris tardigradus* (newly assembled genome in this study), *Nycticebus pygmaeus (32)*, *Nycticebus bengalensis*

(newly assembled genome in this study), *Galago moholi* (newly assembled genome in this study), and *Otolemur garnettii* (GCA_000181295.3)] (Fig. S16A). *P* values for genome size differences of evolutionary clades in primates were calculated by the Mann-Whitney U test.

Evolution of transposable elements

Transposable elements (TEs) have played important roles in the evolution of genomes and phenotypes (*185-187*). In this study, we found significant differences in genome size between the lemurs and other primate evolutionary clades (e.g. Lorisiformes, New World monkeys, Old World monkeys, Hominoidea) by Mann-Whitney U test (*P* < 0.05), due to differences in the relative representation of TEs in the genomes (Figs. S16 and S17). According to TE annotations, we subdivided TEs into the categories DNA, LINE, SINE, LTR, Other and Unknown classes (table S15). The 67 TEs with highest percentage content in 50 primate genomes and the outgroup species genome are given in Fig. S18. The statistical difference analysis of TE percentage content for different primate clades was performed by the Mann-Whitney U test (*P* < 0.05) (Figs. S16 and S17).

According to the TE annotations for each species [Chinese tree shrew (*Tupaia belangeri chinensis*; http://www.treeshrewdb.org) as an outgroup] in this study, we extracted all Alu sequences from the primate and Chinese tree shrew genomes. Next, CD-HIT (*188*) with parameters "-c 0.8 -n 5 -d 0 -M 16000 -T 10" was used for Alu clustering analysis of each species. Finally, where there were three or more Alu sequences in this cluster, we selected the longest Alu sequence as the consensus sequence. For each species, the Alu consensus sequences obtained above were applied as the library, and RepeatMasker (v4.0.6) (http://www.repeatmasker.org/) was used to annotate the Alu transposons in the corresponding species.

We calculated the Alu insertion times for primate (Fig. S19) and tree shrew genomes using the algorithm T = K/2r, where K is the Kimura distance-based copy divergence of TEs and r is the nucleic acid substitution rate. The K-value was obtained from RepeatMasker (v4.0.6) (http://www.repeatmasker.org/). To estimate r-values for primate and tree shrew, we used the alignments of LASTZ (v1.04.00) (*178*) and MULTIZ (v11.2) (*171*) along with genomes in our evolutionary analyses, with the human genome as the reference sequence. With the whole-genome alignments, we used the MSA view tool in PHAST (v1.2.1) (*177*) to extract 4d-site alignments based on the human gene annotations. The phyloFit program in PHAST (v1.2.1) (*177*) was used to estimate the phylogenetic model, with tree topology (our primate phylogeny analyses) as an input parameter. The branch length results were represented in units of substitutions per site. We calculated the root-to-tip substitution rates from the most recent common ancestors in clades to each lineage, and then divided the root-to-tip substitution rates by the divergence time between the Chinese tree shrew and the most recent common ancestor of primates in this study.

Comparative analyses of the nucleotide substitution rates in primates

The nucleotide substitution rates determined in different primate lineages have exhibited heterogeneity as a consequence of the limited genomic data available from primates (*35, 56*). In this study, a considerable number of high-quality primate reference genomes from long-read sequencing platforms (Pacific Biosciences and Nanopore PromethI ON) allowed us to arrive at a robust assessment of the nucleotide substitution rates along diverse primate lineages on a genome-wide scale.

In this study, the nucleotide substitution rates (in units of substitutions per site per million years) in primate genomes were estimated by comparisons of the fourfold degenerate (4d) sites in coding regions and divergence times between primate species (Figs. 3C, S21 and table S16). Then, the nucleotide substitution rates of different evolutionary primate clades were analyzed comparatively by box plotting (Figs. 3C and S21). Our comparative analyses involved different evolutionary clades of primates including great apes (here representing the Hominidae), gibbons, Old World monkeys, New World monkeys, Strepsirrhini, lorises (Lorisidae) (including *Loris tardigradus*, *Nycticebus pygmaeus* and *Nycticebus bengalensis* in this study), Hominoidea, Catarrhini, Simiiformes, Haplorrhini, lemurs (Chiromyiformes+Lemuriformes) and lorisoids (Lorisiformes) (including *Loris tardigradus*, *Nycticebus pygmaeus*, *Nycticebus bengalensis*, *Galago moholi* and *Otolemur garnettii*).

Evolutionary analyses of protein-coding genes in primates

Based on 10,185 identified orthologous genes across 52 species including 50 primates and two outgroup species, Sunda flying lemur (*Galeopterus variegatus*) and Chinese tree shrew (*Tupaia belangeri chinensis*; TreeshrewDB v2.0; http://www.treeshrewdb.org) by methods including RBH, gene synteny and genome synteny (See the section—**SM Text 3. Primate phylogeny analysis and divergence time:** *Identification of pair-wise orthologous genes*) and primate species tree (Fig. 1) in this study, we explored the evolutionary constraints operating on coding regions for each orthologous gene. To this end, we calculated the dN/dS ratio [rates of nonsynonymous (dN) to synonymous (dS) substitutions] of each orthologue by the codeml program in PAML (v4) (*40*) under a free-ratio model. Meanwhile, the codeml algorithm under a free-ratio model could also give other evolutionary parameters (i.e., N, S, dN, dS, N*dN, and S*dS) for each orthologous gene along with all branches in primates.

Next, we characterized the evolutionary rates of tissues by following tissue-specific genes in different evolutionary nodal lineages across the primate phylogeny. Previous studies have indicated that the expression patterns of tissue-specific expressed genes are conserved during mammalian evolution (*63, 64*). Therefore, in this study, we downloaded the gene expression matrix of 30 tissues from 7,862 human samples from the Genotype-Tissue Expression (GTEx) project (*189*) (https://commonfund.nih.gov/GTEx/). We were then able to utilize the tissue-specific expressed genes of human (GRCh38) to infer the evolutionary rates of diverse tissues

across different evolutionary lineages in primates. The evolutionary rate of each tissue for each branch was obtained by the median of the dN/dS values for tissue-specific expressed genes (Fig. 3D).

We identified tissue-specific expressed genes for each tissue based on t-statistics in human (GRCh38) according to a previous study (*63*). Briefly, we performed t-statistics on single-copy orthologous genes between the target tissue samples and all other samples, and then ranked the t-values from high to low. According to the ranking, the top 5% of genes based on *t*-values were defined as tissue-specific expressed genes in human (GRCh38).

SM Text 5. Genomic features and phenotypic adaptations in primate evolution

<u>Identification of positively selected genes in crucial primate evolutionary nodes</u>

The 10,185 orthologous genes among 52 species, including 50 primates and two outgroup species (Sunda flying lemur and Chinese tree shrew), were identified by a combination of methods including RBH, gene synteny and genome synteny (See the section—**SM Text 3. Primate phylogeny analysis and divergence time:** *Identification of pair-wise orthologous genes*).

The orthologous gene coding sequences across these species were aligned by MUSCLE (v3.8.31) (*190*) and the low-quality aligned regions were further trimmed by Gblocks (v0.91b) (*191, 192*) with the parameters "./Gblocks $i -t=c -b4=5". The aligned orthologous genes with CDS length < 100 bp were removed for our downstream evolutionary analyses. Based on our reliably constructed species-guided tree topology (Fig. 1), the branch-site model in PAML (v4) (*40*) and the likelihood rate test (LRT) between null hypothesis and alternative hypothesis were utilized to detect positively selected genes with $P < 0.05$ (χ^2 test) in crucial evolutionary lineages of primates among the single-copy orthologous genes. The Bayes empirical Bayes algorithm in PAML (v4.4) (*40*) was applied to calculate the posterior probabilities of inferred positively selected sites. Thus, we inferred the positively selected genes in 14 crucial evolutionary lineages including the ancestral branch of Strepsirrhini, the ancestral branch of Haplorrhini, the Western tarsier lineage, the ancestral branch of Simiiformes, the ancestral branch of New World monkeys (Platyrrhini), the ancestral branch of Catarrhini, the ancestral branch of Old World monkeys (Cercopithecoidea), the ancestral branch of Colobinae, the ancestral branch of Hominoidea, the ancestral branch of Hominidae, the ancestral branch of Hylobatidae, the ancestral branch of Homininae, the ancestral branch of Hominini (*Homo-Pan*), and the Human lineage.

To infer the positively selected genes of the common ancestral branch of Primates, together with Chinese tree shrew, we also added five extra outgroup genomes, namely Sunda flying lemur (*Galeopterus variegatus*, NCBI project id: PRJNA399345), mouse (*Mus musculus*, GRCm38.p6), domestic cat (*Felis catus*; v9.0.97), domestic dog (*Canis lupus familiaris*; CanFam3.1) and pig (*Sus scrofa*; v11.1.97) from Ensembl 97

(http://jul2019.archive.ensembl.org/index.html) to the identifications of orthologous genes among 56 species based on the identical pipeline including RBH, gene synteny and genome synteny (See the section—**SM Text 3. Primate phylogeny analysis and divergence time:** *Identification of pair-wise orthologous genes*). We identified 9,907 orthologous genes among 56 species including 50 primates and 6 outgroup species. According to the consistent pipeline for the identification of positively selected genes, we inferred the potential positively selected genes along with the common ancestral branch of primates.

In total, we inferred the positively selected genes from 15 evolutionary lineages for primates, and these positively selected genes were checked manually for their sequence alignments.

<u>Evolutionary analyses of gene families in primates</u>

Identification of gene families in primates

DNA and protein data for 5 outgroup species, specifically mouse (*Mus musculus*, GRCm38.p6), dog (*Canis lupus familiaris*; CanFam3.1), pig (*Sus scrofa*; v11.1.97), Sunda flying lemur, and Chinese tree shrew, were obtained from the Ensembl (v97) and TreeshrewDB (v2.0) (*175*) databases. Meanwhile, 27 newly assembled primate genomes (table S1), one long-read genome from *Nycticebus pygmaeus* reported in an accompanying paper (*32*), and 22 downloaded primate genomes (table S8) were integrated into our analyses of gene families. For those genes with alternative splicing variants, the longest transcripts were selected to represent the genes. In this study, we used the TreeFam algorithm (*193*) (http://www.treefam.org) to define a gene family as a group of genes that descended from a single gene in the last common ancestor of the species considered. The specific screening pipeline was similar to a previous study (*194*) as follows: 1) Blastp (*166*) was applied to all protein sequences against a database containing a protein dataset of all species with an e-value of 1e-07 and conjoined fragmental alignments for each gene pair. We assigned a connection (edge) between the two nodes (genes) if more than 1/3 of the region aligned to both genes. An Hscore that ranged from 0 to 100 was used to weigh the similarity (edge). For two genes (G1 and G2), the Hscore was defined as a score (G1G2)/max (score (G1G1), score (G2G2)) (the score here is the raw Blast score). 2) Extraction of gene families (clustered by Hcluster_sg). We utilized the average distance for the hierarchical clustering algorithm requiring the minimum edge weight (Hscore) to be larger than 5, and the minimum edge density (total number of edges/theoretical number of edges) to be larger than 1/3.

In total, we identified 22,174 gene families among 55 species in this study for the downstream analyses.

Expansions of gene families in primates

After using the Treefam algorithm (*193*) to construct gene families from among 55 species, including 50 primates and 5 outgroup species, we further ascertained the

expansions of gene families for some crucial evolutionary nodes in primates. Computational Analysis of gene Family Evolution (CAFE) is a tool for the statistical analysis of the evolution of gene family size (*195*). If the copy number of the gene family in the detected branch lineage was higher than that of its most recent common ancestral branch lineage, then the gene family was defined as being substantially expanded in the detected branch lineage, according to the output of the CAFE (*195*) algorithm.

In order to compare the copy number variations of gene families across different evolutionary clades and to infer the significantly expanded gene families in specific clades, we used an unpaired *t*-test to determine the significance of differences in copy number for different nodes. Each of the target groups (the targeted evolutionary clades) and the contrast groups (all other remaining clades) should contain at least 3 species. Briefly, our identification of significantly expanded gene families in specific nodes had to meet the following rigorous screening procedures:

1) The *P* values of expanded gene families in specific clades using the unpaired *t*-test had to be less than 0.05. Further, the *P* values were corrected by Benjamini-Hochberg false discovery rate (FDR) (*196*) and had to be ≤ 0.05.

2) For the targeted evolutionary clade, the average copy number of gene families from all species had to be more than 4 times that of the control clades (other remaining clade species) (Log_2(targeted clade/control clades)≥ 2). If the average copy number of the gene families in the control clades was equal to zero while the copy number of the gene families in the targeted clade was greater than zero, then the gene families were regarded as potential candidates for expansion in the specific evolutionary clade.

Thus, we inferred the significantly expanded gene families in the 13 evolutionary nodes for primates. These evolutionary nodes included the ancestral branch of the Primates, the ancestral branch of the Haplorrhini, the ancestral branch of the Strepsirrhini, the ancestral branch of the Simiiformes, the ancestral branch of the New World monkeys, the ancestral branch of the Catarrhini, the ancestral branch of Old World monkeys, the ancestral branch of the Colobinae, the ancestral branch of the Hominoidea, the ancestral branch of the Hominidae, the ancestral branch of the Hylobatidae, the ancestral branch of the Homininae, and the ancestral branch of the Hominini (*Homo-Pan*) (Fig. 4A).

Lineage-specific accelerated regions

Lineage-specific accelerated regions were identified from highly conserved elements in the 52-way primate alignments (See **SM Text 3. Primate phylogeny analysis and divergence time:** *Conserved non-exonic element tree*).

In order to find accelerated regions specific to certain subgroups, we estimated acceleration scores for the highly conserved elements using the phyloP program in PHAST (v1.5) (*177*) with the parameters "--method LRT, --branch and --mode ACC" for 12 selected subgroups. The significant acceleration at FDR adjusted p-values ≤ 0.05 was considered in further analyses. Lineage-specific accelerated regions in coding

regions and the Y chromosome were removed. The detailed screening pipeline is to be found in an accompanying paper (*72*).

Finally, according to gene annotations, we obtained the nearest genes to the lineage-specific accelerated regions were all located within 500kb of these genes) for each crucial evolutionary lineage. In all, we inferred the significant lineage-specific accelerated regions in 13 evolutionary nodes for primates. These evolutionary nodes represented the ancestral branch of the Haplorrhini, the ancestral branch of the Strepsirrhini, the ancestral branch of the Simiiformes, the ancestral branch of the New World monkeys, the ancestral branch of the Catarrhini, the ancestral branch of the Old World monkeys, the ancestral branch of the Colobinae, the ancestral branch of the Hominoidea, the ancestral branch of the Hominidae, the ancestral branch of the Hylobatidae, the ancestral branch of the Homininae, and the ancestral branch of the Hominini (*Homo-Pan*) and human (Fig. 4A).

Gene enrichment analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID) (v6.8) (*197-204*) (https://david.ncifcrf.gov/) was used to perform the gene function enrichment analysis, including Functional Category (e.g. UP KEYWORDS), Gene Ontology (e.g. GOTERM BP DIRECT, GOTERM CC DIRECT, GOTERM MF DIRECT), Pathway (e.g., KEGG PATHWAY), and Tissue Expression (e.g., UP_TISSUE) for candidate gene lists under a current background *Homo sapiens* with a two-tailed corrected Fisher's Exact test ($P < 0.05$).

For a gene list (e.g., lineage-specific positively selected genes and genes associated with lineage-specific accelerated regions), the highly expressed genes in specific tissues (or tissue-specific highly expressed genes) were assigned according to the Tissue Expression Database (e.g., UP_TISSUE) in DAVID (v6.8) (*197-204*) under the annotation background of *Homo sapiens*.

SM Text 6. Primate demographic history analyses

We inferred the demographic history for 48 of 50 primate species by applying the pairwise sequentially Markovian coalescent model (PSMC) (*151*). For two species, specifically *Chlorocebus sabaeus* and *Piliocolobus tephrosceles*, the demographic history analyses were excluded because of the lack of short-read genome sequencing data for this study. To ensure the quality of consensus sequences, >46.8x clean short-read sequencing data for newly assembled genomes (table S3) and >30x clean short-read data for downloaded genomes (table S41) were used, respectively. For each species, we first excluded the contigs/scaffolds that were only aligned to the sex chromosomes of human, according to the 52-way whole genome alignments.

For each species, the short-read sequencing data were mapped to their reference genomes using BWA-MEM (v0.7.12) (*205*) with default parameters. Then, SAMtools (v1.3.1) (*206*) was used to sort and remove PCR duplicates. Alignments around indels

were realigned using the indelRealigner program in GATK (v2.6.5) (*207*). After obtaining the bam files, the average depth statistics were calculated for each alignment using BEDTools (v2.26.0) (*208*) and an in-house script. SNPs were identified using BCFtools (v1.9) (https://samtools.github.io/bcftools/) and then filtered for proximity to indels (--SnpGap 6) and depth (1/3 average depth $\leq$DP $\leq$2 * average depth) using the BCFtools filter. Indels were removed from the final output (-v snps) using the BCFtools view. A diploid consensus genome sequence was generated using the "vcfutils.pl vcf2fq" function in BCFtools (v1.9). We transformed the format of the consensus sequence by using fq2psmcfa (with the parameters: fq2psmcfa -q20). The generation times (g) of the different species were taken from the summary listing in table S41. The mutation rates per million years were estimated as given in table S16. The mutation rate per generation was estimated by multiplying the per year mutation rates by the generation time. The population size histories of 48 primates were displayed in Fig. S36. The genome-wide diversity for each species was calculated using VCFtools (v0.1.11) (*209*) (http://vcftools.sourceforge.net) with non-overlapping 50-kb (--window-pi 50000) based on the filtered individual SNP files. In order to reveal the dynamic trends of Ne across multiple species, we used a value of normalized Ne. The normalized Ne was calculated by dividing the real value of Ne at each time point by the maximum value for the same species throughout the timescale. Correlation analyses between Ne and nucleotide diversity were estimated by cor.test function (Pearson's product-moment correlation) in R (v3.5.0) after phylogenetic correction [Phylogenetic Independent Contrast (PIC)] using the ape library (http://ape-package.ird.fr/) in the R software.

Fig. S1. Hi–C interactions of chromosomes. Strong interactions are indicated in dark red, weak interactions in light red. The genomes of four species including *Hylobates pileatus*, *Colobus guereza*, *Saguinus midas* and *Nycticebus bengalensis* were assembled into chromosome levels in this study according to the previous method (*28*).

Hylobates pileatus

Colobus guereza

Saguinus midas

Nycticebus bengalensis
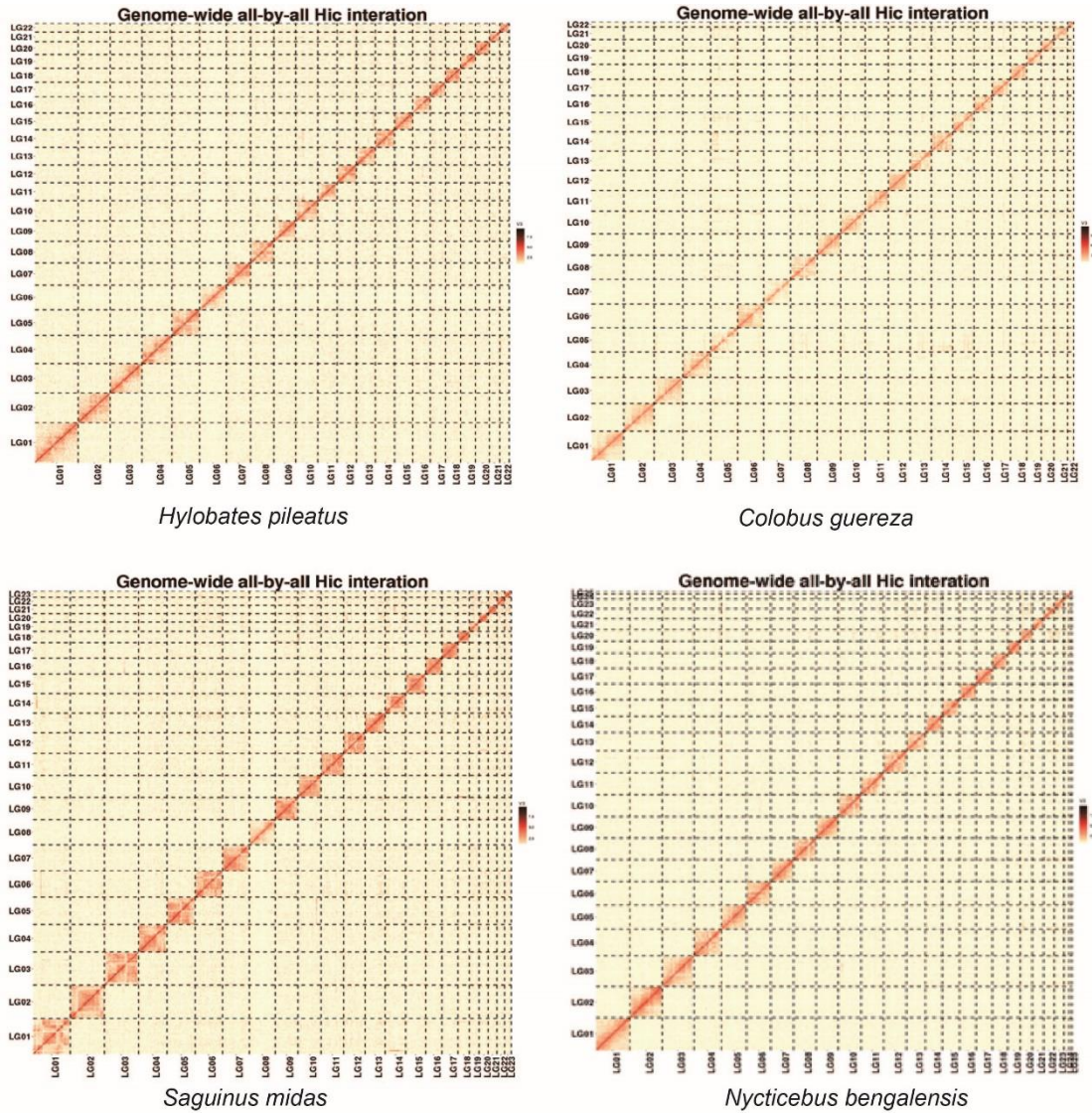
Fig. S2. Evaluation of genome size in primates using 17-Kmer analysis. X-axis represents the depth, whilst the y-axis represents the frequency.
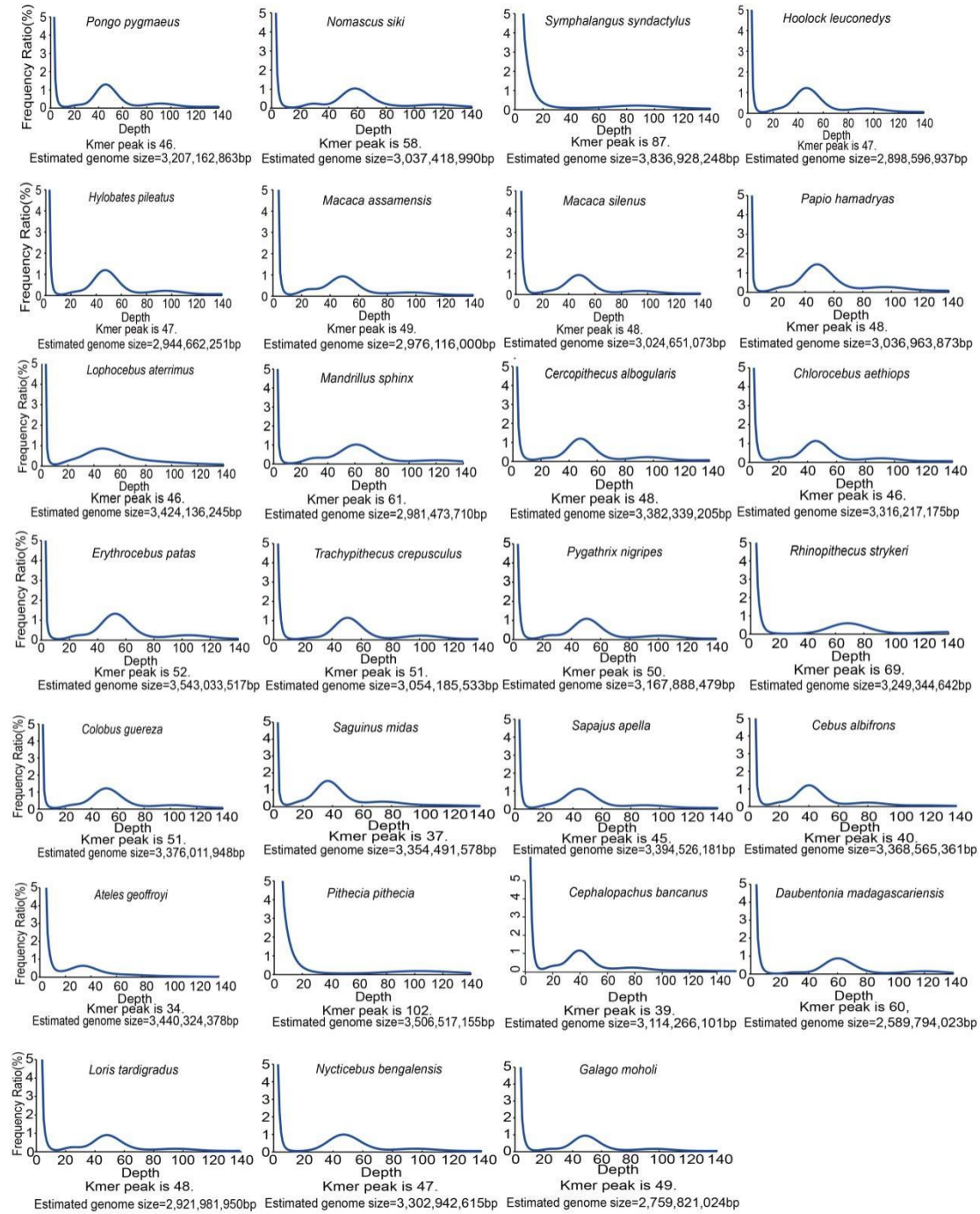
Fig. S3. Summary of primate-specific highly conserved elements. The distribution of primate-specific highly conserved elements is shown by using gene/genomic region partitions. These primate-specific highly conserved elements were divided into two categories (A and B) based on the CONS scores in homologous sequences in the outgroup species. The nine outgroup species were *Galeopterus variegatus*, *Tupaia belangeri*, *Mus musculus*, *Oryctolagus cuniculus*, *Sus scrofa*, *Felis catus*, *Canis familiaris*, *Phascolarctos cinereus* and *Ornithorhynchus anatinus*.
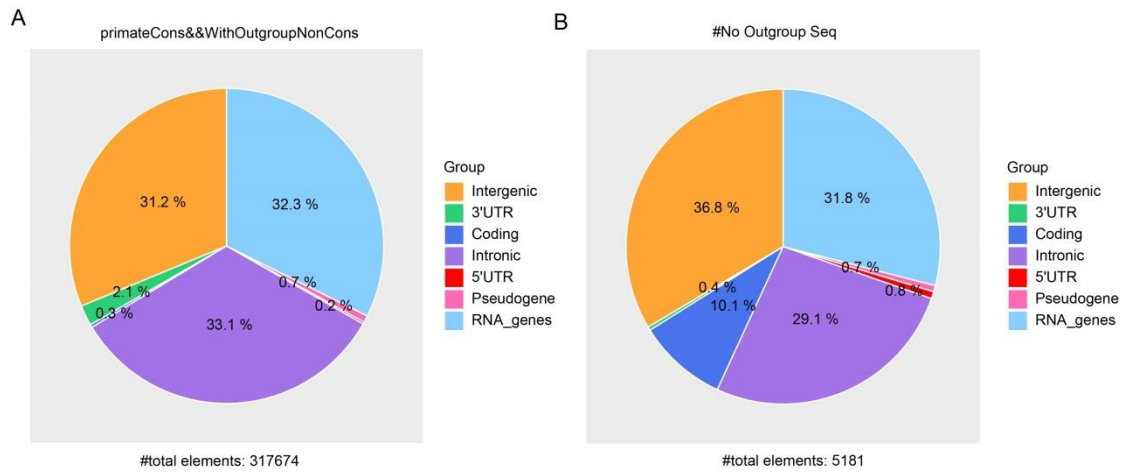
Fig. S4. Maximum likelihood whole-genome nucleotide evidence tree of primates estimated by ExaML. The whole genome alignments (~ 433.5 Mbp) of 52 species (50 primate species + flying lemur [*Galeopterus variegatus*] + tree shrew [*Tupaia belangeri*]) were used to construct a whole genome ML tree under a GTR+GAMMA model. The bootstraps with 100 replicates were applied to produce the nodal supports. The gibbon branches are highlighted by a pink box. The nodal support and branch length are shown respectively. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
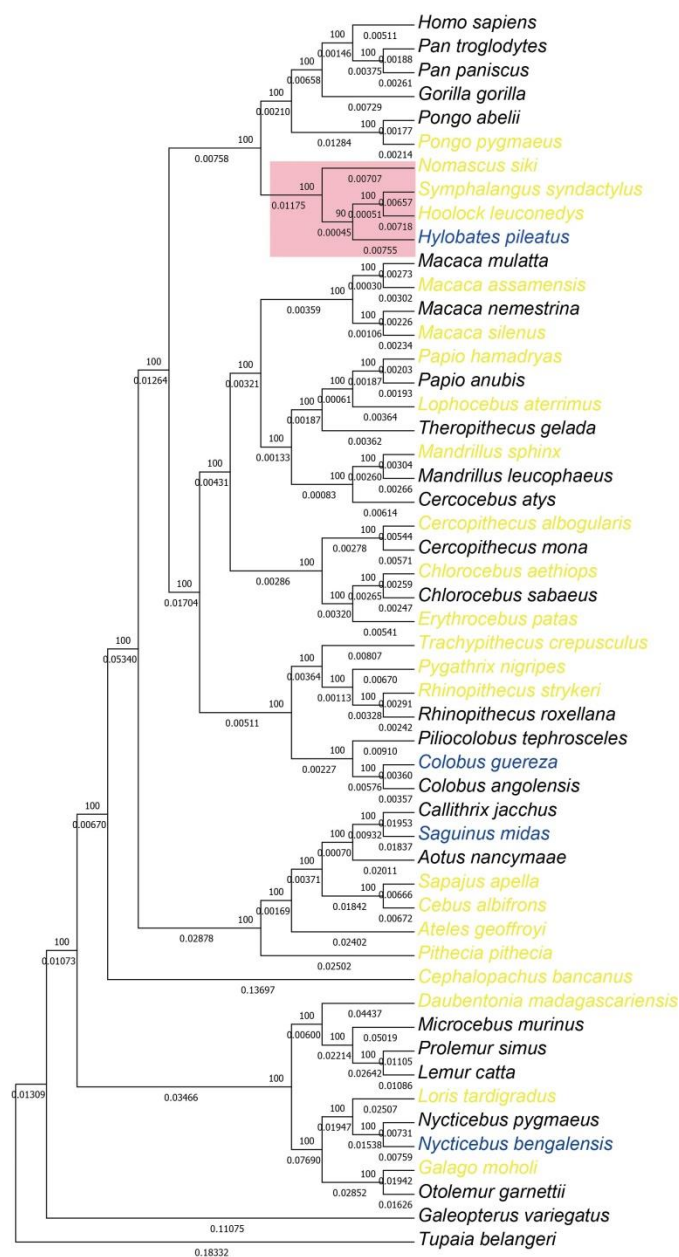
Fig. S5. Coalescent-based whole-genome nucleotide evidence tree estimated by ASTRAL. The whole-genome alignments of 52 species (50 primates+Sunda flying lemur+Chinese tree shrew) were used to construct the whole-genome coalescent-based ASTRAL tree. The bootstrap with 100 replicates was applied to produce nodal supports. The gibbon branches are highlighted by a pink box. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
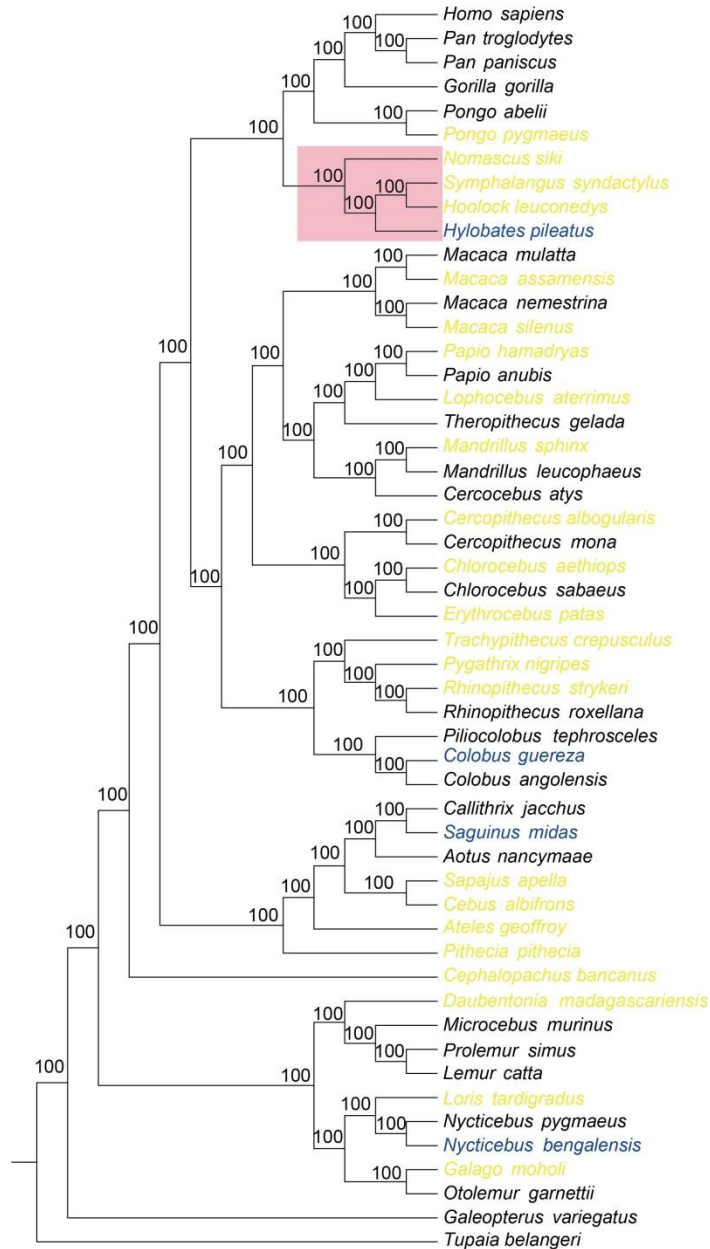
Fig. S6. Maximum likelihood phylogenetic tree of primates estimated using orthologous protein-coding genes. 10,185 orthologous genes among 52 species (50 primate species + flying lemur [*Galeopterus variegatus*] + tree shrew [*Tupaia belangeri*]) were obtained and concatenated into a single super gene to construct a ML tree based on RAxML under a GTR+GAMMA model. The bootstraps with 100 replicates were applied to produce nodal supports. The gibbon branches are highlighted by a pink box. The nodal support and branch length are shown respectively. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
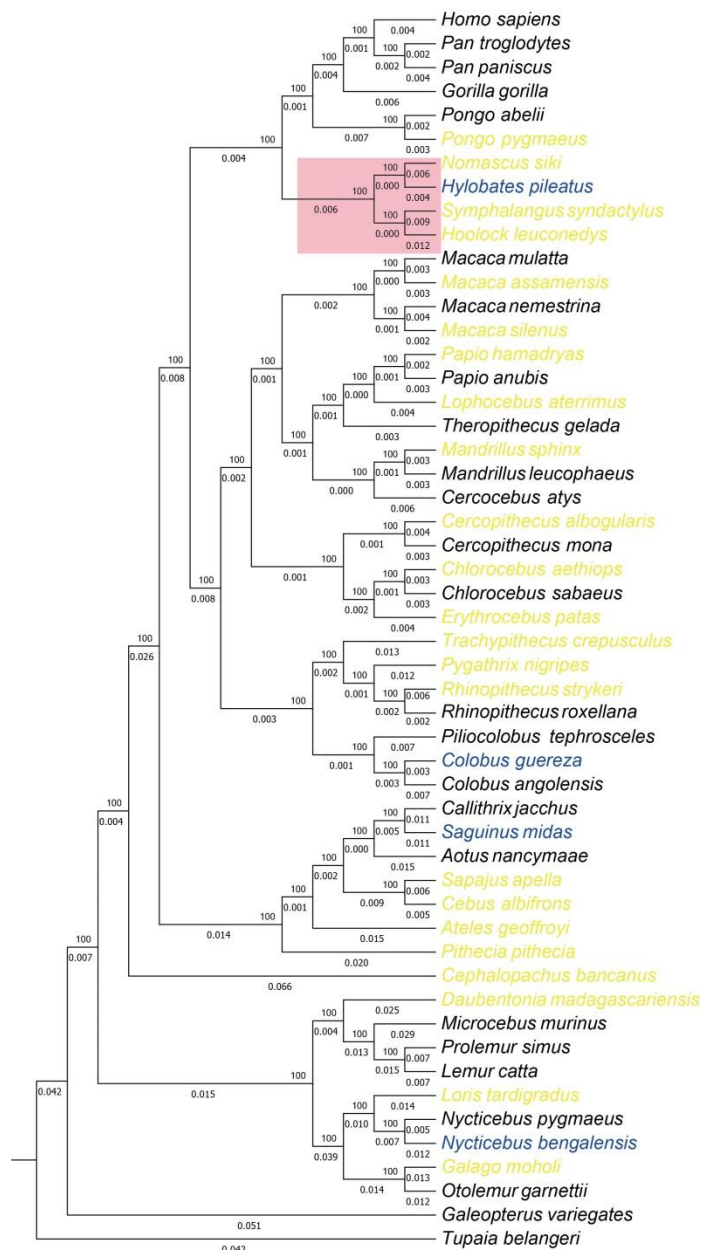
Fig. S7. Maximum likelihood phylogenetic tree of primates estimated using the exon codons with 1st and 2nd positions. The 1st and 2nd codon nucleotides of 10,185 orthologous coding genes were concatenated to construct a ML tree based on RAxML under a GTR+GAMMA model. The bootstraps with 100 replicates were applied to produce the nodal supports. The gibbon branches are highlighted by a pink box. The nodal support and branch length are shown respectively. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
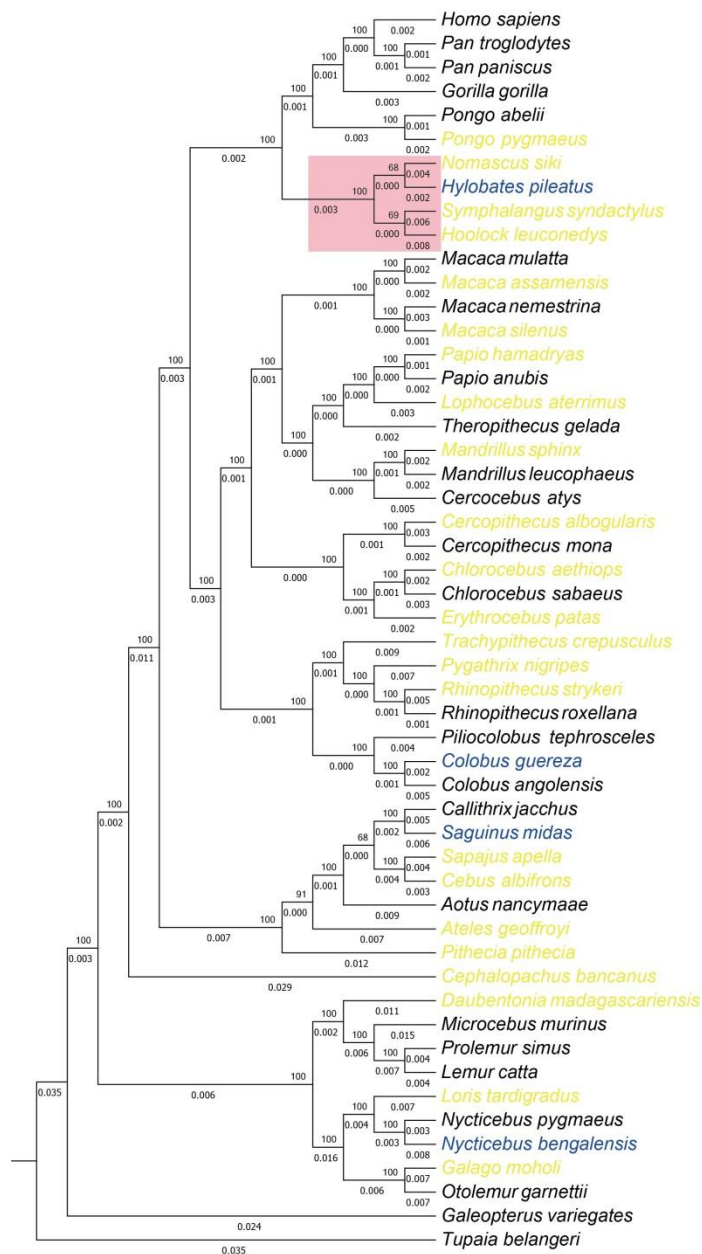
Fig. S8. Maximum likelihood phylogenetic tree of primates estimated using four-fold degenerate sites. The four-fold degenerate sites of 10,185 orthologous coding genes were extracted and concatenated to construct a ML tree based on RAxML with 100 replicates for nodal supports under a GTR+GAMMA model. The gibbon branches are highlighted by a pink box. The nodal support and branch length are shown respectively. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
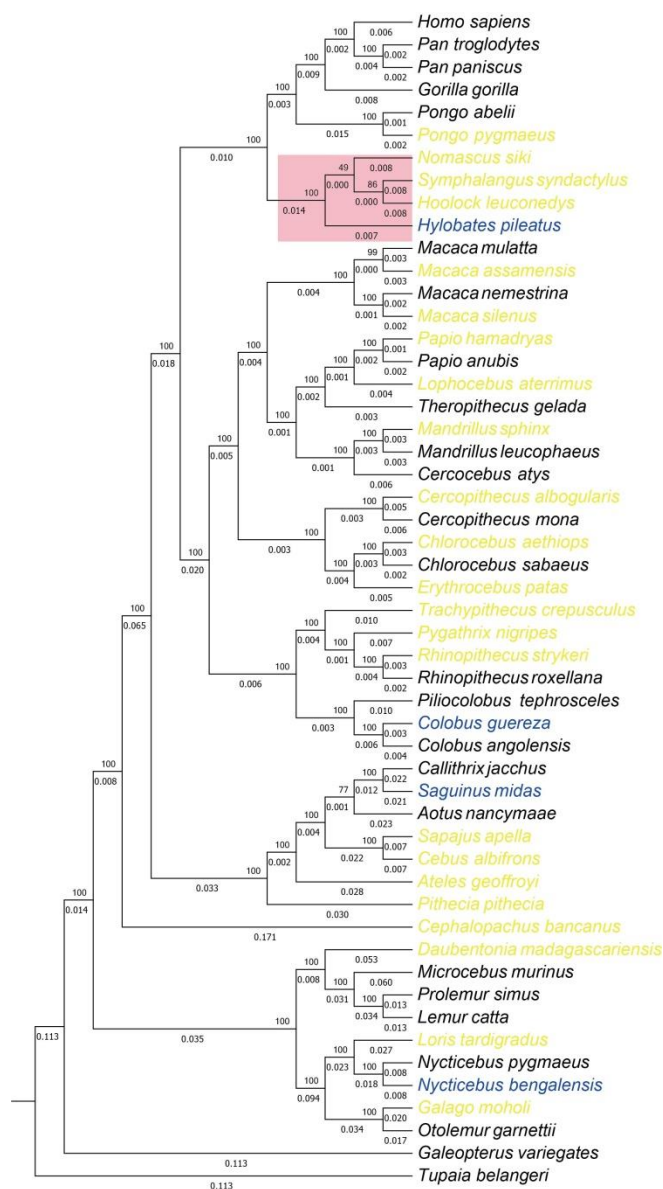
Fig. S9. Maximum likelihood phylogenetic tree of primates estimated using conserved non-exonic elements. The conserved non-exonic elements (~168.3 Mbp) were extracted and concatenated to construct a ML tree based on ExaML with 100 replicates for nodal supports under a GTR+GAMMA model. The gibbon branches are highlighted by a pink box. The nodal support and branch length are shown respectively. The yellow and blue species names represent those genomes newly produced in this study. The genomes of the species marked in blue were assembled at the chromosome level.
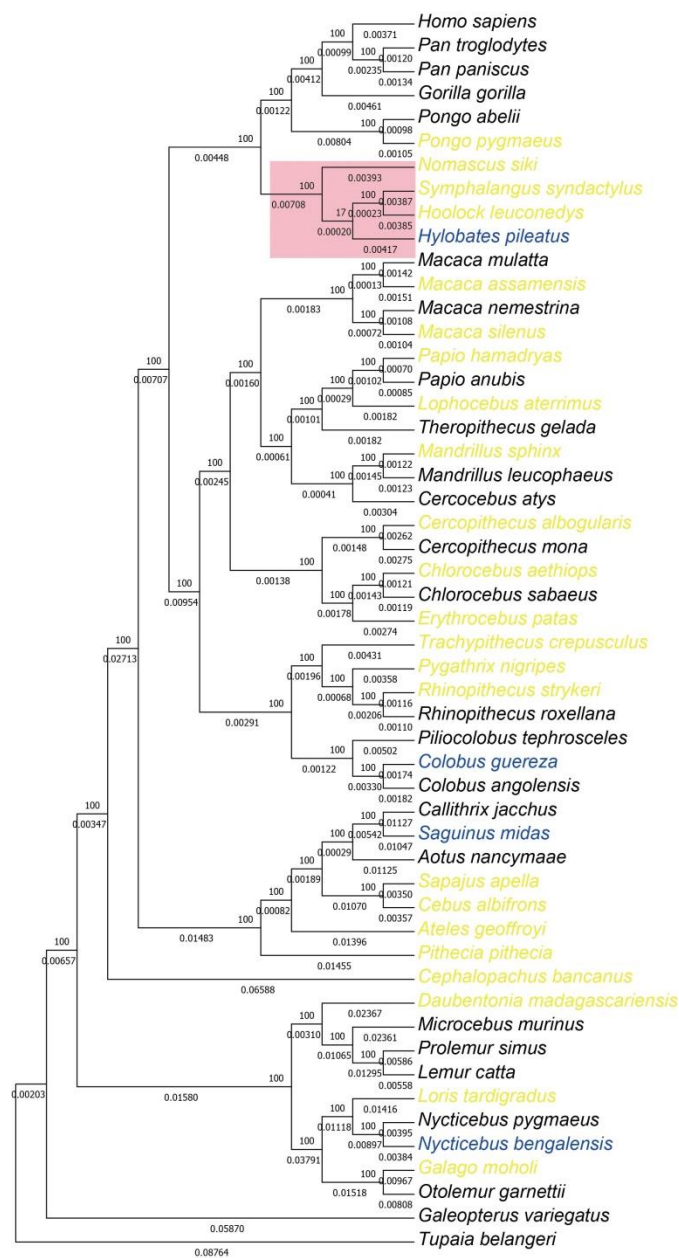
Fig. S10. CoalHMM runs of the two gibbon branches. CoalHMM can be used to infer the proportion of incomplete lineage sorting given three species and an outgroup. CoalHMM requires that the species topology (V0) is specified, so CoalHMM was run for all three possible species topologies as V0 per gibbon branch with human as an outgroup (see Rivas-González et al. for further details) (*38*). Each site in the alignment was then classified as belonging to V0 or to one of the deep coalescence topologies, namely V1 (which follows the same tree as V0), V2 and V3. V2 and V3 thus correspond to topologies that do not follow the species tree. The proportion of sites belonging to each of the four states is represented as bars in the above figure. The exact topology can be consulted in the color legend in Newick format, where species names have been abbreviated by taking the first three letters of the genus and the first three letters of the species name. The horizontal line and number correspond to the level of incomplete lineage sorting for that CoalHMM run, namely V2+V3. The figure shows that the leftmost panel is the one having the least amount of incomplete lineage sorting for both gibbon branches, which corresponds to specifying V0 as the species topology presented in this paper. Moreover, the incomplete lineage sorting proportion estimated by specifying the other topologies as V0 yields asymmetric V2/V3 proportions, with an excess of sites assigned to the topology presented in the paper (red and light blue respectively). All this indicates that incomplete lineage sorting favors the topologies presented in this paper for the gibbon branches, although we cannot discard ancient introgression completely.
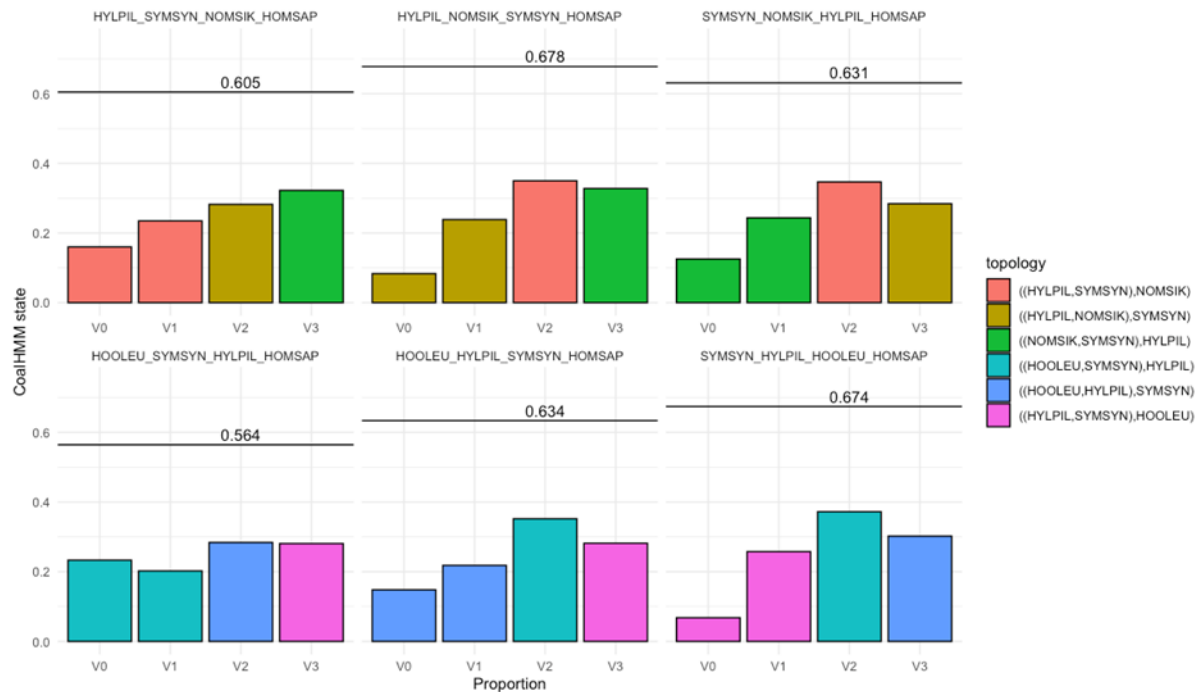
Fig. S11. Information of fossil calibrations in this study. Eleven fossil calibration points (B-L, Mya), colored in purple with confidence intervals provided, were obtained from Vanderpool et al. (*39*) and Perelman *et al*. (*35*). The nodal divergence dating (A colored in red) with a range (76-88 Mya) was obtained from TimeTree (http://www.timetree.org/).
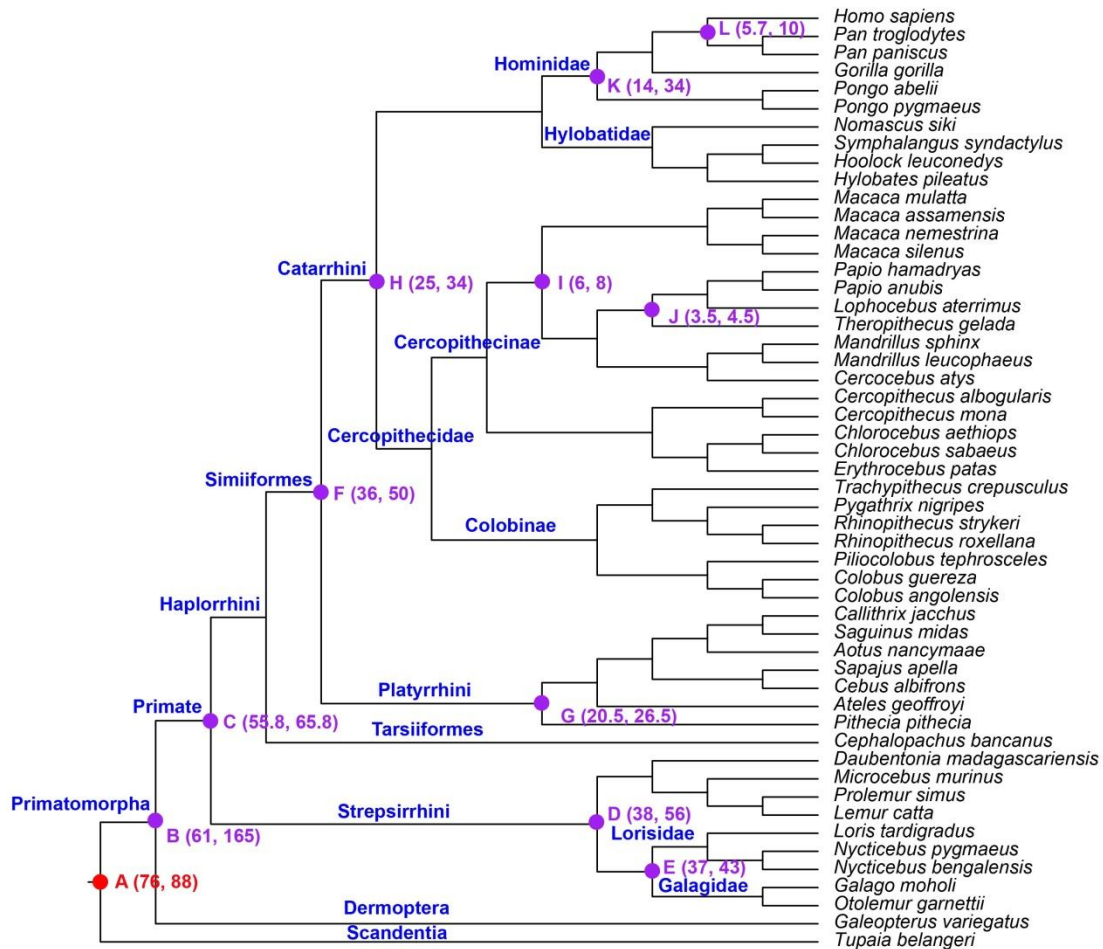
Fig. S12. Divergence dates of primates estimated by the MCMCtree algorithm in PAML4. The blue bars represent 95% credibility intervals. Upper and lower boundaries of 95% credibility intervals are given in parentheses.
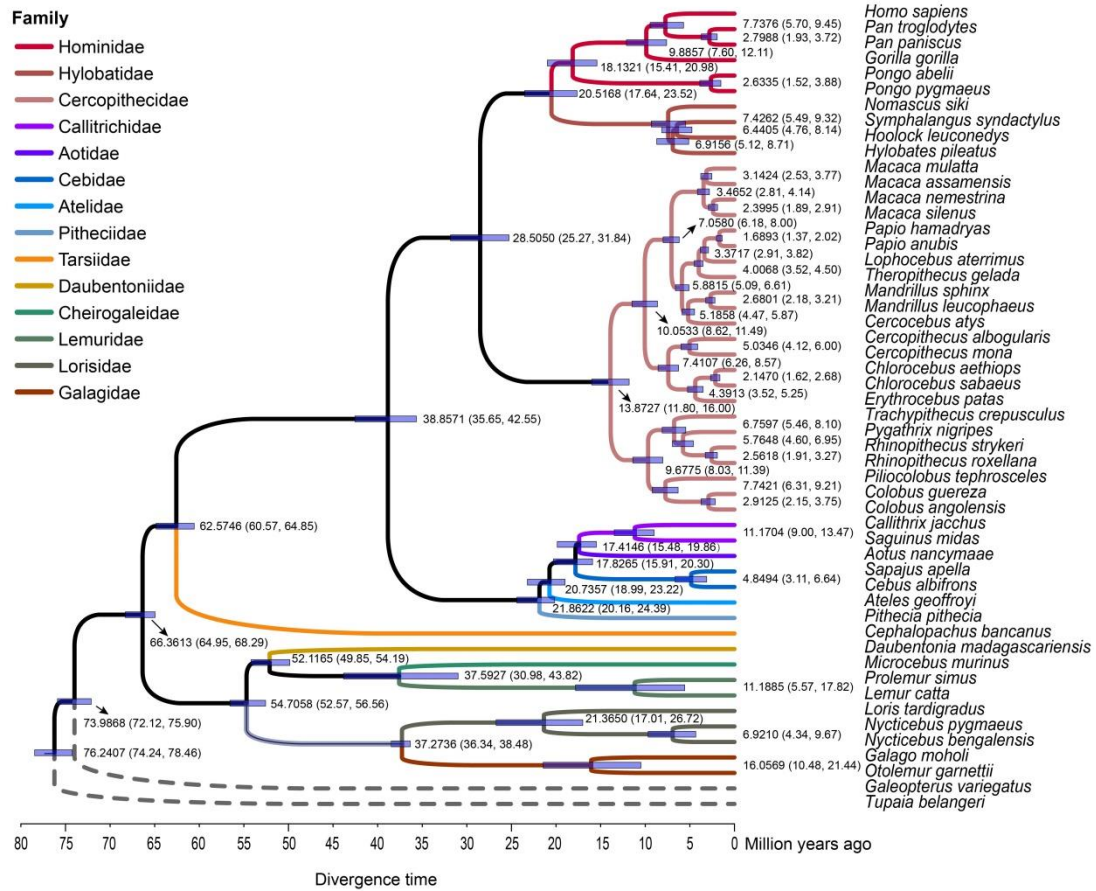
Fig. S13. A fusion event involving chromosome 8 occurred during the emergence of the Catarrhini. (A) Evolutionary pattern of chromosome 8 from the primate common ancestral lineage leading to the human lineage. The ancestral lineage of the Catarrhini is marked by a red asterisk. (B) The evidence for the fusion event involving chromosome 8 from ancestral karyotypes of Simiiformes to ancestral karyotypes of the Catarrhini by synteny alignments. The Chr IDs represent the chromosome numbers of the respective species.
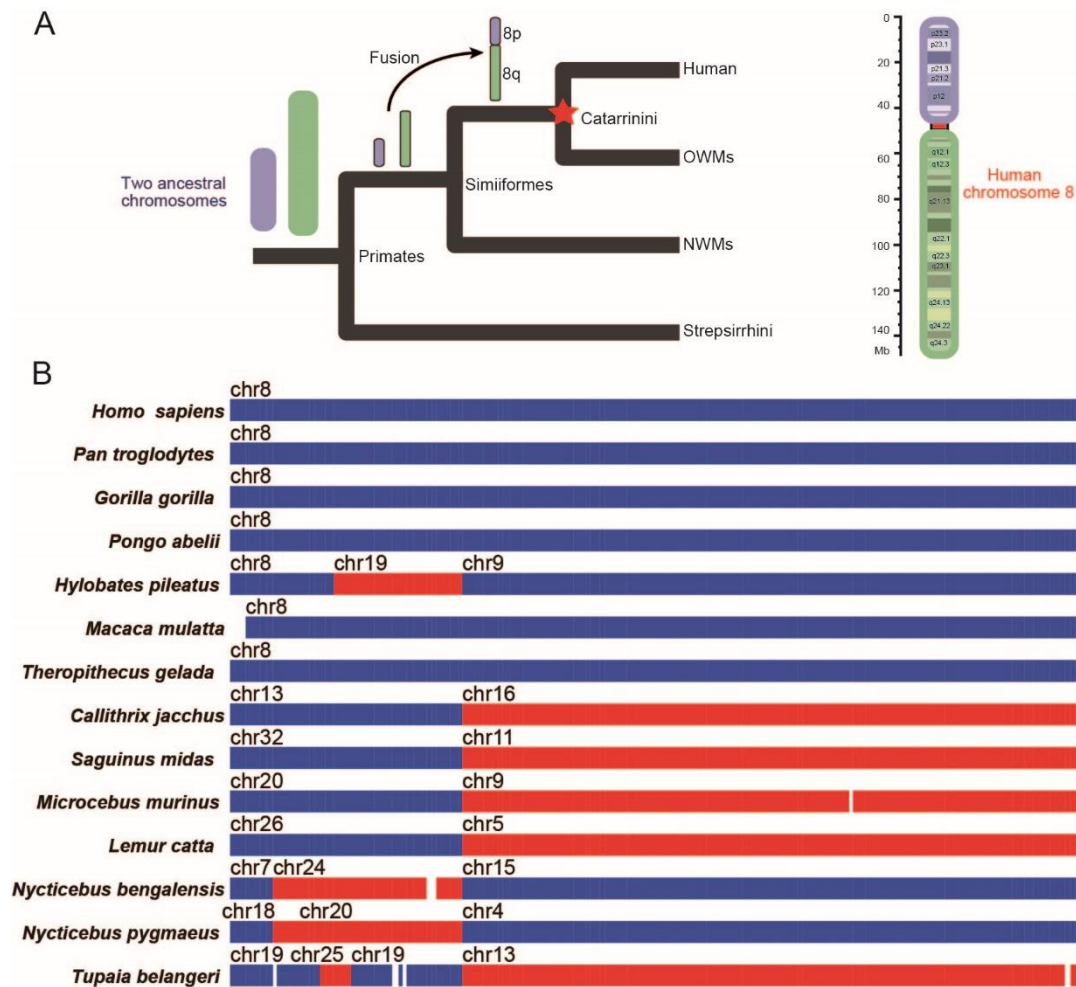
Fig. S14. A sample plot of the identification of segmental duplications for each primate species genome. Here we have used segmental duplications from the *Colobus guereza* genome as an example of a typical plot. LG-number in the outer circle represents the Chromosome number of the species.

Fig. S15. Tissue expression analysis of 57 genes overlapping segmental duplications in the great apes. (A) Proportion of highly expressed genes in tissues. Only those tissues with ≥2 highly expressed genes were retained for this statistic. (B) Highly expressed genes by tissue. Brain has the highest number of genes overlapping segmental duplications. The highly expressed genes in tissues were assigned according to the 'UP_TISSUE' category of Tissue Expression Database in DAVID (https://david.ncifcrf.gov/). SD: segmental duplication.

Term	Highly expression genes
Brain	RANBP2, TENM1, LIMS3, CDK11A, IGSF3, UNC93B1, METTL2A, PRUNE2, PLD5, RGPD5, DACH2, CDK11B, FGF7, SPTLC1, MTHFD1L, LRRTM4, IL1RAPL2, INTS4, POLR2J, DRD5
Cervix carcinoma	RANBP2, TENM1, METTL2B, CDK11A, POLR2J2, UNC93B1, PRUNE2, RGPD5, CDK11B, HERC2, MTHFD1L, OPHN1, TLK2
Testis	LRRC69, LIMS3, METTL2B, PLEKHB2, MTHFD1L, TLK2, EFCAB3, RGPD5, FAM227B, PRKG1, LIMS1
Lung	ADAMTSL1, GYPA, POLR2J2, PLEKHB2, OPHN1, UNC93B1, MMP23B, DACH2, PRKG1
Erythroleukemia	RANBP2, METTL2B, HERC2, MTHFD1L, UNC93B1, METTL2A, TLK2, INTS4, CDK11B
Liver	RANBP2, GYPA, ZNF827, CA5A, IL1RAPL2, PROS1, PRUNE2, TLK2, CDK11B
Kidney	LIMS3, GYPA, SPTLC1, UNC93B1, PRUNE2, COL4A6, POLR2J
Placenta	POLR2J2, SPTLC1, OPHN1, UNC93B1, TLK2, CDK11B, PRKG1
Uterus	MTHFD1L, UNC93B1, PROS1, MMP23B, INTS4
Leukemic T-cell	RANBP2, METTL2B, HERC2, TLK2, CDK11B
Bone marrow	GYPA, DUX4, POLR2J, LIMS1
Blood	RANBP2, GYPA, GYPB, PKD1
Amygdala	MTHFD1L, LRRTM4, KLHL2, DRD5
Trachea	PROS1, PLD5, LIMS1
Ovary	MMP23B, INTS4, DRD5
Eye	METTL2B, MTHFD1L, COL4A6
Cerebellum	PLEKHB2, EFCAB3, KLHL2
Skin	CDK11A, PLEKHB2, TUSC3
Hypothalamus	PRUNE2, KLHL2
Fetal liver	EDA, LIMS1
Cervix	MTHFD1L, PRKG1
Thalamus	IL1RAPL2, PRUNE2
Prostate	ADAMTSL1, COL4A6
Skeletal muscle	AQP7, DUX4
Tongue	ZNF827, LIMS1
Hippocampus	RANBP2, CDK11B
Muscle	MTHFD1L, INTS4
Fetal brain	OPHN1, IL1RAPL2
Pancreas	PLEKHB2, MMP23B
Colon	PLEKHB2, POLR2J

Fig. S16. Comparative analyses of genome sizes and transposable elements in different evolutionary branches among primates. (A) Genome sizes. (B) TE content. (C) *Alu* content. *P* values were calculated by the Mann-Whitney *U* test. The 'lemurs' included *Prolemur simus*, *Lemur catta*, *Microcebus murinus* and *Daubentonia madagascariensis*, whereas the 'lorisoids' included *Loris tardigradus*, *Nycticebus pygmaeus*, *Nycticebus bengalensis*, *Galago moholi* and *Otolemur garnettii* in this study. OWMs: Old World monkeys. NWMs: New World monkeys. TE: transposable element.
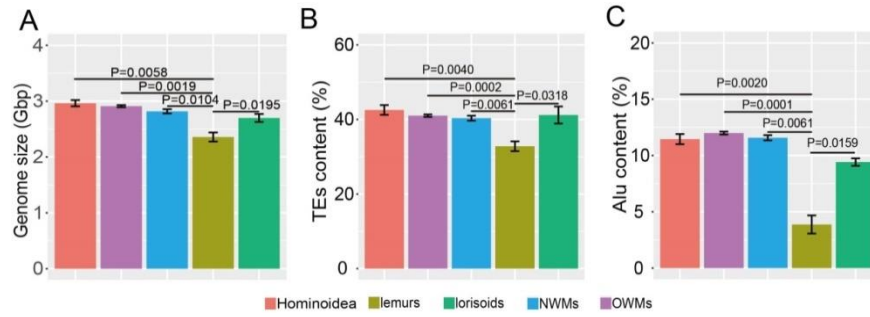
Fig. S17. Transposable elements contributing to variations in primate genome size.
Lemur genome sizes were significantly lower than other evolutionary clades, e.g., lorisoids (Lorisiformes species), NWMs, OWMs, and Hominoidea (including apes and human). *P* values were calculated by the Mann-Whitney *U* test. OWMs: Old World monkeys. NWMs: New World monkeys. TE: transposable element.
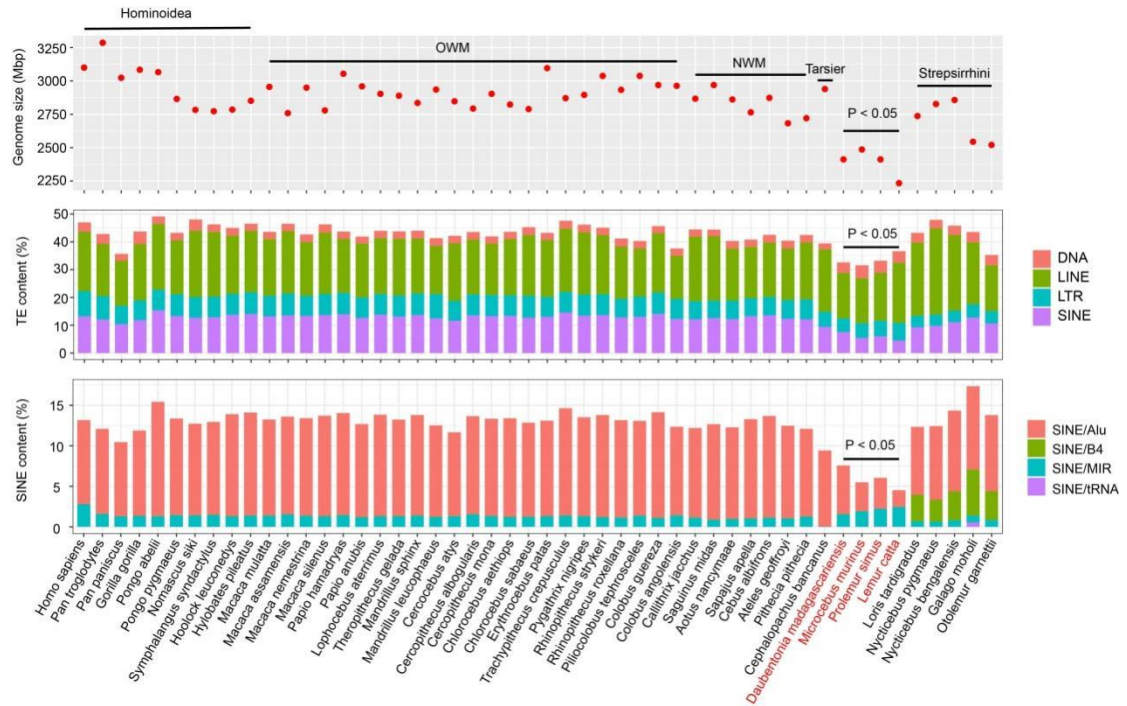
Fig. S18. Comparative analyses of transposable elements with a high proportional genome content in primates. The 67 TE subclasses with the highest genome percentage content in primate genomes are shown. TE: transposable element.
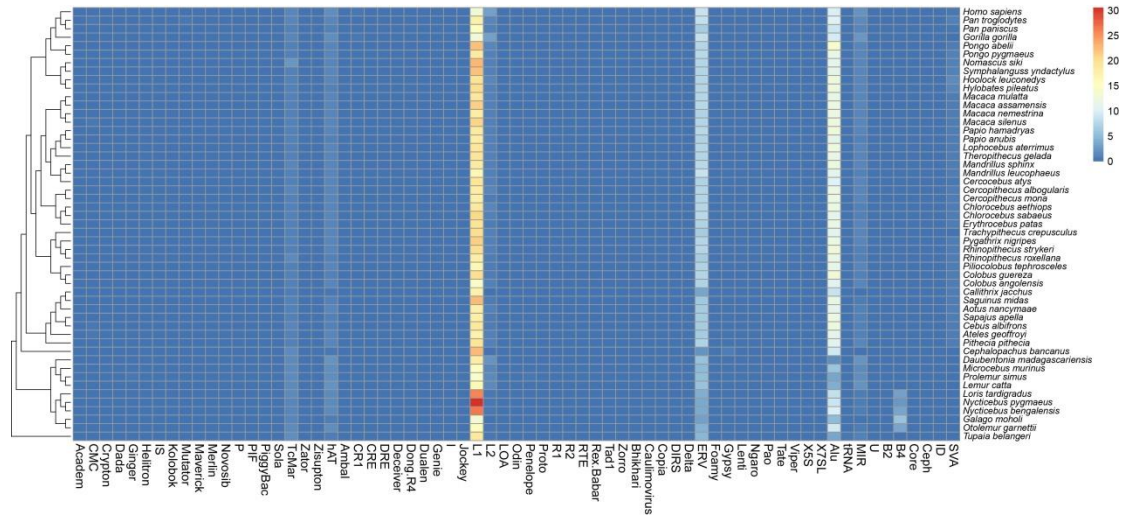
Fig. S19. Comparative analysis of *Alu* insertion times between Simiiformes and lorisoids. (A) *Alu* insertion times in Simiiformes species. (B) *Alu* insertion times in lorisoid species. The insertion time bursts are highlighted by blue bars.
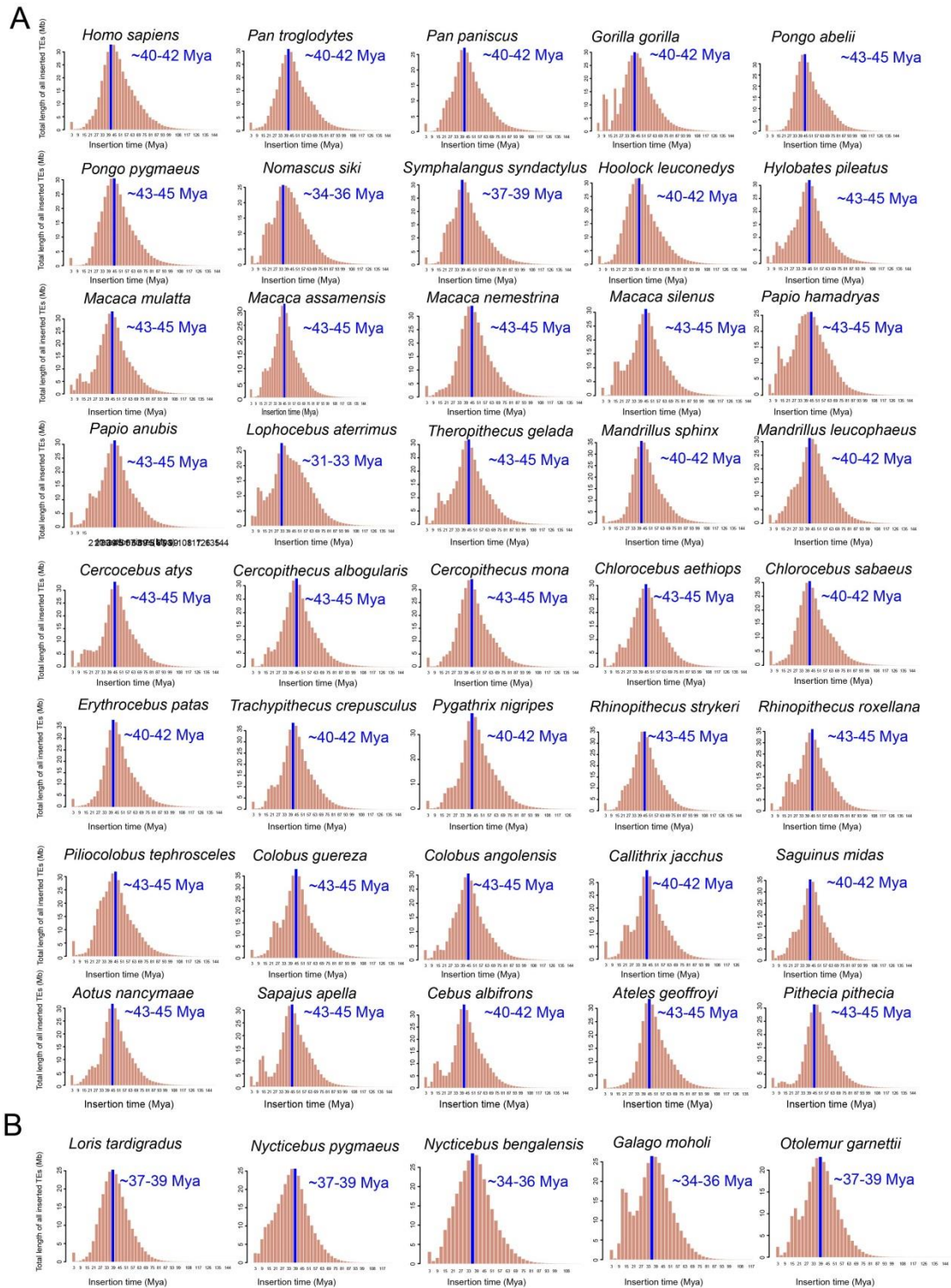
Fig. S20. Proportional analyses of genomic *Alu* subclasses between Simiiformes and Lorisiformes. (A) Proportion of *Alu* subclasses relative to the total length of all Alu elements in each Simiiformes species. In the Simiiformes, for each large evolutionary branch (e.g. great apes, gibbons, Cercopithecinae, Colobinae, NWMs), we selected *Homo sapiens*, *Hylobates pileatus*, *Macaca mulatta*, *Colobus guereza*, and *Saguinus midas* as representative species. (B) Proportion of of *Alu* subclasses relative to the total length of all *Alu* elements in each Lorisiformes species. All 5 extant Lorisiformes species were covered in this study. We observed a dramatic expansion of the *AluS*-related subclass, especially *AluSx* in Simiiformes, which was somewhat divergent from the pattern in Lorisiformes which was dominated by the conspicuous expansion of the *AluJ*-related subclass, especially *AluJb*.
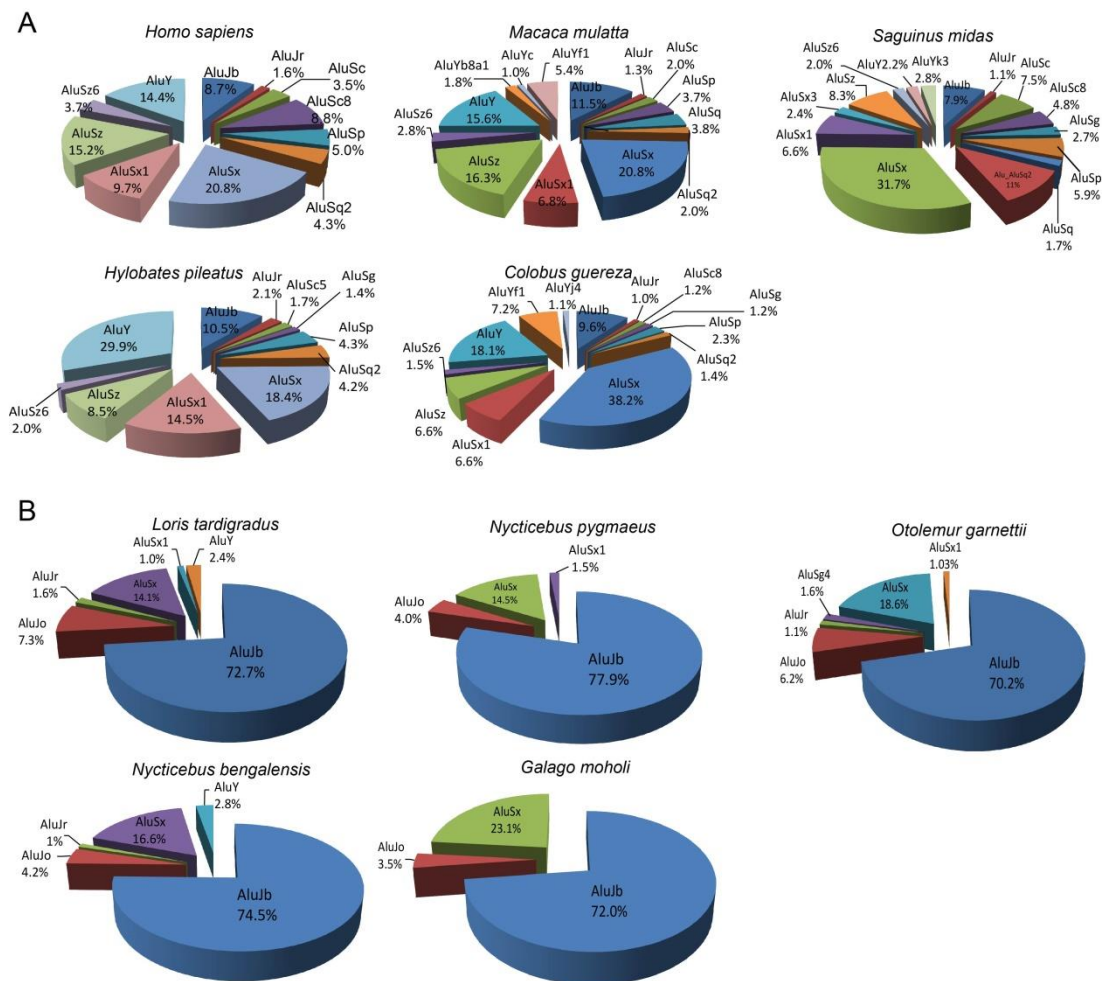
Fig. S21. Nucleotide substitution rates across primate lineages. Substitution rates in lineages were estimated by the comparison of fourfold degenerate (4d) sites in coding regions, in units of substitutions per site per million years. The overall pan-genome background substitution rate in primates is highlighted by a red full line. The nucleotide substitution rate of Western tarsier (*Cephalopachus bancanus*) in Haplorrhini was indicated by a black arrow. OWMs: Old World monkeys. NWMs: New World monkeys.
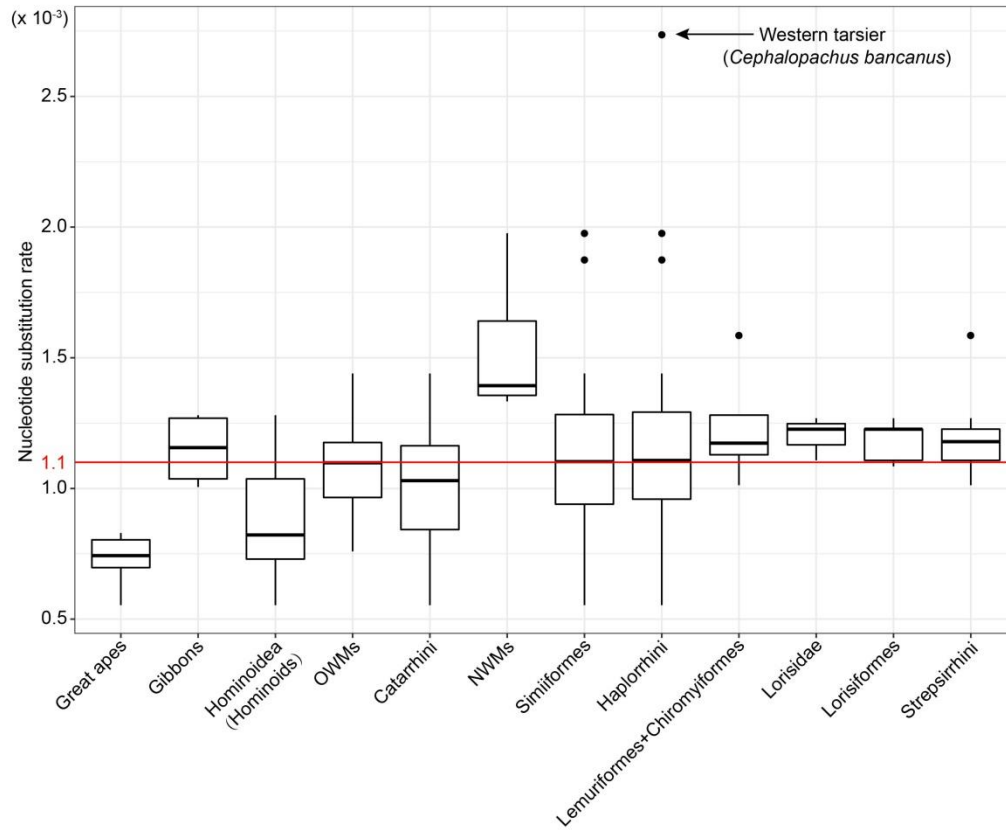
Fig. S22. An example of tissue-specific expressed genes obtained in human. The expression of target adipose samples (n>3) was compared with the expression of all other tissues (n>3) using t-statistics. The t-values were ranked from high to low. The genes with the top 5% t-values were regarded as human adipose-specific expressed genes. The red dotted line shows the top 5% cutoff of t-values (breaks=200). In total, we downloaded the gene expression matrix of 30 tissues representing 7,862 human samples from GTEx (https://commonfund.nih.gov/GTEx) for analyses of tissue-specific expressed genes.
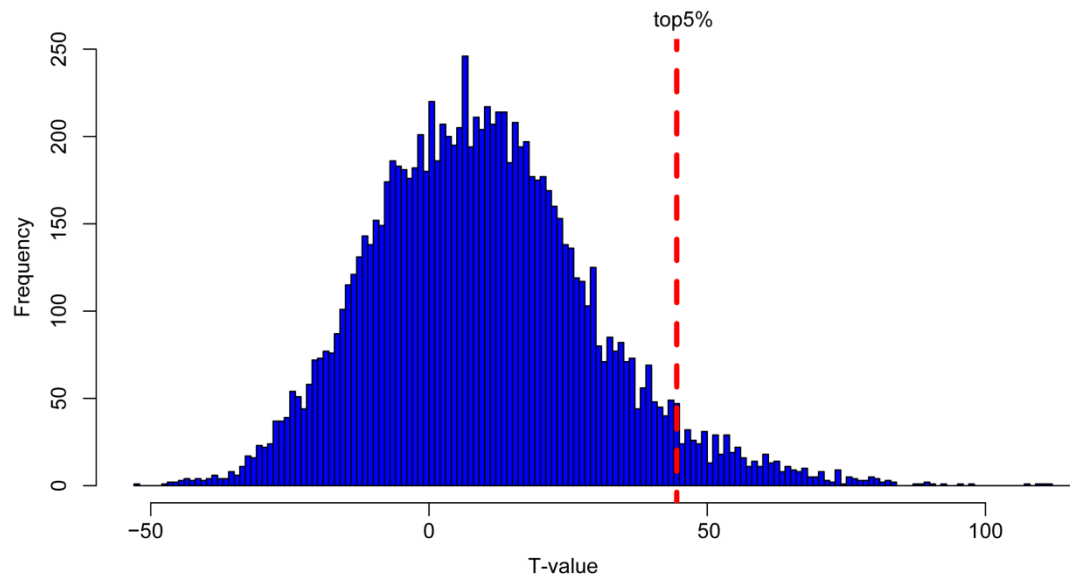
Fig. S23. Comparison of evolutionary constraints in tissues between human and other branches of primates. The X-axis represents d_N/d_S for 30 tissues in human, whereas the Y-axis represents d_N/d_S for 30 tissues for other ancestral branches in primates. Each circle represents a tissue type. The 30 tissues comprised adipose, adrenal gland, bladder, blood, blood vessel, brain, breast, cervix, colon, esophagus, fallopian tube, heart, kidney, liver, lung, muscle, nerve, ovary, pancreas, pituitary, prostate, salivary gland, skin, small intestine, spleen, stomach, testis, thyroid, uterus and vagina.
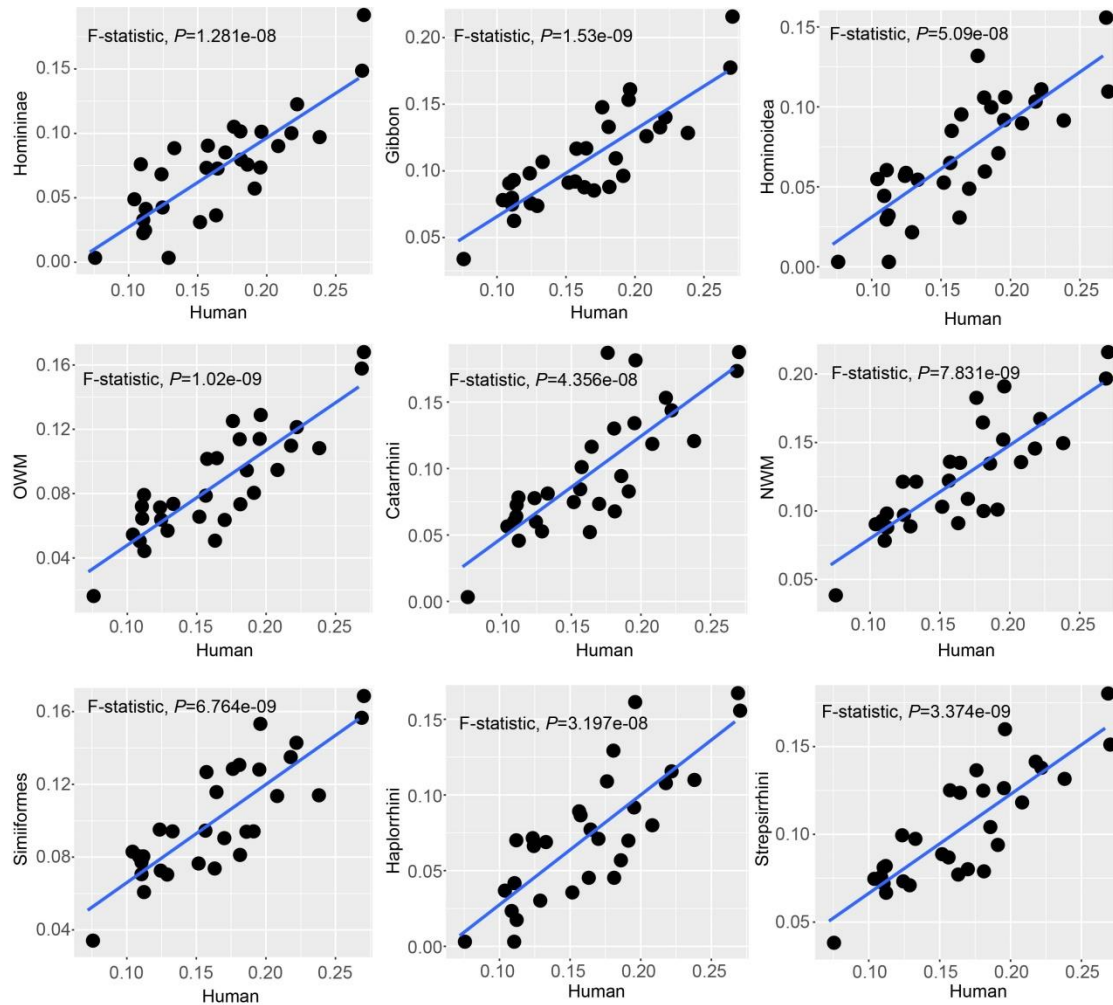
Fig. S24. Sequence alignment analysis of positively selected sites in the *GRHL2* gene in the Simiiformes ancestral lineage. The yellow bar denotes the sequence alignment of positively selected sites. The order of the human protein ENSP00000495564 (Ensembl protein id) is shown as the alignment order in this analysis. The names of the Simiiformes species are highlighted in red.

Fig. S25. Comparative analyses of coat color luminance between Simiiformes and Strepsirrhini/Tarsiiformes. The left hand side represents the comparative analysis of the dorsal luminance between Simiiformes and Strepsirrhini/Tarsiiformes. The right hand side represents comparative analysis of the ventral luminance between Simiiformes and Strepsirrhini/Tarsiiformes. Significance was assessed by means of the Mann-Whitney U test ($P<0.05$). The coat color luminance data of primates were obtained from a previous study (*89*).
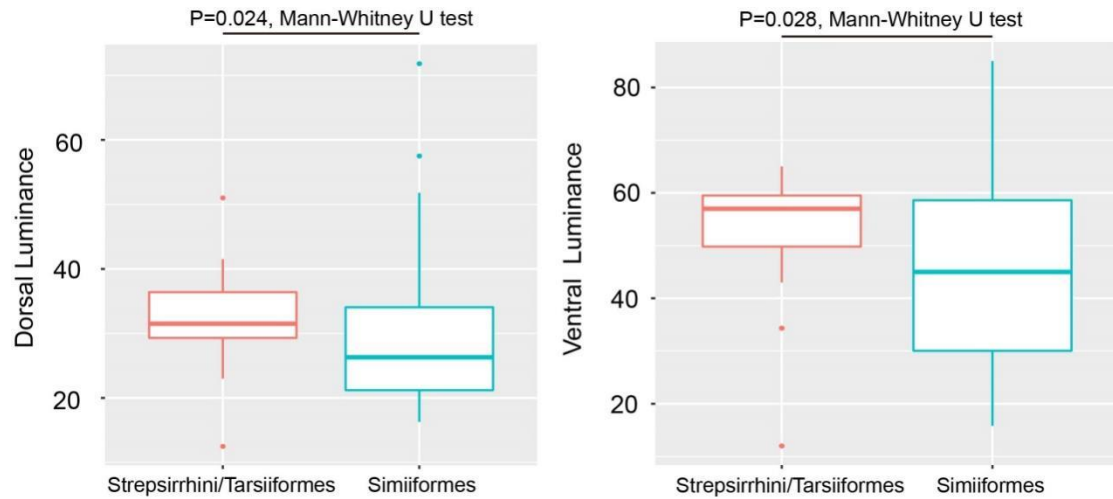
Fig. S26. Sequence alignment analysis of positively selected sites in the *NIPBL* gene in the primate ancestral lineage. The yellow bar denotes the sequence alignment of positively selected sites. The order of the human protein ENSP00000282516 (Ensembl protein id) is shown as the alignment order in this analysis. The names of the primate species are highlighted in red.

638	763	Species
S S E N K L E	H R H D N R R	*Homo sapiens*
S S E N K L E	H R H D N R R	*Pan troglodytes*
S S E N K L E	H R H D N R R	*Pan paniscus*
S S E N K L E	H R H D N R R	*Gorilla gorilla*
S S E N K L E	H R H D N R R	*Pongo abelii*
S S E N K L E	H R H D N R R	*Pongo pygmaeus*
S S E N K L E	H R H D N R R	*Nomascus siki*
S S E N K L E	H R H D N R R	*Symphalangus syndactylus*
S S E N K L E	H R H D N R R	*Hoolock leuconedys*
S S E N K L E	H R H D N R R	*Hylobates pileatus*
S T E N K L E	H R H D N R R	*Macaca mulatta*
S T E N K L E	H R H D N R R	*Macaca assamensis*
S T E N K L E	H R H D N R R	*Macaca nemestrina*
S T E N K L E	H R H D N R R	*Macaca silenus*
S T E N K L E	H R H D N R R	*Papio hamadryas*
S T E N K L E	H R H D N R R	*Papio anubis*
S T E N K L E	H R H D N R R	*Lophocebus aterrimus*
S T E N K L E	H R H D N R R	*Theropithecus gelada*
S T E N K L E	H R H D N R R	*Mandrillus sphinx*
S T E N K L E	H R H D N R R	*Cercocebus atys*
S T E N K L E	H R H D N R R	*Cercopithecus albogularis*
S T E N K L E	H R H D N R R	*Cercopithecus mona*
S T E N K L E	H R H D N R R	*Erythrocebus patas*
S T E N K L E	H R H D N R R	*Trachypithecus crepusculus*
S T E N K L E	H R H D N R R	*Pygathrix nigripes*
S T E N K L E	H R H D N R R	*Rhinopithecus strykeri*
S T E N K L E	H R H D N R R	*Rhinopithecus roxellana*
S T E N K L E	H R H D N R R	*Piliocolobus tephrosceles*
S T E N K L E	H R H D N R R	*Colobus guereza*
S T E N K L E	H R H D N R R	*Colobus angolensis*
S S E N K L E	H R H D N R R	*Callithrix jacchus*
S S E N K L E	H R H D N R R	*Saguinus midas*
S S E N K L E	H R H D N R R	*Aotus nancymaae*
S S E N K L E	H R H D N R R	*Sapajus apella*
S S E N K L E	H R H D N R R	*Cebus albifrons*
S S E N K L E	H R H D N R R	*Ateles geoffroyi*
S S E N K L E	H R H D N R R	*Pithecia pithecia*
S S E N K L E	H R H D N R R	*Cephalopachus bancanus*
S S E N K L E	H R H D N R R	*Daubentonia madagascariensis*
S S E N K I E	H R H D N R R	*Microcebus murinus*
S S E N K I E	H R H D N R R	*Prolemur simus*
S S E N K T E	H R H D N R R	*Lemur catta*
S N E N K I E	H R H D N R R	*Loris tardigradus*
S N E N K I E	H R H D N R R	*Nycticebus pygmaeus*
S N E N K I E	H R H D N R R	*Nycticebus bengalensis*
S N E N K L E	H R H D N R R	*Galago moholi*
S N E N K L E	H R H D N R R	*Otolemur garnettii*
S S E S K L E	H R H E N R R	*Galeopterus variegatus*
S N E S K V E	H K H E N R R	*Tupaia belangeri*
S N E S K L E	H R H E N R R	*Mus musculus*
S N E S K L E	H R H E N R R	*Felis catus*
S S E S K L E	H R H E N R R	*Canis lupus familiaris*
S S E S K L D	H R H E N R R	*Sus scrofa*

42

Fig. S27. Tissue enrichment analyses of genes associated with lineage-specific accelerated regions from the Haplorrhini ancestor leading to the human lineage. Tissue enrichment analyses were performed by DAVID (v6.8) database (https://david.ncifcrf.gov/tools.jsp). Statistical significance was assessed by the Modified Fisher's Exact test with $P < 0.05$. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.
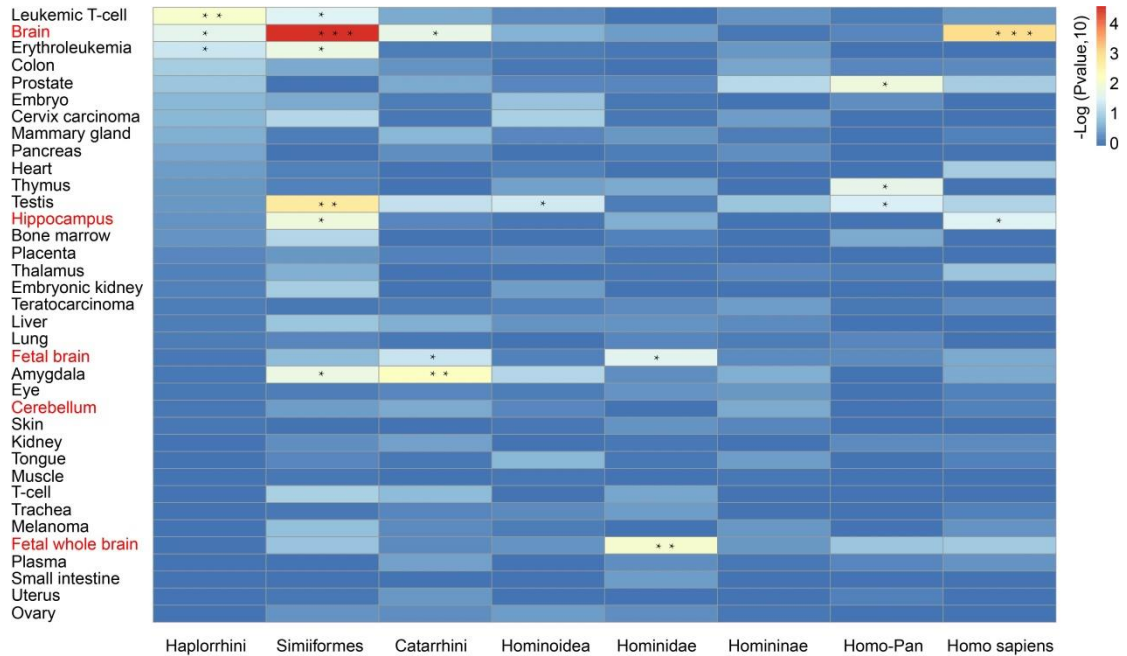
Fig. S28. 39 genes associated with lineage-specific accelerated regions involved in neurotransmitter signaling pathways from the Haplorrhini ancestral lineage leading to the human lineage. The different ancestral lineages are highlighted by different colors. The genes associated with lineage-specific accelerated regions concurrently emerging in different lineages are linked by a solid line.
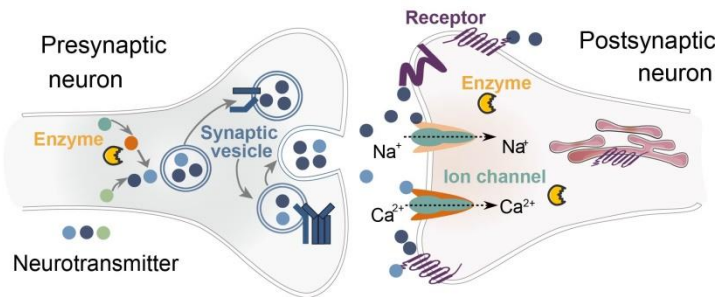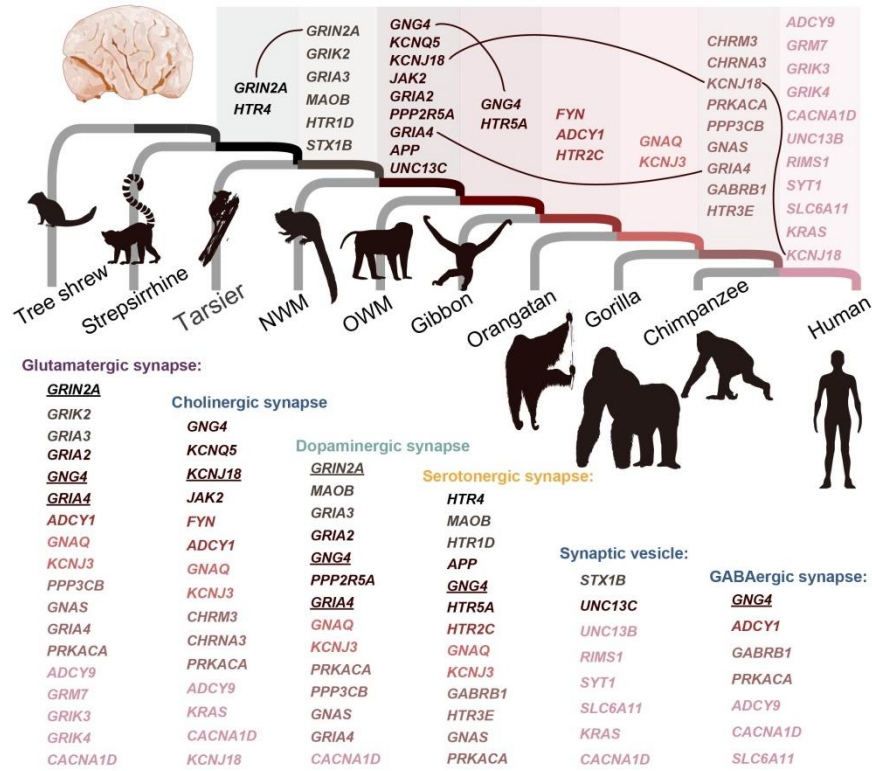
Fig. S29. An ape–specific accelerated region including several highly divergent sites compared to non-ape primate species. The nearest gene (*KIAA1217*) to this ape-specific accelerated region is involved in tail development phenotypes in mouse. The human DNA sequence was used as the reference to determine the sequence order. The names of the great ape (and lesser ape) species are highlighted in red.

```
24,158,524   24,158,531   24,158,582   24,158,781   (Chr10,hg38)
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCGGT  Homo sapiens
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCGGT  Pan troglodytes
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCGGT  Pan paniscus
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCTGGT  Gorilla gorilla
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCAGT  Pongo abelii
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCAGT  Pongo pygmaeus
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCAGT  Hylobates pileatus
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCAGT  Symphalangus syndactylus
TTTAGAG  AATTCAT  GCACTGA  TGTCTGGCCAGT  Hoolock leuconedys
TTTAGAG  AATTCAT  GCACTGA  TGTCCGGCCAGT  Nomascus siki
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Macaca mulatta
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Macaca assamensis
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Macaca nemestrina
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Macaca silenus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Cercocebus atys
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Mandrillus sphinx
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Mandrillus leucophaeus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Papio hamadryas
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Papio anubis
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Lophocebus aterrimus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Theropithecus gelada
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Cercopithecus mona
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Erythrocebus patas
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Chlorocebus sabaeus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Chlorocebus aethiops
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Colobus angolensis
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Colobus guereza
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Piliocolobus tephrosceles
TTTGGAG  AATCAAT  GCAGTGA  TGTTTGGCTGGT  Trachypithecus crepusculus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Pygathrix nigripes
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Rhinopithecus strykeri
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Rhinopithecus roxellana
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Pithecia pithecia
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Sapajus apella
TTTGGAG  AATCCAT  GCAGTGA  TGTTTGGCTGGT  Cebus albifrons
TTTGGAG  AATCCAT  --AGTGG  TGTTTGGCTGGT  Callithrix jacchus
TTTGGAG  AATCCAT  GCAGTGA  TGTTTAGCTGGT  Saguinus midas
TTTGGAG  AATCCAT  GCAGTGA  TGTCTGGC---T  Aotus nancymaae
TTTGGAG  AATGCAT  ACAATGA  TGTTTGGCTGGT  Cephalopachus bancanus
TTTGGAG  AATCCAT  ACAGTGA  TGTTCGGCTGGT  Daubentonia madagascariensis
TTTGGAG  AATCCAT  ACAGTGA  TGTTCGGCTGGT  Lemur catta
TTTGGAG  AATCCAT  ACAGTGA  TGTTCGGCTGGT  Prolemur simus
TTTGGAG  AATCCAT  ACAGTGA  TGTTTGGCTGGT  Microcebus murinus
TTTGGAG  AATCCAT  ACAGTGA  TGTTTGGCTGGT  Otolemur garnettii
TTTGGAG  AATCCAT  ACAGTGA  TGTTTGGCTGGT  Galago moholi
TTTGGAG  AATCCGT  ACAGTGA  TGCTTGGCTGGT  Nycticebus pygmaeus
TTTGGAG  AATCCAT  ACGGTGA  TGTT---CTGGT  Nycticebus bengalensis
TTTGGAG  AATCCAT  ACAGTGA  TGTTTGGCTGGT  Loris tardigradus
TTTGGTG  AATCCAT  ACAGTGA  TGTTCGGCTGGT  Galeopterus variegatus
TTTGGAG  AATCCAT  ATAGTGA  TGTTCGGCTGGT  Tupaia belangeri
```

Fig. S30. Phylogenetic analysis of an ape–specific accelerated region among primate species. The nearest gene to this ape–specific accelerated region (*KIAA1217*) is involved in tail development phenotypes in mouse. Branches of the ape species are highlighted in red. The evolutionary history was inferred by using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model (*210*). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (next to the branches). Evolutionary analyses were conducted in MEGA7 (*211*).
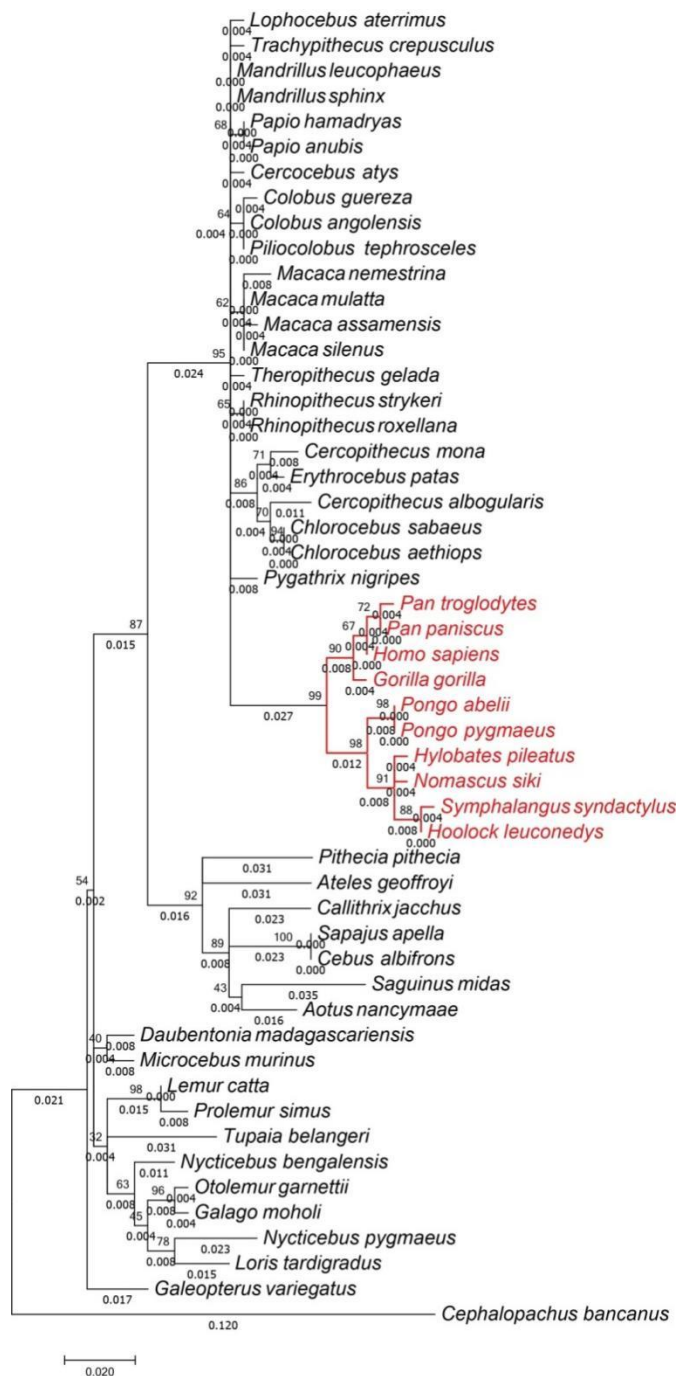
Fig. S31. The lineage–specific accelerated region (Chr10: 24,158,492–24,158,793; total length=301bp) in the great ape lineages overlaps an enhancer EH38E1455433 (pELS). Our analysis suggests that this ape–specific accelerated region may regulate the expression level of *KIAA1217*, the nearest gene to the ape–specific accelerated region.
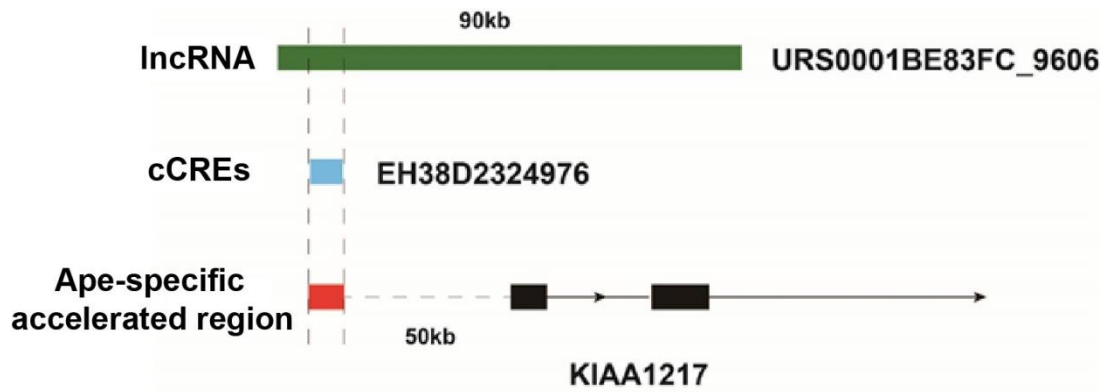
Fig. S32. High-throughput chromosome conformation capture data supporting a strong interaction between the ape-specific accelerated region and its neighboring gene *KIAA1217*. The high-throughput chromosome conformation capture reads for human blood samples were downloaded from Encode (https://www.encodeproject.org/experiments/ENCSR118YWJ/). Reads were sampled by ~30X. Sampled reads were mapped to the human genome with juicer (v1.6) (*212*). The merged_nodups.txt file was used to generate Hi-C contact map employing different levels of resolution (10 Kb, 20 Kb, 50 Kb, 100 Kb, 200 Kb, 500 Kb, 1 Mb, 2 Mb, 5 Mb) with cooler (v0.8.11) (*213*), converted to h5 format and normalized with KR method with hicexplorer (v3.7.2) (*214*). Regional high-throughput chromosome conformation capture maps were visualized by hicexplorer package. This lineage-specific accelerated region and its neighboring gene *KIAA1217* are located within the same topologically associating domain.
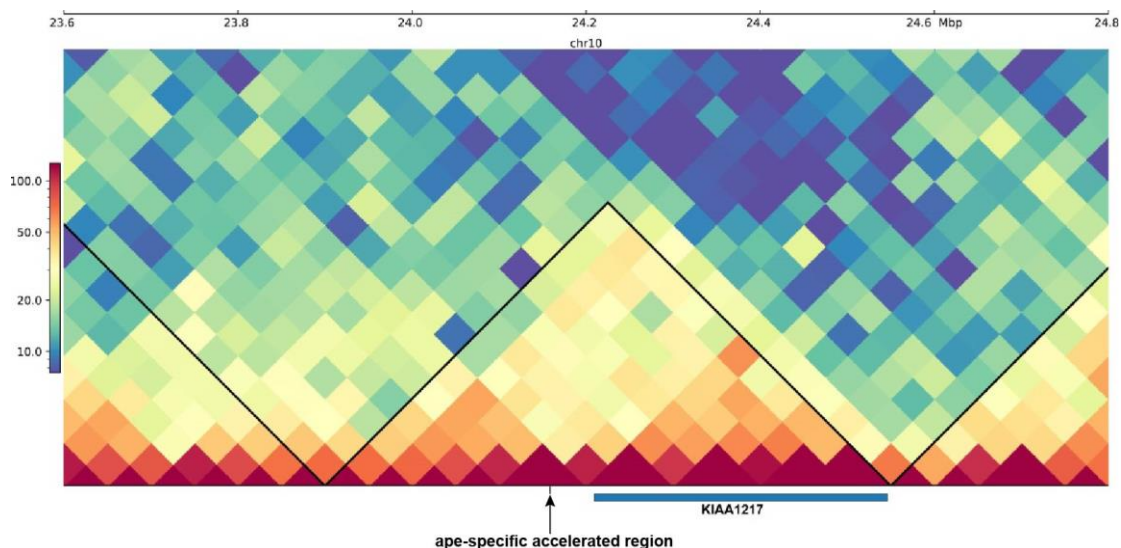
Fig. S33. Sequence alignment analysis of positively selected sites in the *ACADM* gene in the Colobinae ancestral lineage. The Colobinae species are highlighted in red. The multiple sequence alignment is shown across primate species. The human protein sequence (Ensembl protein id: ENSP00000359871) was taken as the reference to determine the sequence order.

75	138	Species
E Y P V P L I	T F D A C L I	*Homo sapiens*
E Y P V P L I	T F D A C L I	*Pan troglodytes*
E Y P V P L I	T F D A C L I	*Pan paniscus*
E Y P V P L I	T F D A C L I	*Gorilla gorilla*
E Y P V P I I	T F D A C L I	*Pongo abelii*
E Y P V P L I	T F D A C L I	*Nomascus siki*
E Y P V P L I	T F D A C L I	*Symphalangus syndactylus*
E Y P V P L I	T F D A C L I	*Hoolock leuconedys*
E Y P V P L I	T F D A C L I	*Hylobates pileatus*
E Y P V P L I	T F D A C L I	*Macaca mulatta*
E Y P V P L I	T F D A C L I	*Macaca assamensis*
E Y P V P L I	T F D A C L I	*Macaca nemestrina*
E Y P V P L I	T F D A C L I	*Macaca silenus*
E Y P V P L I	T F D A C L I	*Papio hamadryas*
E Y P V P L I	T F D A C L I	*Papio anubis*
E Y P V P L I	T F D A C L I	*Lophocebus aterrimus*
E Y P V P L I	T F D A C L I	*Theropithecus gelada*
E Y P V P L I	T F D A C L I	*Mandrillus sphinx*
E Y P V P L I	T F D A C L I	*Mandrillus leucophaeus*
E Y P V P L I	T F D A C L I	*Cercocebus atys*
E Y P V P L I	T F D A C L I	*Cercopithecus albogularis*
E Y P V P L I	T F D A C L I	*Cercopithecus mona*
E Y P V P L I	T F D A C L I	*Chlorocebus aethiops*
E Y P V P L I	T F D A C L I	*Chlorocebus sabaeus*
E Y P V P L I	T F D A C L I	*Erythrocebus patas*
E Y P M P L I	T F D C C L I	*Trachypithecus crepusculus*
E Y P M P L I	T F D C C L I	*Pygathrix nigripes*
E Y P M P L I	T F D C C L I	*Rhinopithecus strykeri*
E Y P M P L I	T F D C C L I	*Rhinopithecus roxellana*
E Y P M P L I	T F D C C L I	*Piliocolobus tephrosceles*
E Y P M P L I	T F D C C L I	*Colobus guereza*
E Y P M P L I	T F D C C L I	*Colobus angolensis*
E Y P V P L I	T F D S C L I	*Saguinus midas*
E Y P V P L I	T F D A C L I	*Aotus nancymaae*
E Y P V P L I	T F D A C L I	*Sapajus apella*
E Y P V P L I	T F D A C L I	*Cebus albifrons*
E Y P V P L I	T F D A C L I	*Ateles geoffroyi*
E Y P V P L I	S F D G C L I	*Pithecia pithecia*
E Y P V P L I	T F D A C L I	*Cephalopachus bancanus*
E Y P V P L I	T L D A C L I	*Daubentonia madagascariensis*
E Y P V P L I	T F D A C L I	*Microcebus murinus*
E Y P V P L I	T L D G C L I	*Prolemur simus*
E Y P V P L I	T L D G C L I	*Lemur catta*
E Y P V P L I	T F D A C L I	*Loris tardigradus*
E Y P V P L I	T F D A C L I	*Nycticebus pygmaeus*
E Y P V P L I	T F D A C L I	*Nycticebus bengalensis*
E Y P V P L I	T F D A C L I	*Otolemur garnettii*
E Y P V P L I	C F D A C L I	*Galeopterus variegatus*
E Y P V P V I	T F D A C L I	*Tupaia belangeri*

Fig. S34. Decrease of members of the *OR52A* gene family in Simiiformes compared to Strepsirrhini. The lineages from Simiiformes and Strepsirrhini are highlighted by orange and blue, respectively.
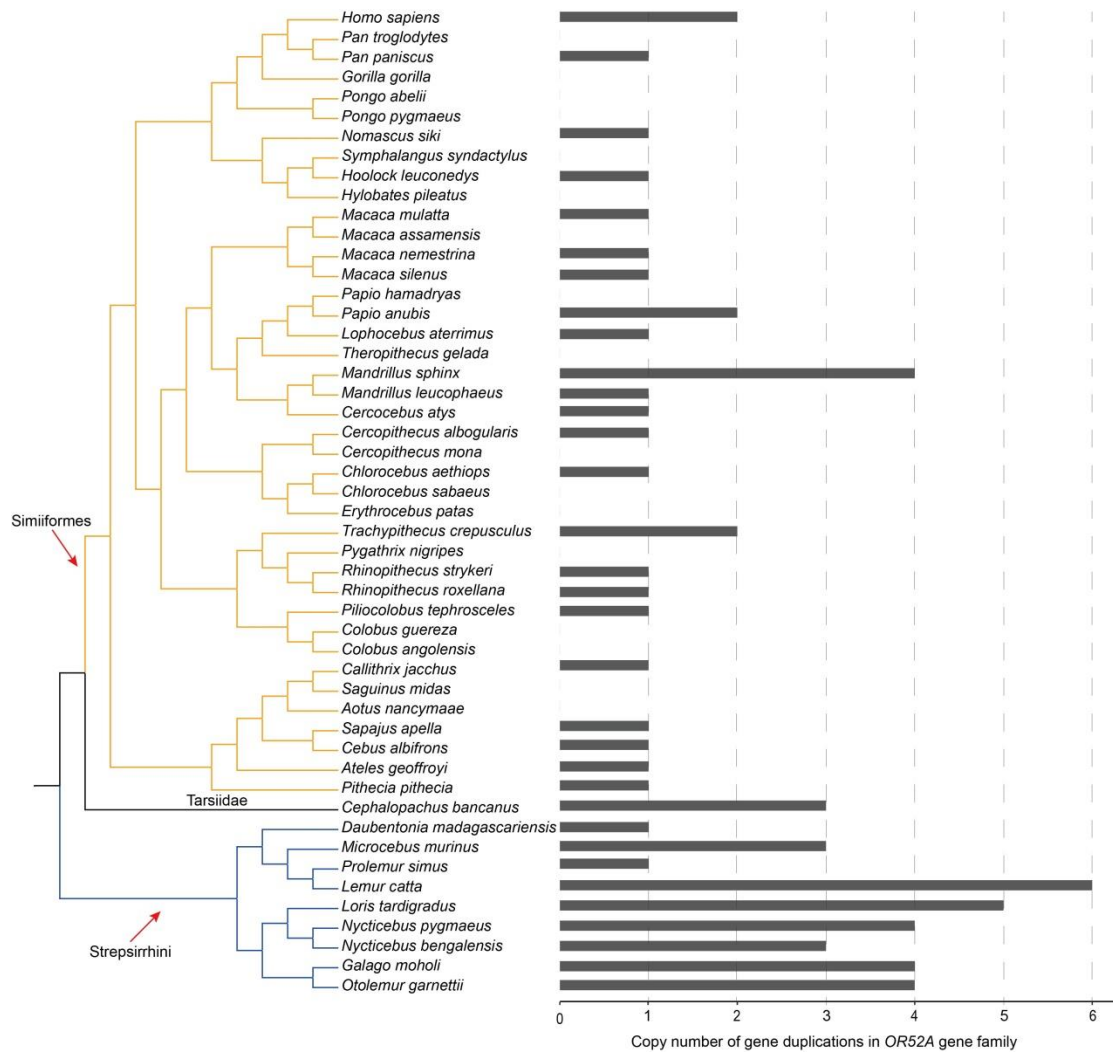
Fig. S35. Decrease of the olfactory receptor gene family *OR52A* in Simiiformes compared to Strepsirrhini. The copy number of each species is shown by point between Strepsirrhini and Simiiformes.
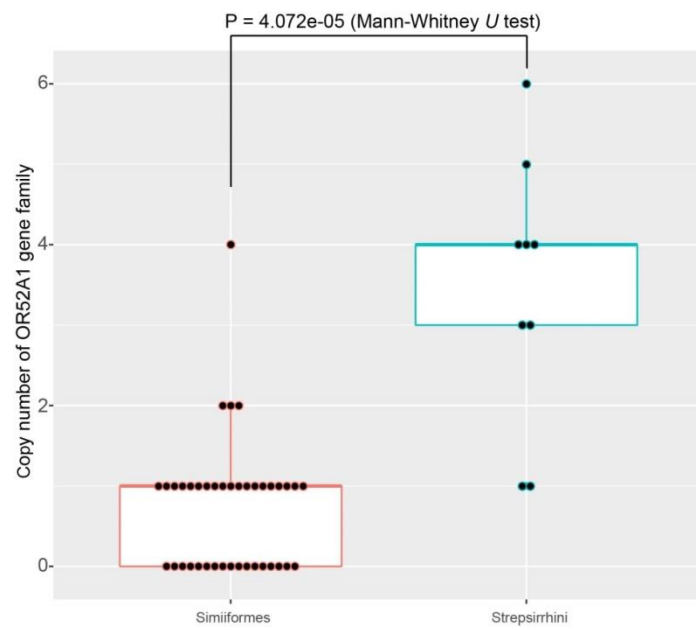
Fig. S36. Population dynamics history of representative primates, including 27 sequenced species and 21 downloaded species, from our primate genome project. The dynamic history of the effective population size (*Ne*) was visualized by using the PSMC algorithm. The X–axis represents years whereas the Y–axis represents *Ne*. The "g" parameter shows the generation time for each primate species. The "μ" values were inferred from the mutation rate per site per year. These parameters were cited and calculated in table S16 and table S41. The downloaded short–read sequencing data list is given in table S41. *Callithrix jacchus* showed aberrant recent population history dynamics, which may be due to admixture between different subspecies within the same species complex. Two species, namely *Chlorocebus sabaeus* and *Piliocolobus tephrosceles,* were excluded due to the lack of short–read data from the NCBI SRA database.

16. *Papio anubis* — Years (g=11, μ=1.1x10⁻⁸)

17. *Lophocebus aterrimus* — Years (g=10, μ=1.1x10⁻⁸)

18. *Theropithecus gelada* — Years (g=10, μ=0.8x10⁻⁸)

19. *Mandrillus sphinx* — Years (g=10, μ=1.0x10⁻⁸)

20. *Mandrillus leucophaeus* — Years (g=10, μ=1.0x10⁻⁸)

21. *Cercocebus atys* — Years (g=10, μ=1.1x10⁻⁸)

22. *Cercopithecus albogularis* — Years (g=8.5, μ=0.9x10⁻⁸)

23. *Cercopithecus albogularis* — Years (g=8.5, μ=0.9x10⁻⁸)

24. *Chlorocebus aethiops* — Years (g=8.5, μ=1.0x10⁻⁸)

25. *Erythrocebus patas* — Years (g=8.5, μ=1.0x10⁻⁸)

26. *Trachypithecus crepusculus* — Years (g=11, μ=1.6x10⁻⁸)

27. *Pygathrix nigripes* — Years (g=10, μ=1.3x10⁻⁸)

28. *Rhinopithecus strykeri* — Years (g=10, μ=1.1x10⁻⁸)

29. *Rhinopithecus strykeri* — Years (g=10, μ=0.8x10⁻⁸)

30. *Colobus guereza* — Years (g=10, μ=1.1x10⁻⁸)

31. *Colobus angolensis* — Years (g=10, μ=1.3x10⁻⁸)

31. *Callithrix jacchus* — Years (g=6, μ=1.2x10⁻⁸)

33. *Saguinus midas* — Years (g=6, μ=1.1x10⁻⁸)

53

34. *Aotus nancymaae*

Ne (x10⁴) vs Years (g=8, μ=1.1x10⁻⁸)

35. *Sapajus apella*

Ne (x10⁴) vs Years (g=15, μ=2.1x10⁻⁸)

36. *Cebus albifrons*

Ne (x10⁴) vs Years (g=15, μ=2.1x10⁻⁸)

37. *Ateles geoffroyi*

Ne (x10⁴) vs Years (g=15, μ=2.0x10⁻⁸)

38. *Pithecia pithecia*

Ne (x10⁴) vs Years (g=9, μ=1.3x10⁻⁸)

39. *Cephalopachus bancanus*

Ne (x10⁴) vs Years (g=7, μ=1.9x10⁻⁸)

40. *Microcebus murinus*

Ne (x10⁴) vs Years (g=3, μ=0.5x10⁻⁸)

41. *Prolemur simus*

Ne (x10⁴) vs Years (g=3, μ=0.4x10⁻⁸)

42. *Lemur catta*

Ne (x10⁴) vs Years (g=3, μ=0.4x10⁻⁸)

43. *Daubentonia madagascariensis*

Ne (x10⁴) vs Years (g=3.5, μ=0.4x10⁻⁸)

44. *Loris tardigradus*

Ne (x10⁴) vs Years (g=2, μ=0.3x10⁻⁸)

45. *Nycticebus pygmaeus*

Ne (x10⁴) vs Years (g=2, μ=0.2x10⁻⁸)

46. *Nycticebus bengalensis*

Ne (x10⁴) vs Years (g=2, μ=0.2x10⁻⁸)

47. *Galago moholi*

Ne (x10⁴) vs Years (g=2, μ=0.2x10⁻⁸)

48. *Otolemur garnettii*

Ne (x10⁴) vs Years (g=2, μ=0.2x10⁻⁸)

Fig. S37. Demographic history of non-human primates. Demographic history of all non-human primate species was fitted according to the geom_smooth (span=0.1) function in the ggplot2 library of R (v3.5.0). The normalized *N*e was inferred by dividing the estimated value of *N*e for each species at each time point with its maximum value. *Callithrix jacchus* was removed from this analysis because the genome was from an inbred family. The time including 100,000 and 50,000 years ago is highlighted by the green and black dotted lines, respectively.
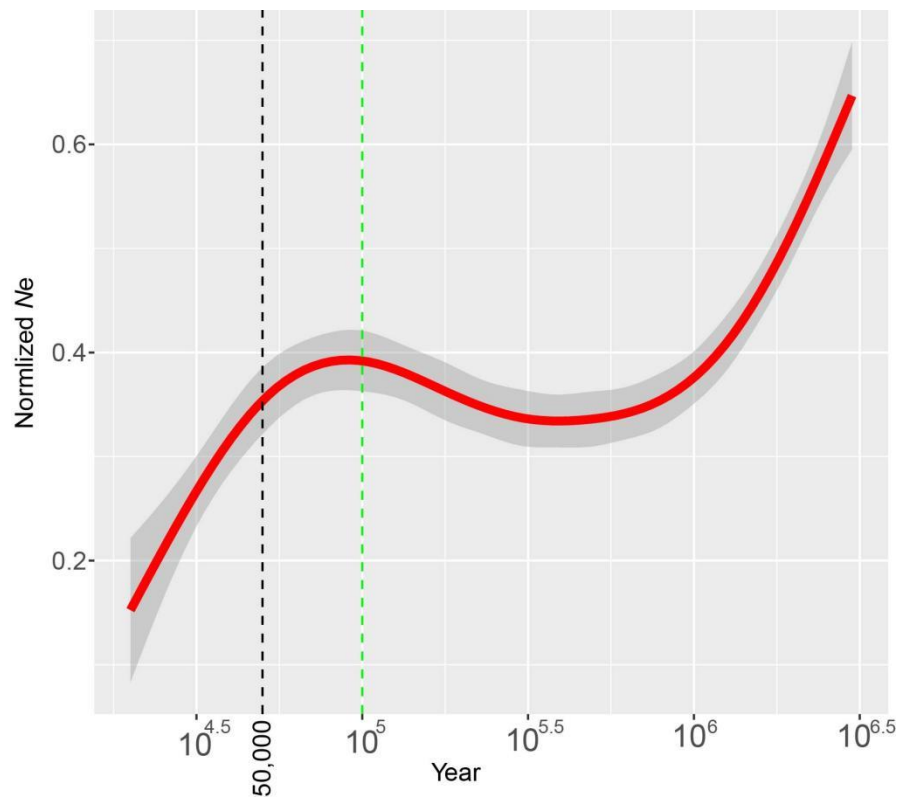
Fig. S38. Comparative analyses of nucleotide diversity in primate species based on IUCN Red List status. The median nucleotide diversity of each primate species was calculated according to a 50-kb non-overlapping sliding window on a whole-genome scale. We downloaded the IUCN Red List status information of each species in this study from THE IUCN RED LIST OF THREATENED SPECIES DATABASE (https://www.iucnredlist.org/). LC: Least Concern; NT: Near Threatened; VU: Vulnerable; EN: Endangered; CR: Critically Endangered. Additional information is provided in table S42.
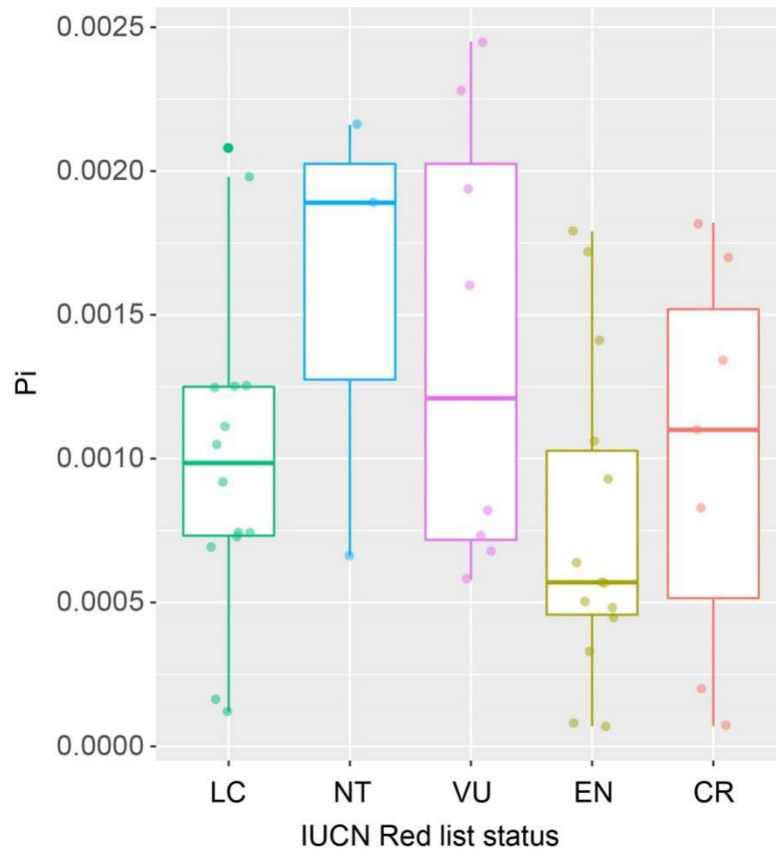
Fig. S39. Population dynamics history of 20 primate species showing a continual decline in *N*e. The demographic histories of primate species over the past million years ago were inferred by using the pairwise sequentially Markovian coalescent model (PSMC) (*151*). We included the IUCN Red List status information for each species. CR=Critically Endangered; EN=Endangered, VU=Vulnerable; NT=Near threatened; LC=Least concern.
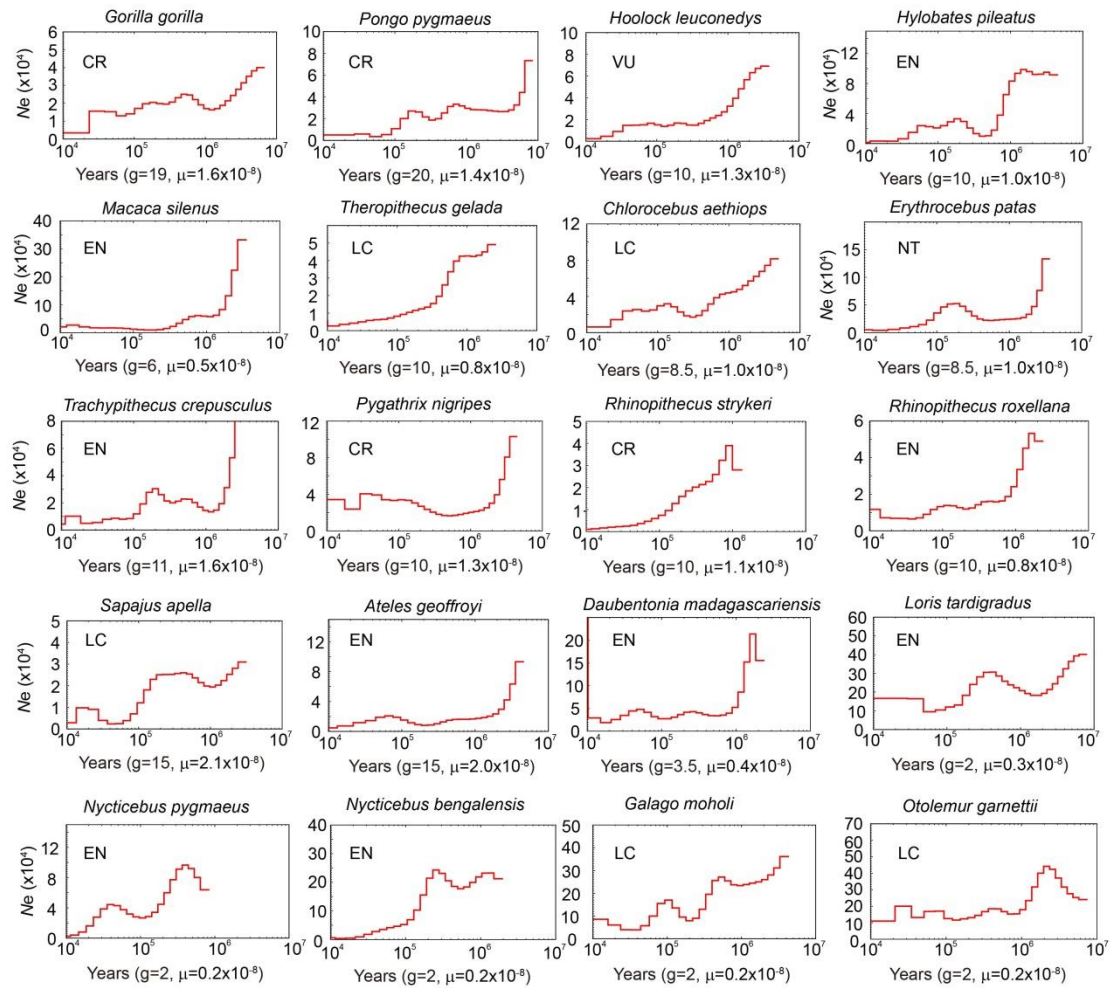
table S1. Single-molecule sequencing raw data statistics for each primate species representing 26 genera from 11 families. Three parameters including long-read sequencing platform, sequencing data and coverage are given.

Family	Genus	Species name	Sequencing platform	Sequencing data (bases)	Coverage (X)
Hominidae	*Pongo*	*Pongo pygmaeus*	Pacbio	151,467,742,407	52.85
Hylobatidae	*Nomascus*	*Nomascus siki*	Pacbio	149,946,567,298	53.88
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	Pacbio	159,077,351,518	57.34
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	ONT	156,590,185,899	56.22
Hylobatidae	*Hylobates*	*Hylobates pileatus*	ONT	150,950,070,274	52.94
Cercopithecidae	*Macaca*	*Macaca assamensis*	ONT	156,970,507,529	56.89
Cercopithecidae	*Macaca*	*Macaca silenus*	ONT	114,522,534,228	41.21
Cercopithecidae	*Papio*	*Papio hamadryas*	ONT	159,646,909,260	52.27
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	Pacbio	158,056,630,696	54.44
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	ONT	165,871,342,996	58.48
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	Pacbio	143,900,291,194	51.52
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	Pacbio	149,295,464,604	52.88
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	ONT	160,101,247,817	51.70
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	ONT	143,786,444,630	50.09
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	ONT	141,205,564,168	48.76
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	ONT	154,627,720,964	52.70
Cercopithecidae	*Colobus*	*Colobus guereza*	ONT	170,750,537,363	57.62
Callitrichidae	*Saguinus*	*Saguinus midas*	ONT	127,496,846,769	42.95
Cebidae	*Sapajus*	*Sapajus apella*	Pacbio	90,284,120,654	32.66
Cebidae	*Cebus*	*Cebus albifrons*	ONT	132,420,433,893	46.07
Atelidae	*Ateles*	*Ateles geoffroyi*	Pacbio	152,574,847,004	56.87
Pitheciidae	*Pithecia*	*Pithecia pithecia*	Pacbio	155,739,771,212	57.23
Tarsiidae	*Cephalopachus*	*Cephalopachus bancanus*	Pacbio	146,507,320,994	49.83
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	Pacbio	149,520,007,402	61.99
Lorisidae	*Loris*	*Loris tardigradus*	Pacbio	154,272,498,770	56.35
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	ONT	144,496,844,856	50.59
Galagidae	*Galago*	*Galago moholi*	Pacbio	156,849,070,599	61.66

table S2. Strategies adopted for the *de novo* genome assembly of the long-read sequencing of 27 primate genomes. Four highlighted species in bold (*Hylobates pileatus*, *Colobus guereza*, *Saguinus midas* and *Nycticebus bengalensis*) were assembled at the chromosome level.

Family	Genus	Species name	Phylogenetic group	Assembly strategy
Hominidae	*Pongo*	*Pongo pygmaeus*	Hominoidea	Wtdbg2
Hylobatidae	*Nomascus*	*Nomascus siki*	Hominoidea	Wtdbg2+scaff10x
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	Hominoidea	Falcon+scaff10x
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	Hominoidea	Wtdbg2
Hylobatidae	***Hylobates***	***Hylobates pileatus***	**Hominoidea**	**NextDenovo+3d-dna**
Cercopithecidae	*Macaca*	*Macaca assamensis*	Old World monkey	NextDenovo
Cercopithecidae	*Macaca*	*Macaca silenus*	Old World monkey	NextDenovo
Cercopithecidae	*Papio*	*Papio hamadryas*	Old World monkey	NextDenovo
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	Old World monkey	Wtdbg2
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	Old World monkey	NextDenovo
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	Old World monkey	Wtdbg2
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	Old World monkey	Wtdbg2
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	Old World monkey	NextDenovo
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	Old World monkey	Wtdbg2
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	Old World monkey	Wtdbg2
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	Old World monkey	Wtdbg2
Cercopithecidae	***Colobus***	***Colobus guereza***	**Old World monkey**	**NextDenovo+3d-dna**
Callitrichidae	***Saguinus***	***Saguinus midas***	**New World monkey**	**NextDenovo+3d-dna**
Cebidae	*Sapajus*	*Sapajus apella*	New World monkey	NextDenovo
Cebidae	*Cebus*	*Cebus albifrons*	New World monkey	NextDenovo
Atelidae	*Ateles*	*Ateles geoffroyi*	New World monkey	Wtdbg2+scaffold10x
Pitheciidae	*Pithecia*	*Pithecia pithecia*	New World monkey	Falcon+scaffold10x
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	Tarsier	Wtdbg2
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	Strepsirrhini	Wtdbg2
Lorisidae	*Loris*	*Loris tardigradus*	Strepsirrhini	Wtdbg2
Lorisidae	***Nycticebus***	***Nycticebus bengalensis***	**Strepsirrhini**	**Wtdbg2+3d-dna**
Galagidae	*Galago*	*Galago moholi*	Strepsirrhini	Wtdbg2

table S3. Statistics of short-read sequencing data for 27 primate species in this study. Sequencing bases, sequencing depth and mapping ratio are listed.

Family	Genus	Species name	Sequencing bases	Sequencing depth	Mapping ratio (%)
Hominidae	*Pongo*	*Pongo pygmaeus*	165,144,953,400	57.62	99.02
Hylobatidae	*Nomascus*	*Nomascus siki*	197,205,561,300	70.86	98.75
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	373,670,997,300	134.70	98.64
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	152,500,809,000	54.76	98.80
Hylobatidae	*Hylobates*	*Hylobates pileatus*	154,924,394,550	54.33	98.33
Cercopithecidae	*Macaca*	*Macaca assamensis*	163,242,183,600	59.16	99.26
Cercopithecidae	*Macaca*	*Macaca silenus*	162,518,565,150	58.49	99.06
Cercopithecidae	*Papio*	*Papio hamadryas*	163,180,148,400	53.43	99.34
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	176,317,463,400	60.73	95.27
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	203,585,704,800	71.78	98.58
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	181,737,628,950	65.07	98.79
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	170,760,436,650	60.48	99.03
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	206,236,279,350	66.59	98.73
Cercopithecidae	*Trachypithecus*	*Pygathrix nigripes*	177,307,191,000	61.22	98.92
Cercopithecidae	*Pygathrix*	*Trachypithecus crepusculus*	174,362,084,550	60.74	98.59
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	256,135,731,710	87.29	92.67
Cercopithecidae	*Colobus*	*Colobus guereza*	192,735,010,500	65.04	98.44
Callitrichidae	*Saguinus*	*Saguinus midas*	138,936,031,800	46.80	98.54
Cebidae	*Sapajus*	*Sapajus apella*	170,992,923,300	61.85	99.24
Cebidae	*Cebus*	*Cebus albifrons*	150,831,284,850	52.47	99.05
Atelidae	*Ateles*	*Ateles geoffroyi*	386,880,162,150	144.20	98.90
Pitheciidae	*Pithecia*	*Pithecia pithecia*	400,370,988,600	147.12	99.20
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	139,605,027,728	47.48	97.40
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	173,969,409,668	72.12	99.00
Lorisidae	*Loris*	*Loris tardigradus*	157,028,756,647	57.36	99.18
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	173,774,219,700	60.84	98.56
Galagidae	*Galago*	*Galago moholi*	151,403,904,696	59.52	98.58

table S4. Statistics of Hi-C sequencing data for 4 primate species in this study. The four species were *Hylobates pileatus*, *Colobus guereza*, *Saguinus midas* and *Nycticebus bengalensis.*

Family	Genus	Species name	Sequencing data (bases)	Coverage (X)
Hylobatidae	*Hylobates*	*Hylobates pileatus*	321,553,728,602	112.77
Cercopithecidae	*Colobus*	*Colobus guereza*	285,772,688,923	96.43
Callitrichidae	*Saguinus*	*Saguinus midas*	434,754,499,557	146.45
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	288,361,915,627	100.94

table S5. Genome coverage statistics of 27 primate genomes newly assembled in this study. Estimated genome sizes for each primate species were obtained from 17 k-mer–based estimations. The assembled genomes were mostly consistent with the k-mer–based estimations, indicating relatively complete genome coverage for all species.

Family	Genus	Species name	Genome size (bp)	Estimated genome sizes (bp)
Hominidae	*Pongo*	*Pongo pygmaeus*	2,866,029,761	3,207,162,863
Hylobatidae	*Nomascus*	*Nomascus siki*	2,783,107,017	3,037,418,990
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	2,774,050,050	3,836,928,248
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	2,785,098,935	2,898,596,937
Hylobatidae	*Hylobates*	*Hylobates pileatus*	2,851,478,483	2,944,662,251
Cercopithecidae	*Macaca*	*Macaca assamensis*	2,759,375,986	2,976,116,000
Cercopithecidae	*Macaca*	*Macaca silenus*	2,778,791,713	3,024,651,073
Cercopithecidae	*Papio*	*Papio hamadryas*	3,054,335,410	3,036,963,873
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	2,903,351,443	3,424,136,245
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	2,836,346,068	2,981,473,710
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	2,793,076,453	3,382,339,205
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	2,823,243,693	3,316,217,175
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	3,096,929,461	3,543,033,517
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	2,870,799,684	3,054,185,533
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	2,896,210,064	3,167,888,479
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	2,934,222,284	3,249,344,642
Cercopithecidae	*Colobus*	*Colobus guereza*	2,963,502,603	3,376,011,948
Callitrichidae	*Saguinus*	*Saguinus midas*	2,968,712,378	3,354,491,578
Cebidae	*Sapajus*	*Sapajus apella*	2,764,443,388	3,394,526,181
Cebidae	*Cebus*	*Cebus albifrons*	2,874,442,237	3,368,565,361
Atelidae	*Ateles*	*Ateles geoffroyi*	2,683,028,796	3,440,324,378
Pitheciidae	*Pithecia*	*Pithecia pithecia*	2,721,439,713	3,506,517,155
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	2,940,271,818	3,114,266,101
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	2,412,055,963	2,589,794,023
Lorisidae	*Loris*	*Loris tardigradus*	2,737,734,828	2,921,981,950
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	2,856,368,736	3,302,942,615
Galagidae	*Galago*	*Galago moholi*	2,543,922,824	2,759,821,024

table S6. Assembly statistics for genomes from 27 primate species in this study.
Three assembly parameters (Contig N50, Scaffold N50 and BUSCO) are listed. All the
assemblies yielded BUSCO complete scores >92%.

Family	Genus	Species name	Contig N50 (bp)	Scaffold N50 (bp)	BUSCO (%)
Hominidae	*Pongo*	*Pongo pygmaeus*	15,810,633	15,810,633	96.3
Hylobatidae	*Nomascus*	*Nomascus siki*	8,523,852	19,995,890	95.2
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	13,408,629	13,408,629	95.5
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	22,338,800	22,338,800	95.0
Hylobatidae	*Hylobates*	*Hylobates pileatus*	17,385,161	131,825,309	95.5
Cercopithecidae	*Macaca*	*Macaca assamensis*	27,348,716	27,348,716	96.6
Cercopithecidae	*Macaca*	*Macaca silenus*	25,691,658	25,691,658	96.2
Cercopithecidae	*Papio*	*Papio hamadryas*	24,989,958	24,989,958	95.8
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	6,802,059	6,802,059	96.5
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	20,999,372	20,999,372	95.2
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	13,914,891	13,914,891	95.9
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	9,639,656	9,639,656	96.4
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	18,747,272	18,747,272	96.4
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	14,781,066	14,781,066	95.0
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	14,260,291	14,260,291	95.1
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	16,766,498	16,766,498	95.2
Cercopithecidae	*Colobus*	*Colobus guereza*	15,612,012	145,964,134	96.2
Callitrichidae	*Saguinus*	*Saguinus midas*	10,987,047	143,273,090	95.2
Cebidae	*Sapajus*	*Sapajus apella*	12,583,278	12,583,278	96.0
Cebidae	*Cebus*	*Cebus albifrons*	18,502,463	18,502,463	95.8
Atelidae	*Ateles*	*Ateles geoffroyi*	29,212,752	29,212,752	96.5
Pitheciidae	*Pithecia*	*Pithecia pithecia*	10,867,430	10,867,430	94.5
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	5,252,243	5,252,243	94.4
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	27,954,151	27,954,151	97.5
Lorisidae	*Loris*	*Loris tardigradus*	7,727,436	7,727,436	96.1
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	19,794,679	130,794,859	94.1
Galagidae	*Galago*	*Galago moholi*	583,552	583,552	92.0

table S7. Gene predictions of protein-coding genes across sequenced primate species in this study. The gene number, average CDS length and average gene length are given for all species.

Family	Genus	Species name	Gene number	Average CDS length (bp)	Average gene length (bp)
Hominidae	*Pongo*	*Pongo pygmaeus*	21,468	1,729	44,298
Hylobatidae	*Nomascus*	*Nomascus siki*	20,362	1,704	42,883
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	21,439	1,624	42,774
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	20,942	1,664	44,844
Hylobatidae	*Hylobates*	*Hylobates pileatus*	20,656	1,697	40,462
Cercopithecidae	*Macaca*	*Macaca assamensis*	20,960	1,740	43,952
Cercopithecidae	*Macaca*	*Macaca silenus*	20,689	1,825	46,755
Cercopithecidae	*Papio*	*Papio hamadryas*	21,226	1,709	44,802
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	21,374	1,746	44,113
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	20,607	1,785	44,758
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	20,668	1,740	42,051
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	21,038	1,728	42,632
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	20,967	1,731	41,633
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	20,368	1,681	44,209
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	20,667	1,700	41,554
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	20,837	1,661	41,159
Cercopithecidae	*Colobus*	*Colobus guereza*	20,768	1,676	39,866
Callitrichidae	*Saguinus*	*Saguinus midas*	20,303	1,739	42,549
Cebidae	*Sapajus*	*Sapajus apella*	20,378	1,801	45,455
Cebidae	*Cebus*	*Cebus albifrons*	20,332	1,785	43,566
Atelidae	*Ateles*	*Ateles geoffroyi*	20,593	1,786	43,717
Pitheciidae	*Pithecia*	*Pithecia pithecia*	20,372	1,600	40,652
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	20,775	1,566	39,334
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	20,066	1,816	39,734
Lorisidae	*Loris*	*Loris tardigradus*	20,423	1,784	41,329
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	21,354	1,687	41,245
Galagidae	*Galago*	*Galago moholi*	20,789	1,663	32,053

table S8. Downloaded genome version for comparative genomics analysis in this study. Genome version id and associated database origins are given.

Family	Genus	Species Name	Genome version	Origin
Hominidae	*Homo*	*Homo sapiens*	GRCh38.p12	ENSEMBL 97
Hominidae	*Gorilla*	*Gorilla gorilla*	GCA_900006655.3	NCBI
Hominidae	*Pan*	*Pan paniscus*	GCA_000258655.2	NCBI
Hominidae	*Pan*	*Pan troglodytes*	GCF_002880755.1	NCBI
Hominidae	*Pongo*	*Pongo abelii*	GCA_002880775.3	NCBI
Cercopithecidae	*Macaca*	*Macaca mulatta*	GCA_008058575.1	NCBI
Cercopithecidae	*Macaca*	*Macaca nemestrina*	GCA_000956065.1	NCBI
Cercopithecidae	*Papio*	*Papio anubis*	GCA_000264685.2	NCBI
Cercopithecidae	*Theropithecus*	*Theropithecus gelada*	GCA_003255815.1	NCBI
Cercopithecidae	*Mandrillus*	*Mandrillus leucophaeus*	PRJNA785018	NCBI
Cercopithecidae	*Cercocebus*	*Cercocebus atys*	GCA_000955945.1	NCBI
Cercopithecidae	*Chlorocebus*	*Chlorocebus sabaeus*	GCA_000409795.2	NCBI
Cercopithecidae	*Cercopithecus*	*Cercopithecus mona*	GCA_014849445.1	NCBI
Cercopithecidae	*Piliocolobus*	*Piliocolobus tephrosceles*	GCA_002776525.2	NCBI
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus roxellana*	GCA_007565055.1	NCBI
Cercopithecidae	*Colobus*	*Colobus angolensis*	GCA_000951035.1	NCBI
Callitrichidae	*Callithrix*	*Callithrix jacchus*	GCA_011100535.2	NCBI
Aotidae	*Aotus*	*Aotus nancymaae*	GCA_000952055.2	NCBI
Cheirogaleidae	*Microcebus*	*Microcebus murinus*	GCA_000165445.3	NCBI
Lemuridae	*Prolemur*	*Prolemur simus*	GCA_003258685.1	NCBI
Galagidae	*Otolemur*	*Otolemur garnettii*	GCA_000181295.3	NCBI
Lemuridae	*Lemur*	*Lemur catta*	PRJNA562215	NCBI
Lorisidae	*Nycticebus*	*Nycticebus pygmaeus*	Our accompanying paper (*32*)	GSA
Cynocephalidae	*Galeopterus*	*Galeopterus variegatus*	PRJNA399345	NCBI
Tupaiidae	*Tupaia*	*Tupaia belangeri*	TreeshrewDB v2.0: http://www.treeshrewdb.org	GSA

table S9. Aligned length statistics for multiple species genome alignments. The genome size, aligned length and alignment rate are given.

Family	Genus	Species name	Genome size	Aligned length	Alignment rate
Hominidae	*Homo*	*Homo sapiens*	3,099,750,718	2,859,711,246	0.92
Hominidae	*Pan*	*Pan troglodytes*	3,024,031,013	2,817,711,027	0.93
Hominidae	*Pan*	*Pan paniscus*	3,286,643,938	2,747,645,364	0.84
Hominidae	*Gorilla*	*Gorilla gorilla*	3,084,595,669	2,794,306,801	0.91
Hominidae	*Pongo*	*Pongo abelii*	3,065,035,716	2,752,899,888	0.9
Hominidae	*Pongo*	*Pongo pygmaeus*	2,866,029,761	2,714,380,444	0.95
Hylobatidae	*Nomascus*	*Nomascus siki*	2,783,107,017	2,603,900,630	0.94
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	2,774,050,050	2,634,493,348	0.95
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	2,785,098,935	2,603,678,188	0.93
Hylobatidae	*Hylobates*	*Hylobates pileatus*	2,851,478,483	2,584,499,780	0.91
Cercopithecidae	*Macaca*	*Macaca mulatta*	2,955,490,605	2,569,702,476	0.87
Cercopithecidae	*Macaca*	*Macaca assamensis*	2,759,375,986	2,557,823,014	0.93
Cercopithecidae	*Macaca*	*Macaca nemestrina*	2,948,703,511	2,555,795,260	0.87
Cercopithecidae	*Macaca*	*Macaca silenus*	2,778,791,713	2,558,701,946	0.92
Cercopithecidae	*Papio*	*Papio hamadryas*	3,054,335,410	2,544,396,386	0.83
Cercopithecidae	*Papio*	*Papio anubis*	2,959,356,508	2,565,346,833	0.87
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	2,903,351,443	2,568,026,562	0.88
Cercopithecidae	*Theropithecus*	*Theropithecus gelada*	2,889,614,139	2,575,616,509	0.89
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	2,836,346,068	2,527,559,704	0.89
Cercopithecidae	*Mandrillus*	*Mandrillus leucophaeus*	2,935,729,887	2,576,476,604	0.88
Cercopithecidae	*Cercocebus*	*Cercocebus atys*	2,848,246,356	2,563,099,972	0.9
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	2,793,076,453	2,534,872,735	0.91
Cercopithecidae	*Cercopithecus*	*Cercopithecus mona*	2,902,804,697	2,517,749,510	0.87
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	2,823,243,693	2,554,745,799	0.9
Cercopithecidae	*Chlorocebus*	*Chlorocebus sabaeus*	2,789,639,778	2,571,549,904	0.92
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	3,096,929,461	2,528,277,858	0.82
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	2,870,799,684	2,513,718,414	0.88
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	2,896,210,064	2,527,627,308	0.87
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	2,934,222,284	2,508,236,626	0.85
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus roxellana*	3,038,460,517	2,564,419,952	0.84
Cercopithecidae	*Piliocolobus*	*Piliocolobus tephrosceles*	3,038,007,327	2,548,107,659	0.84
Cercopithecidae	*Colobus*	*Colobus guereza*	2,963,502,603	2,510,173,070	0.85
Cercopithecidae	*Colobus*	*Colobus angolensis*	2,970,124,662	2,486,046,307	0.84
Callitrichidae	*Callithrix*	*Callithrix jacchus*	2,863,864,636	2,238,835,124	0.78
Callitrichidae	*Saguinus*	*Saguinus midas*	2,968,712,378	2,227,242,188	0.75
Aotidae	*Aotus*	*Aotus nancymaae*	2,861,668,348	2,253,128,038	0.79
Cebidae	*Sapajus*	*Sapajus apella*	2,764,443,388	2,263,802,381	0.82
Cebidae	*Cebus*	*Cebus albifrons*	2,874,442,237	2,251,805,898	0.78
Atelidae	*Ateles*	*Ateles geoffroyi*	2,683,028,796	2,278,151,074	0.85

Pitheciidae	*Pithecia*	*Pithecia pithecia*	2,721,439,713	2,299,935,171	0.85
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	2,940,271,818	1,828,871,241	0.62
Cheirogaleidae	*Microcebus*	*Microcebus murinus*	2,487,392,024	1,722,717,486	0.69
Lemuridae	*Prolemur*	*Prolemur simus*	2,411,593,676	1,706,107,078	0.71
Lemuridae	*Lemur*	*Lemur catta*	2,122,351,751	1,703,765,223	0.8
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	2,412,055,963	1,870,779,156	0.78
Lorisidae	*Loris*	*Loris tardigradus*	2,737,734,828	1,617,637,987	0.59
Lorisidae	*Nycticebus*	*Nycticebus pygmaeus*	2,828,132,058	1,619,133,058	0.57
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	2,856,368,736	1,600,575,412	0.56
Galagidae	*Galago*	*Galago moholi*	2,543,922,824	1,602,622,867	0.63
Galagidae	*Otolemur*	*Otolemur garnettii*	2,519,724,550	1,590,869,944	0.63
Cynocephalidae	*Galeopterus*	*Galeopterus variegatus*	3,349,451,543	1,748,653,931	0.52
Tupaiidae	*Tupaia*	*Tupaia belangeri*	2,667,507,236	1,378,718,279	0.52

table S10. 37 (of 50) representative genomes for the analysis of chromosome evolution and genome rearrangement during primate evolution. Primate genomes with scaffold N50 ≥13 Mb were retained for this analysis. The outgroup species *Tupaia belangeri* was included in this analysis.

Familiy	Genus	Species name	Genome size (bp)	Contig N50 (bp)	Scaffold N50 (bp)
Hominidae	*Homo*	*Homo sapiens*	3,099,750,718	50,761,348	145,138,636
Hominidae	*Pan*	*Pan paniscus*	3,286,643,938	61,424	144,709,823
Hominidae	*Pan*	*Pan troglodytes*	3,024,031,013	12,421,315	130,995,916
Hominidae	*Gorilla*	*Gorilla gorilla*	3,084,595,669	9,406,846	129,679,333
Hominidae	*Pongo*	*Pongo abelii*	3,065,035,716	11,074,009	132,178,492
Hominidae	*Pongo*	*Pongo pygmaeus*	2,866,029,761	15,810,633	15,810,633
Hylobatidae	*Nomascus*	*Nomascus siki*	2,783,107,017	8,523,852	19,995,890
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	2,785,098,935	22,338,800	22,338,800
Hylobatidae	*Hylobates*	*Hylobates pileatus*	2,851,478,483	17,385,161	131,825,309
Cercopithecidae	*Macaca*	*Macaca mulatta*	2,955,490,605	8,353,667	152,195,021
Cercopithecidae	*Macaca*	*Macaca assamensis*	2,759,375,986	27,348,716	27,348,716
Cercopithecidae	*Macaca*	*Macaca nemestrina*	2,948,703,511	106,822	15,219,753
Cercopithecidae	*Papio*	*Papio hamadryas*	3,054,335,410	24,989,958	24,989,958
Cercopithecidae	*Papio*	*Papio anubis*	2,959,356,508	138,819	140,346,614
Cercopithecidae	*Theropithecus*	*Theropithecus gelada*	2,889,614,139	225,709	130,230,028
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	2,836,346,068	20,999,372	20,999,372
Cercopithecidae	*Mandrillus*	*Mandrillus leucophaeus*	2,935,729,887	124,362	26,881,038
Cercopithecidae	*Cercocebus*	*Cercopithecus albogularis*	2,793,076,453	13,914,891	13,914,891
Cercopithecidae	*Cercopithecus*	*Cercopithecus mona*	2,902,804,697	22,791,723	22,791,723
Cercopithecidae	*Chlorocebus*	*Chlorocebus sabaeus*	2,789,639,778	74,669	101,219,884
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	2,870,799,684	14,781,066	14,781,066
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	2,896,210,064	14,260,291	14,260,291
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	2,934,222,284	16,766,498	16,766,498
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus roxellana*	3,038,460,517	5,723,610	144,559,847
Cercopithecidae	*Piliocolobus*	*Piliocolobus tephrosceles*	3,038,007,327	98,446	13,534,873
Cercopithecidae	*Colobus*	*Colobus guereza*	2,963,502,603	15,612,012	145,964,134
Callitrichidae	*Callithrix*	*Callithrix jacchus*	2,863,864,636	13,255,626	129,995,920
Callitrichidae	*Saguinus*	*Saguinus midas*	2,968,712,378	10,987,047	143,273,090
Cebidae	*Cebus*	*Cebus albifrons*	2,874,442,237	18,502,463	18,502,463
Atelidae	*Ateles*	*Ateles geoffroyi*	2,683,028,796	29,212,752	29,212,752
Cheirogaleidae	*Microcebus*	*Microcebus murinus*	2,487,392,024	210,702	108,171,978
Lemuridae	*Lemur*	*Lemur catta*	2,233,154,473	25,671,697	25,671,697
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	2,412,055,963	27,954,151	27,954,151
Lorisidae	*Nycticebus*	*Nycticebus pygmaeus*	2,828,132,058	6,142,509	136,556,743
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	2,856,705,436	19,794,679	130,794,859
Galagidae	*Otolemur*	*Otolemur garnettii*	2,519,724,550	27,100	13,852,661
Tupaiidae	*Tupaia*	*Tupaia belangeri*	2,667,507,236	3,217,288	104,643,080

table S11. Segmental duplications identified in the genomes of 35 primates and 5 outgroup species in this study. Segmental duplications of ≥ 5kb were retained for downstream comparative genomics analysis. The outgroup species were *Cephalopachus bancanus*, *Tupaia belangeri*, *Mus musculus*, *Felis catus* and *Sus scrofa*.

Species name	number	total_len (bp)	ave_len (bp)	min_len (bp)	max_len (bp)
Pan troglodytes	10,124	103,234,954	11,186	5,001	96,148
Gorilla gorilla	7,386	72,980,326	10,410	5,000	91,694
Pongo abelii	14,154	132,584,212	10,283	5,000	90,792
Pongo pygmaeus	1,384	8,490,743	6,265	5,000	35,146
Hylobates pileatus	3,592	23,856,770	7,030	5,000	45,954
Symphalangus syndactylus	5,023	57,209,090	11,961	5,000	146,164
Hoolock leuconedys	1,154	7,200,734	6,310	5,001	21,258
Nomascus siki	781	4,539,606	5,943	5,000	13,786
Macaca mulatta	10,264	131,917,476	13,904	5,000	149,503
Macaca assamensis	2,674	18,528,014	7,363	5,001	83,638
Macaca silenus	2,383	15,059,924	6,662	5,000	42,550
Mandrillus sphinx	3,809	25,968,581	7,252	5,000	61,387
Papio hamadryas	6,364	41,776,416	6,978	5,000	79,026
Lophocebus aterrimus	2,464	14,451,590	6,368	5,001	35,346
Cercopithecus mona	6,695	45,116,068	7,084	5,000	41,154
Cercopithecus albogularis	1,827	11,136,346	6,347	5,000	26,600
Erythrocebus patas	5,636	39,290,548	7,323	5,000	75,941
Chlorocebus aethiops	1,747	10,190,572	6,224	5,000	24,535
Colobus guereza	7,945	51,372,575	7,328	5,000	102,894
Trachypithecus crepusculus	1,654	9,604,991	5,927	5,001	49,667
Pygathrix nigripes	2,110	12,224,016	5,937	5,001	21,684
Rhinopithecus strykeri	2,922	17,907,851	6,349	5,000	57,979
Rhinopithecu roxellana	13,605	135,025,913	10,724	5,000	96,660
Pithecia pithecia	6,117	66,550,338	11,530	5,001	103,216
Ateles geoffroyi	1,502	10,231,427	7,145	5,001	51,758
Sapajus apella	1,069	6,449,939	6,342	5,001	43,391
Cebus albifrons	5,425	37,464,152	7,349	5,001	43,775
Callithrix jacchus	8,202	85,611,181	11,063	5,001	83,467
Saguinus midas	5,990	40,642,413	7,337	5,000	48,231
Cephalopachus bancanus	6,156	65,449,215	10,988	5,001	78,425
Daubentonia madagascariensis	1,771	10,443,328	6,319	5,000	38,891
Lemur catta	2,475	21,469,237	9,006	5,000	84,638
Nycticebus pygmaeus	4,908	27,522,559	6,038	5,000	25,325
Nycticebus bengalensis	3,688	20,697,751	5,853	5,000	21,119
Loris tardigradus	4,095	24,648,780	6,373	5,000	29,931
Galeopterus variegatus	2,911	19,071,037	6,576	5,001	32,550
Tupaia belangeri	3,938	32,827,464	8,971	5,000	67,521
Mus musculus	21,600	165,071,458	11,111	5,000	528,086
Sus scrofa	6,962	71,755,339	11,320	5,000	164,094
Felis catus	2,991	31,776,404	10,908	5,000	217,003

table S12. Screening criteria for lineage-specific segmental duplications in this study.
5 outgroup species (*Galeopterus variegatus*, *Tupaia belangeri*, *Mus musculus*, *Felis catus* and *Sus scrofa*) were included in this analysis.

Lineage	Lineage specific species number cutoff
Great apes	3
Hominoidea	Great apes ≥3 and Gibbons ≥1
Catarrhini	Hominoidea ≥4 and OWMs ≥4
Simiiformes	Catarrhini ≥8 and NWMs ≥1
Haplorrhini	Catarrhini ≥8, NWMs ≥1 and Tarsiers ≥1
Strepsirrhini	2
Primate	Catarrhini ≥8 and NWMs ≥1 and Strepsirrhini ≥1

table S13. 89 genes overlapping lineage–specific segmental duplications in different evolutionary nodes in primates. The 'NA' denotes no overlapping genes with primate-specific segmental duplications.

Lineage	Gene name
Great apes	AC013271.1, AC105052.3, ADAMTSL1, AL445238.1, AQP7, CA5A, CBWD1, CDK11A, CDK11B, COL4A6, DACH2, DRD5, DUX4, EDA, EFCAB3, FAM227B, FGF7, GYPA, GYPB, HERC2, IARS, IGSF3, IL1RAPL2, INTS4, KLHL2, LIMS1, LIMS3, LRRC69, LRRTM4, METTL2A, METTL2B, MMP23B, MTHFD1L, NIPAL2, NUTM2E, OPHN1, PKD1, PLD5, PLEKHB2, POLR2J, POLR2J2, PRAMEF1, PRKG1, PROS1, PRUNE2, RANBP2, RASA4B, RGPD4, RGPD5, SCAI, SPTLC1, TBX20, TENM1, TLK2, TUSC3, UNC93B1, ZNF827
Hominoidea	ALG10, ALG10B, EVA1C, KRT17, RUNDC3B, SMG1
Catarrhini	AC011330.3, CATSPER2, CCL4, CCL4L2, CSH2, CSHL1, DHRS4, DHRS4L2, GGT1, GGTLC2, PDE4DIP, SPANXB1, SPANXC, ZNRF2
Simiiformes	ANKRD18A, ANKRD18B, DEFB131B, FAM95C, GBA, GOLGA6L9, GOLGA8Q, GOLGA8T, GRM5, MTX1, PLEKHB2, SLC9B1, TUBB8
Haplorrhini	NA
Primate	NA

table S14. Biological functions of 47 genes overlapping segmental duplications in different primate lineages. The Human Gene Mutation Database can be searched via the website (http://www.hgmd.cf.ac.uk/ac/index.php).

Gene Symbol	Diseases/Phenotype
ADAMTSL1	Autism spectrum disorder ? Glaucoma, craniofacial and other systemic features
AQP7	No exercise induced glycerol increase, association Obesity, association with
CA5A	Carbonic anhydrase VA deficiency Hyperammonaemia
CDK11A	Diabetes, type 2, association with Retinal folds, microcephaly & mental retardation ?
COL4A6	Chronic kidney disease ? Diffuse leiomyomatosis in Alport syndrome Nonsyndromic hearing loss, X-linked Premature ovarian failure ?
DACH2	Growth retardation, intellectual disability & walking difficulties ?
DRD5	Receptor variant Receptor deficiency Schizophrenia ?
EDA	Ectodermal dysplasia, hypohidrotic Ectodermal dysplasia Tooth agenesis Ectodermal dysplasia, hypohidrotic ? Ectodermal dysplasia, hypohidrotic & inability to sweat Hypodontia Oligodontia Ectodermal dysplasia ? Abnormality of metabolism/homeostasis Autism spectrum disorder with ectodermal dysplasia Hypohidrotic ectodermal dysplasia with bilateral amastia
EFCAB3	Diabetic retinopathy ?
GYPA	Blood group variation Blood group Erik variant Haemolytic disease of the newborn M blood type variant MNS antigen, absence
GYPB	Blood group variation Ss blood group variation
HERC2	Autism spectrum disorder Autism spectrum disorder ? Brain abnormalities, transposition of the great arteries, ventricular septal defect, renal anomaly and hearing loss Global developmental delay and autism spectrum disorder Mental retardation, autosomal recessive 38 Neurological disease ?
IGSF3	Autism spectrum disorder ? Nasolacrimal duct obstruction
LRRC69	Altered deoxycarnitine levels Autism spectrum disorder ?
LRRTM4	Autism spectrum disorder
METTL2B	Autism spectrum disorder ?
MMP23B	Epileptic encephalopathy, early onset, with burst suppression ?
MTHFD1L	Neural tube defects, increased risk, association with Neural tube defects, decreased risk, association with
OPHN1	Mental retardation syndrome, X-linked Intellectual disability Cerebellar hypoplasia Intellectual disability ?

	Abnormality of the nervous system
	Autism spectrum disorder ?
	Developmental delay
	Epilepsy, syndromic
	Intellectual disability & dysmorphic features
	Intellectual disability & hippocampal alterations
	Intellectual disability, autism & myopathy
	Mental retardation and cerebellar hypoplasia
	Mental retardation and epilepsy
	Mental retardation, motor impairment & seizures ?
	Mental retardation, seizures and tall stature
	Mental retardation, seizures, ataxia, hypotonia
	Mental retardation, seizures, hypoplasia & facial dysmorphism
	Mental retardation, X-linked
	Schizophrenia, childhood onset ?
	Seizures, intellectual disability & brain malformations
	Speech delay, learning difficulties & behavioural disorders
PKD1	Polycystic kidney disease 1
	Polycystic kidney disease 1 ?
	Intracranial aneurysms ?
	Polycystic kidney disease, autosomal dominant
	Polycystic kidney disease 1, association with
	Phenotype modifier
	Polycystic kidney disease, autosomal dominant, with male infertility
	Phenotype modifier ?
	Polycystic kidney disease, autosomal dominant & cerebral cavernous malformation
	Polycystic kidney disease, autosomal dominant & testicular germ cell tumour
	Renal cysts ?
	Renal hypoplasia ?
PLD5	Autism spectrum disorder ?
PRAMEF1	Autism spectrum disorder ?
PRKG1	Thoracic aortic aneurysms and dissections ?
	Thoracic aortic aneurysms and dissections
PROS1	Protein S deficiency
	Protein S deficiency ?
	Deep vein thrombosis
	Thrombophilia
	Deep vein thrombosis ?
	Protein S deficiency, type I
	Protein S deficiency, type I/III
	Protein S deficiency, type III
	Autism spectrum disorder ?
	Increased plasma Protein S
	Late-onset thrombosis
	Lipodystrophy, familial partial 3
	Protein S deficiency with thrombophilia
	Pulmonary embolism with deep venous thrombosis
	Recurrent miscarriage ?
	Reduced plasma Protein S
	Thrombotic disease
PRUNE2	Autism spectrum disorder ?
	Schizophrenia ?
RANBP2	Encephalopathy, acute necrotising
	Autism spectrum disorder ?
SPTLC1	Neuropathy, hereditary sensory, type I
	Neuropathy, motor and sensory, type 2 ?
	Sensory and autonomic neuropathy ?
TBX20	Congenital heart disease
	Congenital heart disease ?
	Atrial septal defect ?
	Cardiovascular malformations
	Atrial septal defect

Gene	Disease
	Cardiomyopathy, dilated ?
	Congenital heart defect, protection against
	Tetralogy of Fallot ?
	Atrial septal defects, patent foramen ovale & cardiac valve defec
	Cardiomyopathy, dilated
	Noncompaction, left ventricular ?
	Tetralogy of Fallot/truncus arteriosus
	Ventricular septal defect
TENM1	Anosmia, general congenital
	Autism spectrum disorder ?
	Cerebral palsy ?
	Schizophrenia ?
TLK2	Neurodevelopmental disorder
	Autism spectrum disorder ?
	Intellectual disability ?
	Neotenic complex syndrome ?
	Schizophrenia ?
TUSC3	Mental retardation, non-syndromic, autosomal recessive
	Intellectual disability & autism
	Intellectual disability, nonsyndromic, autosomal recessive
	Attention deficit hyperactivity disorder ?
	Fetal alcohol syndrome, predisposition to ?
	Intellectual disability ?
	Intellectual disability and dysmorphic features
	Intellectual disability with developmental delay
	Intellectual disability, autosomal recessive ?
	Intellectual disability, syndromic
	Mental retardation, non-syndromic, autosomal recessive ?
	Neurodegenerative disease with blindness
UNC93B1	Herpes simplex encephalitis, UNC93B-deficient
ALG10B	Acquired long QT syndrome
	Acquired long QT syndrome, protection against, association
	Fetal alcohol syndrome, predisposition to ?
KRT17	Pachyonychia congenital
	Steatocystoma multiplex
	Diabetes, MODY
	Keratitis-ichthyosis-deafness syndrome, modifier of ?
	Steatocystoma multiplex ?
RUNDC3B	Autism spectrum disorder ?
	Schizophrenia ?
SMG1	Bipolar disorder ?
CATSPER2	Asthenoteratozoospermia & deafness, non-syndromic
	Oligozoospermia
	Asthenozoospermia ?
	Deafness-infertility syndrome
CCL4L2	Haemophilia A, inhibitor development, increased risk
	HIV/AIDS susceptibility, association with
DHRS4L2	Abdominal pain, association with
GGT1	Glutathionuria
PDE4DIP	Lung squamous cancer ?
	Schizophrenia ?
SPANXC	Autism spectrum disorder ?
	Hearing loss ?
ANKRD18A	Thrombocytopaenia ?
GBA	Gaucher disease
	Gaucher disease 1
	Gaucher disease 2
	Parkinson disease ?
	Parkinson disease
	Gaucher disease ?
	Gaucher disease 3

	Parkinson disease, familial
	Parkinson disease, susceptibility ?
	Lewy body dementia
	Gaucher disease 1 ?
	Gaucher disease 3B ?
	Gaucher disease/Parkinson disease ?
	Parkinson disease, modifier of ?
	Gaucher disease/Parkinson disease
	Lewy body dementia ?
	Lewy body dementia/Alzheimer disease
	Parkinson disease with dementia
	Parkinson disease with dementia ?
	Parkinson disease, familial ?
	Parkinson disease, susceptibility
	Reduced activity ?
	Reduced promoter activity
GRM5	Nasopharyngeal carcinoma, association with
	Schizophrenia ?
	Attention deficit hyperactivity disorder
MTX1	Parkinson disease, GBA-associated, modifier of
TUBB8	Oocyte maturation arrest
	Oocyte meiotic I arrest
	Autism spectrum disorder ?
	Infertility, female

table S15. Transposable element content in the 50 primate genomes. The proportion of each species' genome that comprises transposable elements is given.

Species name	DNA (%)	LINE (%)	SINE (%)	LTR (%)	Other (%)	Unknown (%)	Total (%)
Homo sapiens	3.416312	21.37813	13.21746	8.995259	0	0.024094	51.16242
Pan troglodytes	3.495571	18.79533	12.07351	8.422344	0.000004	0.076046	43.50807
Pan paniscus	2.427124	16.20634	10.43182	6.576004	0.000003	0.210771	34.43042
Gorilla gorilla	4.519278	20.2437	11.86604	7.078511	0.000004	0.659051	39.98748
Pongo abelii	2.67973	23.74067	15.37229	7.337878	0.000006	0.134102	44.32494
Pongo pygmaeus	2.702728	19.41397	13.33098	7.78236	0.000007	2.467336	41.91725
Nomascus siki	4.079089	23.81682	12.69031	7.446399	0.000004	0.199346	41.7907
Symphalangus syndactylus	2.740564	23.24009	12.86542	7.437156	0.000004	0.406352	42.00199
Hoolock leuconedys	2.824069	20.91051	13.77718	7.509741	0.000004	0.801514	42.70155
Hylobates pileatus	2.569771	22.11408	14.06666	7.774654	0.000007	0.066315	43.4856
Macaca mulatta	2.627724	20.36664	13.21097	7.390647	0.000005	0.111009	40.33915
Macaca assamensis	2.697511	22.41217	13.5647	7.825183	0.000009	0.201319	41.9095
Macaca nemestrina	2.768158	19.2794	13.36188	7.242853	0.000005	0.038615	39.80063
Macaca silenus	2.997837	22.08491	13.67432	7.527221	0.000005	0.402573	41.66409
Papio hamadryas	2.652947	19.42902	13.962	7.63461	0.000004	0.529461	42.42333
Papio anubis	2.53303	19.37666	12.65983	7.305972	0.000005	0.9019	39.58893
Lophocebus aterrimus	2.626727	20.08604	13.79541	7.437517	0.000025	0.177457	41.38226
Theropithecus gelada	2.706185	20.36382	13.14712	7.545321	0.000005	0.88258	40.73389
Mandrillus sphinx	2.812574	19.77772	13.71045	7.695317	0	0.73101	41.99178
Mandrillus leucophaeus	2.828926	17.33862	12.47235	8.695023	0.000005	0.743809	39.7174
Cercocebus atys	2.672215	20.68596	11.63993	7.148165	0	0.674032	39.11317
Cercopithecus albogularis	2.755068	19.71068	13.60433	7.449134	0.000005	0.330782	41.74052
Cercopithecus mona	2.630284	18.48593	13.30515	7.526644	0.000005	2.183436	41.31828
Chlorocebus aethiops	2.599457	20.17746	13.33488	7.469227	0.000005	0.032981	41.27284
Chlorocebus sabaeus	2.837976	21.76352	12.71812	7.921284	0.000005	2.078198	40.18023
Erythrocebus patas	2.559519	20.5166	13.04474	7.116591	0.000004	0.184734	41.47241
Trachypithecus crepusculus	2.882639	22.74601	14.56216	7.387797	0	0.106141	43.86018
Pygathrix nigripes	2.71409	22.37422	13.46232	7.603127	0.000009	0.317145	42.98065
Rhinopithecus strykeri	2.671557	21.29688	13.6758	7.447362	0.000005	0.638443	42.11801
Rhinopithecus roxellana	2.880717	18.84279	12.81576	6.648548	0.000004	0.902558	40.62366
Piliocolobus tephrosceles	2.669056	17.30729	13.03671	7.287504	0.000171	0.366395	39.43253
Colobus guereza	2.535954	21.34616	14.08538	7.676713	0.000005	0.773977	42.91961
Colobus angolensis	2.629353	15.536	12.29748	7.150639	0.000008	1.120888	36.55863
Callithrix jacchus	2.597564	23.24307	12.15885	6.464473	0.000024	3.456594	41.00548
Saguinus midas	2.357737	23.22166	12.59711	6.240792	0.000004	0.751006	43.14452
Aotus nancymaae	2.680732	18.75946	12.22343	6.62525	0.000004	0.161474	37.78009
Sapajus apella	2.640896	18.49049	13.24118	6.420989	0.000004	0.96467	39.29696
Cebus albifrons	2.774134	19.58234	13.58767	6.555068	0.000004	0.223647	41.45974
Ateles geoffroyi	2.805585	18.53016	12.33164	6.749433	0.000016	0.506796	39.19657
Pithecia pithecia	2.715736	20.48659	12.08776	7.162978	0.000011	1.20784	40.45308
Cephalopachus bancanus	2.139191	22.48067	9.421322	5.381964	0	0.281605	39.70475
Microcebus murinus	3.816077	16.34522	7.552485	4.87733	0.000009	0.317364	31.68335
Prolemur simus	4.641324	16.17629	5.492357	5.251497	0.000011	0.760824	30.8309
Lemur catta	4.288137	17.2709	6.044139	5.60213	0.000016	0.516008	32.10227
Daubentonia madagascariensis	4.122496	21.68452	4.5172	6.286112	0.000005	1.126378	36.61008

Loris tardigradus	3.561639	26.23747	9.291736	4.129729	0	2.26674	42.27465
Nycticebus bengalensis	3.336392	27.28023	11.13431	4.050924	0	1.897464	44.32268
Nycticebus pygmaeus	3.097127	30.95835	9.84234	3.990493	0	1.790004	46.74688
Galago moholi	3.704521	22.27339	12.75717	4.777287	0.000004	1.283037	39.00963
Otolemur garnettii	3.733917	16.47771	10.67751	4.376582	0.000004	1.216316	33.63152

table S16. Substitution rates for all primate species. Substitution rates in different lineages were estimated by comparison of the four-fold degenerate sites within coding regions, in units of substitutions per site per million years. OWMs: Old World monkeys. NWMs: New World monkeys.

Common name	Species name	Substitution rate (×10⁻³)	Lineages
human	*Homo sapiens*	0.769360	Human
Chimpanzee	*Pan troglodytes*	0.690833	great apes
Bonobo	*Pan paniscus*	0.814966	great apes
Western Gorilla	*Gorilla gorilla*	0.829426	great apes
Sumatran Orangutan	*Pongo abelii*	0.553363	great apes
Bornean Orangutan	*Pongo pygmaeus*	0.716367	great apes
Southern White-cheeked Crested Gibbon	*Nomascus siki*	1.04689	gibbons
Siamang	*Symphalangus syndactylus*	1.26521	gibbons
Eastern Hoolock Gibbon	*Hoolock leuconedys*	1.28001	gibbons
Pileated Gibbon	*Hylobates pileatus*	1.00582	gibbons
Rhesus Macaque	*Macaca mulatta*	1.00803	OWMs
Assamese Macaque	*Macaca assamensis*	1.02882	OWMs
Sunda Pig-tailed Macaque	*Macaca nemestrina*	0.959020	OWMs
Lion-tailed Macaque	*Macaca silenus*	0.786217	OWMs
Hamadryas Baboon	*Papio hamadryas*	0.882062	OWMs
Olive Baboon	*Papio anubis*	1.02976	OWMs
Northern Black Crested Mangabey	*Lophocebus aterrimus*	1.14622	OWMs
Gelada	*Theropithecus gelada*	0.842621	OWMs
Mandrill	*Mandrillus sphinx*	0.967419	OWMs
Drill	*Mandrillus leucophaeus*	0.963773	OWMs
Sooty Mangabey	*Cercocebus atys*	1.10719	OWMs
Syke's Monkey	*Cercopithecus albogularis*	1.08521	OWMs
Mona Monkey	*Cercopithecus mona*	1.10985	OWMs
Grivet Monkey	*Chlorocebus aethiops*	1.18807	OWMs
Green Monkey	*Chlorocebus sabaeus*	1.16340	OWMs
Patas Monkey	*Erythrocebus patas*	1.22043	OWMs
Indochinese Gray Langur	*Trachypithecus crepusculus*	1.43968	OWMs
Black-shanked Douc	*Pygathrix nigripes*	1.29194	OWMs
Stryker's Snub-nosed Monkey	*Rhinopithecus strykeri*	1.13282	OWMs
Golden Snub-nosed Monkey	*Rhinopithecus roxellana*	0.759596	OWMs
Ashy Red Colobus	*Piliocolobus tephrosceles*	1.25882	OWMs
Guereza	*Colobus guereza*	1.09806	OWMs
Angolan Colobus	*Colobus angolensis*	1.34288	OWMs
Common Marmoset	*Callithrix jacchus*	1.97639	NWMs
Midas Tamarin	*Saguinus midas*	1.87414	NWMs
Ma's Night Monkey	*Aotus nancymaae*	1.33320	NWMs
Guianan Brown Capuchin	*Sapajus apella*	1.40714	NWMs

Humboldt's White-fronted Capuchin	*Cebus albifrons*	1.37799	NWMs
Central American Spider Monkey	*Ateles geoffroyi*	1.33453	NWMs
White-faced saki	*Pithecia pithecia*	1.39319	NWMs
Western Tarsier	*Cephalopachus bancanus*	2.73656	Tarsiiformes
Aye-aye	*Daubentonia madagascariensis*	1.01236	Chiromyiformes
Gray Mouse Lemur	*Microcebus murinus*	1.58555	lemurs (Lemuriformes)
Greater Bamboo Lemur	*Prolemur simus*	1.17898	lemurs (Lemuriformes)
Ring-tailed Lemur	*Lemur catta*	1.16764	lemurs (Lemuriformes)
Red Slender Loris	*Loris tardigradus*	1.26883	lorisoids (Lorisiformes)
Pygmy Slow Loris	*Nycticebus pygmaeus*	1.10755	lorisoids (Lorisiformes)
Bengal Slow Loris	*Nycticebus bengalensis*	1.22657	lorisoids (Lorisiformes)
Southern Lesser Galago	*Galago moholi*	1.22706	lorisoids (Lorisiformes)
Garnett's Greater Galago	*Otolemur garnettii*	1.08428	lorisoids (Lorisiformes)

table S17. Positively selected genes of the primate ancestor. *P* values were obtained by means of a χ^2 test.

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000168813	*ZNF507*	194.045838	4.16E-44
ENSG00000112333	*NR2E1*	114.063966	1.26E-26
ENSG00000177082	*WDR73*	110.224732	8.75E-26
ENSG00000106012	*IQCE*	104.365488	1.68E-24
ENSG00000131591	*C1orf159*	62.681446	2.43E-15
ENSG00000131013	*PPIL4*	33.652356	6.59E-09
ENSG00000156802	*ATAD2*	31.562434	1.93E-08
ENSG00000134250	*NOTCH2*	31.40739	2.09E-08
ENSG00000103064	*SLC7A6*	29.497584	5.60E-08
ENSG00000042832	*TG*	22.223052	2.43E-06
ENSG00000064607	*SUGP2*	17.063862	3.61E-05
ENSG00000138069	*RAB1A*	14.699082	0.000126108
ENSG00000125831	*CST11*	12.242636	0.000467098
ENSG00000105607	*GCDH*	11.905484	0.000559689
ENSG00000136929	*HEMGN*	11.5726	0.000669308
ENSG00000128815	*WDFY4*	11.14847	0.00084101
ENSG00000166105	*GLB1L3*	10.89862	0.000962359
ENSG00000137076	*TLN1*	10.56693	0.001151288
ENSG00000135090	*TAOK3*	10.397474	0.001261878
ENSG00000039560	*RAI14*	9.699604	0.001843077
ENSG00000259494	*MRPL46*	9.66599	0.001877106
ENSG00000197694	*SPTAN1*	9.390208	0.002181475
ENSG00000144306	*SCRN3*	9.269378	0.002330168
ENSG00000185324	*CDK10*	9.246788	0.002359088
ENSG00000121749	*TBC1D15*	9.222914	0.002390048
ENSG00000113716	*HMGXB3*	9.00032	0.002699323
ENSG00000164190	*NIPBL*	8.938858	0.002791672
ENSG00000172613	*RAD9A*	8.55853	0.00343907
ENSG00000108556	*CHRNE*	8.124474	0.004367165
ENSG00000145777	*TSLP*	7.998236	0.004682294
ENSG00000100034	*PPM1F*	7.928984	0.004864911
ENSG00000166257	*SCN3B*	7.847688	0.005088574
ENSG00000005884	*ITGA3*	7.730874	0.005428447
ENSG00000134882	*UBAC2*	7.53235	0.006060081
ENSG00000197858	*GPAA1*	7.528236	0.006073935
ENSG00000180354	*MTURN*	7.263946	0.007035254
ENSG00000016602	*CLCA4*	7.178636	0.007377678
ENSG00000173714	*WFIKKN2*	7.094918	0.007730283
ENSG00000176974	*SHMT1*	6.95937	0.008338139
ENSG00000147573	*TRIM55*	6.875982	0.008736174
ENSG00000084112	*SSH1*	6.870336	0.008763817
ENSG00000107404	*DVL1*	6.869678	0.008767045
ENSG00000172269	*DPAGT1*	6.864626	0.008791865
ENSG00000153774	*CFDP1*	6.86255	0.008802085
ENSG00000103248	*MTHFSD*	6.858066	0.008824201

ENSG00000212657	KRTAP16-1	6.843184	0.00889801
ENSG00000203908	KHDC3L	6.809798	0.009065902
ENSG00000109805	NCAPG	6.701512	0.009633119
ENSG00000116824	CD2	6.635904	0.009994346
ENSG00000196535	MYO18A	6.36168	0.011661108
ENSG00000168509	HJV	6.316366	0.011962855
ENSG00000151655	ITIH2	6.287248	0.012160973
ENSG00000128710	HOXD10	6.135984	0.013245781
ENSG00000175003	SLC22A1	5.943742	0.014769629
ENSG00000130413	STK33	5.860434	0.015484975
ENSG00000144635	DYNC1LI1	5.84579	0.015614362
ENSG00000135900	MRPL44	5.824368	0.015805646
ENSG00000156172	C8orf37	5.687832	0.017082947
ENSG00000122034	GTF3A	5.559834	0.018377157
ENSG00000167065	DUSP18	5.467248	0.019376113
ENSG00000205442	IZUMO3	5.450048	0.019567789
ENSG00000173157	ADAMTS20	5.449936	0.019569044
ENSG00000164619	BMPER	5.110024	0.023787991
ENSG00000107821	KAZALD1	5.102012	0.023898108
ENSG00000156510	HKDC1	5.064364	0.024422665
ENSG00000148935	GAS2	5.021528	0.025034068
ENSG00000125257	ABCC4	5.013858	0.02514521
ENSG00000102181	CD99L2	4.988836	0.025511364
ENSG00000108576	SLC6A4	4.945932	0.026152135
ENSG00000162623	TYW3	4.724068	0.029743298
ENSG00000168291	PDHB	4.627954	0.031455053
ENSG00000196419	XRCC6	4.49234	0.03404704
ENSG00000177084	POLE	4.456634	0.03476594
ENSG00000066777	ARFGEF1	4.45279	0.034844275
ENSG00000111269	CREBL2	4.358914	0.036815768
ENSG00000103335	PIEZO1	4.349808	0.037013119
ENSG00000186487	MYT1L	4.170962	0.041122444
ENSG00000139597	N4BP2L1	4.133202	0.042049715
ENSG00000117475	BLZF1	4.125438	0.042243084
ENSG00000102125	TAZ	4.027962	0.044751971
ENSG00000132952	USPL1	3.979192	0.046065656
ENSG00000146648	EGFR	3.974956	0.04618166

table S18. Enrichment analyses of tissue-specific expression of positively selected genes in the primate ancestral lineage. *P* value < 0.05 (Modified Fisher's Exact test) are listed.The enrichment analyses were performed by the DAVID Bioinformatics Resources including 8 tissue-specific high expression databases (CGAP_EST_QUARTILE, CGAP_SAGE_QUARTILE, GNF_U133A_QUARTILE, HPA_NORMAL_TISSUE, HPA_NORMAL_TISSUE_CELLTYPE, HPA_RNA_TISSUE, UNIGENE_EST_QUARTILE, and UP_TISSUE).

Category	Term	Count	PValue
UNIGENE_EST_QUARTILE	thymus_normal_3rd	24	0.00226
CGAP_EST_QUARTILE	21853:placenta_normal_3rd	5	0.002702
UNIGENE_EST_QUARTILE	juvenile (< 17 years old)_development_3rd	23	0.0029
CGAP_EST_QUARTILE	38100:bone marrow_neoplasia_3rd	6	0.00602
HPA_NORMAL_TISSUE_CELLTYPE	lung; endothelial cells	10	0.006961
CGAP_EST_QUARTILE	16533:prostate_neoplasia_3rd	4	0.007505
CGAP_EST_QUARTILE	15:colon_neoplasia_3rd	3	0.0087
CGAP_EST_QUARTILE	21544:cerebrum_normal_3rd	3	0.011464
GNF_U133A_QUARTILE	fetallung_3rd	16	0.011464
HPA_NORMAL_TISSUE_CELLTYPE	esophagus; squamous epithelial cells	38	0.01216
HPA_NORMAL_TISSUE	Esophagus	38	0.01216
HPA_NORMAL_TISSUE_CELLTYPE	pancreas; exocrine glandular cells	39	0.012388
UP_TISSUE	Cervix carcinoma	31	0.012815
HPA_NORMAL_TISSUE_CELLTYPE	endometrium 2; glandular cells	38	0.013619
UNIGENE_EST_QUARTILE	infant (< 3 years old)_development_3rd	17	0.014615
UP_TISSUE	Leukemic T-cell	18	0.016363
GNF_U133A_QUARTILE	salivarygland_3rd	53	0.018462
CGAP_SAGE_QUARTILE	168:uncharacterized tissue_mixture of human cancer cell lines_3rd	14	0.018688
HPA_NORMAL_TISSUE_CELLTYPE	epididymis; glandular cells	40	0.018981
HPA_NORMAL_TISSUE	Epididymis	40	0.018981
CGAP_EST_QUARTILE	20376:testi_neoplasia_3rd	6	0.019915
HPA_NORMAL_TISSUE_CELLTYPE	thyroid gland; glandular cells	39	0.022073
HPA_NORMAL_TISSUE	thyroid gland	39	0.022073
CGAP_SAGE_QUARTILE	406:brain_ependymoma_3rd	13	0.022198
UNIGENE_EST_QUARTILE	esophagus_normal_3rd	17	0.023408
CGAP_EST_QUARTILE	26742:uncharacterized tissue_neoplasia_3rd	5	0.02419
HPA_NORMAL_TISSUE_CELLTYPE	tonsil; squamous epithelial cells	37	0.024632
HPA_NORMAL_TISSUE	endometrium 2	38	0.024734
HPA_NORMAL_TISSUE_CELLTYPE	cervix; squamous epithelial cells	32	0.025071
UP_TISSUE	Erythroleukemia	27	0.025138
CGAP_EST_QUARTILE	40108:cerebrum_normal_3rd	4	0.025647
CGAP_SAGE_QUARTILE	1993:retina_central retina_3rd	11	0.027738
CGAP_SAGE_QUARTILE	1645:brain_null_3rd	13	0.029832
UP_TISSUE	Bone marrow	8	0.030151
CGAP_EST_QUARTILE	28506:brain_normal_3rd	5	0.030792
CGAP_EST_QUARTILE	21737:placenta_normal_3rd	3	0.031485
CGAP_EST_QUARTILE	26745:stomach_neoplasia_3rd	5	0.03229
UNIGENE_EST_QUARTILE	esophageal tumor_disease_3rd	17	0.033583
UNIGENE_EST_QUARTILE	abdominal cavity_normal_3rd	19	0.034186

HPA_NORMAL_TISSUE_CELLTYPE	testis; pachytene spermatocytes	16	0.035055
CGAP_EST_QUARTILE	21455:uncharacterized tissue_neoplasia_3rd	3	0.03533
CGAP_SAGE_QUARTILE	351:brain_anaplastic gradeIII, primary, brain_3rd	12	0.036658
HPA_NORMAL_TISSUE_CELLTYPE	endometrium 1; glandular cells	37	0.038555
CGAP_EST_QUARTILE	38092:uncharacterized tissue_uncharacterized histology_3rd	2	0.038981
HPA_NORMAL_TISSUE_CELLTYPE	heart muscle; cardiomyocytes	35	0.039722
HPA_NORMAL_TISSUE	heart muscle	35	0.039722
HPA_NORMAL_TISSUE_CELLTYPE	seminal vesicle; glandular cells	37	0.040065
HPA_NORMAL_TISSUE	seminal vesicle	37	0.040065
HPA_NORMAL_TISSUE_CELLTYPE	small intestine; glandular cells	38	0.040135
CGAP_SAGE_QUARTILE	1906:retina_Bilateral retinoblastoma, poorly differentiated, left orbit_3rd	13	0.040615
CGAP_SAGE_QUARTILE	1604:stem cell_null_3rd	10	0.040945
CGAP_EST_QUARTILE	16479:uncharacterized tissue_uncharacterized histology_3rd	5	0.042522
HPA_NORMAL_TISSUE_CELLTYPE	stomach 2; glandular cells	40	0.043287
HPA_NORMAL_TISSUE	stomach 2	40	0.043287
CGAP_EST_QUARTILE	38079:muscle_normal_3rd	5	0.044327
CGAP_EST_QUARTILE	19639:lung_normal_3rd	2	0.045906
HPA_NORMAL_TISSUE	Testis	45	0.047504
CGAP_SAGE_QUARTILE	174:mammary gland_breast carcinoma_3rd	13	0.048418
CGAP_EST_QUARTILE	21736:placenta_normal_3rd	3	0.048758
HPA_NORMAL_TISSUE_CELLTYPE	salivary gland; glandular cells	37	0.04889
HPA_NORMAL_TISSUE	salivary gland	37	0.04889
UP_TISSUE	Liver	29	0.049474
HPA_NORMAL_TISSUE	Pancreas	39	0.049964

table S19. 30 positively selected genes along with primate ancestral lineage exhibiting the biased expression in brain. The tissue-biased expression genes were assigned according to eight expression databases including CGAP_EST_QUARTILE, CGAP_SAGE_QUARTILE, GNF_U133A_QUARTILE, HPA_NORMAL_TISSUE, HPA_NORMAL_TISSUE_CELLTYPE, HPA_RNA_TISSUE, UNIGENE_EST_QUARTILE, and UP_TISSUE in the DAVID Bioinformatics Resources (https://david.ncifcrf.gov/summary.jsp).

Ensembl Gene ID	Gene Name	Gene Description
ENSG00000186487	*MYT1L*	myelin transcription factor 1 like
ENSG00000197694	*SPTAN1*	spectrin alpha, non-erythrocytic 1
ENSG00000180354	*MTURN*	maturin, neural progenitor differentiation regulator homolog
ENSG00000136929	*HEMGN*	hemogen
ENSG00000172269	*DPAGT1*	dolichyl-phosphate N-acetylglucosaminephosphotransferase 1
ENSG00000132952	*USPL1*	ubiquitin specific peptidase like 1
ENSG00000106012	*IQCE*	IQ motif containing E
ENSG00000102181	*CD99L2*	CD99 molecule like 2
ENSG00000146648	*EGFR*	epidermal growth factor receptor
ENSG00000135900	*MRPL44*	mitochondrial ribosomal protein L44
ENSG00000103248	*MTHFSD*	methenyltetrahydrofolate synthetase domain containing
ENSG00000130413	*STK33*	serine/threonine kinase 33
ENSG00000168813	*ZNF507*	zinc finger protein 507
ENSG00000107404	*DVL1*	dishevelled segment polarity protein 1
ENSG00000166257	*SCN3B*	sodium voltage-gated channel beta subunit 3
ENSG00000100034	*PPM1F*	protein phosphatase, Mg2+/Mn2+ dependent 1F
ENSG00000138069	*RAB1A*	RAB1A, member RAS oncogene family
ENSG00000196419	*XRCC6*	X-ray repair cross complementing 6
ENSG00000162623	*TYW3*	tRNA-yW synthesizing protein 3 homolog
ENSG00000144635	*DYNC1LI1*	dynein cytoplasmic 1 light intermediate chain 1
ENSG00000131591	*C1ORF159*	chromosome 1 open reading frame 159
ENSG00000064607	*SUGP2*	SURP and G-patch domain containing 2
ENSG00000121749	*TBC1D15*	TBC1 domain family member 15
ENSG00000039560	*RAI14*	retinoic acid induced 14
ENSG00000109805	*NCAPG*	non-SMC condensin I complex subunit G
ENSG00000117475	*BLZF1*	basic leucine zipper nuclear factor 1
ENSG00000145777	*TSLP*	thymic stromal lymphopoietin
ENSG00000176974	*SHMT1*	serine hydroxymethyltransferase 1
ENSG00000122034	*GTF3A*	general transcription factor IIIA
ENSG00000147573	*TRIM55*	tripartite motif containing 55

table S20. Positively selected genes of the Simiiformes ancestor. Orthologous sequences of 50 primates and two outgroup species (flying lemur and tree shrew) were utilized to obtain positively selected genes in the Simiiformes ancestor lineage using a branch-site model in PAML4. P values were calculated by means of a χ^2 test.

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000167037	SGSM1	449.3512	9.98E-100
ENSG00000134709	HOOK1	194.746826	2.93E-44
ENSG00000215271	HOMEZ	118.725922	1.20E-27
ENSG00000187955	COL14A1	112.06168	3.46E-26
ENSG00000129521	EGLN3	99.378944	2.09E-23
ENSG00000144674	GOLGA4	90.36033	1.99E-21
ENSG00000144668	ITGA9	85.376384	2.47E-20
ENSG00000075391	RASAL2	85.267824	2.61E-20
ENSG00000243710	CFAP57	80.786632	2.51E-19
ENSG00000188089	PLA2G4E	74.502796	6.06E-18
ENSG00000124743	KLHL31	69.900668	6.24E-17
ENSG00000006715	VPS41	65.227618	6.67E-16
ENSG00000174348	PODN	57.165606	4.01E-14
ENSG00000077254	USP33	45.49202	1.53E-11
ENSG00000094880	CDC23	43.07724	5.26E-11
ENSG00000018189	RUFY3	42.075758	8.78E-11
ENSG00000033100	CHPF2	37.896662	7.46E-10
ENSG00000152382	TADA1	33.028134	9.08E-09
ENSG00000165471	MBL2	29.562052	5.42E-08
ENSG00000102575	ACP5	26.667634	2.42E-07
ENSG00000144550	CPNE9	25.762424	3.86E-07
ENSG00000177606	JUN	25.614606	4.17E-07
ENSG00000090097	PCBP4	22.5868	2.01E-06
ENSG00000175203	DCTN2	19.515158	9.98E-06
ENSG00000176248	ANAPC2	18.61186	1.60E-05
ENSG00000197969	VPS13A	17.867988	2.37E-05
ENSG00000152270	PDE3B	16.856954	4.03E-05
ENSG00000164961	WASHC5	16.086088	6.05E-05
ENSG00000103353	UBFD1	15.883054	6.74E-05
ENSG00000170175	CHRNB1	15.635368	7.68E-05
ENSG00000128652	HOXD3	15.313044	9.11E-05
ENSG00000116337	AMPD2	14.881692	0.000114469
ENSG00000070367	EXOC5	13.781746	0.000205322
ENSG00000163072	NOSTRIN	13.611044	0.000224859
ENSG00000177683	THAP5	12.844554	0.000338463
ENSG00000125656	CLPP	12.652206	0.000375123
ENSG00000113845	TIMMDC1	12.469068	0.000413746
ENSG00000187642	PERM1	12.254928	0.000464031
ENSG00000122679	RAMP3	11.90888	0.00055867
ENSG00000197272	IL27	11.861378	0.000573101
ENSG00000104177	MYEF2	11.73929	0.000611944

ENSG00000173141	*MRPL57*	11.248472	0.000796886
ENSG00000160712	*IL6R*	11.228806	0.000805374
ENSG00000183035	*CYLC1*	11.008464	0.000906968
ENSG00000108344	*PSMD3*	10.913936	0.000954435
ENSG00000162891	*IL20*	10.166388	0.001430242
ENSG00000165935	*SMCO2*	9.922508	0.001632698
ENSG00000243927	*MRPS6*	9.81793	0.001728187
ENSG00000180878	*C11orf42*	9.62165	0.001922969
ENSG00000168016	*TRANK1*	9.534342	0.002016625
ENSG00000196843	*ARID5A*	9.404968	0.002163983
ENSG00000114942	*EEF1B2*	9.13835	0.002503067
ENSG00000083307	*GRHL2*	9.11578	0.002534139
ENSG00000023445	*BIRC3*	9.071566	0.002596146
ENSG00000237452	*BHMG1*	8.843928	0.002940656
ENSG00000185974	*GRK1*	8.689602	0.003200305
ENSG00000111254	*AKAP3*	8.682594	0.003212634
ENSG00000117505	*DR1*	8.5147	0.00352289
ENSG00000182150	*ERCC6L2*	8.508168	0.003535558
ENSG00000170419	*VSTM2A*	8.439194	0.00367219
ENSG00000141741	*MIEN1*	8.387434	0.003778239
ENSG00000185909	*KLHDC8B*	8.159626	0.00428332
ENSG00000154639	*CXADR*	8.073828	0.004490914
ENSG00000146670	*CDCA5*	8.045894	0.004560691
ENSG00000188991	*SLC15A5*	7.997822	0.004683365
ENSG00000242485	*MRPL20*	7.86438	0.005041811
ENSG00000100867	*DHRS2*	7.848354	0.0050867
ENSG00000196811	*CHRNG*	7.813086	0.005186925
ENSG00000162949	*CAPN13*	7.77842	0.005287402
ENSG00000012171	*SEMA3B*	7.599362	0.005838896
ENSG00000102904	*TSNAXIP1*	7.573866	0.005922063
ENSG00000110693	*SOX6*	7.557326	0.005976661
ENSG00000167656	*LY6D*	7.551428	0.005996254
ENSG00000176155	*CCDC57*	7.405518	0.006502412
ENSG00000123243	*ITIH5*	7.302212	0.006886977
ENSG00000283428	*CCDC195*	7.298964	0.006899438
ENSG00000148290	*SURF1*	7.138004	0.007546711
ENSG00000121410	*A1BG*	7.06324	0.007868157
ENSG00000154252	*GAL3ST2*	7.038236	0.007978758
ENSG00000178752	*ERFE*	7.026466	0.008031369
ENSG00000107897	*ACBD5*	6.976502	0.008258688
ENSG00000153802	*TMPRSS11D*	6.883142	0.008701247
ENSG00000172167	*MTBP*	6.784054	0.009197576
ENSG00000165730	*STOX1*	6.694496	0.0096711
ENSG00000136868	*SLC31A1*	6.692318	0.009682922
ENSG00000144410	*CPO*	6.557756	0.010442797
ENSG00000125779	*PANK2*	6.509746	0.010728483
ENSG00000203872	*C6orf163*	6.501334	0.010779359
ENSG00000149548	*CCDC15*	6.43685	0.011177668
ENSG00000131808	*FSHB*	6.328406	0.011881906

ENSG00000106436	*MYL10*	6.30312	0.012052566
ENSG00000100023	*PPIL2*	6.287638	0.012158297
ENSG00000165695	*AK8*	6.197936	0.012789938
ENSG00000182492	*BGN*	6.169942	0.012993887
ENSG00000177646	*ACAD9*	6.148424	0.013152926
ENSG00000133858	*ZFC3H1*	6.119822	0.013367426
ENSG00000129226	*CD68*	6.104664	0.013482558
ENSG00000006555	*TTC22*	6.087322	0.01361553
ENSG00000030066	*NUP160*	6.069492	0.013753651
ENSG00000171855	*IFNB1*	6.018916	0.014153337
ENSG00000004799	*PDK4*	5.983534	0.01444004
ENSG00000155307	*SAMSN1*	5.980414	0.014465606
ENSG00000120802	*TMPO*	5.927846	0.014903448
ENSG00000168487	*BMP1*	5.916176	0.015002485
ENSG00000157884	*CIB4*	5.909386	0.01506042
ENSG00000123454	*DBH*	5.8726	0.015378323
ENSG00000139946	*PELI2*	5.816662	0.015875045
ENSG00000151093	*OXSM*	5.786518	0.016149546
ENSG00000074964	*ARHGEF10L*	5.777486	0.016232743
ENSG00000104375	*STK3*	5.773854	0.016266323
ENSG00000132879	*FBXO44*	5.722368	0.016750125
ENSG00000142185	*TRPM2*	5.706242	0.01690469
ENSG00000060140	*STYK1*	5.686276	0.017098102
ENSG00000168314	*MOBP*	5.679996	0.017159407
ENSG00000171204	*TMEM126B*	5.673824	0.017219879
ENSG00000171060	*C8orf74*	5.628776	0.017667964
ENSG00000162763	*LRRC52*	5.611454	0.017843451
ENSG00000115368	*WDR75*	5.605486	0.017904328
ENSG00000155850	*SLC26A2*	5.59047	0.01805845
ENSG00000100422	*CERK*	5.523092	0.018767059
ENSG00000110448	*CD5*	5.436524	0.019719876
ENSG00000187210	*GCNT1*	5.4056	0.020072248
ENSG00000131697	*NPHP4*	5.39571	0.020186311
ENSG00000139540	*SLC39A5*	5.37829	0.020388853
ENSG00000145545	*SRD5A1*	5.356528	0.020644836
ENSG00000188163	*FAM166A*	5.34222	0.020814949
ENSG00000204444	*APOM*	5.323088	0.021044686
ENSG00000168356	*SCN11A*	5.275364	0.021629265
ENSG00000040487	*PQLC2*	5.261008	0.021808384
ENSG00000070182	*SPTB*	5.236584	0.02211666
ENSG00000170854	*RIOX2*	5.221002	0.022315689
ENSG00000157404	*KIT*	5.213524	0.022411864
ENSG00000087299	*L2HGDH*	5.205536	0.022515072
ENSG00000198885	*ITPRIPL1*	5.200184	0.022584497
ENSG00000183808	*RBM12B*	5.199092	0.02259869
ENSG00000149658	*YTHDF1*	5.196914	0.022627025
ENSG00000178828	*RNF186*	5.195598	0.022644163
ENSG00000104427	*ZC2HC1A*	5.180342	0.022843827
ENSG00000132600	*PRMT7*	5.170256	0.02297683

ENSG00000188389	*PDCD1*	5.169508	0.022986725
ENSG00000181418	*DDN*	5.153974	0.023193234
ENSG00000159261	*CLDN14*	5.144154	0.023324773
ENSG00000149308	*NPAT*	5.133838	0.023463789
ENSG00000105664	*COMP*	5.083712	0.02415161
ENSG00000008838	*MED24*	5.073708	0.024291369
ENSG00000119318	*RAD23B*	5.071646	0.02432028
ENSG00000123496	*IL13RA2*	5.067454	0.024379165
ENSG00000162738	*VANGL2*	5.036854	0.024813514
ENSG00000131773	*KHDRBS3*	4.973408	0.025739882
ENSG00000126218	*F10*	4.931624	0.026369525
ENSG00000164236	*ANKRD33B*	4.88775	0.027047914
ENSG00000144791	*LIMD1*	4.885796	0.027078546
ENSG00000173662	*TAS1R1*	4.842906	0.027760077
ENSG00000137831	*UACA*	4.817134	0.028178161
ENSG00000115841	*RMDN2*	4.710282	0.029982733
ENSG00000068650	*ATP11A*	4.659032	0.030890602
ENSG00000139496	*NUP58*	4.626058	0.031489835
ENSG00000100249	*C22orf31*	4.618984	0.031619962
ENSG00000197323	*TRIM33*	4.58952	0.032168022
ENSG00000132854	*KANK4*	4.572368	0.032491623
ENSG00000188313	*PLSCR1*	4.570416	0.032528665
ENSG00000004455	*AK2*	4.5162	0.033575326
ENSG00000166260	*COX11*	4.511038	0.033676798
ENSG00000116151	*MORN1*	4.485256	0.034188426
ENSG00000113296	*THBS4*	4.472648	0.034441575
ENSG00000183665	*TRMT12*	4.443008	0.035044451
ENSG00000115363	*EVA1A*	4.422854	0.035460678
ENSG00000039523	*RIPOR1*	4.394094	0.036063616
ENSG00000174943	*KCTD13*	4.284406	0.03846352
ENSG00000115091	*ACTR3*	4.27827	0.03860262
ENSG00000108244	*KRT23*	4.24545	0.039355639
ENSG00000113810	*SMC4*	4.241152	0.039455386
ENSG00000064933	*PMS1*	4.188884	0.040689875
ENSG00000078081	*LAMP3*	4.183464	0.040820187
ENSG00000185880	*TRIM69*	4.115186	0.042499852
ENSG00000055732	*MCOLN3*	4.092196	0.043081634
ENSG00000164405	*UQCRQ*	4.091046	0.043110955
ENSG00000180263	*FGD6*	4.082468	0.043330326
ENSG00000166582	*CENPV*	4.081766	0.04334833
ENSG00000127184	*COX7C*	4.076508	0.043483436
ENSG00000183161	*FANCF*	4.04233	0.044372518
ENSG00000106477	*CEP41*	4.024186	0.044852259
ENSG00000167775	*CD320*	4.022642	0.044893335
ENSG00000143436	*MRPL9*	4.014564	0.045108885
ENSG00000076344	*RGS11*	3.980968	0.046017111
ENSG00000140950	*MEAK7*	3.974056	0.046206346
ENSG00000151092	*NGLY1*	3.956014	0.046704172
ENSG00000164171	*ITGA2*	3.944738	0.047018175

ENSG00000143578	*CREB3L4*	3.939496	0.047164906
ENSG00000108826	*MRPL27*	3.922366	0.047647777
ENSG00000157593	*SLC35B2*	3.920568	0.047698762
ENSG00000133069	*TMCC2*	3.918826	0.047748214
ENSG00000126787	*DLGAP5*	3.918074	0.047769578
ENSG00000166860	*ZBTB39*	3.854156	0.049622887

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S21 (as a separate Excel file). Identification of lineage-specific accelerated regions along with the Simiiformes branch. Lineage-specific accelerated regions significance was determined by *FDR*≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S22 (as a separate Excel file). Significantly expanded gene families in the Simiiformes lineage in contrast to all other primate species. Unpaired t test P values were adjusted by Benjamini-Hochberg *FDR* (*FDR* < 0.05). Gene copy numbers in each species are listed. Red represents copy numbers in the Simiiformes species, whereas black represents copy numbers in other primate species. The species names (see Fig. 1) are shown by first three letter abbreviations of the Latin names. Here, human gene names of expanded gene families with at least one copy in human are used.

table S23. Functional enrichment analysis of the gene-set including positively selected genes, genes associated with lineage-specific accelerated regions, and expanded gene families in the Simiiformes lineage. Categories with *P* value < 0.05 (Modified Fisher's Exact test) are listed.

Category	Term	Count	P
GOTERM_CC_DIRECT	GO:0005829~cytosol	181	2.65E-08
GOTERM_CC_DIRECT	GO:0005654~nucleoplasm	131	5.31E-06
UP_SEQ_FEATURE	COMPBIAS:Basic and acidic residues	179	1.56E-05
UP_SEQ_FEATURE	COMPBIAS:Polar residues	210	3.51E-05
GOTERM_CC_DIRECT	GO:0005911~cell-cell junction	15	0.000151
GOTERM_MF_DIRECT	GO:0005515~protein binding	343	0.000184
KEGG_PATHWAY	hsa04510:Focal adhesion	16	0.000237
GOTERM_BP_DIRECT	GO:0001736~establishment of planar polarity	5	0.000404
GOTERM_MF_DIRECT	GO:0003712~transcription cofactor activity	12	0.000664
GOTERM_BP_DIRECT	GO:0008360~regulation of cell shape	12	0.00094
GOTERM_CC_DIRECT	GO:0016020~membrane	83	0.000943
UP_SEQ_FEATURE	REGION:Disordered	363	0.001003
GOTERM_CC_DIRECT	GO:0014069~postsynaptic density	16	0.001398
GOTERM_CC_DIRECT	GO:0005856~cytoskeleton	25	0.00177
KEGG_PATHWAY	hsa04151:PI3K-Akt signaling pathway	20	0.002232
GOTERM_BP_DIRECT	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	44	0.002714
UP_KW_CELLULAR_COMPONENT	KW-0539~Nucleus	168	0.003355
UP_KW_CELLULAR_COMPONENT	KW-0965~Cell junction	36	0.003462
UP_SEQ_FEATURE	REPEAT:ANK 5	11	0.003769
GOTERM_MF_DIRECT	GO:0017056~structural constituent of nuclear pore	5	0.005163
GOTERM_CC_DIRECT	GO:0005769~early endosome	16	0.00539
GOTERM_CC_DIRECT	GO:0005635~nuclear envelope	12	0.005437
UP_KW_BIOLOGICAL_PROCESS	KW-0804~Transcription	83	0.006561
UP_KW_BIOLOGICAL_PROCESS	KW-0805~Transcription regulation	81	0.006565
GOTERM_BP_DIRECT	GO:0030335~positive regulation of cell migration	14	0.006585
GOTERM_MF_DIRECT	GO:0030165~PDZ domain binding	8	0.006932
GOTERM_BP_DIRECT	GO:0060011~Sertoli cell proliferation	3	0.00745
GOTERM_BP_DIRECT	GO:0048193~Golgi vesicle transport	4	0.007631
UP_KW_CELLULAR_COMPONENT	KW-0770~Synapse	23	0.007874
UP_KW_CELLULAR_COMPONENT	KW-0206~Cytoskeleton	48	0.00823
GOTERM_BP_DIRECT	GO:0006897~endocytosis	12	0.008965
GOTERM_BP_DIRECT	GO:0016477~cell migration	14	0.009474
GOTERM_MF_DIRECT	GO:0051015~actin filament binding	13	0.009621
GOTERM_BP_DIRECT	GO:0030036~actin cytoskeleton organization	11	0.009653
UP_SEQ_FEATURE	REPEAT:ANK 1	14	0.010011
UP_SEQ_FEATURE	REPEAT:ANK 2	14	0.010325
GOTERM_CC_DIRECT	GO:0005796~Golgi lumen	8	0.011297
GOTERM_BP_DIRECT	GO:0007160~cell-matrix adhesion	8	0.011839
UP_SEQ_FEATURE	REPEAT:ANK 6	8	0.012984
GOTERM_CC_DIRECT	GO:0009986~cell surface	25	0.013084
UP_SEQ_FEATURE	REPEAT:ANK 4	11	0.013947
GOTERM_BP_DIRECT	GO:0045184~establishment of protein localization	5	0.014301

UP_KW_BIOLOGICAL_PROCESS	KW-0221~Differentiation	31	0.014739
GOTERM_BP_DIRECT	GO:0051220~cytoplasmic sequestering of protein	3	0.017081
UP_SEQ_FEATURE	REPEAT:ANK 3	12	0.017689
UP_SEQ_FEATURE	MOTIF:LXXLL motif 1	4	0.017692
UP_SEQ_FEATURE	MOTIF:LXXLL motif 2	4	0.017692
GOTERM_CC_DIRECT	GO:0034451~centriolar satellite	8	0.018591
GOTERM_MF_DIRECT	GO:0001540~beta-amyloid binding	7	0.018829
UP_SEQ_FEATURE	DOMAIN:GAGE	3	0.019016
GOTERM_BP_DIRECT	GO:0060976~coronary vasculature development	4	0.019107
GOTERM_BP_DIRECT	GO:0090090~negative regulation of canonical Wnt signaling pathway	9	0.019124
GOTERM_CC_DIRECT	GO:0005634~nucleus	158	0.020576
UP_SEQ_FEATURE	REPEAT:TPR 7	6	0.020953
UP_KW_CELLULAR_COMPONENT	KW-0963~Cytoplasm	157	0.021229
GOTERM_BP_DIRECT	GO:0048661~positive regulation of smooth muscle cell proliferation	6	0.022482
UP_SEQ_FEATURE	DOMAIN:DAD	3	0.023391
GOTERM_MF_DIRECT	GO:1990247~N6-methyladenosine-containing RNA binding	3	0.023579
UP_SEQ_FEATURE	REPEAT:Spectrin 2	4	0.02462
UP_SEQ_FEATURE	REPEAT:TPR 10	4	0.02462
UP_SEQ_FEATURE	REPEAT:Spectrin 1	4	0.02462
GOTERM_BP_DIRECT	GO:0046037~GMP metabolic process	3	0.025316
GOTERM_MF_DIRECT	GO:0005518~collagen binding	6	0.025332
UP_SEQ_FEATURE	REPEAT:TPR 3	9	0.025777
GOTERM_BP_DIRECT	GO:0050673~epithelial cell proliferation	5	0.025947
UP_KW_MOLECULAR_FUNCTION	KW-0009~Actin-binding	14	0.026433
GOTERM_BP_DIRECT	GO:0051649~establishment of localization in cell	7	0.026785
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	7	0.027221
KEGG_PATHWAY	hsa04810:Regulation of actin cytoskeleton	12	0.027247
GOTERM_MF_DIRECT	GO:0008139~nuclear localization sequence binding	4	0.027507
UP_KW_MOLECULAR_FUNCTION	KW-0010~Activator	27	0.027996
UP_SEQ_FEATURE	DOMAIN:GBD/FH3	3	0.028134
GOTERM_BP_DIRECT	GO:0051897~positive regulation of protein kinase B signaling	8	0.02865
UP_KW_BIOLOGICAL_PROCESS	KW-0833~Ubl conjugation pathway	29	0.028823
GOTERM_BP_DIRECT	GO:0010628~positive regulation of gene expression	20	0.02906
GOTERM_BP_DIRECT	GO:0046718~viral entry into host cell	7	0.029176
GOTERM_CC_DIRECT	GO:0001726~ruffle	7	0.029204
KEGG_PATHWAY	hsa04024:cAMP signaling pathway	12	0.029634
KEGG_PATHWAY	hsa05132:Salmonella infection	13	0.029753
GOTERM_BP_DIRECT	GO:0009101~glycoprotein biosynthetic process	3	0.029924
GOTERM_BP_DIRECT	GO:0038180~nerve growth factor signaling pathway	3	0.029924
GOTERM_BP_DIRECT	GO:0045842~positive regulation of mitotic metaphase/anaphase transition	3	0.029924
GOTERM_CC_DIRECT	GO:0043005~neuron projection	16	0.030172
GOTERM_BP_DIRECT	GO:0010592~positive regulation of lamellipodium assembly	4	0.031011
GOTERM_BP_DIRECT	GO:0070979~protein K11-linked ubiquitination	4	0.031011
GOTERM_CC_DIRECT	GO:0005814~centriole	9	0.031031
GOTERM_BP_DIRECT	GO:0051301~cell division	16	0.031381
GOTERM_BP_DIRECT	GO:0043410~positive regulation of MAPK cascade	9	0.031991
GOTERM_BP_DIRECT	GO:0001525~angiogenesis	12	0.032259
GOTERM_BP_DIRECT	GO:0043484~regulation of RNA splicing	6	0.033858
UP_KW_CELLULAR_COMPONENT	KW-0628~Postsynaptic cell membrane	9	0.03396

GOTERM_BP_DIRECT	GO:0010468~regulation of gene expression	12	0.0343
GOTERM_BP_DIRECT	GO:0071872~cellular response to epinephrine stimulus	3	0.034835
UP_SEQ_FEATURE	DNA_BIND:HMG box	5	0.035308
UP_KW_CELLULAR_COMPONENT	KW-0906~Nuclear pore complex	5	0.035323
GOTERM_MF_DIRECT	GO:0005102~receptor binding	17	0.035456
UP_SEQ_FEATURE	REPEAT:TPR 4	7	0.035457
GOTERM_CC_DIRECT	GO:0005737~cytoplasm	147	0.03582
GOTERM_CC_DIRECT	GO:0032991~macromolecular complex	25	0.036309
GOTERM_CC_DIRECT	GO:0015629~actin cytoskeleton	12	0.0372
GOTERM_CC_DIRECT	GO:0005925~focal adhesion	17	0.037435
UP_SEQ_FEATURE	REPEAT:TPR 6	6	0.037836
GOTERM_BP_DIRECT	GO:0019228~neuronal action potential	4	0.039636
GOTERM_BP_DIRECT	GO:0043388~positive regulation of DNA binding	4	0.039636
GOTERM_BP_DIRECT	GO:0006515~misfolded or incompletely synthesized protein catabolic process	3	0.040033
GOTERM_BP_DIRECT	GO:0060996~dendritic spine development	3	0.040033
GOTERM_CC_DIRECT	GO:0043240~Fanconi anaemia nuclear complex	3	0.04004
GOTERM_BP_DIRECT	GO:0007015~actin filament organization	8	0.040568
UP_SEQ_FEATURE	REPEAT:TPR 2	9	0.041128
UP_SEQ_FEATURE	REPEAT:TPR 1	9	0.041128
GOTERM_CC_DIRECT	GO:0005762~mitochondrial large ribosomal subunit	5	0.042015
GOTERM_MF_DIRECT	GO:0015276~ligand-gated ion channel activity	4	0.042676
GOTERM_CC_DIRECT	GO:0046658~anchored component of plasma membrane	4	0.042768
UP_SEQ_FEATURE	DOMAIN:L27	3	0.044372
UP_SEQ_FEATURE	REPEAT:Spectrin 7	3	0.044372
UP_SEQ_FEATURE	REPEAT:Spectrin 8	3	0.044372
UP_SEQ_FEATURE	REPEAT:Spectrin 9	3	0.044372
GOTERM_CC_DIRECT	GO:0012505~endomembrane system	8	0.044737
KEGG_PATHWAY	hsa04145:Phagosome	9	0.045402
GOTERM_BP_DIRECT	GO:0032707~negative regulation of interleukin-23 production	2	0.045484
GOTERM_BP_DIRECT	GO:0003210~cardiac atrium formation	2	0.045484
GOTERM_BP_DIRECT	GO:0006895~Golgi to endosome transport	3	0.045503
GOTERM_BP_DIRECT	GO:0006904~vesicle docking involved in exocytosis	4	0.04599
GOTERM_BP_DIRECT	GO:0031532~actin cytoskeleton reorganization	5	0.046711
GOTERM_MF_DIRECT	GO:0005200~structural constituent of cytoskeleton	7	0.047567
GOTERM_BP_DIRECT	GO:0007010~cytoskeleton organization	8	0.047582
UP_SEQ_FEATURE	REPEAT:TPR 5	6	0.047637
UP_SEQ_FEATURE	REGION:Cisternal side	2	0.048103
UP_SEQ_FEATURE	REGION:Required for targeting to the nucleus and nuclear pore complex	2	0.048103
UP_SEQ_FEATURE	REGION:Pore side	2	0.048103
GOTERM_MF_DIRECT	GO:0031267~small GTPase binding	13	0.0482
GOTERM_MF_DIRECT	GO:0038177~death receptor agonist activity	2	0.048311
UP_KW_MOLECULAR_FUNCTION	KW-9996~Developmental protein	34	0.048854
UP_KW_MOLECULAR_FUNCTION	KW-0217~Developmental protein	34	0.048854
UP_KW_CELLULAR_COMPONENT	KW-0967~Endosome	24	0.049146
GOTERM_BP_DIRECT	GO:0022900~electron transport chain	5	0.049176
GOTERM_CC_DIRECT	GO:0016235~aggresome	4	0.04936
GOTERM_CC_DIRECT	GO:0005912~adherens junction	9	0.049855

table S24. Tissue-specific expression enrichment analysis of the genes including positively selected genes, genes associated with lineage-specific accelerated regions, and expanded gene families in the Simiiformes lineage. Categories with *P* value < 0.05 (Modified Fisher's Exact test) are listed.

Category	Term	Count	*P*
CGAP_EST_QUARTILE	33578:brain_neoplasia_3rd	23	3.85E-06
CGAP_EST_QUARTILE	21078:genitourinary_neoplasia_3rd	11	1.10E-05
UP_TISSUE	Testis	166	1.43E-04
GNF_U133A_QUARTILE	Prostate_3rd	268	1.44E-04
CGAP_EST_QUARTILE	20008:placenta_normal_3rd	8	3.14E-04
CGAP_EST_QUARTILE	16366:cerebrum_normal_3rd	13	4.88E-04
UP_TISSUE	Brain	250	5.50E-04
GNF_U133A_QUARTILE	Cerebellum_3rd	292	5.97E-04
GNF_U133A_QUARTILE	fetalliver_3rd	112	6.01E-04
CGAP_EST_QUARTILE	21766:placenta_normal_3rd	11	6.81E-04
CGAP_EST_QUARTILE	379:heart_normal_3rd	19	7.28E-04
CGAP_SAGE_QUARTILE	656:brain_null_3rd	49	7.30E-04
GNF_U133A_QUARTILE	Smooth Muscle_3rd	134	7.73E-04
CGAP_EST_QUARTILE	10713:lymph node_normal_3rd	14	8.30E-04
CGAP_EST_QUARTILE	6372:uncharacterized tissue_uncharacterized histology_3rd	6	8.90E-04
CGAP_EST_QUARTILE	23995:uncharacterized tissue_uncharacterized histology_3rd	15	0.001453
CGAP_EST_QUARTILE	10749:colon_neoplasia_3rd	5	0.001656
GNF_U133A_QUARTILE	TONGUE_3rd	117	0.001681
UNIGENE_EST_QUARTILE	vascular_normal_3rd	100	0.001932
CGAP_EST_QUARTILE	10956:colon_neoplasia_3rd	8	0.002047
HPA_NORMAL_TISSUE_CELLTYPE	caudate; neuronal cells	201	0.002237
GNF_U133A_QUARTILE	BM-CD105+Endothelial_3rd	130	0.002368
CGAP_EST_QUARTILE	19994:placenta_normal_3rd	11	0.002648
CGAP_EST_QUARTILE	20249:colon_neoplasia_3rd	7	0.002745
CGAP_SAGE_QUARTILE	410:cerebellum_medulloblastoma_3rd	66	0.002865
CGAP_EST_QUARTILE	674:placenta_normal_3rd	9	0.003423
GNF_U133A_QUARTILE	WHOLE BLOOD_3rd	141	0.003478
CGAP_EST_QUARTILE	20574:placenta_normal_3rd	8	0.003642
CGAP_EST_QUARTILE	16016:uteru_uncharacterized histology_3rd	13	0.003682
CGAP_EST_QUARTILE	16593:mammary gland_normal_3rd	14	0.003731
HPA_NORMAL_TISSUE_CELLTYPE	cerebellum; cells in molecular layer	151	0.003883
GNF_U133A_QUARTILE	Uterus_3rd	90	0.003951
HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; neuropil	167	0.003966
CGAP_EST_QUARTILE	19998:placenta_normal_3rd	10	0.003969
UNIGENE_EST_QUARTILE	laryngeal cancer_disease_3rd	99	0.003986
HPA_NORMAL_TISSUE	caudate	221	0.004434
HPA_NORMAL_TISSUE_CELLTYPE	caudate; glial cells	154	0.004439
HPA_NORMAL_TISSUE	hippocampus	217	0.004893
CGAP_SAGE_QUARTILE	359:brain_Glioblastoma_3rd	47	0.004929
CGAP_EST_QUARTILE	19268:cerebrum_normal_3rd	3	0.005406
UNIGENE_EST_QUARTILE	larynx_normal_3rd	96	0.006266
CGAP_SAGE_QUARTILE	144:brain_astrocytoma grade II_3rd	48	0.006305
CGAP_EST_QUARTILE	20280:placenta_normal_3rd	7	0.006422

CGAP_EST_QUARTILE	13801:colon_normal_3rd	4	0.007133
CGAP_EST_QUARTILE	11116:head and neck_neoplasia_3rd	3	0.007982
CGAP_EST_QUARTILE	19377:placenta_normal_3rd	9	0.008055
UNIGENE_EST_QUARTILE	muscle_normal_3rd	119	0.00834
HPA_NORMAL_TISSUE_CELLTYPE	lymph node; non-germinal center cells	194	0.008416
HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; glial cells	162	0.008569
GNF_U133A_QUARTILE	Occipital Lobe_3rd	84	0.008877
CGAP_EST_QUARTILE	23547:placenta_normal_3rd	8	0.008973
CGAP_EST_QUARTILE	21477:uncharacterized tissue_neoplasia_3rd	6	0.009119
CGAP_EST_QUARTILE	492:thymu_normal_3rd	11	0.009715
HPA_NORMAL_TISSUE_CELLTYPE	placenta; trophoblastic cells	202	0.009798
UNIGENE_EST_QUARTILE	pituitary gland_normal_3rd	97	0.00989
CGAP_EST_QUARTILE	16884:mammary gland_neoplasia_3rd	5	0.010265
CGAP_EST_QUARTILE	11260:head and neck_neoplasia_3rd	6	0.010599
CGAP_SAGE_QUARTILE	655:vascular_normal liver_3rd	34	0.01094
CGAP_EST_QUARTILE	25:uncharacterized tissue_uncharacterized histology_3rd	9	0.011146
CGAP_EST_QUARTILE	19520:uncharacterized tissue_normal_3rd	5	0.011298
UNIGENE_EST_QUARTILE	pharynx_normal_3rd	63	0.011311
UNIGENE_EST_QUARTILE	ear_normal_3rd	81	0.012223
CGAP_EST_QUARTILE	20729:genitourinary_neoplasia_3rd	6	0.012235
CGAP_EST_QUARTILE	19072:colon_neoplasia_3rd	6	0.012235
CGAP_SAGE_QUARTILE	283:brain_anaplastic astrocytoma grade III_3rd	47	0.01236
HPA_NORMAL_TISSUE_CELLTYPE	hippocampus; glial cells	138	0.012769
UNIGENE_EST_QUARTILE	mammary gland_normal_3rd	128	0.012795
CGAP_EST_QUARTILE	40265:synovium_uncharacterized histology_3rd	12	0.013012
HPA_NORMAL_TISSUE_CELLTYPE	lung; alveolar cells type II	59	0.013192
HPA_NORMAL_TISSUE_CELLTYPE	parathyroid gland; glandular cells	188	0.013221
HPA_NORMAL_TISSUE	parathyroid gland	188	0.013221
CGAP_EST_QUARTILE	21531:cerebrum_normal_3rd	5	0.013558
CGAP_EST_QUARTILE	16631:mammary gland_normal_3rd	3	0.014436
CGAP_EST_QUARTILE	6259:mammary gland_neoplasia_3rd	3	0.014436
GNF_U133A_QUARTILE	PLACENTA_3rd	298	0.014903
CGAP_SAGE_QUARTILE	430:brain_null_3rd	57	0.015094
GNF_U133A_QUARTILE	Olfactory Bulb_3rd	74	0.015513
CGAP_EST_QUARTILE	686:cervix_neoplasia_3rd	11	0.015939
CGAP_EST_QUARTILE	23545:placenta_normal_3rd	7	0.015942
UNIGENE_EST_QUARTILE	juvenile (< 17 years old)_development_3rd	100	0.016062
CGAP_EST_QUARTILE	553:mammary gland_normal_3rd	8	0.016234
HPA_NORMAL_TISSUE_CELLTYPE	cerebellum; cells in granular layer	149	0.016828
HPA_NORMAL_TISSUE	placenta	232	0.01687
CGAP_EST_QUARTILE	26805:placenta_normal_3rd	16	0.017197
HPA_NORMAL_TISSUE_CELLTYPE	spleen; cells in red pulp	171	0.017318
HPA_NORMAL_TISSUE_CELLTYPE	lung; alveolar cells type I	44	0.017664
CGAP_SAGE_QUARTILE	604:cartilage_chondrosarcoma grade 2._3rd	32	0.017788
CGAP_EST_QUARTILE	16373:cerebrum_normal_3rd	6	0.018147
CGAP_EST_QUARTILE	32281:liver_normal_3rd	5	0.018892
HPA_NORMAL_TISSUE	cerebellum	223	0.019729
CGAP_EST_QUARTILE	16695:mammary gland_normal_3rd	9	0.019827
HPA_NORMAL_TISSUE_CELLTYPE	hippocampus; neuronal cells	196	0.019859

UNIGENE_EST_QUARTILE	tongue_normal_3rd	108	0.020126
CGAP_EST_QUARTILE	21106:genitourinary_neoplasia_3rd	8	0.020232
UP_TISSUE	Erythroleukemia	137	0.020339
HPA_NORMAL_TISSUE_CELLTYPE	soft tissue 2; fibroblasts	136	0.020461
CGAP_EST_QUARTILE	16290:cerebrum_normal_3rd	6	0.020473
CGAP_EST_QUARTILE	11528:stomach_neoplasia_3rd	6	0.020473
CGAP_EST_QUARTILE	20739:genitourinary_neoplasia_3rd	7	0.020578
CGAP_EST_QUARTILE	16329:cerebrum_normal_3rd	7	0.020578
HPA_NORMAL_TISSUE	fallopian tube	230	0.020632
CGAP_EST_QUARTILE	22090:nervou_normal_3rd	15	0.020775
HPA_NORMAL_TISSUE	lymph node	204	0.021038
GNF_U133A_QUARTILE	Pons_3rd	173	0.021359
CGAP_EST_QUARTILE	21491:uncharacterized tissue_neoplasia_3rd	10	0.021618
HPA_NORMAL_TISSUE	spleen	186	0.022012
GNF_U133A_QUARTILE	skin_3rd	57	0.022243
CGAP_EST_QUARTILE	19178:head and neck_neoplasia_3rd	3	0.022482
CGAP_SAGE_QUARTILE	646:mammary gland_invasive breast cancer ER+, PR+, Her2-, grade II_3rd	26	0.023385
HPA_NORMAL_TISSUE_CELLTYPE	endometrium 2; glandular cells	209	0.023539
CGAP_EST_QUARTILE	20591:uteru_neoplasia_3rd	5	0.023637
HPA_NORMAL_TISSUE_CELLTYPE	skin 1; eccrine glands	14	0.023766
CGAP_EST_QUARTILE	21747:placenta_normal_3rd	6	0.024314
CGAP_EST_QUARTILE	20571:placenta_normal_3rd	9	0.024726
HPA_NORMAL_TISSUE_CELLTYPE	cerebellum; Purkinje cells	177	0.025691
HPA_NORMAL_TISSUE	endometrium 2	214	0.025773
CGAP_SAGE_QUARTILE	1443:stomach_Adenocarcinoma_3rd	35	0.026107
CGAP_EST_QUARTILE	11470:stomach_neoplasia_3rd	8	0.026913
CGAP_EST_QUARTILE	20003:placenta_normal_3rd	3	0.027051
CGAP_EST_QUARTILE	1528:cerebrum_uncharacterized histology_3rd	9	0.027469
HPA_NORMAL_TISSUE	soft tissue 2	166	0.028199
CGAP_SAGE_QUARTILE	645:mammary gland_invasive breast cancer, ER+, PR+, Her-, grade II_3rd	28	0.028264
CGAP_EST_QUARTILE	13843:uncharacterized tissue_neoplasia_3rd	9	0.028429
HPA_NORMAL_TISSUE_CELLTYPE	endometrium 1; glandular cells	211	0.028445
CGAP_EST_QUARTILE	20062:uncharacterized tissue_neoplasia_3rd	7	0.028489
UP_TISSUE	Hippocampus	29	0.028881
CGAP_SAGE_QUARTILE	70:prostate_carcinoma_3rd	63	0.029364
CGAP_EST_QUARTILE	21725:placenta_normal_3rd	6	0.030115
CGAP_EST_QUARTILE	16496:lung_normal_3rd	6	0.030115
CGAP_EST_QUARTILE	11090:kidney_uncharacterized histology_3rd	8	0.030189
HPA_NORMAL_TISSUE_CELLTYPE	cerebral cortex; neuronal cells	208	0.030341
CGAP_EST_QUARTILE	21863:placenta_normal_3rd	7	0.031072
UP_TISSUE	Leukemic T-cell	79	0.031074
CGAP_EST_QUARTILE	32:colon_neoplasia_3rd	14	0.031094
GNF_U133A_QUARTILE	Appendix_3rd	118	0.031112
HPA_NORMAL_TISSUE_CELLTYPE	tonsil; non-germinal center cells	191	0.031126
CGAP_EST_QUARTILE	20123:uteru_neoplasia_3rd	6	0.03169
CGAP_EST_QUARTILE	16321:cerebrum_normal_3rd	3	0.031957
CGAP_EST_QUARTILE	23453:colon_neoplasia_3rd	3	0.031957
HPA_NORMAL_TISSUE_CELLTYPE	salivary gland; glandular cells	213	0.031974

HPA_NORMAL_TISSUE	salivary gland	213	0.031974
HPA_NORMAL_TISSUE_CELLTYPE	epididymis; glandular cells	225	0.032156
HPA_NORMAL_TISSUE	epididymis	225	0.032156
CGAP_EST_QUARTILE	20721:uncharacterized tissue_normal_3rd	10	0.032573
CGAP_EST_QUARTILE	937:gastrointestinal tract_normal_3rd	13	0.03349
CGAP_SAGE_QUARTILE	1906:retina_Bilateral retinoblastoma, poorly differentiated, left orbit_3rd	58	0.033925
CGAP_EST_QUARTILE	847:ovary_normal_3rd	15	0.034755
CGAP_EST_QUARTILE	21720:placenta_normal_3rd	5	0.035125
CGAP_EST_QUARTILE	21715:placenta_normal_3rd	9	0.035798
CGAP_SAGE_QUARTILE	263:cerebellum_medulloblastoma_3rd	57	0.036652
CGAP_EST_QUARTILE	21451:uncharacterized tissue_neoplasia_3rd	7	0.036684
CGAP_EST_QUARTILE	23731:genitourinary_normal_3rd	3	0.037182
CGAP_SAGE_QUARTILE	1649:brain__3rd	52	0.038069
CGAP_EST_QUARTILE	19862:testi_normal_3rd	10	0.038089
CGAP_SAGE_QUARTILE	420:mammary gland_normal breast tissue from a breast cancer patient (corresponding to IDC7)_3rd	39	0.038165
CGAP_EST_QUARTILE	23542:placenta_normal_3rd	6	0.038498
CGAP_EST_QUARTILE	40304:thymu_uncharacterized histology_3rd	8	0.038848
HPA_NORMAL_TISSUE_CELLTYPE	adipose tissue; adipocytes	128	0.039208
HPA_NORMAL_TISSUE	adipose tissue	128	0.039208
UNIGENE_EST_QUARTILE	uterine tumor_disease_3rd	121	0.039307
GNF_U133A_QUARTILE	spinalcord_3rd	69	0.039772
CGAP_SAGE_QUARTILE	1883:retina_Bilateral retinoblastoma, poorly differentiated, left orbit_3rd	48	0.039968
GNF_U133A_QUARTILE	TemporalLobe_3rd	93	0.040124
CGAP_EST_QUARTILE	32326:eye_normal_3rd	16	0.040321
CGAP_EST_QUARTILE	852:nervou_normal_3rd	14	0.040775
CGAP_SAGE_QUARTILE	350:brain_anaplastic gradeIII, primary, brain_3rd	46	0.040778
HPA_NORMAL_TISSUE	cerebral cortex	245	0.041106
HPA_NORMAL_TISSUE_CELLTYPE	skin 1; cells in basal layer	35	0.041643
CGAP_EST_QUARTILE	6163:mammary gland_neoplasia_3rd	3	0.042708
CGAP_EST_QUARTILE	40588:endocrine_uncharacterized histology_3rd	7	0.042907
CGAP_EST_QUARTILE	673:nervou_normal_3rd	8	0.04301
HPA_NORMAL_TISSUE	endometrium 1	216	0.0432
HPA_NORMAL_TISSUE_CELLTYPE	pancreas; exocrine glandular cells	213	0.04373
HPA_NORMAL_TISSUE	liver	195	0.04389
CGAP_SAGE_QUARTILE	356:brain_Glioblastoma_3rd	35	0.044292
UNIGENE_EST_QUARTILE	lymph node_normal_3rd	123	0.045889
HPA_RNA_TISSUE	testis	61	0.047302
CGAP_SAGE_QUARTILE	124:cerebellum_medulloblastoma, cerebellum_3rd	56	0.047322
CGAP_EST_QUARTILE	27218:ovary_neoplasia_3rd	7	0.047985
CGAP_EST_QUARTILE	21525:cerebrum_normal_3rd	10	0.04834
CGAP_EST_QUARTILE	16821:mammary gland_neoplasia_3rd	10	0.04834
CGAP_EST_QUARTILE	10715:lymph node_normal_3rd	10	0.04834
CGAP_EST_QUARTILE	20006:placenta_normal_3rd	3	0.048517
CGAP_EST_QUARTILE	19196:head and neck_neoplasia_3rd	3	0.048517
CGAP_SAGE_QUARTILE	506:white blood cells_invasive breast tumor (~2 cm), extensive high-grade comedo DCIS (~4 cm), positive lymph nodes_3rd	50	0.049104
CGAP_SAGE_QUARTILE	607:cartilage_Dedifferentiated chondrosarcoma lung metastasis_3rd	37	0.049192
CGAP_EST_QUARTILE	21660:brain_neoplasia_3rd	5	0.049339

CGAP_SAGE_QUARTILE	362:pancreas_null_3rd	35	0.049509
CGAP_SAGE_QUARTILE	647:mammary gland_normal breast tissue from a breast cancer patient (corresponding to IDC7)_3rd	24	0.049887

table S25. Five candidate genes including four genes associated with lineage-specific accelerated regions and one positively selected gene in the Simiiformes ancestral lineage in the pathway 'axon guidance' with the high expression in the human brain. The 5 candidates were identified as highly expressed genes in brain by the DAVID Tissue_Expression database (https://david.ncifcrf.gov/home.jsp). The positively selected gene is displayed in bold font.

Ensembl Gene ID	Gene Name	Tissue Category
ENSG00000044524	*EPHA3*	Cerebellum_3rd/brain_anaplastic gradeIII, primary, brain_3rd
ENSG00000136238	*RAC1*	cerebral cortex; neuropil/hippocampus/brain_astrocytoma grade II_3rd/ brain_anaplastic astrocytoma grade III_3rd/ hippocampus; glial cells/ cerebellum_medulloblastoma_3rd/ brain_3rd/ brain_anaplastic gradeIII, primary, brain_3rd/ cerebral cortex/ brain_Glioblastoma_3rd/ brain_neoplasia_3rd
ENSG00000196358	*NTNG2*	Brain
ENSG00000012171	***SEMA3B***	Brain/ Cerebellum_3rd
ENSG00000153993	*SEMA3D*	Cerebellum_3rd/ brain_anaplastic gradeIII, primary, brain_3rd

table S26. Positively selected genes involved in brain development from the primate common ancestor leading to the human lineage. 772 genes involved in 'GO:0007420~brain development' were retrieved from g:Profiler (104v) (https://biit.cs.ut.ee/gprofiler/gost). 772 genes overlapped positively selected genes from the primate ancestral lineage leading to the human lineage in order to identify PSGs involved in 'brain development' in each crucial node.

Lineage	Positively selected genes involved in 'brain development' (GO:0007420)
Primate ancestor	*SLC6A4, NR2E1, NIPBL, XRCC6*
Haplorrhini ancestor	*ATAT1, FRS2, PFKFB3, TCTN1*
Simiiformes ancestor	*GRHL2, SRD5A1, UQCRQ, AK8*
Catarrhini ancestor	*KDM1A, UBA6, SEMA3A, NCOA1, XRN2, PPP3CA, GDF7, CDK5*
Hominoidea ancestor	*DPYSL2, MAST1, SCYL2*
Great ape ancestor	*DUOX2*
Homininae ancestor	*SUN2, CDH22, RBPJ, LDB1*
Hominini ancestor	*TNR, TTC21B, CRH*
Human	*NRP1, NIN, ROBO1*

table S27. Positively selected gene involved in 'Microcephaly' from the Primate ancestor leading to the human lineage. 1,133 genes putatively involved in microcephaly (HP:0000252) were obtained from g:Profiler (v104) (https://biit.cs.ut.ee/gprofiler/gost).

Lineage	PSGs involved in 'Microcephaly' (HP:0000252)
Primate ancestor	*ITGA3, NIPBL, DPAGT1, WDR73, POLE, CDK10, MYT1L, SPTAN1*
Haplorrhini ancestor	*HCCS, UPB1, MOCS1, THOC6, SLC5A6, LINS1, GTPBP2, DALRD3, SIX6*
Simiiformes ancestor	*ACBD5, TIMMDC1, AMPD2, CLPP, PRMT7, NGLY1, KIT, TMEM126B, ERCC6L2, FANCF, CHRNG*
Catarrhini ancestor	*PNKP, ERCC8, SNAP29, ESS2, DOCK8, TDP2, DNMT3A, PAX3, PPP3CA, PEX10, CDK5, TP53RK, NDUFA6, PI4KA*
Hominoidea ancestor	*SALL4, RPL5, AGT, SCYL2, JMJD1C, BLM*
Great ape ancestor	*SLC18A2, IFT140*
Homininae ancestor	*POLR1A, TMEM165, XPR1, CPT2, RBPJ, LARP7, RUSC2*
Hominini ancestor	*KMT2E, MTFMT*
Human	*TRIT1, NIN, FARSB, GATA6, SMAD4, PISD*

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S28 (as a separate Excel file). Identification of lineage-specific accelerated regions along the Catarrhini branch. Lineage-specific accelerated regions significance was determined by FDR≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S29 (as a separate Excel file). Identification of lineage-specific accelerated regions along the great ape (Hominidae) branch. Lineage-specific accelerated regions significance was determined by FDR≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S30 (as a separate Excel file). Identification of lineage-specific accelerated regions along the human branch. Lineage-specific accelerated regions significance was determined by FDR≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

table S31. 15 genes associated with lineage-specific accelerated regions with a high level of expression in fetal brain for the great ape ancestral lineage. The 15 candidate genes were identified as being highly expressed in fetal brain by the DAVID Tissue_Expression database (UP_TISSUE) (https://david.ncifcrf.gov/home.jsp).

Ensembl Gene ID	Gene Name	Gene Description	UP_TISSUE
ENSG00000081189	MEF2C	myocyte enhancer factor 2C	Fetal brain
ENSG00000072315	TRPC5	transient receptor potential cation channel subfamily C member 5	Fetal brain
ENSG00000182985	CADM1	cell adhesion molecule 1	Fetal brain
ENSG00000163590	PPM1L	protein phosphatase, Mg2+/Mn2+ dependent 1L	Fetal brain
ENSG00000113645	WWC1	WW and C2 domain containing 1	Fetal brain
ENSG00000187905	LRRC74B	leucine rich repeat containing 74B	Fetal brain
ENSG00000058668	ATP2B4	ATPase plasma membrane Ca2+ transporting 4	Fetal brain
ENSG00000258818	RNASE4	ribonuclease A family member 4	Fetal brain
ENSG00000164742	ADCY1	adenylate cyclase 1	Fetal brain
ENSG00000170502	NUDT9	nudix hydrolase 9	Fetal brain
ENSG00000090776	EFNB1	ephrin B1	Fetal brain
ENSG00000126858	RHOT1	ras homolog family member T1	Fetal brain
ENSG00000089091	DZANK1	double zinc ribbon and ankyrin repeat domains	Fetal brain
ENSG00000151208	DLG5	discs large MAGUK scaffold protein 5	Fetal brain
ENSG00000147162	OGT	O-linked N-acetylglucosamine (GlcNAc) transferase	Fetal brain

table S32. 5 positively selected genes and 33 genes associated with lineage-specific accelerated regions showing a high level of expression in the human brain along with primate ancestral lineage leading to the human lineage. Candidate genes were identified as being highly expressed in the human brain-related tissues by the DAVID Tissue_Expression database (UP_TISSUE) (https://david.ncifcrf.gov/home.jsp).

Category	Gene Symbol
positively selected genes	*PPP3CA, ADCY2, SLC1A6, SLC18A2, SLC6A4*
genes associated with lineage-specific accelerated regions	*GRIA2, APP, CHRM3, GABRB1, CHRNA3, MAOB, GRIK4, HTR2C, CACNA1D, GRIK2, HTR4, RIMS1, GRIN2A, GRM7, GNG4, GRIA3, GRIA4, KCNJ3, UNC13C, SYT1, PPP2R5A, SLC6A11, KCNJ18, HTR5A, STX1B, ADCY9, GNAQ, GNAS, KRAS, KCNQ5, GRIK3, ADCY1, JAK2*

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S33 (as a separate Excel file). Identification of lineage-specific accelerated regions along the ape (Hominoidea) branch. Lineage-specific accelerated regions significance was determined by FDR≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

table S34. Positively selected genes of the gibbon ancestral lineage. Identification of positively selected genes in the gibbon ancestral lineage was used the branch-site model in PAML4. *P* values were calculated by means of a χ^2 test. Four positively selected genes (*LONP1, BRCA2*, *NEK1* and *SLC25A24*) were involved in abnormal upper limb bone morphology (HP:0040070) according to Human Phenotype Ontology in the g:Profiler database (https://biit.cs.ut.ee/gprofiler/gost).

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000137501	*SYTL2*	364.154914	3.51E-81
ENSG00000136542	*GALNT5*	266.771874	5.73E-60
ENSG00000130724	*CHMP2A*	235.329284	4.10E-53
ENSG00000164323	*CFAP97*	153.6108	2.82E-35
ENSG00000125107	*CNOT1*	107.103322	4.23E-25
ENSG00000124074	*ENKD1*	64.818516	8.21E-16
ENSG00000135378	*PRRG4*	64.717368	8.64E-16
ENSG00000172426	*RSPH9*	58.023872	2.59E-14
ENSG00000162300	*ZFPL1*	38.868894	4.53E-10
ENSG00000175785	*PRIMA1*	37.141416	1.10E-09
ENSG00000123307	*NEUROD4*	35.959902	2.01E-09
ENSG00000134369	*NAV1*	35.94184	2.03E-09
ENSG00000148985	*PGAP2*	31.912158	1.61E-08
ENSG00000196365	*LONP1*	25.110698	5.41E-07
ENSG00000180628	*PCGF5*	23.928804	1.00E-06
ENSG00000106290	*TAF6*	20.806838	5.08E-06
ENSG00000180739	*S1PR5*	20.433454	6.17E-06
ENSG00000157551	*KCNJ15*	19.417698	1.05E-05
ENSG00000070748	*CHAT*	19.300442	1.12E-05
ENSG00000139618	*BRCA2*	19.080764	1.25E-05
ENSG00000160867	*FGFR4*	18.803404	1.45E-05
ENSG00000057019	*DCBLD2*	18.073392	2.13E-05
ENSG00000105697	*HAMP*	17.98844	2.22E-05
ENSG00000170242	*USP47*	17.331206	3.14E-05
ENSG00000175556	*LONRF3*	16.491402	4.89E-05
ENSG00000173040	*EVC2*	16.239878	5.58E-05
ENSG00000133119	*RFC3*	14.55125	0.000136398
ENSG00000065675	*PRKCQ*	14.399382	0.000147851
ENSG00000115318	*LOXL3*	13.674512	0.000217385
ENSG00000143171	*RXRG*	13.632332	0.000222324
ENSG00000187848	*P2RX2*	12.302396	0.000452377
ENSG00000173020	*GRK2*	12.150848	0.000490656
ENSG00000081026	*MAGI3*	12.121408	0.000498463
ENSG00000110841	*PPFIBP1*	12.112154	0.000500942
ENSG00000089558	*KCNH4*	11.692894	0.000627392
ENSG00000284862	*CCDC39*	11.184462	0.000824852
ENSG00000077684	*JADE1*	11.026462	0.000898204
ENSG00000187772	*LIN28B*	10.743642	0.001046382
ENSG00000147548	*NSD3*	10.61213	0.001123482
ENSG00000076356	*PLXNA2*	10.572942	0.00114755

ENSG00000139151	*PLCZ1*	10.561626	0.001154596
ENSG00000254470	*AP5B1*	10.188478	0.001413209
ENSG00000162073	*PAQR4*	9.863488	0.001685912
ENSG00000167543	*TP53I13*	9.73626	0.001806681
ENSG00000162869	*PPP1R21*	9.673682	0.001869263
ENSG00000102383	*ZDHHC15*	9.395724	0.002174921
ENSG00000162344	*FGF19*	9.359302	0.002218566
ENSG00000104549	*SQLE*	8.93113	0.002803508
ENSG00000137672	*TRPC6*	8.86454	0.002907635
ENSG00000186795	*KCNK18*	8.768724	0.003064398
ENSG00000213722	*DDAH2*	8.71389	0.003157946
ENSG00000156414	*TDRD9*	8.564324	0.003428142
ENSG00000132740	*IGHMBP2*	8.500424	0.003550637
ENSG00000142599	*RERE*	8.47771	0.003595244
ENSG00000155761	*SPAG17*	8.391164	0.003770494
ENSG00000115361	*ACADL*	8.375528	0.003803071
ENSG00000168268	*NT5DC2*	8.227782	0.004125392
ENSG00000172197	*MBOAT1*	8.164908	0.004270864
ENSG00000136270	*TBRG4*	8.145464	0.004316901
ENSG00000108669	*CYTH1*	8.064678	0.00451365
ENSG00000274523	*RCC1L*	7.784724	0.005268984
ENSG00000127445	*PIN1*	7.684184	0.005570686
ENSG00000169605	*GKN1*	7.678632	0.005587851
ENSG00000135636	*DYSF*	7.665486	0.00562871
ENSG00000184436	*THAP7*	7.65978	0.005646539
ENSG00000165688	*PMPCA*	7.549944	0.006001194
ENSG00000213949	*ITGA1*	7.516216	0.006114599
ENSG00000113492	*AGXT2*	7.435096	0.006396395
ENSG00000135018	*UBQLN1*	7.409268	0.006488872
ENSG00000126216	*TUBGCP3*	7.38097	0.006591761
ENSG00000147255	*IGSF1*	7.303328	0.006882701
ENSG00000160472	*TMEM190*	7.190066	0.007330828
ENSG00000139410	*SDSL*	7.150844	0.007492872
ENSG00000079337	*RAPGEF3*	7.110632	0.007662809
ENSG00000105894	*PTN*	7.108974	0.0076699
ENSG00000163630	*SYNPR*	7.008828	0.008110876
ENSG00000153029	*MR1*	6.909976	0.008571614
ENSG00000140379	*BCL2A1*	6.838598	0.008920883
ENSG00000100416	*TRMU*	6.82763	0.008975829
ENSG00000096063	*SRPK1*	6.80745	0.009077831
ENSG00000137601	*NEK1*	6.568646	0.010379088
ENSG00000150782	*IL18*	6.52527	0.01063524
ENSG00000187806	*TMEM202*	6.42877	0.011228631
ENSG00000109674	*NEIL3*	6.42795	0.011233816
ENSG00000117280	*RAB29*	6.373028	0.011586772
ENSG00000131480	*AOC2*	6.264768	0.01231623
ENSG00000237441	*RGL2*	6.250534	0.012415587
ENSG00000169994	*MYO7B*	6.210048	0.012702719
ENSG00000151304	*SRFBP1*	6.139516	0.013219349

ENSG00000160094	*ZNF362*	6.073346	0.013723674
ENSG00000138483	*CCDC54*	6.047518	0.013925859
ENSG00000104312	*RIPK2*	5.973464	0.014522725
ENSG00000151779	*NBAS*	5.913754	0.015023124
ENSG00000089048	*ESF1*	5.738964	0.016592577
ENSG00000136546	*SCN7A*	5.73693	0.016611803
ENSG00000103994	*ZNF106*	5.669664	0.017260762
ENSG00000153446	*C16orf89*	5.449654	0.019572203
ENSG00000147570	*DNAJC5B*	5.400272	0.020133614
ENSG00000167861	*HID1*	5.375732	0.020418771
ENSG00000132359	*RAP1GAP2*	5.368004	0.020509434
ENSG00000166035	*LIPC*	5.273922	0.021647188
ENSG00000164532	*TBX20*	5.199398	0.022594712
ENSG00000182400	*TRAPPC6B*	5.124974	0.023583924
ENSG00000214376	*VSTM5*	5.119642	0.023656497
ENSG00000168769	*TET2*	5.110472	0.023781849
ENSG00000136573	*BLK*	5.04416	0.024709085
ENSG00000124785	*NRN1*	4.997492	0.025384076
ENSG00000138867	*GUCD1*	4.978124	0.025669805
ENSG00000164037	*SLC9B1*	4.90694	0.026748989
ENSG00000134917	*ADAMTS8*	4.837132	0.027853181
ENSG00000171840	*NINJ2*	4.793242	0.0285716
ENSG00000152464	*RPP38*	4.75275	0.029251512
ENSG00000134905	*CARS2*	4.750132	0.029296047
ENSG00000138944	*SHISAL1*	4.629372	0.031429066
ENSG00000012983	*MAP4K5*	4.613724	0.031717084
ENSG00000100418	*DESI1*	4.48855	0.034122608
ENSG00000102189	*EEA1*	4.483832	0.03421692
ENSG00000112031	*MTRF1L*	4.478106	0.034331747
ENSG00000183559	*C10orf120*	4.433806	0.035233856
ENSG00000128482	*RNF112*	4.433124	0.035247936
ENSG00000174485	*DENND4A*	4.431694	0.035277478
ENSG00000174827	*PDZK1*	4.407648	0.035778138
ENSG00000088682	*COQ9*	4.313376	0.037813848
ENSG00000116667	*C1orf21*	4.240852	0.039462359
ENSG00000092421	*SEMA6A*	4.06231	0.043850475
ENSG00000121270	*ABCC11*	4.00205	0.045444959
ENSG00000166479	*TMX3*	3.987282	0.045844961
ENSG00000074696	*HACD3*	3.98035	0.046033997
ENSG00000085491	*SLC25A24*	3.938498	0.047192896

table S35. Positively selected genes of the great ape ancestral lineage. Identification of positively selected genes in the Hominidae ancestral lineage was used by the branch-site model in PAML4. *P* values were calculated by means of a χ^2 test.

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000104237	*RP1*	1365.451372	6.76E-299
ENSG00000137106	*GRHPR*	140.016026	2.64E-32
ENSG00000181744	*DIPK2A*	54.241744	1.77E-13
ENSG00000036530	*CYP46A1*	49.068068	2.47E-12
ENSG00000179271	*GADD45GIP1*	46.22338	1.06E-11
ENSG00000179388	*EGR3*	38.939902	4.37E-10
ENSG00000048540	*LMO3*	27.891554	1.28E-07
ENSG00000025770	*NCAPH2*	23.456698	1.28E-06
ENSG00000177728	*TMEM94*	20.592338	5.68E-06
ENSG00000137500	*CCDC90B*	17.90543	2.32E-05
ENSG00000075415	*SLC25A3*	16.973928	3.79E-05
ENSG00000182836	*PLCXD3*	16.031252	6.23E-05
ENSG00000187535	*IFT140*	15.164406	9.85E-05
ENSG00000105143	*SLC1A6*	12.455108	0.00041685
ENSG00000107902	*LHPP*	11.046678	0.000888463
ENSG00000157330	*C1orf158*	9.895464	0.001656867
ENSG00000058056	*USP13*	9.594442	0.001951672
ENSG00000164080	*RAD54L2*	8.653064	0.003265118
ENSG00000006747	*SCIN*	8.30123	0.003961824
ENSG00000198911	*SREBF2*	8.190026	0.004212132
ENSG00000171551	*ECEL1*	8.059504	0.004526557
ENSG00000125755	*SYMPK*	8.053118	0.004542541
ENSG00000097021	*ACOT7*	7.768796	0.005315647
ENSG00000096696	*DSP*	7.41257	0.006476974
ENSG00000137693	*YAP1*	6.402814	0.011393963
ENSG00000125744	*RTN2*	6.097962	0.013533788
ENSG00000166833	*NAV2*	6.007836	0.014242484
ENSG00000150756	*ATPSCKMT*	5.989346	0.014392538
ENSG00000064692	*SNCAIP*	5.845306	0.015618658
ENSG00000083097	*DOP1A*	5.815804	0.015882791
ENSG00000140279	*DUOX2*	5.607628	0.017882454
ENSG00000148143	*ZNF462*	5.420178	0.019905331
ENSG00000136279	*DBNL*	5.252712	0.021912593
ENSG00000165646	*SLC18A2*	5.197328	0.022621636
ENSG00000155158	*TTC39B*	4.791232	0.02860496
ENSG00000214595	*EML6*	4.558784	0.032750318
ENSG00000137504	*CREBZF*	4.526448	0.033374821
ENSG00000020577	*SAMD4A*	4.305688	0.037985128
ENSG00000145868	*FBXO38*	4.20396	0.040329697
ENSG00000154269	*ENPP3*	4.18356	0.040817875
ENSG00000166532	*RIMKLB*	3.970096	0.046315132
ENSG00000070915	*SLC12A3*	3.951398	0.046832446

table S36. KEGG pathway enrichment analysis of combined genes including positively selected genes and genes associated with lineage-specific accelerated regions in the great ape ancestral lineage. KEGG pathways with P values < 0.05 (Modified Fisher's Exact test) are given.

Category	Term	Count	P
KEGG_PATHWAY	hsa00534:Glycosaminoglycan biosynthesis - heparan sulfate / heparin	3	0.031987
KEGG_PATHWAY	hsa04740:Olfactory transduction	11	0.033139
KEGG_PATHWAY	hsa04392:Hippo signaling pathway - multiple species	3	0.045311

table S37. Positively selected genes of the Colobinae ancestral lineage. Identification of positively selected genes in the Colobinae ancestral lineage was used by the branch-site model in PAML4. P values were calculated by means of a χ^2 test.

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000269964	MEI4	98.581512	3.12E-23
ENSG00000162999	DUSP19	97.283538	6.01E-23
ENSG00000188493	C19orf54	72.684878	1.52E-17
ENSG00000215187	FAM166B	59.811832	1.04E-14
ENSG00000133398	MED10	57.399812	3.56E-14
ENSG00000133740	E2F5	50.895054	9.74E-13
ENSG00000007952	NOX1	49.21334	2.30E-12
ENSG00000198836	OPA1	34.717522	3.81E-09
ENSG00000107262	BAG1	32.455514	1.22E-08
ENSG00000008405	CRY1	29.17048	6.63E-08
ENSG00000120094	HOXB1	25.433198	4.58E-07
ENSG00000135457	TFCP2	24.350108	8.03E-07
ENSG00000145362	ANK2	22.669638	1.92E-06
ENSG00000108846	ABCC3	20.402486	6.27E-06
ENSG00000233436	BTBD18	17.441824	2.96E-05
ENSG00000183476	SH2D7	16.985252	3.77E-05
ENSG00000188816	HMX2	16.70217	4.37E-05
ENSG00000117054	ACADM	16.521876	4.81E-05
ENSG00000179361	ARID3B	16.081094	6.07E-05
ENSG00000159131	GART	15.052248	0.000104576
ENSG00000149476	TKFC	14.187868	0.000165434
ENSG00000103148	NPRL3	13.985484	0.000184227
ENSG00000105851	PIK3CG	11.857502	0.000574295
ENSG00000117425	PTCH2	11.455252	0.000712922
ENSG00000128191	DGCR8	11.433356	0.000721373
ENSG00000122965	RBM19	11.401142	0.00073399
ENSG00000173166	RAPH1	11.13651	0.00084645
ENSG00000170955	CAVIN3	11.131152	0.000848899
ENSG00000079156	OSBPL6	10.936678	0.000942791
ENSG00000119711	ALDH6A1	10.55033	0.001161673
ENSG00000174948	GPR149	10.491368	0.001199335
ENSG00000137507	LRRC32	9.601798	0.001943869
ENSG00000196091	MYBPC1	8.749114	0.003097525
ENSG00000126790	L3HYPDH	7.977488	0.004736262
ENSG00000181722	ZBTB20	7.866442	0.005036065
ENSG00000140807	NKD1	7.696486	0.005532844
ENSG00000126749	EMG1	7.393782	0.006544973
ENSG00000117643	MAN1C1	7.37091	0.006628738
ENSG00000112378	PERP	7.28951	0.006935839
ENSG00000141425	RPRD1A	7.274298	0.006994822
ENSG00000175104	TRAF6	7.23014	0.007168962
ENSG00000151379	MSGN1	6.539992	0.010547582
ENSG00000186417	GLDN	6.4298	0.011222121
ENSG00000283654	LMLN2	6.388102	0.011488781

ENSG00000100364	KIAA0930	6.337098	0.011823817
ENSG00000074211	PPP2R2C	5.754212	0.016449172
ENSG00000166510	CCDC68	5.5596	0.018379613
ENSG00000197548	ATG7	5.443194	0.019644715
ENSG00000198171	DDRGK1	5.393746	0.020209042
ENSG00000127334	DYRK2	5.26192	0.02179696
ENSG00000114062	UBE3A	5.020984	0.025041934
ENSG00000151322	NPAS3	4.892078	0.026980194
ENSG00000186919	ZACN	4.870608	0.027317874
ENSG00000204574	ABCF1	4.394166	0.036062093
ENSG00000148516	ZEB1	4.32845	0.037480362
ENSG00000204186	ZDBF2	4.287796	0.038386896
ENSG00000170271	FAXDC2	4.233564	0.039632136
ENSG00000102580	DNAJC3	4.140248	0.041875034
ENSG00000162852	CNST	4.098982	0.042909042
ENSG00000077984	CST7	4.02863	0.044734254

table S38. Positively selected genes and genes associated with lineage-specific accelerated regions in the ancestral lineage of the Colobinae with a high expression level in the stomach, colon, pancreas and small intestine. These genes were identified as highly expressed genes in digestive organs by the DAVID Tissue_Expression database (UP_TISSUE) (https://david.ncifcrf.gov/home.jsp).

Highly Expressed Tissue	Gene Name	Gene Category
Pancreas	MYBPC1, PERP, PIK3CG	positively selected genes
Small intestine	ZBTB20	positively selected genes
Colon	ABCC3, ZBTB20, ACADM, GLDN, NOX1	positively selected genes
Stomach	BCHE	genes associated with lineage-specific accelerated regions
Pancreas	CASZ1, RNASE4, SLC4A4, ZPBP	genes associated with lineage-specific accelerated regions
Small intestine	AHI1, THEMIS, EIF4E	genes associated with lineage-specific accelerated regions
Colon	SHF, TMEM267, CYP4A11, MTMR8, PEX26, SOD1, PAFAH1B1	genes associated with lineage-specific accelerated regions

The following supplementary table is presented as a separate Excel file, as the table is very large.

table S39 (as a separate Excel file). Identification of lineage-specific accelerated regions along the Colobinae ancestral lineage. Lineage-specific accelerated regions significance was determined by *FDR*≤0.05. Nearest genes were obtained by reference to the distance to the lineage-specific accelerated regions in 500kbp.

table S40. Positively selected genes in the Strepsirrhini ancestor lineage.
Identification of positively selected genes in the Strepsirrhini ancestral lineage was used
by the branch-site model in PAML4. *P* values were calculated by means of a χ^2 test.

Ensembl Gene ID	Gene Name	2ΔLNL	P
ENSG00000114735	*HEMK1*	141.525484	1.23E-32
ENSG00000186765	*FSCN2*	85.299248	2.56E-20
ENSG00000100284	*TOM1*	82.569186	1.02E-19
ENSG00000198399	*ITSN2*	49.740946	1.75E-12
ENSG00000090686	*USP48*	32.245068	1.36E-08
ENSG00000165209	*STRBP*	29.629326	5.23E-08
ENSG00000157827	*FMNL2*	29.403862	5.88E-08
ENSG00000139865	*TTC6*	28.990602	7.27E-08
ENSG00000166886	*NAB2*	26.219928	3.05E-07
ENSG00000168101	*NUDT16L1*	26.110156	3.22E-07
ENSG00000197217	*ENTPD4*	25.511488	4.40E-07
ENSG00000123213	*NLN*	24.750996	6.52E-07
ENSG00000116984	*MTR*	23.905984	1.01E-06
ENSG00000177302	*TOP3A*	23.256428	1.42E-06
ENSG00000213347	*MXD3*	19.99643	7.76E-06
ENSG00000170871	*KIAA0232*	19.589466	9.60E-06
ENSG00000171823	*FBXL14*	18.506026	1.69E-05
ENSG00000083635	*NUFIP1*	17.981726	2.23E-05
ENSG00000018408	*WWTR1*	17.627836	2.69E-05
ENSG00000164031	*DNAJB14*	17.343768	3.12E-05
ENSG00000133895	*MEN1*	15.685886	7.48E-05
ENSG00000106829	*TLE4*	14.923122	0.000111982
ENSG00000072041	*SLC6A15*	14.71282	0.000125192
ENSG00000127311	*HELB*	13.738278	0.000210128
ENSG00000184345	*IQCF2*	13.651138	0.000220108
ENSG00000143190	*POU2F1*	13.56612	0.000230305
ENSG00000143442	*POGZ*	13.484488	0.000240544
ENSG00000103710	*RASL12*	13.23324	0.000275028
ENSG00000136518	*ACTL6A*	13.13785	0.00028939
ENSG00000152359	*POC5*	13.12923	0.000290724
ENSG00000038219	*BOD1L1*	11.95479	0.000545071
ENSG00000164056	*SPRY1*	11.789776	0.000595569
ENSG00000157060	*SHCBP1L*	11.703302	0.000623893
ENSG00000089101	*CFAP61*	11.238458	0.000801197
ENSG00000148248	*SURF4*	10.809582	0.001009761
ENSG00000055813	*CCDC85A*	10.757528	0.00103856
ENSG00000155100	*OTUD6B*	10.638918	0.001107324
ENSG00000160714	*UBE2Q1*	10.517752	0.001182332
ENSG00000050730	*TNIP3*	10.318172	0.001317267
ENSG00000112282	*MED23*	10.00073	0.001564782
ENSG00000113441	*LNPEP*	9.954804	0.001604302
ENSG00000186009	*ATP4B*	9.69278	0.001849934
ENSG00000088888	*MAVS*	9.681742	0.001861081
ENSG00000188676	*IDO2*	9.587644	0.001958911

ENSG00000127083	*OMD*	9.547224	0.00200252
ENSG00000132321	*IQCA1*	9.513708	0.002039426
ENSG00000044459	*CNTLN*	9.432446	0.002131798
ENSG00000112996	*MRPS30*	9.13207	0.002511674
ENSG00000143970	*ASXL2*	8.953982	0.002768654
ENSG00000126870	*WDR60*	8.74762	0.003100064
ENSG00000169499	*PLEKHA2*	8.550788	0.003453727
ENSG00000112149	*CD83*	8.530206	0.003493002
ENSG00000072778	*ACADVL*	8.507412	0.003537028
ENSG00000198162	*MAN1A2*	8.482844	0.003585112
ENSG00000197563	*PIGN*	8.477702	0.00359526
ENSG00000105982	*RNF32*	8.463044	0.003624348
ENSG00000162621	*LRRC53*	8.370052	0.003814547
ENSG00000186827	*TNFRSF4*	8.299018	0.003966652
ENSG00000130584	*ZBTB46*	8.264992	0.004041684
ENSG00000085365	*SCAMP1*	8.163744	0.004273606
ENSG00000005893	*LAMP2*	7.965534	0.004767642
ENSG00000105856	*HBP1*	7.91357	0.004906535
ENSG00000158488	*CD1E*	7.664748	0.005631012
ENSG00000117500	*TMED5*	7.470494	0.006271834
ENSG00000184281	*TSSC4*	7.3731	0.00662067
ENSG00000173588	*CEP83*	7.363026	0.006657865
ENSG00000108349	*CASC3*	7.32772	0.006789913
ENSG00000112893	*MAN2A1*	7.246276	0.007104821
ENSG00000186834	*HEXIM1*	7.12267	0.007611527
ENSG00000130643	*CALY*	7.099018	0.00771262
ENSG00000107742	*SPOCK2*	7.08294	0.007782123
ENSG00000146530	*VWDE*	7.040056	0.007970654
ENSG00000134013	*LOXL2*	6.997842	0.008160804
ENSG00000237524	*TEX51*	6.884606	0.008694123
ENSG00000182185	*RAD51B*	6.711616	0.009578688
ENSG00000175806	*MSRA*	6.646118	0.00993721
ENSG00000171522	*PTGER4*	6.586174	0.010277381
ENSG00000103174	*NAGPA*	6.566768	0.010390047
ENSG00000172803	*SNX32*	6.55571	0.010454811
ENSG00000239779	*WBP1*	6.549898	0.010489017
ENSG00000127952	*STYXL1*	6.547634	0.010502373
ENSG00000141579	*ZNF750*	6.522076	0.010654356
ENSG00000104731	*KLHDC4*	6.46091	0.011027316
ENSG00000187097	*ENTPD5*	6.437372	0.011174384
ENSG00000198870	*STKLD1*	6.404706	0.011381827
ENSG00000162711	*NLRP3*	6.339114	0.011810386
ENSG00000137563	*GGH*	6.310662	0.012001402
ENSG00000143740	*SNAP47*	6.283776	0.01218482
ENSG00000154342	*WNT3A*	6.21876	0.012640361
ENSG00000115425	*PECR*	6.193688	0.012820673
ENSG00000127720	*METTL25*	6.054942	0.013867431
ENSG00000129636	*ITFG1*	6.02319	0.014119104
ENSG00000131389	*SLC6A6*	5.910126	0.015054095

ENSG00000177238	TRIM72	5.891954	0.015210214
ENSG00000140057	AK7	5.84708	0.01560292
ENSG00000214842	RAD51AP2	5.825848	0.015792353
ENSG00000227729	RD3L	5.767494	0.016325298
ENSG00000154734	ADAMTS1	5.683774	0.017122499
ENSG00000127419	TMEM175	5.657242	0.017383438
ENSG00000155833	CYLC2	5.621062	0.017745893
ENSG00000189079	ARID2	5.562164	0.018352715
ENSG00000197024	ZNF398	5.466428	0.019385207
ENSG00000137860	SLC28A2	5.407258	0.020053191
ENSG00000166503	HDGFL3	5.387556	0.020280857
ENSG00000147113	DIPK2B	5.309378	0.021210928
ENSG00000095321	CRAT	5.278032	0.021596145
ENSG00000136933	RABEPK	5.238302	0.022094829
ENSG00000107651	SEC23IP	5.221372	0.022310942
ENSG00000115548	KDM3A	5.150958	0.023233552
ENSG00000080572	PIH1D3	5.1201	0.023650254
ENSG00000135124	P2RX4	5.098456	0.02394715
ENSG00000011566	MAP4K3	5.09675	0.023970716
ENSG00000116981	NT5C1A	5.056098	0.024539429
ENSG00000226763	SRRM5	5.015044	0.025127991
ENSG00000147853	AK3	5.010198	0.025198425
ENSG00000115966	ATF2	4.985442	0.025561454
ENSG00000078900	TP73	4.91436	0.026634329
ENSG00000131148	EMC8	4.885066	0.027089999
ENSG00000240021	TEX35	4.876474	0.02722518
ENSG00000008283	CYB561	4.832564	0.027927069
ENSG00000006116	CACNG3	4.787818	0.028661713
ENSG00000275342	PRAG1	4.759668	0.029134167
ENSG00000219545	UMAD1	4.75676	0.029183433
ENSG00000113594	LIFR	4.740496	0.029460576
ENSG00000143278	F13B	4.674916	0.0306062
ENSG00000127863	TNFRSF19	4.633296	0.031357269
ENSG00000174173	TRMT10C	4.629904	0.031419322
ENSG00000106689	LHX2	4.613796	0.031715753
ENSG00000114757	PEX5L	4.60436	0.031890755
ENSG00000214954	LRRC69	4.591566	0.032129646
ENSG00000135374	ELF5	4.504958	0.033796726
ENSG00000170915	PAQR8	4.470852	0.034477795
ENSG00000189280	GJB5	4.391888	0.036110305
ENSG00000124574	ABCC10	4.363386	0.036719252
ENSG00000011600	TYROBP	4.360382	0.036784056
ENSG00000132530	XAF1	4.347404	0.037065404
ENSG00000136425	CIB2	4.262288	0.038967404
ENSG00000129925	TMEM8A	4.248942	0.039274791
ENSG00000105926	MPP6	4.226916	0.039787674
ENSG00000204099	NEU4	4.179098	0.040925477
ENSG00000147896	IFNK	4.174898	0.041027032
ENSG00000091127	PUS7	4.164328	0.041283787

ENSG00000176915	*ANKLE2*	4.129522	0.042141253
ENSG00000102683	*SGCG*	4.12693	0.042205852
ENSG00000162522	*KIAA1522*	4.073226	0.043567992
ENSG00000124243	*BCAS4*	4.050164	0.044167054
ENSG00000163638	*ADAMTS9*	4.035122	0.044562453
ENSG00000112297	*CRYBG1*	4.00633	0.04532972
ENSG00000118412	*CASP8AP2*	3.9742	0.046202396
ENSG00000119686	*FLVCR2*	3.961318	0.046557237
ENSG00000088970	*KIZ*	3.924582	0.047585019

table S41. Generation times and short-read data from each primate species according to previous studies. 'NA' indicates that the data were not obtained from prior studies and that in this study we used the generation times of their mostly related species based on our species tree as succedanea.

Family	Genus	Species name	Generation time (year)	Illumina short-read data (NCBI sra/EBI database)
Hominidae	*Homo*	*Homo sapiens*	29 (*215*)	ERR004098-ERR004577; ERR004580; ERR004582; ERR004584-ERR004621;ERR004623-ERR004640; ERR004643-ERR004811
Hominidae	*Pan*	*Pan troglodytes*	25 (*215*)	SRX8439961
Hominidae	*Pan*	*Pan paniscus*	25 (*216*)	SRX242681, SRX242682
Hominidae	*Gorilla*	*Gorilla gorilla*	19 (*215*)	SRX243505
Hominidae	*Pongo*	*Pongo abelii*	20 (*217*)	SRX243481, SRX243482
Hominidae	*Pongo*	*Pongo pygmaeus*	20 (*218*)	In this study
Hylobatidae	*Nomascus*	*Nomascus siki*	10 (*24, 219*)	In this study
Hylobatidae	*Symphalangus*	*Symphalangus syndactylus*	10 (*24, 219*)	In this study
Hylobatidae	*Hoolock*	*Hoolock leuconedys*	10 (*24, 219*)	In this study
Hylobatidae	*Hylobates*	*Hylobates pileatus*	10 (*24, 219*)	In this study
Cercopithecidae	*Macaca*	*Macaca mulatta*	6 (*220*)	SRX6606578
Cercopithecidae	*Macaca*	*Macaca assamensis*	6 (*220, 221*)	In this study
Cercopithecidae	*Macaca*	*Macaca nemestrina*	NA	SRX3106039
Cercopithecidae	*Macaca*	*Macaca silenus*	6 (*220, 221*)	In this study
Cercopithecidae	*Papio*	*Papio hamadryas*	11 (*222*)	In this study
Cercopithecidae	*Papio*	*Papio anubis*	11 (*222*)	SRX4503011, SRX4503012, SRX4503015
Cercopithecidae	*Lophocebus*	*Lophocebus aterrimus*	10 (*223*)	In this study
Cercopithecidae	*Theropithecus*	*Theropithecus gelada*	10 (*224*)	SRX2521053, SRX2521054
Cercopithecidae	*Mandrillus*	*Mandrillus sphinx*	10 (*225*)	In this study
Cercopithecidae	*Mandrillus*	*Mandrillus leucophaeus*	10 (*226*)	SRX794703—SRX794706
Cercopithecidae	*Cercocebus*	*Cercocebus atys*	NA	SRX790327—SRX790334
Cercopithecidae	*Cercopithecus*	*Cercopithecus albogularis*	NA	In this study
Cercopithecidae	*Cercopithecus*	*Cercopithecus mona*	7 (*227*)	SRX7990421
Cercopithecidae	*Chlorocebus*	*Chlorocebus aethiops*	8.5 (*228*)	In this study
Cercopithecidae	*Chlorocebus*	*Chlorocebus sabaeus*	8.5 (*228*)	NA
Cercopithecidae	*Erythrocebus*	*Erythrocebus patas*	NA	In this study
Cercopithecidae	*Trachypithecus*	*Trachypithecus crepusculus*	11 (*229*)	In this study
Cercopithecidae	*Pygathrix*	*Pygathrix nigripes*	NA	In this study
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus strykeri*	10 (*26*)	In this study
Cercopithecidae	*Rhinopithecus*	*Rhinopithecus roxellana*	10 (*26*)	SRX828307
Cercopithecidae	*Piliocolobus*	*Piliocolobus tephrosceles*	5 (*230, 231*)	NA
Cercopithecidae	*Colobus*	*Colobus guereza*	5 (*230, 231*)	In this study
Cercopithecidae	*Colobus*	*Colobus angolensis*	5 (*230, 231*)	SRX792356—SRX792361
Callitrichidae	*Callithrix*	*Callithrix jacchus*	6 (*232*)	In this study
Callitrichidae	*Saguinus*	*Saguinus midas*	6 (*215*)	In this study
Aotidae	*Aotus*	*Aotus nancymaae*	8 (*215*)	SRX4473017

Cebidae	*Sapajus*	*Sapajus apella*	15 (*215*)	In this study
Cebidae	*Cebus*	*Cebus albifrons*	15 (*215*)	In this study
Atelidae	*Ateles*	*Ateles geoffroyi*	15 (*215*)	In this study
Pitheciidae	*Pithecia*	*Pithecia pithecia*	9 (*215*)	In this study
Tarsidae	*Cephalopachus*	*Cephalopachus bancanus*	7 (*62*)	In this study
Cheirogaleidae	*Microcebus*	*Microcebus murinus*	3 (*233*)	SRX670567
Lemuridae	*Prolemur*	*Prolemur simus*	3.5 (*234*)	SRX4465422
Lemuridae	*Lemur*	*Lemur catta*	3 (*235*)	In this study
Daubentoniidae	*Daubentonia*	*Daubentonia madagascariensis*	3.5 (*236*)	In this study
Lorisidae	*Loris*	*Loris tardigradus*	NA	In this study
Lorisidae	*Nycticebus*	*Nycticebus pygmaeus*	2 (*237*)	In this study
Lorisidae	*Nycticebus*	*Nycticebus bengalensis*	2 (*237*)	In this study
Galagidae	*Galago*	*Galago moholi*	NA	In this study
Galagidae	*Otolemur*	*Otolemur garnettii*	NA	SRX4465433

table S42. Summary information of conservation genetics for primate species. In total, 46 non-human primate species were analyzed in this analysis. The species *Chlorocebus sabaeus* and *Piliocolobus tephrosceles* were excluded because the high coverage short-read data are missing in the NCBI SRA database (https://www.ncbi.nlm.nih.gov/). *Callithrix jacchus* was removed from this analysis because the genome was from an inbred individual. The IUCN Red List species status was obtained from the website (https://www.iucnredlist.org/).

Species name	Pi$_{median}$	IUCN Red List status	Ne (x 10^4 in 20,000 years ago)
Pan troglodytes	0.00057	EN	0.241449682
Pan paniscus	0.00057	EN	0.648794739
Gorilla gorilla	0.00083	CR	1.559425526
Pongo abelii	0.00182	CR	0.717874511
Pongo pygmaeus	0.0011	CR	0.503474723
Nomascus siki	0.0017	CR	1.359667189
Symphalangus syndactylus	0.00106	EN	5.274632286
Hoolock leuconedys	0.00058	VU	0.450462482
Hylobates pileatus	0.0005	EN	0.310245424
Macaca mulatta	0.00208	LC	1.868278975
Macaca assamensis	0.00216	NT	3.208116315
Macaca nemestrina	0.00194	VU	3.009434564
Macaca silenus	0.00045	EN	2.07890067
Papio hamadryas	0.00074	LC	2.564670298
Papio anubis	0.00012	LC	1.263479978
Lophocebus aterrimus	0.0016	VU	5.499862681
Theropithecus gelada	0.00016	LC	0.420439062
Mandrillus sphinx	0.00245	VU	3.028259057
Mandrillus leucophaeus	0.00064	EN	1.811095063
Cercocebus atys	0.00073	VU	6.567033139
Cercopithecus albogularis	0.00125	LC	1.363557634
Cercopithecus mona	0.00189	NT	2.888423855
Chlorocebus aethiops	0.00111	LC	1.435820132
Erythrocebus patas	0.00066	NT	0.602580514
Trachypithecus crepusculus	0.00093	EN	0.499043814
Pygathrix nigripes	0.00134	CR	2.380222025

Rhinopithecus strykeri	0.00007	CR	0.202554603
Rhinopithecus roxellana	0.00033	EN	0.675578962
Colobus guereza	0.00073	LC	1.594377736
Colobus angolensis	0.00082	VU	0.607909735
Saguinus midas	0.00125	LC	1.289943937
Aotus nancymaae	0.00068	VU	3.061421186
Sapajus apella	0.00074	LC	0.905683755
Cebus albifrons	0.00092	LC	0.366957549
Ateles geoffroyi	0.00141	EN	1.163167684
Pithecia pithecia	0.00105	LC	0.931025333
Cephalopachus bancanus	0.00228	VU	1.766781487
Daubentonia madagascariensis	0.00048	EN	2.551512884
Microcebus murinus	0.00198	LC	5.354912039
Prolemur simus	0.0002	CR	3.599630087
Lemur catta	0.00172	EN	7.150895406
Loris tardigradus	0.00179	EN	16.62657852
Nycticebus pygmaeus	0.00007	EN	1.608037778
Nycticebus bengalensis	0.00008	EN	0.949696297
Galago moholi	0.00069	LC	6.055302869
Otolemur garnettii	0.00125	LC	19.97383041

References and Notes

1. A. Estrada, P. A. Garber, A. B. Rylands, C. Roos, E. Fernandez-Duque, A. Di Fiore, K. A. 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. 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). doi:10.1126/sciadv.1600946 Medline

2. C. Roos, K. M. Helgen, R. P. Miguez, N. M. L. Thant, N. Lwin, A. K. Lin, A. Lin, K. M. Yi, P. Soe, Z. M. Hein, M. N. N. Myint, T. Ahmed, D. Chetry, M. Urh, E. G. Veatch, N. Duncan, P. Kamminga, M. A. H. Chua, L. Yao, C. Matauschek, D. Meyer, Z. J. Liu, M. Li, T. Nadler, P. F. Fan, L. K. Quyet, M. Hofreiter, D. Zinner, F. Momberg, Mitogenomic phylogeny of the Asian colobine genus *Trachypithecus* with special focus on *Trachypithecus phayrei* (Blyth, 1847) and description of a new species. *Zool. Res.* **41**, 656–669 (2020). doi:10.24272/j.issn.2095-8137.2020.254 Medline

3. 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**, 3487–3498.e10 (2017). [doi:10.1016/j.cub.2017.09.047](doi:10.1016/j.cub.2017.09.047) Medline

4. P. F. Fan, K. He, X. Chen, A. Ortiz, B. Zhang, C. Zhao, Y. Q. Li, H. B. Zhang, C. Kimock, W. Z. Wang, C. Groves, S. T. Turvey, C. Roos, K. M. Helgen, X. L. Jiang, Description of a new species of *Hoolock* gibbon (Primates: Hylobatidae) based on integrative taxonomy. *Am. J. Primatol.* **79**, e22631 (2017). [doi:10.1002/ajp.22631](doi:10.1002/ajp.22631) Medline

5. C. Li, C. Zhao, P. F. Fan, White-cheeked macaque (*Macaca leucogenys*): A new macaque species from Medog, southeastern Tibet. *Am. J. Primatol.* **77**, 753–766 (2015). [doi:10.1002/ajp.22394](doi:10.1002/ajp.22394) Medline

6. J. Rogers, R. A. Gibbs, Comparative primate genomics: Emerging patterns of genome content and dynamics. *Nat. Rev. Genet.* **15**, 347–359 (2014). [doi:10.1038/nrg3707](doi:10.1038/nrg3707) Medline

7. B. Rockx, T. Kuiken, S. Herfst, T. Bestebroer, M. M. Lamers, B. B. Oude Munnink, D. de Meulder, G. van Amerongen, J. van den Brand, N. M. A. Okba, D. Schipper, P. van Run, L. Leijten, R. Sikkema, E. Verschoor, B. Verstrepen, W. Bogers, J. Langermans, C. Drosten, M. Fentener van Vlissingen, R. Fouchier, R. de Swart, M. Koopmans, B. L. Haagmans, Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. *Science* **368**, 1012–1015 (2020). doi:10.1126/science.abb7314 Medline

8. A. Chandrashekar, J. Liu, A. J. Martinot, K. McMahan, N. B. Mercado, L. Peter, L. H. Tostanoski, J. Yu, Z. Maliga, M. Nekorchuk, K. Busman-Sahay, M. Terry, L. M. Wrijil, S. Ducat, D. R. Martinez, C. Atyeo, S. Fischinger, J. S. Burke, M. D. Slein, L. Pessaint, A. Van Ry, J. Greenhouse, T. Taylor, K. Blade, A. Cook, B. Finneyfrock, R. Brown, E. Teow, J. Velasco, R. Zahn, F. Wegmann, P. Abbink, E. A. Bondzie, G. Dagotto, M. S. Gebre, X. He, C. Jacob-Dolan, N. Kordana, Z. Li, M. A. Lifton, S. H. Mahrokhian, L. F. Maxfield, R. Nityanandam, J. P. Nkolola, A. G. Schmidt, A. D. Miller, R. S. Baric, G. Alter, P. K. Sorger, J. D. Estes, H. Andersen, M. G. Lewis, D. H. Barouch, SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science* **369**, 812–817 (2020). doi:10.1126/science.abc4776 Medline

9. Q. Gao, L. Bao, H. Mao, L. Wang, K. Xu, M. Yang, Y. Li, L. Zhu, N. Wang, Z. Lv, H. Gao, X. Ge, B. Kan, Y. Hu, J. Liu, F. Cai, D. Jiang, Y. Yin, C. Qin, J. Li, X. Gong, X. Lou, W. Shi, D. Wu, H. Zhang, L. Zhu, W. Deng, Y. Li, J. Lu, C. Li, X. Wang, W. Yin, Y. Zhang, C. Qin, Development of an inactivated vaccine candidate for SARS-CoV-2. *Science* **369**, 77–81 (2020). doi:10.1126/science.abc1932 Medline

10. J. Yu, L. H. Tostanoski, L. Peter, N. B. Mercado, K. McMahan, S. H. Mahrokhian, J. P. Nkolola, J. Liu, Z. Li, A. Chandrashekar, D. R. Martinez, C. Loos, C. Atyeo, S. Fischinger, J. S. Burke, M. D. Slein, Y. Chen, A. Zuiani, F. J. N. Lelis, M. Travers, S. Habibi, L. Pessaint, A. Van Ry, K. Blade, R. Brown, A. Cook, B. Finneyfrock, A. Dodson, E. Teow, J. Velasco, R. Zahn, F. Wegmann, E. A. Bondzie, G. Dagotto, M. S. Gebre, X. He, C. Jacob-Dolan, M. Kirilova, N. Kordana, Z. Lin, L. F. Maxfield, F. Nampanya, R. Nityanandam, J. D. Ventura, H. Wan, Y. Cai, B. Chen, A. G. Schmidt, D. R. Wesemann, R. S. Baric, G. Alter, H. Andersen, M. G. Lewis, D. H. Barouch, DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science* **369**, 806–811 (2020). doi:10.1126/science.abc6284 Medline

11. V. J. Munster, F. Feldmann, B. N. Williamson, N. van Doremalen, L. Pérez-Pérez, J. Schulz, K. Meade-White, A. Okumura, J. Callison, B. Brumbaugh, V. A. Avanzato, R. Rosenke, P. W. Hanley, G. Saturday, D. Scott, E. R. Fischer, E. de Wit, Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature* **585**, 268–272 (2020). [doi:10.1038/s41586-020-2324-7](doi:10.1038/s41586-020-2324-7) Medline

12. N. B. Mercado, R. Zahn, F. Wegmann, C. Loos, A. Chandrashekar, J. Yu, J. Liu, L. Peter, K. McMahan, L. H. Tostanoski, X. He, D. R. Martinez, L. Rutten, R. Bos, D. van Manen, J. Vellinga, J. Custers, J. P. Langedijk, T. Kwaks, M. J. G. Bakkers, D. Zuijdgeest, S. K. Rosendahl Huber, C. Atyeo, S. Fischinger, J. S. Burke, J. Feldman, B. M. Hauser, T. M. Caradonna, E. A. Bondzie, G. Dagotto, M. S. Gebre, E. Hoffman, C. Jacob-Dolan, M. Kirilova, Z. Li, Z. Lin, S. H. Mahrokhian, L. F. Maxfield, F. Nampanya, R. Nityanandam, J. P. Nkolola, S. Patel, J. D. Ventura, K. Verrington, H. Wan, L. Pessaint, A. Van Ry, K. Blade, A. Strasbaugh, M. Cabus, R. Brown, A. Cook, S. Zouantchangadou, E. Teow, H. Andersen, M. G. Lewis, Y. Cai, B. Chen, A. G. Schmidt, R. K. Reeves, R. S. Baric, D. A. Lauffenburger, G. Alter, P. Stoffels, M. Mammen, J. Van Hoof, H. Schuitemaker, D. H. Barouch, Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583–588 (2020). [doi:10.1038/s41586-020-2607-z](doi:10.1038/s41586-020-2607-z) Medline

13. K. S. Corbett, B. Flynn, K. E. Foulds, J. R. Francica, S. Boyoglu-Barnum, A. P. Werner, B. Flach, S. O'Connell, K. W. Bock, M. Minai, B. M. Nagata, H. Andersen, D. R. Martinez, A. T. Noe, N. Douek, M. M. Donaldson, N. N. Nji, G. S. Alvarado, D. K. Edwards, D. R. Flebbe, E. Lamb, N. A. Doria-Rose, B. C. Lin, M. K. Louder, S. O'Dell, S. D. Schmidt, E. Phung, L. A. Chang, C. Yap, J. M. Todd, L. Pessaint, A. Van Ry, S. Browne, J. Greenhouse, T. Putman-Taylor, A. Strasbaugh, T. A. Campbell, A. Cook, A. Dodson, K. Steingrebe, W. Shi, Y. Zhang, O. M. Abiona, L. Wang, A. Pegu, E. S. Yang, K. Leung, T. Zhou, I. T. Teng, A. Widge, I. Gordon, L. Novik, R. A. Gillespie, R. J. Loomis, J. I. Moliva, G. Stewart-Jones, S. Himansu, W. P. Kong, M. C. Nason, K. M. Morabito, T. J. Ruckwardt, J. E. Ledgerwood, M. R. Gaudinski, P. D. Kwong, J. R. Mascola, A. Carfi, M. G. Lewis, R. S. Baric, A. McDermott, I. N. Moore, N. J. Sullivan, M. Roederer, R. A. Seder, B. S. Graham, Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in nonhuman primates. *N. Engl. J. Med.* **383**, 1544–1555 (2020). doi:10.1056/NEJMoa2024671 Medline

14. N. van Doremalen, T. Lambe, A. Spencer, S. Belij-Rammerstorfer, J. N. Purushotham, J. R. Port, V. A. Avanzato, T. Bushmaker, A. Flaxman, M. Ulaszewska, F. Feldmann, E. R. Allen, H. Sharpe, J. Schulz, M. Holbrook, A. Okumura, K. Meade-White, L. Pérez-Pérez, N. J. Edwards, D. Wright, C. Bissett, C. Gilbride, B. N. Williamson, R. Rosenke, D. Long, A. Ishwarbhai, R. Kailath, L. Rose, S. Morris, C. Powers, J. Lovaglio, P. W. Hanley, D. Scott, G. Saturday, E. de Wit, S. C. Gilbert, V. J. Munster, ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **586**, 578–582 (2020). [doi:10.1038/s41586-020-2608-y](doi:10.1038/s41586-020-2608-y) Medline

15. B. N. Williamson, F. Feldmann, B. Schwarz, K. Meade-White, D. P. Porter, J. Schulz, N. van Doremalen, I. Leighton, C. K. Yinda, L. Pérez-Pérez, A.

Okumura, J. Lovaglio, P. W. Hanley, G. Saturday, C. M. Bosio, S. Anzick, K. Barbian, T. Cihlar, C. Martens, D. P. Scott, V. J. Munster, E. de Wit, Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature* **585**, 273–276 (2020). [doi:10.1038/s41586-020-2423-5](doi:10.1038/s41586-020-2423-5) Medline

16. T. Z. Song, H. Y. Zheng, J. B. Han, L. Jin, X. Yang, F. L. Liu, R. H. Luo, R. R. Tian, H. R. Cai, X. L. Feng, C. Liu, M. H. Li, Y. T. Zheng, Delayed severe cytokine storm and immune cell infiltration in SARS-CoV-2-infected aged Chinese rhesus macaques. *Zool. Res.* **41**, 503–516 (2020). doi:10.24272/j.issn.2095-8137.2020.202 Medline

17. W. Enard, S. Pääbo, Comparative primate genomics. *Annu. Rev. Genomics Hum. Genet.* **5**, 351–378 (2004). doi:10.1146/annurev.genom.5.061903.180040 Medline

18. 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**, eaar6343 (2018). doi:10.1126/science.aar6343 Medline

19. Chimpanzee Sequencing and Analysis Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005). doi:10.1038/nature04072 Medline

20. 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. Messina, Y. Shen, H. X. Song, T. Wylie, L. Zhang, E. Birney, K. Han, M. K. Konkel, J. Lee, A. F. 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. Zwieg; Rhesus Macaque Genome Sequencing and Analysis Consortium, Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222–234 (2007). doi:10.1126/science.1139247 Medline

21. 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). doi:10.1038/nature10842 Medline

22. Marmoset Genome Sequencing and Analysis Consortium, The common marmoset genome provides insight into primate biology and evolution. *Nat. Genet.* **46**, 850–857 (2014). [doi:10.1038/ng.3042](doi:10.1038/ng.3042) Medline

23. 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. 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). doi:10.1038/nature09687 Medline

24. L. Carbone, R. A. Harris, S. Gnerre, K. R. Veeramah, B. Lorente-Galdos, J. Huddleston, T. J. Meyer, J. Herrero, C. Roos, B. Aken, F. Anaclerio, N. Archidiacono, C. Baker, D. Barrell, M. A. Batzer, K. Beal, A. Blancher, C. L. Bohrson, M. Brameier, M. S. Campbell, O. Capozzi, C. Casola, G. Chiatante, A. Cree, A. Damert, P. J. de Jong, L. Dumas, M. Fernandez-Callejo, P. Flicek, N. V. Fuchs, I. Gut, M. Gut, M. W. Hahn, J. Hernandez-Rodriguez, L. W. Hillier, R. Hubley, B. Ianc, Z. Izsvák, N. G. Jablonski, L. M. Johnstone, A. Karimpour-Fard, M. K. Konkel, D. Kostka, N. H. Lazar, S. L. Lee, L. R. Lewis, Y. Liu, D. P. Locke, S. Mallick, F. L. Mendez, M. Muffato, L. V. Nazareth, K. A. Nevonen, M. O'Bleness, C. Ochis, D. T. Odom, K. S. Pollard, J. Quilez, D. Reich, M. Rocchi, G. G. Schumann, S. Searle, J. M. Sikela, G. Skollar, A. Smit, K. Sonmez, B. ten Hallers, E. Terhune, G. W. Thomas, B. Ullmer, M. Ventura, J. A. Walker, J. D. Wall, L. Walter, M. C. Ward, S. J. Wheelan, C. W. Whelan, S. White, L. J. Wilhelm, A. E. Woerner, M. Yandell, B. Zhu, M. F. Hammer, T. Marques-Bonet, E. E. Eichler, L. Fulton, C. Fronick, D. M. Muzny, W. C. Warren, K. C. Worley, J. Rogers, R. K. Wilson, R. A. Gibbs, Gibbon genome and the fast karyotype evolution of small apes. *Nature* **513**, 195–201 (2014). doi:10.1038/nature13679 Medline

25. 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). [doi:10.1038/ng.3615](doi:10.1038/ng.3615) Medline

26. X. Zhou, B. Wang, Q. Pan, J. Zhang, S. Kumar, X. Sun, Z. Liu, H. Pan, Y. Lin, G. Liu, W. Zhan, M. Li, B. Ren, X. Ma, H. Ruan, C. Cheng, D. Wang, F. Shi, Y.

Hui, Y. Tao, C. Zhang, P. Zhu, Z. Xiang, W. Jiang, J. Chang, H. Wang, Z. Cao, Z. Jiang, B. Li, G. Yang, C. Roos, P. A. Garber, M. W. Bruford, R. Li, M. Li, Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. *Nat. Genet.* **46**, 1303–1310 (2014). doi:10.1038/ng.3137 Medline

27. A. O. Ayoola, B. L. Zhang, R. P. Meisel, L. M. Nneji, Y. Shao, O. B. Morenikeji, A. C. Adeola, S. I. Ng'ang'a, B. G. Ogunjemite, A. O. Okeyoyin, C. Roos, D. D. Wu, Population genomics reveals incipient speciation, introgression, and adaptation in the African mona monkey (*Cercopithecus mona*). *Mol. Biol. Evol.* **38**, 876–890 (2021). doi:10.1093/molbev/msaa248 Medline

28. D. M. Bickhart, B. D. Rosen, S. Koren, B. L. Sayre, A. R. Hastie, S. Chan, J. Lee, E. T. Lam, I. Liachko, S. T. Sullivan, J. N. Burton, H. J. Huson, J. C. Nystrom, C. M. Kelley, J. L. Hutchison, Y. Zhou, J. Sun, A. Crisà, F. A. Ponce de León, J. C. Schwartz, J. A. Hammond, G. C. Waldbieser, S. G. Schroeder, G. E. Liu, M. J. Dunham, J. Shendure, T. S. Sonstegard, A. M. Phillippy, C. P. Van Tassell, T. P. Smith, Single-molecule sequencing and chromatin conformation capture enable *de novo* reference assembly of the domestic goat genome. *Nat. Genet.* **49**, 643–650 (2017). doi:10.1038/ng.3802 Medline

29. B.-L. Zhang, W. Chen, Z. Wang, W. Pang, M.-T. Luo, S. Wang, Y. Shao, W.-Q. He, Y. Deng, L. Zhou, J. Chen, M. Yang, Y. Wu, L. Wang, H. Fernandez, S. Molloy, H. Meunier, F. Wanert, L. Kuderna, T. Marques-Bonet, C. Roos, X. Qi, M. Li, Z.-J. Liu, M. H. Schierup, D. N. Cooper, J. Liu, Y.-T. Zheng, G. Zhang, D.-D. Wu, Comparative genomics reveals the hybrid origin of a macaque group. *Sci. Adv.* **9**, eadd3580 (2023). doi:10.1126/sciadv.abn7153

30. H. Wu, Z. Wang, Y. Zhang, L. Frantz, C. Roos, D. M. Irwin, C. Zhang, X. Liu, D. Wu, S. Huang, T. Gu, J. Liu, L. Yu, Hybrid origin of a primate, the gray snub-nosed monkey. *Science* **380**, eabl4997 (2023) doi:10.1126/science.abl4997.

31. X.-G. Qi, J. Wu, L. Zhao, L. Wang, X. Guang, P. A. Garber, C. Opie, Y. Yuan, R. Diao, G. Li, K. Wang, R. Pan, W. Ji, H. Sun, Z.-P. Huang, C. Xu, A. B. Witarto, R. Jia, C. Zhang, C. Deng, Q. Qiu, G. Zhang, C. C. Grueter, D. Wu, B. Li, Adaptations to a cold climate promoted social evolution in Asian colobine primates. *Science* **380**, eabl8621 (2023). doi:10.1126/science.abl8621

32. M.-L. Li, S. Wang, P. Xu, H.-Y. Tian, M. Bai, Y.-P. Zhang, Y. Shao, Z.-J. Xiong, X.-G. Qi, D. N. Cooper, G. Zhang, H. H. Zhu, D.-D. Wu, Functional genomics analysis reveals the evolutionary adaptation and demographic history of pygmy lorises. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2123030119 (2022). doi:10.1073/pnas.2123030119 Medline

33. M. S. Ye, J. Y. Zhang, D. D. Yu, M. Xu, L. Xu, L. B. Lv, Q. Y. Zhu, Y. Fan, Y. G. Yao, Comprehensive annotation of the Chinese tree shrew genome by

large-scale RNA sequencing and long-read isoform sequencing. *Zool. Res.* **42**, 692–709 (2021). [doi:10.24272/j.issn.2095-8137.2021.272](#) [Medline](#)

34. A. M. Kozlov, A. J. Aberer, A. Stamatakis, ExaML version 3: A tool for phylogenomic analyses on supercomputers. *Bioinformatics* **31**, 2577–2579 (2015). [doi:10.1093/bioinformatics/btv184](#) [Medline](#)

35. P. Perelman, W. E. Johnson, C. Roos, H. N. Seuánez, J. E. Horvath, M. A. Moreira, B. Kessing, J. Pontius, M. Roelke, Y. Rumpler, M. P. Schneider, A. Silva, S. J. O'Brien, J. Pecon-Slattery, A molecular phylogeny of living primates. *PLOS Genet.* **7**, e1001342 (2011). [doi:10.1371/journal.pgen.1001342](#) [Medline](#)

36. C. M. Shi, Z. Yang, Coalescent-based analyses of genomic sequence data provide a robust resolution of phylogenetic relationships among major groups of gibbons. *Mol. Biol. Evol.* **35**, 159–179 (2018). [doi:10.1093/molbev/msx277](#) [Medline](#)

37. A. Hobolth, O. F. Christensen, T. Mailund, M. H. Schierup, Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. *PLOS Genet.* **3**, e7 (2007). [doi:10.1371/journal.pgen.0030007](#) [Medline](#)

38. I. Rivas-González, M. Rousselle, F. Li, L. Zhou, J. Y. Dutheil, K. Munch, Y. Shao, D. Wu, M. H. Schierup, G. Zhang, Pervasive incomplete lineage sorting illuminates speciation and selection processes in primates. *Science* **380**, eabn4409 (2022). [doi:10.1126/science.abn4409](#)

39. 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). [doi:10.1371/journal.pbio.3000954](#) [Medline](#)

40. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). [doi:10.1093/molbev/msm088](#) [Medline](#)

41. S. Álvarez-Carretero, A. U. Tamuri, M. Battini, F. F. Nascimento, E. Carlisle, R. J. Asher, Z. Yang, P. C. J. Donoghue, M. Dos Reis, A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* **602**, 263–267 (2022). [doi:10.1038/s41586-021-04341-1](#) [Medline](#)

42. C. Liu, J. Gao, X. Cui, Z. Li, L. Chen, Y. Yuan, Y. Zhang, L. Mei, L. Zhao, D. Cai, M. Hu, B. Zhou, Z. Li, T. Qin, H. Si, G. Li, Z. Lin, Y. Xu, C. Zhu, Y. Yin, C. Zhang, W. Xu, Q. Li, K. Wang, M. T. P. Gilbert, R. Heller, W. Wang, J. Huang, Q. Qiu, A towering genome: Experimentally validated adaptations to high blood pressure and extreme stature in the giraffe. *Sci. Adv.* **7**, eabe9459 (2021). [doi:10.1126/sciadv.abe9459](#) [Medline](#)

43. E. E. Eichler, D. Sankoff, Structural dynamics of eukaryotic chromosome evolution. *Science* **301**, 793–797 (2003). doi:10.1126/science.1086132 Medline

44. Y. Yin, H. Fan, B. Zhou, Y. Hu, G. Fan, J. Wang, F. Zhou, W. Nie, C. Zhang, L. Liu, Z. Zhong, W. Zhu, G. Liu, Z. Lin, C. Liu, J. Zhou, G. Huang, Z. Li, J. Yu, Y. Zhang, Y. Yang, B. Zhuo, B. Zhang, J. Chang, H. Qian, Y. Peng, X. Chen, L. Chen, Z. Li, Q. Zhou, W. Wang, F. Wei, Molecular mechanisms and topological consequences of drastic chromosomal rearrangements of muntjac deer. *Nat. Commun.* **12**, 6858 (2021). doi:10.1038/s41467-021-27091-0 Medline

45. R. Stanyon, M. Rocchi, O. Capozzi, R. Roberto, D. Misceo, M. Ventura, M. F. Cardone, F. Bigoni, N. Archidiacono, Primate chromosome evolution: Ancestral karyotypes, marker order and neocentromeres. *Chromosome Res.* **16**, 17–39 (2008). doi:10.1007/s10577-007-1209-z Medline

46. T. Marques-Bonet, J. M. Kidd, M. Ventura, T. A. Graves, Z. Cheng, L. W. Hillier, Z. Jiang, C. Baker, R. Malfavon-Borja, L. A. Fulton, C. Alkan, G. Aksay, S. Girirajan, P. Siswara, L. Chen, M. F. Cardone, A. Navarro, E. R. Mardis, R. K. Wilson, E. E. Eichler, A burst of segmental duplications in the genome of the African great ape ancestor. *Nature* **457**, 877–881 (2009). doi:10.1038/nature07744 Medline

47. P. D. Stenson, M. Mort, E. V. Ball, M. Chapman, K. Evans, L. Azevedo, M. Hayden, S. Heywood, D. S. Millar, A. D. Phillips, D. N. Cooper, The Human Gene Mutation Database (HGMD®): Optimizing its use in a clinical diagnostic or research setting. *Hum. Genet.* **139**, 1197–1207 (2020). doi:10.1007/s00439-020-02199-3 Medline

48. L. Chen, Q. Qiu, Y. Jiang, K. Wang, Z. Lin, Z. Li, F. Bibi, Y. Yang, J. Wang, W. Nie, W. Su, G. Liu, Q. Li, W. Fu, X. Pan, C. Liu, J. Yang, C. Zhang, Y. Yin, Y. Wang, Y. Zhao, C. Zhang, Z. Wang, Y. Qin, W. Liu, B. Wang, Y. Ren, R. Zhang, Y. Zeng, R. R. da Fonseca, B. Wei, R. Li, W. Wan, R. Zhao, W. Zhu, Y. Wang, S. Duan, Y. Gao, Y. E. Zhang, C. Chen, C. Hvilsom, C. W. Epps, L. G. Chemnick, Y. Dong, S. Mirarab, H. R. Siegismund, O. A. Ryder, M. T. P. Gilbert, H. A. Lewin, G. Zhang, R. Heller, W. Wang, Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. *Science* **364**, eaav6202 (2019). doi:10.1126/science.aav6202 Medline

49. J. D. Smith, J. W. Bickham, T. R. Gregory, Patterns of genome size diversity in bats (order Chiroptera). *Genome* **56**, 457–472 (2013). doi:10.1139/gen-2013-0046 Medline

50. S. Shen, L. Lin, J. J. Cai, P. Jiang, E. J. Kenkel, M. R. Stroik, S. Sato, B. L. Davidson, Y. Xing, Widespread establishment and regulatory impact of Alu

exons in human genes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2837–2842 (2011). doi:10.1073/pnas.1012834108 Medline

51. G. E. Liu, C. Alkan, L. Jiang, S. Zhao, E. E. Eichler, Comparative analysis of Alu repeats in primate genomes. *Genome Res.* **19**, 876–885 (2009). doi:10.1101/gr.083972.108 Medline

52. T. Hayakawa, Y. Satta, P. Gagneux, A. Varki, N. Takahata, Alu-mediated inactivation of the human CMP- *N*-acetylneuraminic acid hydroxylase gene. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11399–11404 (2001). doi:10.1073/pnas.191268198 Medline

53. P. Kuehnen, M. Mischke, S. Wiegand, C. Sers, B. Horsthemke, S. Lau, T. Keil, Y. A. Lee, A. Grueters, H. Krude, An Alu element-associated hypermethylation variant of the *POMC* gene is associated with childhood obesity. *PLOS Genet.* **8**, e1002543 (2012). doi:10.1371/journal.pgen.1002543 Medline

54. J. Jurka, Evolutionary impact of human Alu repetitive elements. *Curr. Opin. Genet. Dev.* **14**, 603–608 (2004). doi:10.1016/j.gde.2004.08.008 Medline

55. G. Zhang, C. Li, Q. Li, B. Li, D. M. Larkin, C. Lee, J. F. Storz, A. Antunes, M. J. Greenwold, R. W. Meredith, A. Ödeen, J. Cui, Q. Zhou, L. Xu, H. Pan, Z. Wang, L. Jin, P. Zhang, H. Hu, W. Yang, J. Hu, J. Xiao, Z. Yang, Y. Liu, Q. Xie, H. Yu, J. Lian, P. Wen, F. Zhang, H. Li, Y. Zeng, Z. Xiong, S. Liu, L. Zhou, Z. Huang, N. An, J. Wang, Q. Zheng, Y. Xiong, G. Wang, B. Wang, J. Wang, Y. Fan, R. R. da Fonseca, A. Alfaro-Núñez, M. Schubert, L. Orlando, T. Mourier, J. T. Howard, G. Ganapathy, A. Pfenning, O. Whitney, M. V. Rivas, E. Hara, J. Smith, M. Farré, J. Narayan, G. Slavov, M. N. Romanov, R. Borges, J. P. Machado, I. Khan, M. S. Springer, J. Gatesy, F. G. Hoffmann, J. C. Opazo, O. Håstad, R. H. Sawyer, H. Kim, K. W. Kim, H. J. Kim, S. Cho, N. Li, Y. Huang, M. W. Bruford, X. Zhan, A. Dixon, M. F. Bertelsen, E. Derryberry, W. Warren, R. K. Wilson, S. Li, D. A. Ray, R. E. Green, S. J. O'Brien, D. Griffin, W. E. Johnson, D. Haussler, O. A. Ryder, E. Willerslev, G. R. Graves, P. Alström, J. Fjeldså, D. P. Mindell, S. V. Edwards, E. L. Braun, C. Rahbek, D. W. Burt, P. Houde, Y. Zhang, H. Yang, J. Wang, E. D. Jarvis, M. T. Gilbert, J. Wang; Avian Genome Consortium, Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014). doi:10.1126/science.1251385 Medline

56. P. Moorjani, C. E. Amorim, P. F. Arndt, M. Przeworski, Variation in the molecular clock of primates. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10607–10612 (2016). doi:10.1073/pnas.1600374113 Medline

57. E. Fontanillas, J. J. Welch, J. A. Thomas, L. Bromham, The influence of body size and net diversification rate on molecular evolution during the radiation of animal phyla. *BMC Evol. Biol.* **7**, 95 (2007). doi:10.1186/1471-2148-7-95 Medline

58. A. Wong, Covariance between testes size and substitution rates in primates. *Mol. Biol. Evol.* **31**, 1432–1436 (2014). doi:10.1093/molbev/msu091 Medline

59. W. H. Li, M. Tanimura, The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326**, 93–96 (1987). [doi:10.1038/326093a0](doi:10.1038/326093a0) Medline

60. M. E. Steiper, N. M. Young, Primate molecular divergence dates. *Mol. Phylogenet. Evol.* **41**, 384–394 (2006). doi:10.1016/j.ympev.2006.05.021 Medline

61. S. H. Kim, N. Elango, C. Warden, E. Vigoda, S. V. Yi, Heterogeneous genomic molecular clocks in primates. *PLOS Genet.* **2**, e163 (2006). doi:10.1371/journal.pgen.0020163 Medline

62. J. Schmitz, A. Noll, C. A. Raabe, G. Churakov, R. Voss, M. Kiefmann, T. Rozhdestvensky, J. Brosius, R. Baertsch, H. Clawson, C. Roos, A. Zimin, P. Minx, M. J. Montague, R. K. Wilson, W. C. Warren, Genome sequence of the basal haplorrhine primate *Tarsius syrichta* reveals unusual insertions. *Nat. Commun.* **7**, 12997 (2016). doi:10.1038/ncomms12997 Medline

63. L. Fang, W. Cai, S. Liu, O. Canela-Xandri, Y. Gao, J. Jiang, K. Rawlik, B. Li, S. G. Schroeder, B. D. Rosen, C. J. Li, T. S. Sonstegard, L. J. Alexander, C. P. Van Tassell, P. M. VanRaden, J. B. Cole, Y. Yu, S. Zhang, A. Tenesa, L. Ma, G. E. Liu, Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. *Genome Res.* **30**, 790–801 (2020). [doi:10.1101/gr.250704.119](doi:10.1101/gr.250704.119) Medline

64. B. Y. Liao, J. Zhang, Low rates of expression profile divergence in highly expressed genes and tissue-specific genes during mammalian evolution. *Mol. Biol. Evol.* **23**, 1119–1128 (2006). doi:10.1093/molbev/msj119 Medline

65. G. J. Wyckoff, W. Wang, C. I. Wu, Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304–309 (2000). [doi:10.1038/35002070](doi:10.1038/35002070) Medline

66. T. Boehm, Evolution of vertebrate immunity. *Curr. Biol.* **22**, R722–R732 (2012). [doi:10.1016/j.cub.2012.07.003](doi:10.1016/j.cub.2012.07.003) Medline

67. H. Y. Wang, H. C. Chien, N. Osada, K. Hashimoto, S. Sugano, T. Gojobori, C. K. Chou, S. F. Tsai, C. I. Wu, C. K. Shen, Rate of evolution in brain-expressed genes in humans and other primates. *PLOS Biol.* **5**, e13 (2007). doi:10.1371/journal.pbio.0050013 Medline

68. Materials and methods are available as supplementary materials.

69. J. Tohyama, M. Nakashima, S. Nabatame, C. Gaik-Siew, R. Miyata, Z. Rener-Primec, M. Kato, N. Matsumoto, H. Saitsu, *SPTAN1* encephalopathy: Distinct phenotypes and genotypes. *J. Hum. Genet.* **60**, 167–173 (2015). [doi:10.1038/jhg.2015.5](doi:10.1038/jhg.2015.5) Medline

70. P. Mansfield, J. N. Constantino, D. Baldridge, *MYT1L*: A systematic review of genetic variation encompassing schizophrenia and autism. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **183**, 227–233 (2020). doi:10.1002/ajmg.b.32781 Medline

71. M. Maekawa, T. Ohnishi, K. Hashimoto, A. Watanabe, Y. Iwayama, H. Ohba, E. Hattori, K. Yamada, T. Yoshikawa, Analysis of strain-dependent prepulse inhibition points to a role for Shmt1 (*SHMT1*) in mice and in schizophrenia. *J. Neurochem.* **115**, 1374–1385 (2010). doi:10.1111/j.1471-4159.2010.07039.x Medline

72. X. Bi, L. Zhou, J.-J. Zhang, S. Feng, M. Hu, D. N. Cooper, J. Lin, J. Li, D.-D. Wu, G. Zhang, Lineage-specific accelerated sequences underlying primate evolution. *Sci. Adv.* 10.1126/sciadv.adc9507 (2023).

73. J. K. Rilling, T. R. Insel, Differential expansion of neural projection systems in primate brain evolution. *Neuroreport* **10**, 1453–1459 (1999). doi:10.1097/00001756-199905140-00012 Medline

74. K. Isler, E. Christopher Kirk, J. M. Miller, G. A. Albrecht, B. R. Gelvin, R. D. Martin, Endocranial volumes of primate species: Scaling analyses using a comprehensive and reliable data set. *J. Hum. Evol.* **55**, 967–978 (2008). doi:10.1016/j.jhevol.2008.08.004 Medline

75. C. Plachez, L. J. Richards, Mechanisms of axon guidance in the developing nervous system. *Curr. Top. Dev. Biol.* **69**, 267–346 (2005). doi:10.1016/S0070-2153(05)69010-2 Medline

76. M. A. Robichaux, C. W. Cowan, Signaling mechanisms of axon guidance and early synaptogenesis. *Curr. Top. Behav. Neurosci.* **16**, 19–48 (2014). doi:10.1007/978-3-662-45758-0_255 Medline

77. J. Falk, A. Bechara, R. Fiore, H. Nawabi, H. Zhou, C. Hoyo-Becerra, M. Bozon, G. Rougon, M. Grumet, A. W. Püschel, J. R. Sanes, V. Castellani, Dual functional activity of semaphorin 3B is required for positioning the anterior commissure. *Neuron* **48**, 63–75 (2005). doi:10.1016/j.neuron.2005.10.024 Medline

78. M. A. Wolman, Y. Liu, H. Tawarayama, W. Shoji, M. C. Halloran, Repulsion and attraction of axons by semaphorin3D are mediated by different neuropilins in vivo. *J. Neurosci.* **24**, 8428–8435 (2004). doi:10.1523/JNEUROSCI.2349-04.2004 Medline

79. C. Kudo, I. Ajioka, Y. Hirata, K. Nakajima, Expression profiles of *EphA3* at both the RNA and protein level in the developing mammalian forebrain. *J. Comp. Neurol.* **487**, 255–269 (2005). doi:10.1002/cne.20551 Medline

80. M. V. Tejada-Simon, Modulation of actin dynamics by *Rac1* to target cognitive function. *J. Neurochem.* **133**, 767–779 (2015). doi:10.1111/jnc.13100 Medline

81. S. L. Eastwood, P. J. Harrison, Decreased mRNA expression of netrin-G1 and netrin-G2 in the temporal lobe in schizophrenia and bipolar disorder. *Neuropsychopharmacology* **33**, 933–945 (2008). doi:10.1038/sj.npp.1301457 Medline

82. D. Pan, Hippo signaling in organ size control. *Genes Dev.* **21**, 886–897 (2007). doi:10.1101/gad.1536007 Medline

83. S. H. Patel, F. D. Camargo, D. Yimlamai, Hippo signaling in the liver regulates organ size, cell fate, and carcinogenesis. *Gastroenterology* **152**, 533–545 (2017). doi:10.1053/j.gastro.2016.10.047 Medline

84. R. H. Gokhale, A. W. Shingleton, Size control: The developmental physiology of body and organ size regulation. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 335–356 (2015). doi:10.1002/wdev.181 Medline

85. E. C. Kirk, Comparative morphology of the eye in primates. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **281**, 1095–1103 (2004). doi:10.1002/ar.a.20115 Medline

86. A. C. Wiik, C. Wade, T. Biagi, E. O. Ropstad, E. Bjerkås, K. Lindblad-Toh, F. Lingaas, A deletion in nephronophthisis 4 (*NPHP4*) is associated with recessive cone-rod dystrophy in standard wire-haired dachshund. *Genome Res.* **18**, 1415–1421 (2008). doi:10.1101/gr.074302.107 Medline

87. P. Liskova, L. Dudakova, C. J. Evans, K. E. Rojas Lopez, N. Pontikos, D. Athanasiou, H. Jama, J. Sach, P. Skalicka, V. Stranecky, S. Kmoch, C. Thaung, M. Filipec, M. E. Cheetham, A. E. Davidson, S. J. Tuft, A. J. Hardcastle, Ectopic *GRHL2* expression due to non-coding mutations promotes cell state transition and causes posterior polymorphous corneal dystrophy 4. *Am. J. Hum. Genet.* **102**, 447–459 (2018). doi:10.1016/j.ajhg.2018.02.002 Medline

88. Y. Toda, T. Hayakawa, A. Itoigawa, Y. Kurihara, T. Nakagita, M. Hayashi, R. Ashino, A. D. Melin, Y. Ishimaru, S. Kawamura, H. Imai, T. Misaka, Evolution of the primate glutamate taste sensor from a nucleotide sensor. *Curr. Biol.* **31**, 4641–4649.e5 (2021). doi:10.1016/j.cub.2021.08.002 Medline

89. J. M. Kamilar, B. J. Bradley, Interspecific variation in primate coat colour supports Gloger's rule. *J. Biogeogr.* **38**, 2270–2277 (2011). doi:10.1111/j.1365-2699.2011.02587.x

90. S. Hu, Y. Chen, B. Zhao, N. Yang, S. Chen, J. Shen, G. Bao, X. Wu, *KIT* is involved in melanocyte proliferation, apoptosis and melanogenesis in the Rex Rabbit. *PeerJ* **8**, e9402 (2020). doi:10.7717/peerj.9402 Medline

91. M. C. Garrido, B. C. Bastian, *KIT* as a therapeutic target in melanoma. *J. Invest. Dermatol.* **130**, 20–27 (2010). doi:10.1038/jid.2009.334 Medline

92. J. M. Grichnik, Kit and melanocyte migration. *J. Invest. Dermatol.* **126**, 945–947 (2006). [doi:10.1038/sj.jid.5700164](doi:10.1038/sj.jid.5700164) Medline

93. Y. Mizutani, N. Hayashi, M. Kawashima, G. Imokawa, A single UVB exposure increases the expression of functional *KIT* in human melanocytes by up-regulating *MITF* expression through the phosphorylation of p38/CREB. *Arch. Dermatol. Res.* **302**, 283–294 (2010). [doi:10.1007/s00403-009-1007-x](doi:10.1007/s00403-009-1007-x) Medline

94. R. Kitamura, K. Tsukamoto, K. Harada, A. Shimizu, S. Shimada, T. Kobayashi, G. Imokawa, Mechanisms underlying the dysfunction of melanocytes in vitiligo epidermis: Role of SCF/KIT protein interactions and the downstream effector, MITF-M. *J. Pathol.* **202**, 463–475 (2004). doi:10.1002/path.1538 Medline

95. B. Wen, Y. Chen, H. Li, J. Wang, J. Shen, A. Ma, J. Qu, K. Bismuth, J. Debbache, H. Arnheiter, L. Hou, Allele-specific genetic interactions between Mitf and Kit affect melanocyte development. *Pigment Cell Melanoma Res.* **23**, 441–447 (2010). [doi:10.1111/j.1755-148X.2010.00699.x](doi:10.1111/j.1755-148X.2010.00699.x) Medline

96. D. L. C. van den Berg, R. Azzarelli, K. Oishi, B. Martynoga, N. Urbán, D. H. W. Dekkers, J. A. Demmers, F. Guillemot, *NIPBL* interacts with *ZFP609* and the integrator complex to regulate cortical neuron migration. *Neuron* **93**, 348–361 (2017). doi:10.1016/j.neuron.2016.11.047 Medline

97. G. H. Mochida, C. A. Walsh, Molecular genetics of human microcephaly. *Curr. Opin. Neurol.* **14**, 151–156 (2001). [doi:10.1097/00019052-200104000-00003](doi:10.1097/00019052-200104000-00003) Medline

98. S. H. Montgomery, I. Capellini, C. Venditti, R. A. Barton, N. I. Mundy, Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Mol. Biol. Evol.* **28**, 625–638 (2011). doi:10.1093/molbev/msq237 Medline

99. L. Shi, M. Li, Q. Lin, X. Qi, B. Su, Functional divergence of the brain-size regulating gene *MCPH1* during primate evolution and the origin of humans. *BMC Biol.* **11**, 62 (2013). [doi:10.1186/1741-7007-11-62](doi:10.1186/1741-7007-11-62) Medline

100. L. Shi, B. Su, A transgenic monkey model for the study of human brain evolution. *Zool. Res.* **40**, 236–238 (2019). doi:10.24272/j.issn.2095-8137.2019.031 Medline

101. J. Rogers, P. Kochunov, K. Zilles, W. Shelledy, J. Lancaster, P. Thompson, R. Duggirala, J. Blangero, P. T. Fox, D. C. Glahn, On the genetic architecture of cortical folding and brain volume in primates. *Neuroimage* **53**, 1103–1108 (2010). doi:10.1016/j.neuroimage.2010.02.020 Medline

102. S. V. Puram, A. Riccio, S. Koirala, Y. Ikeuchi, A. H. Kim, G. Corfas, A. Bonni, A TRPC5-regulated calcium signaling pathway controls dendrite patterning in

the mammalian brain. *Genes Dev.* **25**, 2659–2673 (2011). doi:10.1101/gad.174060.111 Medline

103. A. Yamada, E. Inoue, M. Deguchi-Tawarada, C. Matsui, A. Togawa, T. Nakatani, Y. Ono, Y. Takai, *Necl-2/CADM1* interacts with *ErbB4* and regulates its activity in GABAergic neurons. *Mol. Cell. Neurosci.* **56**, 234–243 (2013). doi:10.1016/j.mcn.2013.06.003 Medline

104. R. Kusano, K. Fujita, Y. Shinoda, Y. Nagaura, H. Kiyonari, T. Abe, T. Watanabe, Y. Matsui, M. Fukaya, H. Sakagami, T. Sato, J. I. Funahashi, M. Ohnishi, S. Tamura, T. Kobayashi, Targeted disruption of the mouse protein phosphatase *ppm1l* gene leads to structural abnormalities in the brain. *FEBS Lett.* **590**, 3606–3615 (2016). doi:10.1002/1873-3468.12429 Medline

105. M. Talarowska, J. Szemraj, M. Kowalczyk, P. Gałecki, Serum *KIBRA* mRNA and protein expression and cognitive functions in depression. *Med. Sci. Monit.* **22**, 152–160 (2016). doi:10.12659/MSM.895200 Medline

106. A. K. Pandey, L. Lu, X. Wang, R. Homayouni, R. W. Williams, Functionally enigmatic genes: A case study of the brain ignorome. *PLOS ONE* **9**, e88889 (2014). doi:10.1371/journal.pone.0088889 Medline

107. A. Graziano, G. Foffani, E. B. Knudsen, J. Shumsky, K. A. Moxon, Passive exercise of the hind limbs after complete thoracic transection of the spinal cord promotes cortical reorganization. *PLOS ONE* **8**, e54350 (2013). doi:10.1371/journal.pone.0054350 Medline

108. H. Li, J. C. Radford, M. J. Ragusa, K. L. Shea, S. R. McKercher, J. D. Zaremba, W. Soussou, Z. Nie, Y. J. Kang, N. Nakanishi, S. Okamoto, A. J. Roberts, J. J. Schwarz, S. A. Lipton, Transcription factor *MEF2C* influences neural stem/progenitor cell differentiation and maturation *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9397–9402 (2008). doi:10.1073/pnas.0802876105 Medline

109. Y. Chang, O. Klezovitch, R. S. Walikonis, V. Vasioukhin, J. J. LoTurco, Discs large 5 is required for polarization of citron kinase in mitotic neural precursors. *Cell Cycle* **9**, 1990–1997 (2010). doi:10.4161/cc.9.10.11730 Medline

110. M. R. Sarkisian, Discs large 5: a new regulator of Citron kinase localization in developing neocortex: comment on: Chang Y, et al. Cell Cycle 2010; 9:1990-7. *Cell Cycle* **9**, 1876 (2010). doi:10.4161/cc.9.10.11923 Medline

111. D. A. Berg, L. Belnoue, H. Song, A. Simon, Neurotransmitter-mediated control of neurogenesis in the adult vertebrate brain. *Development* **140**, 2548–2561 (2013). doi:10.1242/dev.088005 Medline

112. P. Levitt, J. A. Harvey, E. Friedman, K. Simansky, E. H. Murphy, New evidence for neurotransmitter influences on brain development. *Trends Neurosci.* **20**, 269–274 (1997). doi:10.1016/S0166-2236(96)01028-4 Medline

113. L. Wang, X. You, S. Lotinun, L. Zhang, N. Wu, W. Zou, Mechanical sensing protein PIEZO1 regulates bone homeostasis via osteoblast-osteoclast crosstalk. *Nat. Commun.* **11**, 282 (2020). doi:10.1038/s41467-019-14146-6 Medline

114. M. Linder, M. Hecking, E. Glitzner, K. Zwerina, M. Holcmann, L. Bakiri, M. G. Ruocco, J. Tuckermann, G. Schett, E. F. Wagner, M. Sibilia, *EGFR* controls bone development by negatively regulating mTOR-signaling during osteoblast differentiation. *Cell Death Differ.* **25**, 1094–1106 (2018). doi:10.1038/s41418-017-0054-7 Medline

115. F. Xiao, C. Wang, C. Wang, Y. Gao, X. Zhang, X. Chen, *BMPER* enhances bone formation by promoting the osteogenesis-angiogenesis coupling process in mesenchymal stem cells. *Cell. Physiol. Biochem.* **45**, 1927–1939 (2018). doi:10.1159/000487969 Medline

116. S. Zanotti, E. Canalis, *Notch1* and *Notch2* expression in osteoblast precursors regulates femoral microarchitecture. *Bone* **62**, 22–28 (2014). doi:10.1016/j.bone.2014.01.023 Medline

117. J. M. Kim, C. Lin, Z. Stavre, M. B. Greenblatt, J. H. Shim, Osteoblast-osteoclast communication and bone homeostasis. *Cells* **9**, 2073 (2020). doi:10.3390/cells9092073 Medline

118. S. Zanotti, E. Canalis, Notch signaling and the skeleton. *Endocr. Rev.* **37**, 223–253 (2016). doi:10.1210/er.2016-1002 Medline

119. M. Schmidt, Locomotion and postural behaviour. *Adv. Sci. Res.* **5**, 23–39 (2011). doi:10.5194/asr-5-23-2010

120. Y. He, X. Luo, B. Zhou, T. Hu, X. Meng, P. A. Audano, Z. N. Kronenberg, E. E. Eichler, J. Jin, Y. Guo, Y. Yang, X. Qi, B. Su, Long-read assembly of the Chinese rhesus macaque genome and identification of ape-specific structural variants. *Nat. Commun.* **10**, 4233 (2019). doi:10.1038/s41467-019-12174-w Medline

121. S. A. Williams, G. A. Russo, Evolution of the hominoid vertebral column: The long and the short of it. *Evol. Anthropol.* **24**, 15–32 (2015). doi:10.1002/evan.21437 Medline

122. K. Semba, K. Araki, Z. Li, K. Matsumoto, M. Suzuki, N. Nakagata, K. Takagi, M. Takeya, K. Yoshinobu, M. Araki, K. Imai, K. Abe, K. Yamamura, A novel murine gene, Sickle tail, linked to the Danforth's short tail locus, is required for normal development of the intervertebral disc. *Genetics* **172**, 445–456 (2006). doi:10.1534/genetics.105.048934 Medline

123. N. Al Dhaheri, N. Wu, S. Zhao, Z. Wu, R. D. Blank, J. Zhang, C. Raggio, M. Halanski, J. Shen, K. Noonan, G. Qiu, B. Nemeth, S. Sund, S. L. Dunwoodie, G. Chapman, I. Glurich, R. D. Steiner, E. Wohler, R. Martin, N. L. Sobreira,

P. F. Giampietro, *KIAA1217*: A novel candidate gene associated with isolated and syndromic vertebral malformations. *Am. J. Med. Genet. A.* **182**, 1664–1672 (2020). doi:10.1002/ajmg.a.61607 Medline

124. J. R. Usherwood, J. E. Bertram, Understanding brachiation: Insight from a collisional perspective. *J. Exp. Biol.* **206**, 1631–1642 (2003). doi:10.1242/jeb.00306 Medline

125. J. R. Usherwood, S. G. Larson, J. E. Bertram, Mechanisms of force and power production in unsteady ricochetal brachiation. *Am. J. Phys. Anthropol.* **120**, 364–372 (2003). doi:10.1002/ajpa.10133 Medline

126. S. M. Cheyne, "Gibbon locomotion research in the field: Problems, possibilities, and benefits for conservation," in *Primate Locomotion: Linking Field and Laboratory Research*, K. D'Août, E. E. Vereecke, Eds. (Springer, NY, 2011), pp. 201–213.

127. C. Thiel, K. Kessler, A. Giessl, A. Dimmler, S. A. Shalev, S. von der Haar, M. Zenker, D. Zahnleiter, H. Stöss, E. Beinder, R. Abou Jamra, A. B. Ekici, N. Schröder-Kress, T. Aigner, T. Kirchner, A. Reis, J. H. Brandstätter, A. Rauch, *NEK1* mutations cause short-rib polydactyly syndrome type majewski. *Am. J. Hum. Genet.* **88**, 106–114 (2011). doi:10.1016/j.ajhg.2010.12.004 Medline

128. J. El Hokayem, C. Huber, A. Couvé, J. Aziza, G. Baujat, R. Bouvier, D. P. Cavalcanti, F. A. Collins, M.-P. Cordier, A.-L. Delezoide, M. Gonzales, D. Johnson, M. Le Merrer, A. Levy-Mozziconacci, P. Loget, D. Martin-Coignard, J. Martinovic, G. R. Mortier, M.-J. Perez, J. Roume, G. Scarano, A. Munnich, V. Cormier-Daire, *NEK1* and *DYNC2H1* are both involved in short rib polydactyly Majewski type but not in Beemer Langer cases. *J. Med. Genet.* **49**, 227–233 (2012). doi:10.1136/jmedgenet-2011-100717 Medline

129. J. M. Vazquez, V. J. Lynch, Pervasive duplication of tumor suppressors in Afrotherians during the evolution of large bodies and reduced cancer risk. *eLife* **10**, e65041 (2021). doi:10.7554/eLife.65041 Medline

130. J. G. M. Thewissen, L. N. Cooper, J. C. George, S. Bajpai, From land to water: The origin of whales, dolphins, and porpoises. *Evolution (N. Y.)* **2**, 272–288 (2009). doi:10.1007/s12052-009-0135-2

131. W. L. Jungers, "Body size and scaling of limb proportions in primates," in *Size and Scaling in Primate Biology*, W. L. Jungers, Ed. (Springer, 1985), pp. 345–381.

132. A. M. Rudolf, Q. Wu, L. Li, J. Wang, Y. Huang, J. Togo, C. Liechti, M. Li, C. Niu, Y. Nie, F. Wei, J. R. Speakman, A single nucleotide mutation in the dual-oxidase 2 (*DUOX2*) gene causes some of the panda's unique metabolic phenotypes. *Natl. Sci. Rev.* **9**, nwab125 (2021). doi:10.1093/nsr/nwab125 Medline

133. K. R. Johnson, C. C. Marden, P. Ward-Bailey, L. H. Gagnon, R. T. Bronson, L. R. Donahue, Congenital hypothyroidism, dwarfism, and hearing impairment caused by a missense mutation in the mouse dual oxidase 2 gene, Duox2. *Mol. Endocrinol.* **21**, 1593–1602 (2007). doi:10.1210/me.2007-0085 Medline

134. J. G. Fleagle, *Primate Adaptation and Evolution* (Academic, 2013).

135. K. Milton, Physiological ecology of howlers (*Alouatta*): Energetic and digestive considerations and comparison with the Colobinae. *Int. J. Primatol.* **19**, 513–548 (1998). doi:10.1023/A:1020364523213

136. I. Matsuda, C. A. Chapman, M. Clauss, Colobine forestomach anatomy and diet. *J. Morphol.* **280**, 1608–1616 (2019). doi:10.1002/jmor.21052 Medline

137. M. C. Janiak, Digestive enzymes of human and nonhuman primates. *Evol. Anthropol.* **25**, 253–266 (2016). doi:10.1002/evan.21498 Medline

138. J. J. Kim, R. Miura, Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences. *Eur. J. Biochem.* **271**, 483–493 (2004). doi:10.1046/j.1432-1033.2003.03948.x Medline

139. C. Matziouridou, S. D. C. Rocha, O. A. Haabeth, K. Rudi, H. Carlsen, A. Kielland, iNOS- and NOX1-dependent ROS production maintains bacterial homeostasis in the ileum of mice. *Mucosal Immunol.* **11**, 774–784 (2018). doi:10.1038/mi.2017.106 Medline

140. C.-J. Li, R. W. Li, R. L. Baldwin Vi, Assembly and analysis of changes in transcriptomes of dairy cattle rumen epithelia during lactation and dry periods. *Agric. Sci.* **9**, 619–638 (2018).

141. M. C. Janiak, A. S. Burrell, J. D. Orkin, T. R. Disotell, Duplication and parallel evolution of the pancreatic ribonuclease gene (*RNASE1*) in folivorous non-colobine primates, the howler monkeys (*Alouatta spp.*). *Sci. Rep.* **9**, 20366 (2019). doi:10.1038/s41598-019-56941-7 Medline

142. P. Pontarotti, *Evolutionary Biology: Mechanisms and Trends* (Springer, 2012).

143. N. J. Dominy, P. W. Lucas, Ecological importance of trichromatic vision to primates. *Nature* **410**, 363–366 (2001). doi:10.1038/35066567 Medline

144. N. G. Caine, N. I. Mundy, Demonstration of a foraging advantage for trichromatic marmosets (*Callithrix geoffroyi*) dependent on food colour. *Proc. Biol. Sci.* **267**, 439–444 (2000). doi:10.1098/rspb.2000.1019 Medline

145. A. C. Smith, H. M. Buchanan-Smith, A. K. Surridge, D. Osorio, N. I. Mundy, The effect of colour vision status on the detection and selection of fruits by tamarins (*Saguinus spp.*). *J. Exp. Biol.* **206**, 3159–3165 (2003). doi:10.1242/jeb.00536 Medline

146. S. Heritage, Modeling olfactory bulb evolution through primate phylogeny. *PLOS ONE* **9**, e113904 (2014). doi:10.1371/journal.pone.0113904 Medline

147. A. Matsui, Y. Go, Y. Niimura, Degeneration of olfactory receptor gene repertories in primates: No direct link to full trichromatic vision. *Mol. Biol. Evol.* **27**, 1192–1200 (2010). doi:10.1093/molbev/msq003 Medline

148. T. D. Smith, K. P. Bhatnagar, Microsmatic primates: Reconsidering how and when size matters. *Anat. Rec. B New Anat.* **279**, 24–31 (2004). doi:10.1002/ar.b.20026 Medline

149. A. Berghard, A. C. Hägglund, S. Bohm, L. Carlsson, *Lhx2*-dependent specification of olfactory sensory neurons is required for successful integration of olfactory, vomeronasal, and GnRH neurons. *FASEB J.* **26**, 3464–3472 (2012). doi:10.1096/fj.12-206193 Medline

150. J. Hirota, P. Mombaerts, The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8751–8755 (2004). doi:10.1073/pnas.0400940101 Medline

151. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011). doi:10.1038/nature10231 Medline

152. A. D. Barnosky, P. L. Koch, R. S. Feranec, S. L. Wing, A. B. Shabel, Assessing the causes of late Pleistocene extinctions on the continents. *Science* **306**, 70–75 (2004). doi:10.1126/science.1101476 Medline

153. X. Luo, Y. Liu, D. Dang, T. Hu, Y. Hou, X. Meng, F. Zhang, T. Li, C. Wang, M. Li, H. Wu, Q. Shen, Y. Hu, X. Zeng, X. He, L. Yan, S. Zhang, C. Li, B. Su 3rd, 3D Genome of macaque fetal brain reveals evolutionary innovations during primate corticogenesis. *Cell* **184**, 723–740.e21 (2021). doi:10.1016/j.cell.2021.01.001 Medline

154. 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). doi:10.1038/s41586-021-03535-x Medline

155. G. Dumas, S. Malesys, T. Bourgeron, Systematic detection of brain protein-coding genes under positive selection during primate evolution and their roles in cognition. *Genome Res.* **31**, 484–496 (2021). doi:10.1101/gr.262113.120 Medline

156. J. K. Rilling, Human and nonhuman primate brains: Are they allometrically scaled versions of the same design? *Evol. Anthropol.* **15**, 65–77 (2006). doi:10.1002/evan.20095

157. H. Stephan, H. Frahm, G. Baron, New and revised data on volumes of brain structures in insectivores and primates. *Folia Primatol. (Basel)* **35**, 1–29 (1981). [doi:10.1159/000155963](doi:10.1159/000155963) Medline

158. Genome annotation GFF files at Mendeley Data for: Y. Shao, L. Zhou, F. Li, L. Zhao, B.-L. Zhang, F. Shao, J.-W. Chen, C.-Y. Chen, X.-P. Bi, X.-L. Zhuang, H.-L. Zhu, J. Hu, Z. Sun, X. Li, D. Wang, I. Rivas-González, S. Wang, Y.-M. Wang, W. Chen, G. Li, H.-M. Lu, Y. Liu, L. F. K. Kuderna, K. K.-H. Farh, P.-F. Fan, L. Yu, M. Li, Z.-J. Liu, G. P. Tiley, A. D. Yoder, C. Roos, T. Hayakawa, T. Marques-Bonet, J. Rogers, P. D. Stenson, D. N. Cooper, M. H. Schierup, Y.-G. Yao, Y.-P. Zhang, W. Wang, X.-G. Qi, G. Zhang, D.-D. Wu, Phylogenomic analyses provide insights into primate evolution, Mendeley (2023); https://doi.org/10.17632/87bh23zxj2.1.

159. Genome annotation GFF files at Figshare for: Y. Shao, L. Zhou, F. Li, L. Zhao, B.-L. Zhang, F. Shao, J.-W. Chen, C.-Y. Chen, X.-P. Bi, X.-L. Zhuang, H.-L. Zhu, J. Hu, Z. Sun, X. Li, D. Wang, I. Rivas-González, S. Wang, Y.-M. Wang, W. Chen, G. Li, H.-M. Lu, Y. Liu, L. F. K. Kuderna, K. K.-H. Farh, P.-F. Fan, L. Yu, M. Li, Z.-J. Liu, G. P. Tiley, A. D. Yoder, C. Roos, T. Hayakawa, T. Marques-Bonet, J. Rogers, P. D. Stenson, D. N. Cooper, M. H. Schierup, Y.-G. Yao, Y.-P. Zhang, W. Wang, X.-G. Qi, G. Zhang, D.-D. Wu, Phylogenomic analyses provide insights into primate evolution, Figshare (2023); https://doi.org/10.6084/m9.figshare.21692894.v1.

160. Gene sequences for: Y. Shao, L. Zhou, F. Li, L. Zhao, B.-L. Zhang, F. Shao, J.-W. Chen, C.-Y. Chen, X.-P. Bi, X.-L. Zhuang, H.-L. Zhu, J. Hu, Z. Sun, X. Li, D. Wang, I. Rivas-González, S. Wang, Y.-M. Wang, W. Chen, G. Li, H.-M. Lu, Y. Liu, L. F. K. Kuderna, K. K.-H. Farh, P.-F. Fan, L. Yu, M. Li, Z.-J. Liu, G. P. Tiley, A. D. Yoder, C. Roos, T. Hayakawa, T. Marques-Bonet, J. Rogers, P. D. Stenson, D. N. Cooper, M. H. Schierup, Y.-G. Yao, Y.-P. Zhang, W. Wang, X.-G. Qi, G. Zhang, D.-D. Wu, Phylogenomic analyses provide insights into primate evolution, Dryad (2023); https://doi.org/10.5061/dryad.8w9ghx3qj.

161. J. Ruan, H. Li, Fast and accurate long-read assembly with wtdbg2. *Nat. Methods* **17**, 155–158 (2020). [doi:10.1038/s41592-019-0669-3](doi:10.1038/s41592-019-0669-3) Medline

162. C.-S. Chin, P. Peluso, F. J. Sedlazeck, M. Nattestad, G. T. Concepcion, A. Clum, C. Dunn, R. O'Malley, R. Figueroa-Balderas, A. Morales-Cruz, G. R. Cramer, M. Delledonne, C. Luo, J. R. Ecker, D. Cantu, D. R. Rank, M. C. Schatz, Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* **13**, 1050–1054 (2016). doi:10.1038/nmeth.4035 Medline

163. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi:10.1093/bioinformatics/btp324 Medline

164. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015). doi:10.1093/bioinformatics/btv351 Medline

165. G. Benson, Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999). doi:10.1093/nar/27.2.573 Medline

166. J. Ye, S. McGinnis, T. L. Madden, BLAST: Improvements for better sequence analysis. *Nucleic Acids Res.* **34**, W6–W9 (2006). doi:10.1093/nar/gkl164 Medline

167. R. She, J. S. Chu, K. Wang, J. Pei, N. Chen, GenBlastA: Enabling BLAST to identify homologous gene sequences. *Genome Res.* **19**, 143–149 (2009). doi:10.1101/gr.082081.108 Medline

168. E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise. *Genome Res.* **14**, 988–995 (2004). doi:10.1101/gr.1865504 Medline

169. M. Stanke, S. Waack, Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003). doi:10.1093/bioinformatics/btg1080 Medline

170. S. Hunter, R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, U. Das, L. Daugherty, L. Duquenne, R. D. Finn, J. Gough, D. Haft, N. Hulo, D. Kahn, E. Kelly, A. Laugraud, I. Letunic, D. Lonsdale, R. Lopez, M. Madera, J. Maslen, C. McAnulla, J. McDowall, J. Mistry, A. Mitchell, N. Mulder, D. Natale, C. Orengo, A. F. Quinn, J. D. Selengut, C. J. A. Sigrist, M. Thimma, P. D. Thomas, F. Valentin, D. Wilson, C. H. Wu, C. Yeats, InterPro: The integrative protein signature database. *Nucleic Acids Res.* **37**, D211–D215 (2009). doi:10.1093/nar/gkn785 Medline

171. M. Blanchette, W. J. Kent, C. Riemer, L. Elnitski, A. F. Smit, K. M. Roskin, R. Baertsch, K. Rosenbloom, H. Clawson, E. D. Green, D. Haussler, W. Miller, Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004). doi:10.1101/gr.1933104 Medline

172. J. Gatesy, R. H. Baker, Hidden likelihood support in genomic data: Can forty-five wrongs make a right? *Syst. Biol.* **54**, 483–492 (2005). doi:10.1080/10635150590945368 Medline

173. L. S. Kubatko, J. H. Degnan, Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* **56**, 17–24 (2007). doi:10.1080/10635150601146041 Medline

174. 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). doi:10.1186/s12859-018-2129-y Medline

175. Y. Fan, M. S. Ye, J. Y. Zhang, L. Xu, D. D. Yu, T. L. Gu, Y. L. Yao, J. Q. Chen, L. B. Lv, P. Zheng, D. D. Wu, G. J. Zhang, Y. G. Yao, Chromosomal level assembly and population sequencing of the Chinese tree shrew genome. *Zool. Res.* **40**, 506–521 (2019). doi:10.24272/j.issn.2095-8137.2019.063 Medline

176. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014). doi:10.1093/bioinformatics/btu033 Medline

177. M. J. Hubisz, K. S. Pollard, A. Siepel, PHAST and RPHAST: Phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**, 41–51 (2011). [doi:10.1093/bib/bbq072](doi:10.1093/bib/bbq072) Medline

178. R. S. Harris, "Improved pairwise alignment of genomic DNA," The Pennsylvania State University, University Park, PA (2007).

179. B. R. Jones, A. Rajaraman, E. Tannier, C. Chauve, ANGES: Reconstructing ANcestral GEnomeS maps. *Bioinformatics* **28**, 2388–2390 (2012). doi:10.1093/bioinformatics/bts457 Medline

180. G. Tesler, GRIMM: Genome rearrangements web server. *Bioinformatics* **18**, 492–493 (2002). doi:10.1093/bioinformatics/18.3.492 Medline

181. B. S. Emanuel, T. H. Shaikh, Segmental duplications: An 'expanding' role in genomic instability and disease. *Nat. Rev. Genet.* **2**, 791–800 (2001). [doi:10.1038/35093500](doi:10.1038/35093500) Medline

182. R. V. Samonte, E. E. Eichler, Segmental duplications and the evolution of the primate genome. *Nat. Rev. Genet.* **3**, 65–72 (2002). [doi:10.1038/nrg705](doi:10.1038/nrg705) Medline

183. J. A. Bailey, Z. Gu, R. A. Clark, K. Reinert, R. V. Samonte, S. Schwartz, M. D. Adams, E. W. Myers, P. W. Li, E. E. Eichler, Recent segmental duplications in the human genome. *Science* **297**, 1003–1007 (2002). doi:10.1126/science.1072047 Medline

184. F. Delehelle, S. Cussat-Blanc, J. M. Alliot, H. Luga, P. Balaresque, ASGART: Fast and parallel genome scale segmental duplications mapping. *Bioinformatics* **34**, 2708–2714 (2018). doi:10.1093/bioinformatics/bty172 Medline

185. E. B. Chuong, N. C. Elde, C. Feschotte, Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017). [doi:10.1038/nrg.2016.139](doi:10.1038/nrg.2016.139) Medline

186. S. J. Thomson, F. G. Goh, H. Banks, T. Krausgruber, S. V. Kotenko, B. M. Foxwell, I. A. Udalova, The role of transposable elements in the regulation of IFN-lambda1 gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11564–11569 (2009). doi:10.1073/pnas.0904477106 Medline

187. P. É. Jacques, J. Jeyakani, G. Bourque, The majority of primate-specific regulatory sequences are derived from transposable elements. *PLOS Genet.* **9**, e1003504 (2013). doi:10.1371/journal.pgen.1003504 Medline

188. Y. Huang, B. Niu, Y. Gao, L. Fu, W. Li, CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010). doi:10.1093/bioinformatics/btq003 Medline

189. GTEx Consortium, The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013). [doi:10.1038/ng.2653](doi:10.1038/ng.2653) Medline

190. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004). [doi:10.1093/nar/gkh340](doi:10.1093/nar/gkh340) Medline

191. J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000). doi:10.1093/oxfordjournals.molbev.a026334 Medline

192. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007). [doi:10.1080/10635150701472164](doi:10.1080/10635150701472164) Medline

193. H. Li, A. Coghlan, J. Ruan, L. J. Coin, J. K. Hériché, L. Osmotherly, R. Li, T. Liu, Z. Zhang, L. Bolund, G. K. Wong, W. Zheng, P. Dehal, J. Wang, R. Durbin, TreeFam: A curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res.* **34**, D572–D580 (2006). [doi:10.1093/nar/gkj118](doi:10.1093/nar/gkj118) Medline

194. E. B. Kim, X. Fang, A. A. Fushan, Z. Huang, A. V. Lobanov, L. Han, S. M. Marino, X. Sun, A. A. Turanov, P. Yang, S. H. Yim, X. Zhao, M. V. Kasaikina, N. Stoletzki, C. Peng, P. Polak, Z. Xiong, A. Kiezun, Y. Zhu, Y. Chen, G. V. Kryukov, Q. Zhang, L. Peshkin, L. Yang, R. T. Bronson, R. Buffenstein, B. Wang, C. Han, Q. Li, L. Chen, W. Zhao, S. R. Sunyaev, T. J. Park, G. Zhang, J. Wang, V. N. Gladyshev, Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* **479**, 223–227 (2011). doi:10.1038/nature10533 Medline

195. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: A computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269–1271 (2006). doi:10.1093/bioinformatics/btl097 Medline

196. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995). [doi:10.1111/j.2517-6161.1995.tb02031.x](doi:10.1111/j.2517-6161.1995.tb02031.x)

197. X. Jiao, B. T. Sherman, W. Huang, R. Stephens, M. W. Baseler, H. C. Lane, R. A. Lempicki, DAVID-WS: A stateful web service to facilitate gene/protein list

analysis. *Bioinformatics* **28**, 1805–1806 (2012). doi:10.1093/bioinformatics/bts251 Medline

198. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009). doi:10.1038/nprot.2008.211 Medline

199. W. Huang, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009). doi:10.1093/nar/gkn923 Medline

200. W. Huang, B. T. Sherman, X. Zheng, J. Yang, T. Imamichi, R. Stephens, R. A. Lempicki, Extracting biological meaning from large gene lists with DAVID. *Curr. Protoc. Bioinformatics* **Chapter 13**, Unit 13.11 (2009). doi:10.1002/0471250953.bi1311s27 Medline

201. W. Huang, B. T. Sherman, R. Stephens, M. W. Baseler, H. C. Lane, R. A. Lempicki, DAVID gene ID conversion tool. *Bioinformation* **2**, 428–430 (2008). doi:10.6026/97320630002428 Medline

202. B. T. Sherman, W. Huang, Q. Tan, Y. Guo, S. Bour, D. Liu, R. Stephens, M. W. Baseler, H. C. Lane, R. A. Lempicki, DAVID Knowledgebase: A gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. *BMC Bioinformatics* **8**, 426 (2007). doi:10.1186/1471-2105-8-426 Medline

203. G. Dennis Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, R. A. Lempicki, DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3 (2003). doi:10.1186/gb-2003-4-5-p3 Medline

204. D. A. Hosack, G. Dennis Jr., B. T. Sherman, H. C. Lane, R. A. Lempicki, Identifying biological themes within lists of genes with EASE. *Genome Biol.* **4**, R70 (2003). doi:10.1186/gb-2003-4-10-r70 Medline

205. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN] (2013).

206. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011). doi:10.1093/bioinformatics/btr509 Medline

207. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). doi:10.1038/ng.806 Medline

208. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). doi:10.1093/bioinformatics/btq033 Medline

209. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi:10.1093/bioinformatics/btr330 Medline

210. M. Hasegawa, H. Kishino, T. Yano, Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985). doi:10.1007/BF02101694 Medline

211. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016). doi:10.1093/molbev/msw054 Medline

212. N. C. Durand, M. S. Shamim, I. Machol, S. S. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016). doi:10.1016/j.cels.2016.07.002 Medline

213. N. Abdennur, L. A. Mirny, Cooler: Scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2020). doi:10.1093/bioinformatics/btz540 Medline

214. F. Ramírez, V. Bhardwaj, L. Arrigoni, K. C. Lam, B. A. Grüning, J. Villaveces, B. Habermann, A. Akhtar, T. Manke, High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018). doi:10.1038/s41467-017-02525-w Medline

215. S. I. Perez, M. F. Tejedor, N. M. Novo, L. Aristide, Divergence times and the evolutionary radiation of New World monkeys (Platyrrhini, Primates): An analysis of fossil and molecular data. *PLOS ONE* **8**, e68029 (2013). doi:10.1371/journal.pone.0068029 Medline

216. J. Eriksson, H. Siedel, D. Lukas, M. Kayser, A. Erler, C. Hashimoto, G. Hohmann, C. Boesch, L. Vigilant, Y-chromosome analysis confirms highly sex-biased dispersal and suggests a low male effective population size in bonobos (*Pan paniscus*). *Mol. Ecol.* **15**, 939–949 (2006). doi:10.1111/j.1365-294X.2006.02845.x Medline

217. M. E. Steiper, Population history, biogeography, and taxonomy of orangutans (Genus: *Pongo*) based on a population genetic meta-analysis of multiple loci. *J. Hum. Evol.* **50**, 509–522 (2006). doi:10.1016/j.jhevol.2005.12.005 Medline

218. Y. Zhang, O. A. Ryder, Y. Zhang, Genetic divergence of orangutan subspecies (*Pongo pygmaeus*). *J. Mol. Evol.* **52**, 516–526 (2001). doi:10.1007/s002390010182 Medline

219. S. K. Kim, L. Carbone, C. Becquet, A. R. Mootnick, D. J. Li, P. J. de Jong, J. D. Wall, Patterns of genetic variation within and between Gibbon species. *Mol. Biol. Evol.* **28**, 2211–2218 (2011). doi:10.1093/molbev/msr033 Medline

220. Z. Fan, G. Zhao, P. Li, N. Osada, J. Xing, Y. Yi, L. Du, P. Silva, H. Wang, R. Sakate, X. Zhang, H. Xu, B. Yue, J. Li, Whole-genome sequencing of tibetan macaque (*Macaca Thibetana*) provides new insight into the macaque evolutionary history. *Mol. Biol. Evol.* **31**, 1475–1489 (2014). doi:10.1093/molbev/msu104 Medline

221. X. Zhang, Q. Zhang, B. Su, Emergence and evolution of inter-specific segregating retrocopies in cynomolgus monkey (*Macaca fascicularis*) and rhesus macaque (*Macaca mulatta*). *Sci. Rep.* **6**, 32598 (2016). doi:10.1038/srep32598 Medline

222. J. Rogers, M. Raveendran, R. A. Harris, T. Mailund, K. Leppälä, G. Athanasiadis, M. H. Schierup, J. Cheng, K. Munch, J. A. Walker, M. K. Konkel, V. Jordan, C. J. Steely, T. O. Beckstrom, C. Bergey, A. Burrell, D. Schrempf, A. Noll, M. Kothe, G. H. Kopp, Y. Liu, S. Murali, K. Billis, F. J. Martin, M. Muffato, L. Cox, J. Else, T. Disotell, D. M. Muzny, J. Phillips-Conroy, B. Aken, E. E. Eichler, T. Marques-Bonet, C. Kosiol, M. A. Batzer, M. W. Hahn, J. Tung, D. Zinner, C. Roos, C. J. Jolly, R. A. Gibbs, K. C. Worley; Baboon Genome Analysis Consortium, The comparative genomics and complex population history of *Papio* baboons. *Sci. Adv.* **5**, eaau6947 (2019). doi:10.1126/sciadv.aau6947 Medline

223. S. K. Patterson, A. A. Sandel, J. A. Miller, J. C. Mitani, Data quality and the comparative method: The case of primate group size. *Int. J. Primatol.* **35**, 990–1003 (2014). doi:10.1007/s10764-014-9777-1

224. D. Zinner, A. Atickem, J. C. Beehner, A. Bekele, T. J. Bergman, R. Burke, S. Dolotovskaya, P. J. Fashing, S. Gippoliti, S. Knauf, Y. Knauf, A. Mekonnen, A. Moges, N. Nguyen, N. C. Stenseth, C. Roos, Phylogeography, mitochondrial DNA diversity, and demographic history of geladas (*Theropithecus gelada*). *PLOS ONE* **13**, e0202303 (2018). doi:10.1371/journal.pone.0202303 Medline

225. Y. Yin, T. Yang, H. Liu, Z. Huang, Y. Zhang, Y. Song, W. Wang, X. Guang, S. K. Sahu, K. Kristiansen, The draft genome of mandrill (*Mandrillus sphinx*): An Old World monkey. *Sci. Rep.* **10**, 2431 (2020). doi:10.1038/s41598-020-59110-3 Medline

226. N. Ting, C. Astaras, G. Hearn, S. Honarvar, J. Corush, A. S. Burrell, N. Phillips, B. J. Morgan, E. L. Gadsby, R. Raaum, C. Roos, Genetic signatures of a demographic collapse in a large-bodied forest dwelling primate (*Mandrillus leucophaeus*). *Ecol. Evol.* **2**, 550–561 (2012). doi:10.1002/ece3.98 Medline

227. 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). doi:10.1007/s10764-018-0029-7

228. W. C. Warren, A. J. Jasinska, R. García-Pérez, H. Svardal, C. Tomlinson, M. Rocchi, N. Archidiacono, O. Capozzi, P. Minx, M. J. Montague, K. Kyung, L. W. Hillier, M. Kremitzki, T. Graves, C. Chiang, J. Hughes, N. Tran, Y. Huang, V. Ramensky, O. W. Choi, Y. J. Jung, C. A. Schmitt, N. Juretic, J. Wasserscheid, T. R. Turner, R. W. Wiseman, J. J. Tuscher, J. A. Karl, J. E. Schmitz, R. Zahn, D. H. O'Connor, E. Redmond, A. Nisbett, B. Jacquelin, M. C. Müller-Trutwin, J. M. Brenchley, M. Dione, M. Antonio, G. P. Schroth, J. R. Kaplan, M. J. Jorgensen, G. W. Thomas, M. W. Hahn, B. J. Raney, B. Aken, R. Nag, J. Schmitz, G. Churakov, A. Noll, R. Stanyon, D. Webb, F. Thibaud-Nissen, M. Nordborg, T. Marques-Bonet, K. Dewar, G. M. Weinstock, R. K. Wilson, N. B. Freimer, The genome of the vervet (*Chlorocebus aethiops sabaeus*). *Genome Res.* **25**, 1921–1933 (2015). doi:10.1101/gr.192922.115 Medline

229. C. Roos, R. Liedigk, V. N. Thinh, T. Nadler, D. Zinner, The hybrid origin of the Indochinese gray langur *Trachypithecus crepusculus*. *Int. J. Primatol.* **40**, 9–27 (2019). doi:10.1007/s10764-017-0008-4

230. T. Minhós, L. Chikhi, C. Sousa, L. M. Vicente, M. Ferreira da Silva, R. Heller, C. Casanova, M. W. Bruford, Genetic consequences of human forest exploitation in two *Colobus* monkeys in Guinea Bissau. *Biol. Conserv.* **194**, 194–208 (2016). doi:10.1016/j.biocon.2015.12.019

231. J. M. Allen, M. M. Miyamoto, C.-H. Wu, T. E Carter, J. Ungvari-Martin, K. Magrini, C. A. Chapman, Primate DNA suggests long-term stability of an African rainforest. *Ecol. Evol.* **2**, 2829–2842 (2012). doi:10.1002/ece3.395 Medline

232. M. K. Konkel, B. Ullmer, E. L. Arceneaux, S. Sanampudi, S. A. Brantley, R. Hubley, A. F. Smit, M. A. Batzer, Discovery of a new repeat family in the *Callithrix jacchus* genome. *Genome Res.* **26**, 649–659 (2016). doi:10.1101/gr.199075.115 Medline

233. U. Radespiel, H. Lutermann, B. Schmelting, E. Zimmermann, An empirical estimate of the generation time of mouse lemurs. *Am. J. Primatol.* **81**, e23062 (2019). doi:10.1002/ajp.23062 Medline

234. C. L. Frasier, J.-N. Rakotonirina, L. G. Razanajatovo, T. S. Nasolonjanahary, S. B. Rasolonileniraka, J. F. Mamiaritiana, E. E. Ramarolahy, Louis Jr., Expanding knowledge on life history traits and infant development in the greater bamboo lemur (*Prolemur simus*): Contributions from Kianjavato, Madagascar. *Primate Conserv.* **2015**, 75–86 (2015). [doi:10.1896/052.029.0110](doi:10.1896/052.029.0110)

235. M. J. E. Charpentier, C. V. Williams, C. M. Drea, Inbreeding depression in ring-tailed lemurs (*Lemur catta*): Genetic diversity predicts parasitism, immunocompetence, and survivorship. *Conserv. Genet.* **9**, 1605–1615 (2008). [doi:10.1007/s10592-007-9499-4](doi:10.1007/s10592-007-9499-4)

236. G. H. Perry, D. Reeves, P. Melsted, A. Ratan, W. Miller, K. Michelini, E. E. Louis Jr., J. K. Pritchard, C. E. Mason, Y. Gilad, A genome sequence resource for the aye-aye (*Daubentonia madagascariensis*), a nocturnal lemur from Madagascar. *Genome Biol. Evol.* **4**, 126–135 (2012). [doi:10.1093/gbe/evr132](doi:10.1093/gbe/evr132) Medline

237. D. Pan, J.-H. Chen, C. Groves, Y.-X. Wang, E. Narushima, H. Fitch-Snyder, P. Crow, X. Jinggong, V. N. Thanh, O. Ryder, L. Chemnick, H.-W. Zhang, Y.-X. Fu, Y.-P. Zhang, Mitochondrial control region and population genetic patterns of *Nycticebus bengalensis* and *N. pygmaeus*. *Int. J. Primatol.* **28**, 791–799 (2007). [doi:10.1007/s10764-007-9157-1](doi:10.1007/s10764-007-9157-1)