

PRIMATE GENOMES

Phylogenomic analyses provide insights into primate evolution

Yong Shao^{1†}, Long Zhou^{2†}, Fang Li^{3,4}, Lan Zhao⁵, Bao-Lin Zhang¹, Feng Shao⁶, Jia-Wei Chen⁷, Chun-Yan Chen⁸, Xupeng Bi², Xiao-Lin Zhuang^{1,9}, Hong-Liang Zhu⁷, Jiang Hu¹⁰, Zongyi Sun¹⁰, Xin Li¹⁰, Depeng Wang¹⁰, Iker Rivas-González¹¹, Sheng Wang¹, Yun-Mei Wang¹, Wu Chen¹², Gang Li¹³, Hui-Meng Lu¹⁴, Yang Liu¹³, Lukas F. K. Kuderna^{15,16}, Kyle Kai-How Farh¹⁶, Peng-Fei Fan¹⁷, Li Yu¹⁸, Ming Li¹⁹, Zhi-Jin Liu²⁰, George P. Tiley²¹, Anne D. Yoder²¹, Christian Roos²², Takashi Hayakawa^{23,24}, Tomas Marques-Bonet^{15,25,33,34}, Jeffrey Rogers²⁶, Peter D. Stenson²⁷, David N. Cooper²⁷, Mikkel Heide Schierup¹¹, 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 $\sim 2.4 \times 10^9$ base pairs (Gbp) (*Daubentonia madagascariensis*) to ~ 3.1 Gbp (*Erythrocebus patas*), which were mostly consistent with the k-mer-based estimations (fig. S2 and table S5), with a high average contig N50 length of $\sim 15.9 \times 10^6$ 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).

¹State Key Laboratory of Genetic Resources and Evolution, Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China.

²Center of Evolutionary & Organismal Biology, and Women's Hospital at Zhejiang University School of Medicine, Hangzhou 310058, China. ³Section for Ecology and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. ⁴Institute of Animal Sex and Development, Zhejiang Wanli University, Ningbo 315100, China. ⁵Shaanxi Key Laboratory for Animal Conservation, College of Life Sciences, Northwest University, Xi'an 710069, China. ⁶Key Laboratory of Freshwater Fish Reproduction and Development (Ministry of Education), Southwest University School of Life Sciences, Chongqing 400715, China. ⁷BGI-Shenzhen, Shenzhen 518083, China. ⁸School of Ecology and Environment, Northwestern Polytechnical University, Xi'an 710072, China. ⁹Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming 650204, China. ¹⁰Grindomics Biosciences, Beijing 102206, China. ¹¹Bioinformatics Research Centre, Aarhus University, DK-8000 Aarhus, Denmark. ¹²Guangzhou Zoo & Guangzhou Wildlife Research Center, Guangzhou 510070, China. ¹³College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China. ¹⁴School of Life Sciences, Northwestern Polytechnical University, Xi'an 710072, China. ¹⁵Institute of Evolutionary Biology (UPF-CSIC), PRBB, 08003 Barcelona, Spain. ¹⁶Illumina Artificial Intelligence Laboratory, Illumina Inc, San Diego, CA 92122, USA. ¹⁷School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275, China. ¹⁸CAS Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, School of Life Sciences, Yunnan University, Kunming 650091, China. ¹⁹CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. ²⁰College of Life Sciences, Capital Normal University, Beijing 100048, China. ²¹Department of Biology, Duke University, Durham, NC 27708, USA. ²²Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center Leibniz Institute for Primate Research, 37077 Göttingen, Germany. ²³Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan. ²⁴Japan Monkey Centre, Inuyama, Aichi 484-0081, Japan. ²⁵Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain. ²⁶Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. ²⁷Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK. ²⁸Key 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. ²⁹Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650201, China. ³⁰National 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. ³¹Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou 311121, China. ³²KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650204, China. ³³CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain. ³⁴Institut 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@zjhu.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 synteny orthologous sequences were retrieved

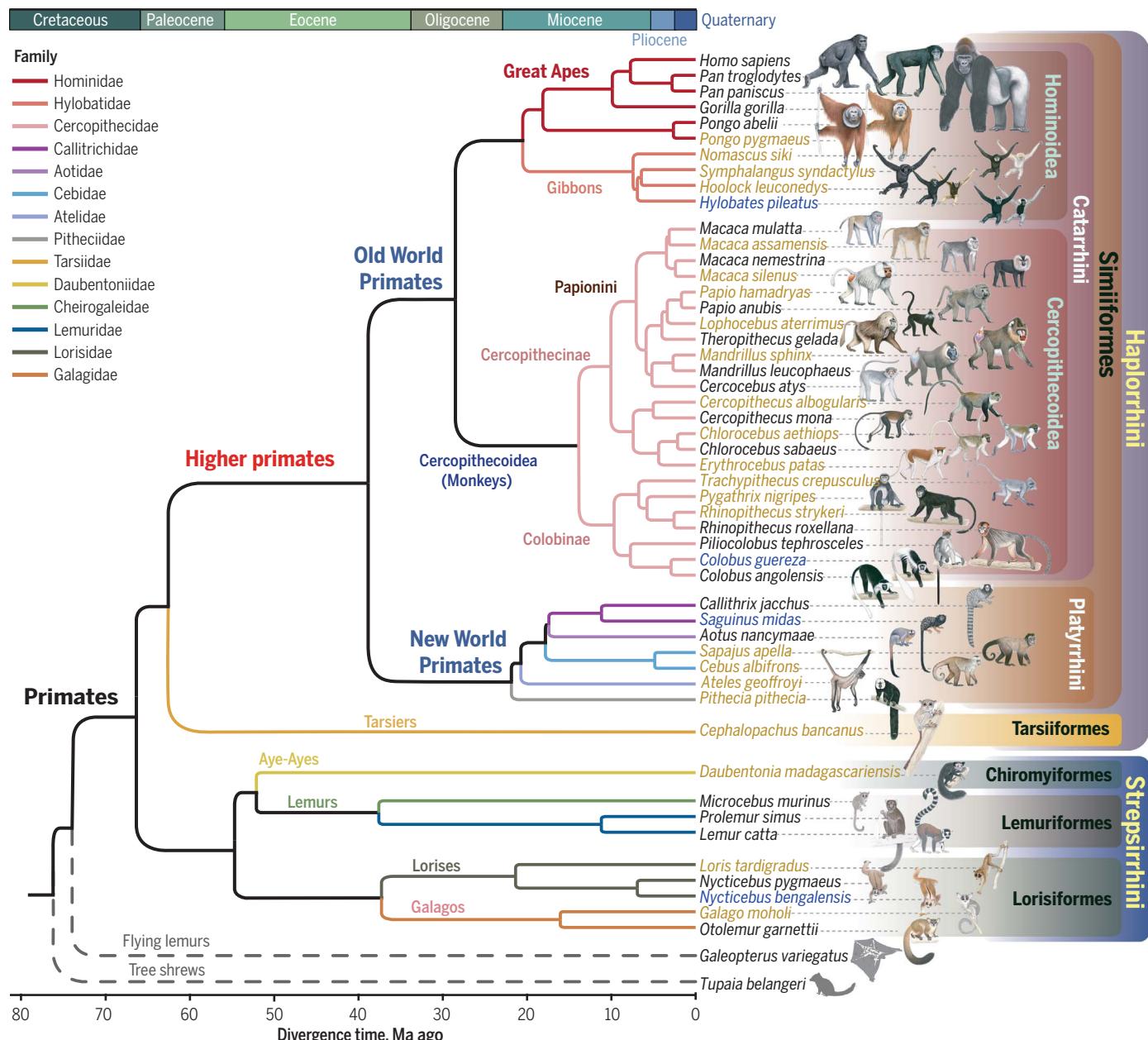


Fig. 1. Genomic phylogeny of primates. The maximum likelihood method was used to infer the primate species tree from whole-genome sequences across 52 species, including 50 primate species and two outgroup species (the Sunda flying lemur and the Chinese tree shrew) with 100 bootstraps under a GTR+GAMMA 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.

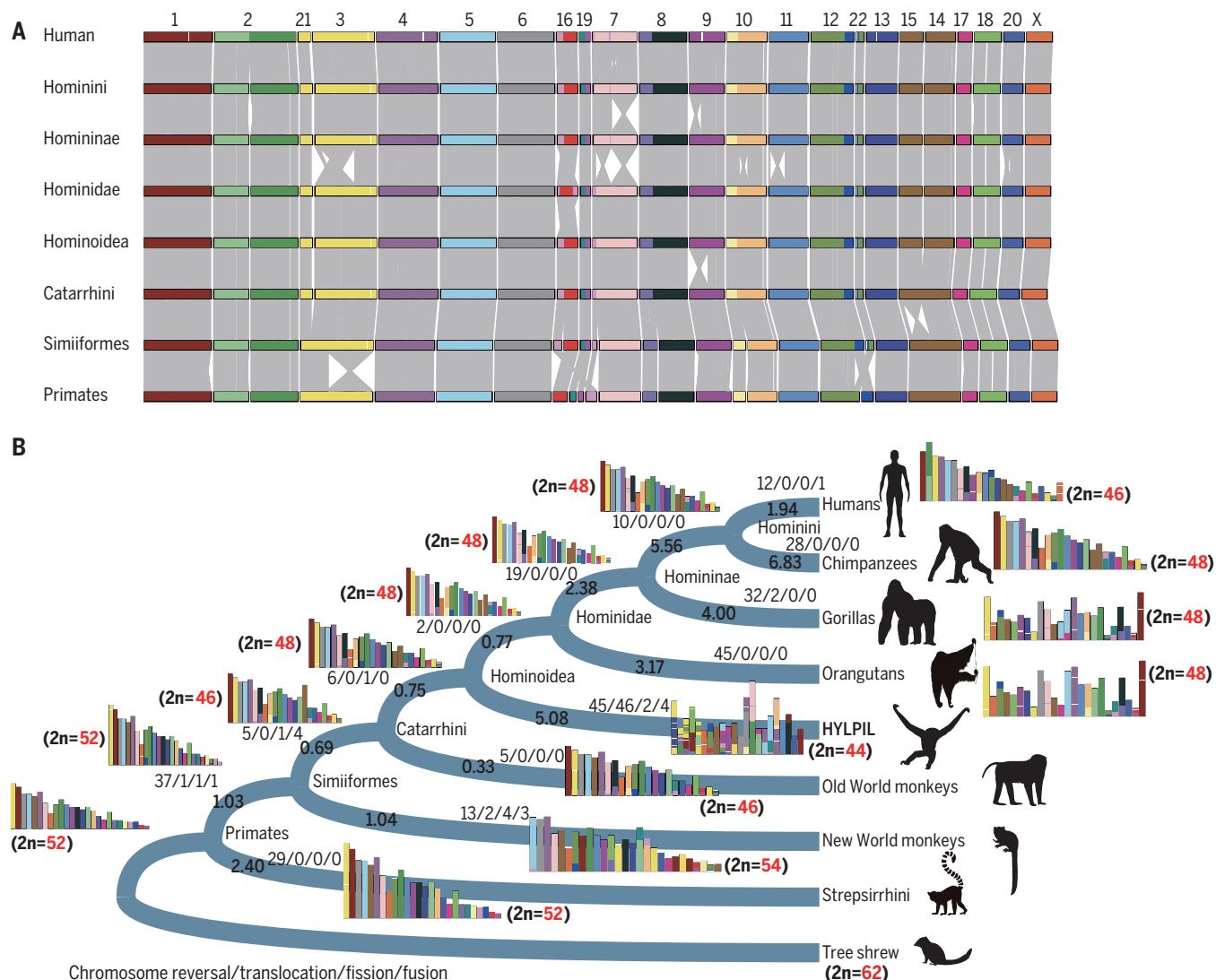


Fig. 2. Reconstruction of primate ancestral chromosomes. (A) Chromosome evolution patterns from the primate common ancestral lineage leading to the human lineage. Chromosomes are colored on the basis of human homologies. (B) Karyotype evolution and genome rearrangement. The rates of genomic rearrangement 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*), (*Sympthalangus*, *Hoolock*)) and ((*Nomascus*, (*Sympthalangus*, *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: Lorisidae 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

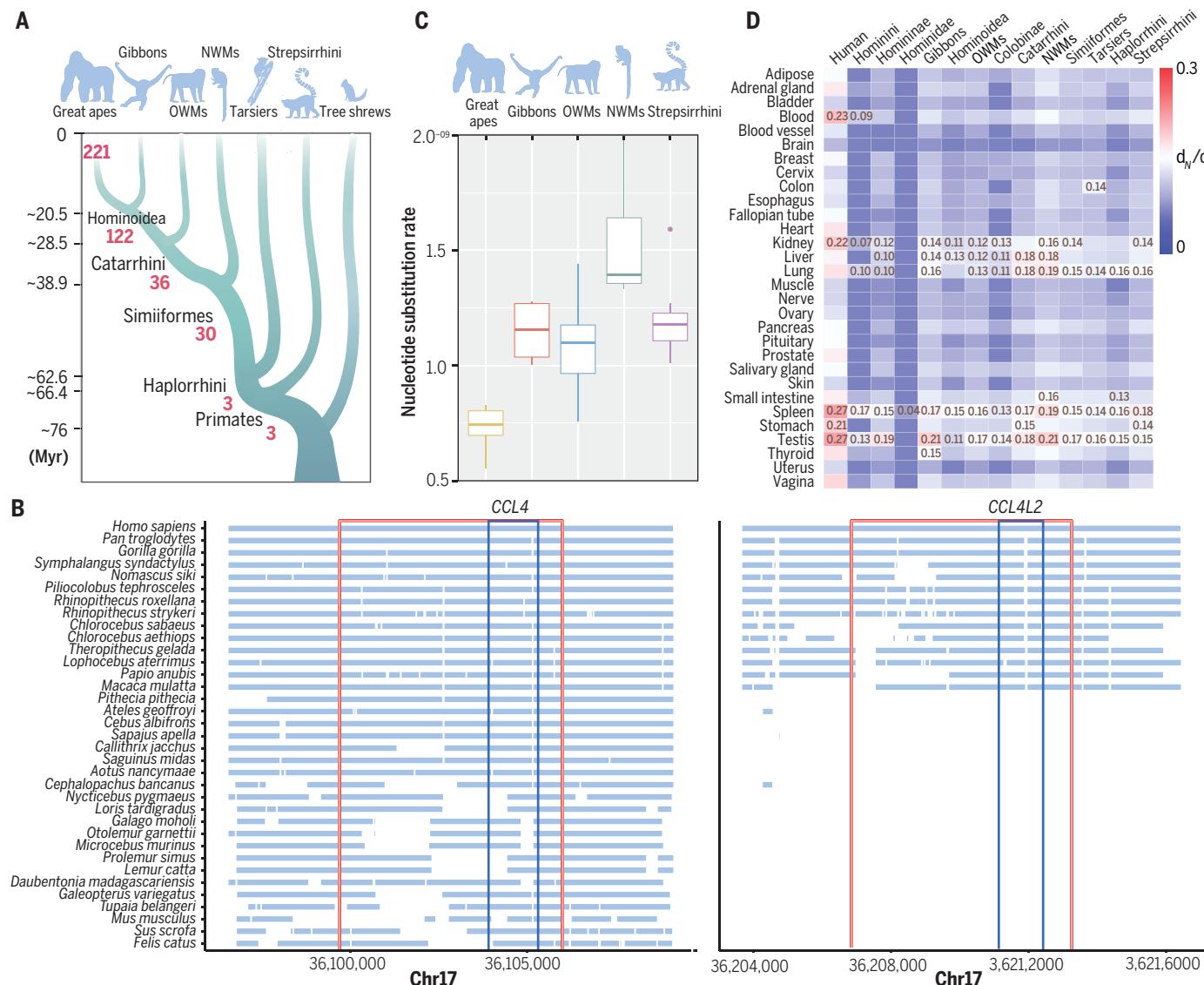


Fig. 3. Structural evolution in primate genomes. (A) Evolutionary pattern of lineage-specific segmental duplications in primates. The numbers of lineage-specific segmental duplications are shown in red. The largest number of segmental duplications was found in the great ape lineage. OWMs, Old World monkeys; 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

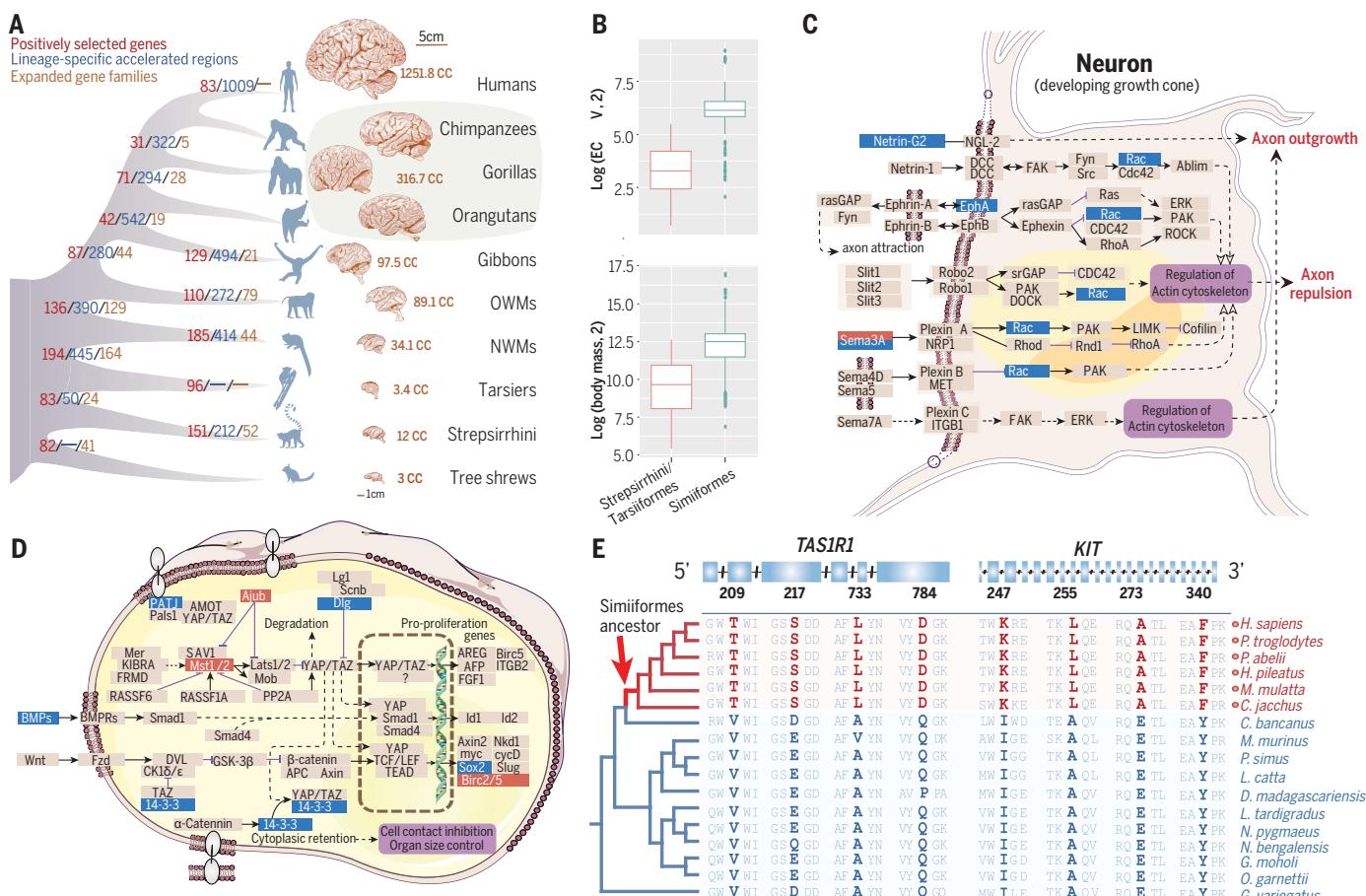


Fig. 4. Genomic changes and phenotype evolution in the ancestor of the Simiiformes. (A) Increased level of genomic evolutionary change, including positively selected genes, lineage-specific accelerated regions, and significantly expanded gene families, seen in the Simiiformes ancestral lineage. The brain 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 (*PATJ*, *SOX2*, *BMP2*, *DLG2*, and *YWHAQ*) in the Simiiformes ancestral lineage are shown in red, and the products of genes associated with lineage-specific accelerated regions (*SEMA3B*, *SEMA3D*, *EPHA3*, *RAC1*, *NTNG2*, and *SEMA3D*) are shown in blue. (E) Multiple sequence alignments of two positively selected genes, *TASIR1* 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 $\sim 2\text{ cm}^3$ in the mouse lemur to $\sim 1300\text{ cm}^3$ 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 I113 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, *PIEZ01*, *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

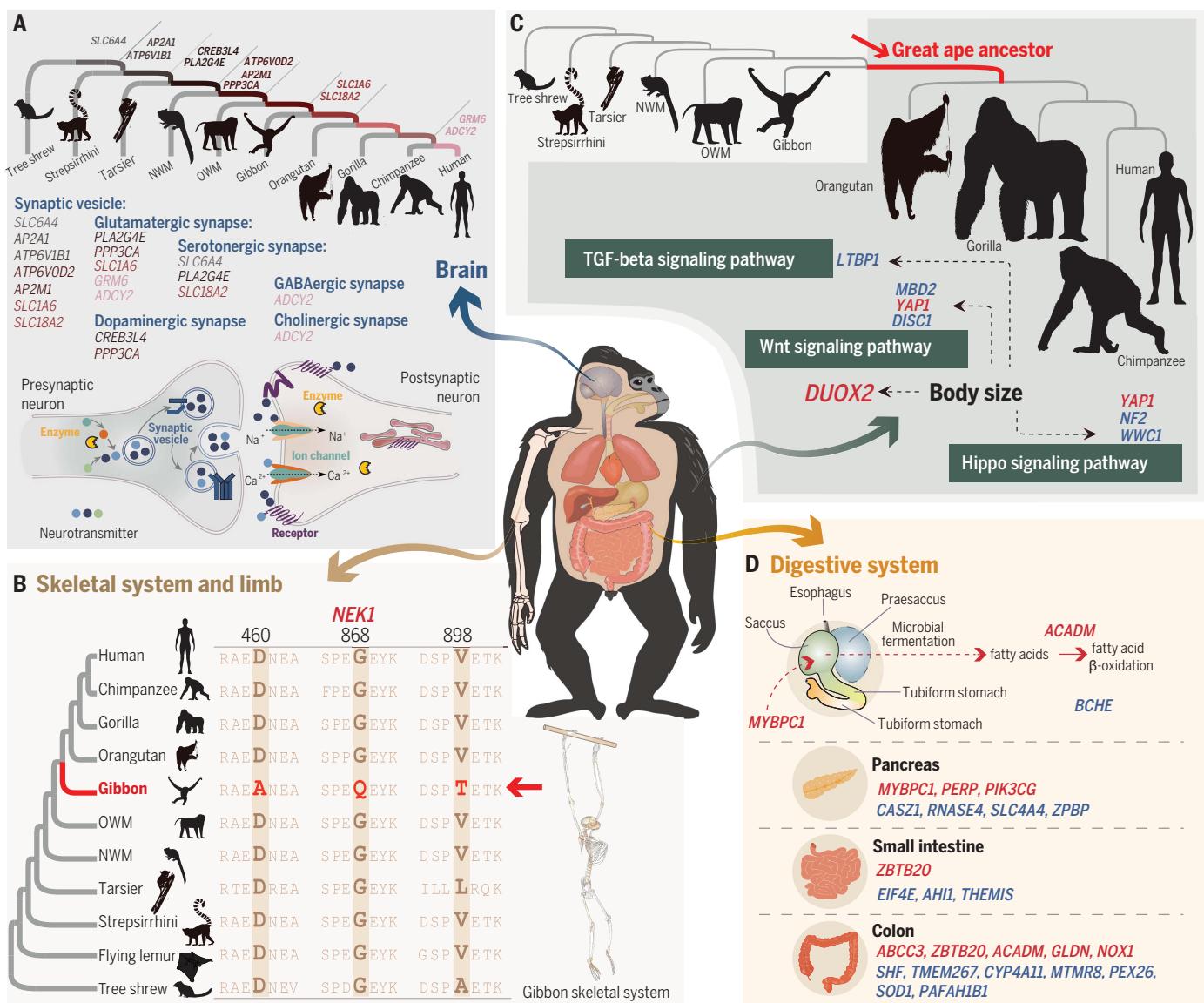


Fig. 5. Associations between genomic evolutionary characteristics and phenotypic traits in primates. (A) Positively selected genes and genes associated with lineage-specific accelerated regions from the primate ancestral lineage leading to the human lineage that are involved in transport, release, and 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,

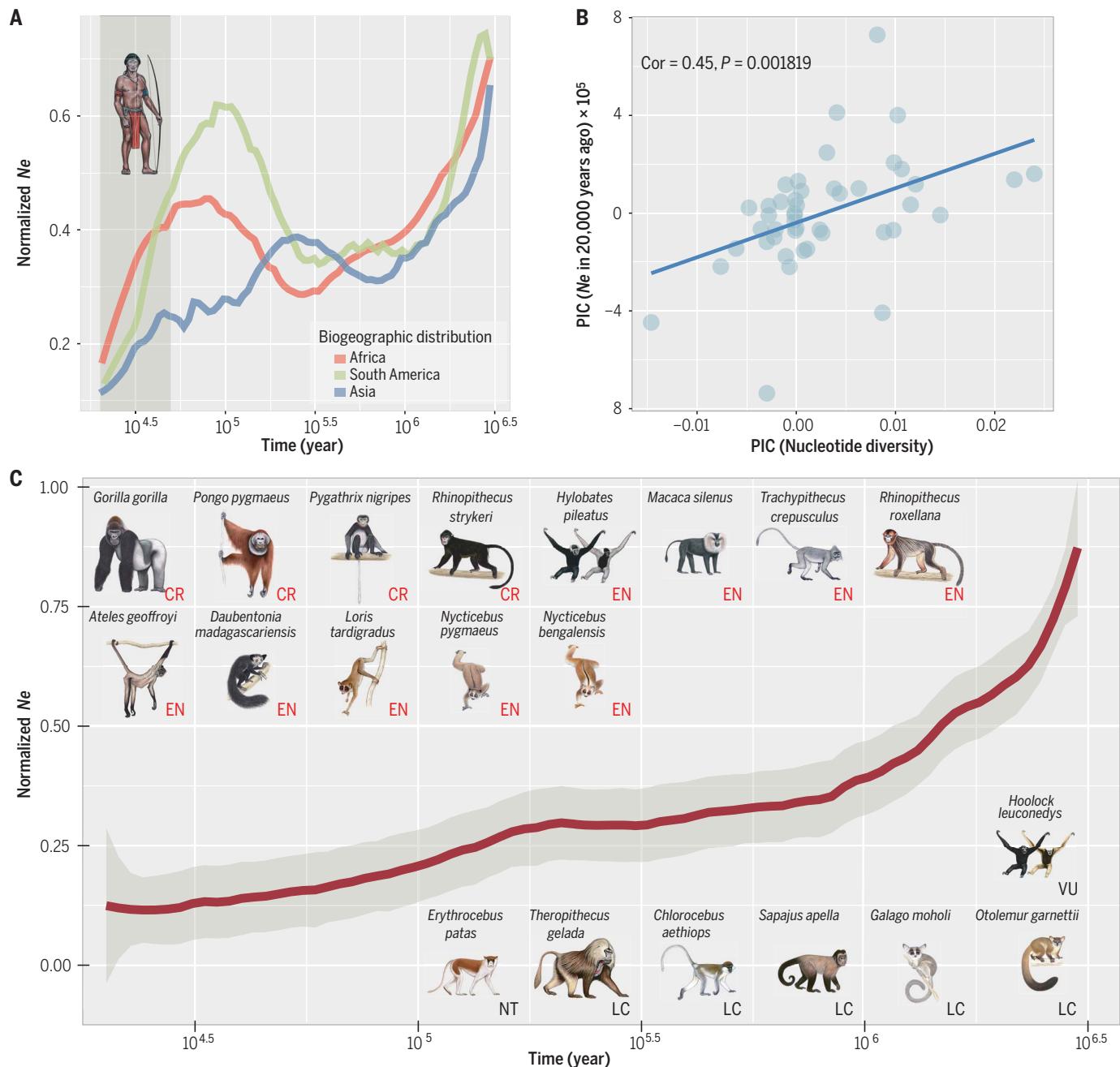


Fig. 6. Demographic history of nonhuman primates. (A) Primate species grouped according to their biogeographic distribution (Africa, Asia, or South America). The plot shows the normalized demographic history of all species within each biogeographic region. The normalized N_e was inferred by dividing 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 corection 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*, *YAPI*, 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 information

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

- A. Estrada et al., *Sci. Adv.* **3**, e1600946 (2017).
- C. Roos et al., *Zool. Res.* **41**, 656–669 (2020).
- A. Nater et al., *Curr. Biol.* **27**, 3487–3498.e10 (2017).
- P. F. Fan et al., *Am. J. Primatol.* **79**, e22631 (2017).
- C. Li, C. Zhao, P. F. Fan, *Am. J. Primatol.* **77**, 753–766 (2015).
- J. Rogers, R. A. Gibbs, *Nat. Rev. Genet.* **15**, 347–359 (2014).
- B. Rockx et al., *Science* **368**, 1012–1015 (2020).
- A. Chandrasekhar et al., *Science* **369**, 812–817 (2020).
- Q. Gao et al., *Science* **369**, 77–81 (2020).
- J. Yu et al., *Science* **369**, 806–811 (2020).
- V. J. Munster et al., *Nature* **585**, 268–272 (2020).
- N. B. Mercado et al., *Nature* **586**, 583–588 (2020).
- K. S. Corbett et al., *N. Engl. J. Med.* **383**, 1544–1555 (2020).
- N. van Doremalen et al., *Nature* **586**, 578–582 (2020).
- B. N. Williamson et al., *Nature* **585**, 273–276 (2020).
- T. Z. Song et al., *Zool. Res.* **41**, 503–516 (2020).
- W. Enard, S. Pääbo, *Annu. Rev. Genomics Hum. Genet.* **5**, 351–378 (2004).
- Z. N. Kronenberg et al., *Science* **360**, eaar6343 (2018).
- Chimpanzee Sequencing and Analysis Consortium, *Nature* **437**, 69–87 (2005).
- R. A. Gibbs et al., *Science* **316**, 222–234 (2007).
- A. Scally et al., *Nature* **483**, 169–175 (2012).
- Marmoset Genome Sequencing and Analysis Consortium, *Nat. Genet.* **46**, 850–857 (2014).
- D. P. Locke et al., *Nature* **469**, 529–533 (2011).
- L. Carbone et al., *Nature* **513**, 195–201 (2014).
- L. Yu et al., *Nat. Genet.* **48**, 947–952 (2016).
- X. Zhou et al., *Nat. Genet.* **46**, 1303–1310 (2014).
- A. O. Ayoola et al., *Mol. Biol. Evol.* **38**, 876–890 (2021).
- D. M. Bickhart et al., *Nat. Genet.* **49**, 643–650 (2017).
- B.-L. Zhang et al., *Sci. Adv.* **9**, eadd3580 (2023).
- H. Wu et al., *Science* **380**, eabl4997 (2023).
- X.-G. Qi et al., *Science* **380**, eabl8621 (2023).
- M.-L. Li et al., *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2123030119 (2022).
- M. S. Ye et al., *Zool. Res.* **42**, 692–709 (2021).
- A. M. Kozlov, A. J. Aberer, A. Stamatakis, *Bioinformatics* **31**, 2577–2579 (2015).
- P. Perelman et al., *PLOS Genet.* **7**, e1001342 (2011).
- C. M. Shi, Z. Yang, *Mol. Biol. Evol.* **35**, 159–179 (2018).
- A. Hobolth, O. F. Christensen, T. Mailund, M. H. Schierup, *PLOS Genet.* **3**, e7 (2007).
- I. Rivas-González et al., *Science* **380**, eabn4409 (2022).
- D. Vanderpool et al., *PLOS Biol.* **18**, e3000954 (2020).
- Z. Yang, *Mol. Biol. Evol.* **20**, 1586–1591 (2007).
- S. Álvarez-Carretero et al., *Nature* **602**, 263–267 (2022).
- C. Liu et al., *Sci. Adv.* **7**, eabe9459 (2021).
- E. E. Eichler, D. Sankoff, *Science* **301**, 793–797 (2003).
- Y. Yin et al., *Nat. Commun.* **12**, 6858 (2021).
- R. Stanyon et al., *Chromosome Res.* **16**, 17–39 (2008).
- T. Marques-Bonet et al., *Nature* **457**, 877–881 (2009).
- P. D. Stenson et al., *Hum. Genet.* **139**, 1197–1207 (2020).
- L. Chen et al., *Science* **364**, eaav6202 (2019).
- J. D. Smith, J. W. Bickham, T. R. Gregory, *Genome* **56**, 457–472 (2013).
- S. Shen et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2837–2842 (2011).
- G. E. Liu, C. Alkan, L. Jiang, S. Zhao, E. E. Eichler, *Genome Res.* **19**, 876–885 (2009).
- T. Hayakawa, Y. Satta, P. Gagneux, A. Varki, N. Takahata, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11399–11404 (2001).
- P. Kuehnert et al., *PLOS Genet.* **8**, e1002543 (2012).
- J. Jurka, *Curr. Opin. Genet. Dev.* **14**, 603–608 (2004).
- G. Zhang et al., *Science* **346**, 1311–1320 (2014).
- P. Moorjani, C. E. Amorim, P. F. Arndt, M. Przeworski, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10607–10612 (2016).
- E. Fontanillas, J. J. Welch, J. A. Thomas, L. Bromham, *BMC Evol. Biol.* **7**, 95 (2007).
- A. Wong, *Mol. Biol. Evol.* **31**, 1432–1436 (2014).
- W. H. Li, M. Tanimura, *Nature* **326**, 93–96 (1987).
- M. E. Steiper, N. M. Young, *Mol. Phylogenet. Evol.* **41**, 384–394 (2006).
- S. H. Kim, N. Elango, C. Warden, E. Vigoda, S. V. Yi, *PLOS Genet.* **2**, e163 (2006).
- J. Schmitz et al., *Nat. Commun.* **7**, 12997 (2016).
- L. Fang et al., *Genome Res.* **30**, 790–801 (2020).
- B. Y. Liao, J. Zhang, *Mol. Biol. Evol.* **23**, 1119–1128 (2006).
- G. J. Wyckoff, W. Wang, C. I. Wu, *Nature* **403**, 304–309 (2000).
- T. Boehm, *Curr. Biol.* **22**, R722–R732 (2012).
- H. Y. Wang et al., *PLOS Biol.* **5**, e13 (2007).
- Materials and methods are available as supplementary materials.
- J. Tohyama et al., *J. Hum. Genet.* **60**, 167–173 (2015).
- P. Mansfield, J. N. Constantino, D. Baldwin, *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **183**, 227–233 (2020).
- M. Maekawa et al., *J. Neurochem.* **115**, 1374–1385 (2010).
- X. Bi et al., *Sci. Adv.* 10.1126/sciadv.95057 (2023).
- J. K. Rilling, T. R. Insel, *Neuroreport* **10**, 1453–1459 (1999).
- K. Isler et al., *J. Hum. Evol.* **55**, 967–978 (2008).
- C. Plachez, L. J. Richards, *Curr. Top. Dev. Biol.* **69**, 267–346 (2005).
- M. A. Robichaux, C. W. Cowan, *Curr. Top. Behav. Neurosci.* **16**, 19–48 (2014).
- J. Falk et al., *Neuron* **48**, 63–75 (2005).
- M. A. Wolman, Y. Liu, H. Tawarayama, W. Shoji, M. C. Halloran, *J. Neurosci.* **24**, 8428–8435 (2004).
- C. Kudo, I. Ajioka, Y. Hirata, K. Nakajima, *J. Comp. Neurol.* **487**, 255–269 (2005).
- M. V. Tejada-Simon, *J. Neurochem.* **133**, 767–779 (2015).
- S. L. Eastwood, P. J. Harrison, *Neuropsychopharmacology* **33**, 933–945 (2008).
- D. Pan, *Genes Dev.* **21**, 886–897 (2007).
- S. H. Patel, F. D. Camargo, D. Yilmaz, *Gastroenterology* **152**, 533–545 (2017).
- R. H. Gokhale, A. W. Shingleton, *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 335–356 (2015).
- E. C. Kirk, *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **281**, 1095–1103 (2004).
- A. C. Wilk et al., *Genome Res.* **18**, 1415–1421 (2008).
- P. Liskova et al., *Am. J. Hum. Genet.* **102**, 447–459 (2018).
- Y. Toda et al., *Curr. Biol.* **31**, 4641–4649.e5 (2021).
- J. M. Kamilar, B. J. Bradley, *J. Biogeogr.* **38**, 2270–2277 (2011).
- S. Hu et al., *PeerJ* **8**, e9402 (2020).
- M. C. Garrido, B. C. Bastian, *J. Invest. Dermatol.* **130**, 20–27 (2010).
- J. M. Grichnik, *J. Invest. Dermatol.* **126**, 945–947 (2006).
- Y. Mizutani, N. Hayashi, M. Kawashima, G. Imokawa, *Arch. Dermatol. Res.* **302**, 283–294 (2010).
- R. Kitamura et al., *J. Pathol.* **202**, 463–475 (2004).
- B. Wen et al., *Pigment Cell Melanoma Res.* **23**, 441–447 (2010).
- D. L. C. van den Berg et al., *Neuron* **93**, 348–361 (2017).
- G. H. Mochida, C. A. Walsh, *Curr. Opin. Neurol.* **14**, 151–156 (2001).
- S. H. Montgomery, I. Capellini, C. Venditti, R. A. Barton, N. I. Mundy, *Mol. Biol. Evol.* **28**, 625–638 (2011).
- L. Shi, M. Li, Q. Lin, X. Qi, B. Su, *BMC Biol.* **11**, 62 (2013).
- L. Shi, B. Su, *Zool. Res.* **40**, 236–238 (2019).
- J. Rogers et al., *Neuroimage* **53**, 1103–1108 (2010).
- S. V. Puram et al., *Genes Dev.* **25**, 2659–2673 (2011).
- A. Yamada et al., *Mol. Cell. Neurosci.* **56**, 234–243 (2013).
- R. Kusano et al., *FEBS Lett.* **590**, 3606–3615 (2016).
- M. Talarowska, J. Szemraj, M. Kowalczyk, P. Gafecki, *Med. Sci. Monit.* **22**, 152–160 (2016).
- A. K. Pandey, L. Lu, X. Wang, R. Homayouni, R. W. Williams, *PLOS ONE* **9**, e88889 (2014).
- A. Graziani, G. Foffani, E. B. Knudsen, J. Shumsky, K. A. Moxon, *PLOS ONE* **8**, e54350 (2013).
- H. Li et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9397–9402 (2008).
- Y. Chang, O. Klezovitch, R. S. Walikonis, V. Vasioukhin, J. J. LoTurco, *Cell Cycle* **9**, 1990–1997 (2010).
- M. R. Sarkisian, *Cell Cycle* **9**, 1876 (2010).
- D. A. Berg, L. Belnoue, H. Song, A. Simon, *Development* **140**, 2548–2561 (2013).
- P. Levitt, J. A. Harvey, E. Friedman, K. Simansky, E. H. Murphy, *Trends Neurosci.* **20**, 269–274 (1997).
- L. Wang et al., *Nat. Commun.* **11**, 282 (2020).
- M. Lindner et al., *Cell Death Differ.* **25**, 1094–1106 (2018).
- F. Xiao et al., *Cell. Physiol. Biochem.* **45**, 1927–1939 (2018).
- S. Zanotti, E. Canalis, *Bone* **62**, 22–28 (2014).
- J. M. Kim, C. Lin, Z. Stavre, M. B. Greenblatt, J. H. Shim, *Cells* **9**, 2073 (2020).
- S. Zanotti, E. Canalis, *Endocr. Rev.* **37**, 223–253 (2016).
- M. Schmidt, *Adv. Sci. Res.* **5**, 23–39 (2011).
- Y. He et al., *Nat. Commun.* **10**, 4233 (2019).
- S. A. Williams, G. A. Russo, *Evol. Anthropol.* **24**, 15–32 (2015).
- K. Semba et al., *Genetics* **172**, 445–456 (2006).
- N. Al Dhaheri et al., *Am. J. Med. Genet. A* **182**, 1664–1672 (2020).
- J. R. Usherwood, J. E. Bertram, *J. Exp. Biol.* **206**, 1631–1642 (2003).
- J. R. Usherwood, S. G. Larson, J. E. Bertram, *Am. J. Phys. Anthropol.* **120**, 364–372 (2003).
- S. M. Cheyne, in *Primate Locomotion: Linking Field and Laboratory Research*, K. D'Août, E. E. Vereecke, Eds. (Springer, NY, 2011), pp. 201–213.
- C. Thiel et al., *Am. J. Hum. Genet.* **88**, 106–114 (2011).
- J. El Hokayem et al., *J. Med. Genet.* **49**, 227–233 (2012).
- J. M. Vasquez, V. J. Lynch, *eLife* **10**, e65041 (2021).
- J. G. M. Thewissen, L. N. Cooper, J. C. George, S. Bajpai, *Evolution (N. Y.)* **2**, 272–288 (2009).
- W. L. Jungers, in *Size and Scaling in Primate Biology*, W. L. Jungers, Ed. (Springer, 1985), pp. 345–381.
- A. M. Rudolf et al., *Natl. Sci. Rev.* **9**, nwab125 (2021).
- K. R. Johnson et al., *Mol. Endocrinol.* **21**, 1593–1602 (2007).
- J. G. Fleagle, *Primate Adaptation and Evolution* (Academic, 2013).
- K. Milton, *Int. J. Primatol.* **19**, 513–548 (1998).
- I. Matsuda, C. A. Chapman, M. Clauss, *J. Morphol.* **280**, 1608–1616 (2019).
- M. C. Janiak, *Evol. Anthropol.* **25**, 253–266 (2016).
- J. J. Kim, R. Miura, *Eur. J. Biochem.* **271**, 483–493 (2004).

139. C. Matzouridou *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

We are grateful to the many individuals in our host institutions who provided support for this project. **Funding:** This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grants XDPB17 and XDB31020000); the National Natural Science Foundation of China (grants 31822048 and 32270500); the CAS Light of West China Program (grant xbzg-zdys-202213); the Yunnan Fundamental Research Project (grant 2019FJ010); the Animal Branch of the Germplasm Bank of Wild Species of Chinese Academy of Science (Large Research Infrastructure Funding); the International Partnership Program of Chinese Academy of Sciences (grant 152453KYSB20170002); a Villum Investigator Grant (25900 to G.Z.); the Japan Society for the Promotion of Science (JSPS) KAKENHI grants 16K18630, 19K16241, 20H04987, 21H04919, and 21KK0106; Hokkaido University Sousei Tokutei Research; and JSPS Bilateral Joint Research Project (JPJSBP grant 120219902 to T.H.). T.M.B. was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant 864203), PID2021-126004NB-100 (MICIIN/FEDER, UE), and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2021 SGR 00177). **Author contributions:** D.D.W. and G.J.Z. led the project. D.D.W., G.J.Z., and X.G.Q. conceived and designed the research. Y.S., L.Z., F.L., L.Z., B.L.Z., F.S., J.W.C., C.Y.C., X.P.B., X.L.Z., H.L.Z., I.R.G., S.W., Y.M.W., L.K., G.L., H.M.L., Y.L., and P.D.S. performed comparative genomics analysis. L.Z., J.H., Z.Y.S., X.L., D.P.W., and K.F. contributed genome sequencing, assembly, and annotation. P.F.F., M.L., Z.J.L., G.P.T., A.D.Y., C.R., T.H., T.M.B., and J.R. collected samples. J.R. and T.M.B. generated

some genome assemblies for comparative genomics analysis. C.R., G.P.T., J.R., L.Y., M.H.S., D.N.C., Y.G.Y., Y.P.Z., W.W., and X.G.Q. provided comments for improving the manuscript. Y.S., X.G.Q., and L.Z. plotted and revised the figures. Y.S. drafted the manuscript. D.D.W., G.J.Z., and Y.S. wrote the manuscript. D.N.C. edited the manuscript. All authors approved the final manuscript.

Competing interests: J.R. is also a core scientist at the Wisconsin National Primate Research Center, University of Wisconsin, Madison. Employees of Illumina, Inc., are indicated in the list of author affiliations. The authors declare no competing financial interests.

Data and materials availability: All 27 primate genome assemblies and the raw genome long- and short-read sequencing data have been deposited at the NCBI Assembly Database (<https://www.ncbi.nlm.nih.gov/assembly/>) and the Sequence Read Archive Database (<https://www.ncbi.nlm.nih.gov/sra/>) under accessible BioProject accession codes PRJNA785018 and PRJNA911016. All genome annotation GFF files have been uploaded to the Mendeley Data database (158) and the Figshare database (159). The positively selected genes and their sequence alignments have been uploaded to a public Dryad dataset (160). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

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

Submitted 16 December 2021; accepted 26 January 2023
10.1126/science.abn6919