

Population Genomics of Japanese Macaques (*Macaca fuscata*): Insights Into Deep Population Divergence and Multiple Merging Histories

Atsunori Higashino ¹, Katsuki Nakamura ¹, Naoki Osada ^{2,*}

¹Center for the Evolutionary Origins of Human Behavior, Kyoto University, Inuyama, Aichi 484-8506, Japan

²Faculty of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan

*Corresponding author: E-mail: nosada@ist.hokudai.ac.jp.

Accepted: December 27, 2024

Abstract

The influence of long-term climatic changes such as glacial cycles on the history of living organisms has been a subject of research for decades, but the detailed population dynamics during the environmental fluctuations and their effects on genetic diversity and genetic load are not well understood on a genome-wide scale. The Japanese macaque (*Macaca fuscata*) is a unique primate adapted to the cold environments of the Japanese archipelago. Despite the past intensive research for the Japanese macaque population genetics, the genetic background of Japanese macaques at the whole-genome level has been limited to a few individuals, and the comprehensive demographic history and genetic differentiation of Japanese macaques have been underexplored. We conducted whole-genome sequencing of 64 Japanese macaque individuals from 5 different regions, revealing significant genetic differentiation and functional variant diversity across populations. In particular, Japanese macaques have low genetic diversity and harbor many shared and population-specific gene loss, which might contribute to population-specific phenotypes. Our estimation of population demography using phased haplotypes suggested that, after the strong population bottleneck shared among all populations around 400 to 500 kya, the divergence among populations initiated around 150 to 200 kya, but there has been the time with strong gene flow between some populations after the split, indicating multiple population split and merge events probably due to habitat fragmentation and fusion during glacial cycles. These findings not only present a complex population history of Japanese macaques but also enhance their value as research models, particularly in neuroscience and behavioral studies. This comprehensive genomic analysis sheds light on the adaptation and evolution of Japanese macaques, contributing valuable insights to both evolutionary biology and biomedical research.

Key words: Japanese macaques, genetic diversity, population dynamics.

Significance

In this study, we analyze high-coverage whole-genome data of 64 Japanese macaque (*Macaca fuscata*) individuals across the Japanese archipelago to elucidate their population history and its genomic consequences. We revealed strong genetic differentiation among local populations with considerably small effective population sizes, and each population harbors a number of characteristic loss-of-function mutations affecting hundreds of genes, a part of which may contribute to the population-specific phenotypes of this species. The initial divergence among populations was estimated to have occurred around 150 to 200 kya based on cross species coalescent analysis, which further suggested the subsequent glacial cycles served as factors for both merging and splitting of local populations, highlighting the importance of environmental fluctuation for shaping the pattern of genomic diversity in nonhuman primates.

Introduction

The influence of long-term climatic changes on the history of living organisms has been a subject of research for decades (e.g. Willis et al. 2004; Cooper et al. 2015). Glacial cycles are among the most significant factors, having extensively altered global species distributions and impacted population sizes (Haffer 1969; Provan and Bennett 2008). However, the detailed population dynamics during these glacial cycles and their effects on genetic diversity and genetic load have been explored on a genome-wide scale in a limited number of organisms (e.g. Lorenzana et al. 2022; Lozier et al. 2023; Feng et al. 2024). Recent studies in population genomics have unveiled that actual population histories are far more complex than previously thought, often involving multiple splits and merges, as suggested by research on archaic and modern humans (Bergström et al. 2020; Hubisz et al. 2020; Ragsdale et al. 2023). Studying species that have endured harsh environments over extended periods, particularly those forming multiple structured populations, can offer valuable insights into how environmental changes have shaped genomic diversity of species.

The Japanese macaque (*Macaca fuscata*), a species of the family Cercopithecidae, endemic to the Japanese archipelago, belongs to the *mulatta* group within the genus *Macaca* (Zinner et al. 2013). The *mulatta* group consists of three species: *M. mulatta*, *M. cyclopis*, and *M. fuscata*, with *M. fuscata* being the most distantly related to the other two. The divergence time of *M. fuscata* to the others has been estimated to be around 1.0 to 2.3 Ma using whole-genome datasets (Kuderna et al. 2023; Zhang et al. 2023; Zhou et al. 2024). However, Osada et al. (2021) demonstrated that *M. fuscata* is more closely related to Chinese *M. mulatta* than to Indian *M. mulatta*, suggesting past gene flow between *M. fuscata* and Chinese *M. mulatta* after the divergence of Chinese and Indian *M. mulatta* that was estimated to have occurred around 162 kya (Hernandez et al. 2007).

As the nonhuman primate inhabiting the northernmost regions, the Japanese macaque has uniquely adapted to cold environments, having large body size, long fur, and extremely short tail. They distribute wide range of the Japanese archipelago except for the Hokkaido Island and the Ryukyu Islands. A population in the Yaku Islands, which locates the southernmost area Japanese macaques inhabiting, is often categorized as a different subspecies, *M. fuscata yakui*, although genetic studies using nuclear genome markers rejected the monophyly of *M. f. yakui* (Ito et al. 2021). Fossil records suggested that they have inhabited in the Japanese archipelago at least 430 kya, but their evolutionary history is poorly understood (Aimi 2002). Studies in population genetics have extensively explored the lineage of Japanese macaques. Analyses

comparing mitochondrial DNA haplotypes reveal genetic differentiation between the Western and Eastern populations, with the Eastern populations exhibiting lower nucleotide diversity than their Western counterparts (Kawamoto et al. 2007). Furthermore, among the Eastern populations, the Shimokita Peninsula population exhibits lower genetic diversity, attributed to the restricted gene flow with neighboring populations (Kawamoto et al. 2008). Ito et al. (2021) performed double-digested restriction-site associated DNA sequencing (ddRAD-seq) including a wide range of Japanese macaque samples across the Japanese archipelago and clarified their fine-scale genetic structure (Ito et al. 2021). They also performed ecological niche modeling and showed that their habitat was severely restricted to south coast of Japanese archipelago during the Last Glacial Maximum (LGM). The ddRAD-seq analysis by Ito et al. (2021) and other phylogeographic studies based on mitochondrial DNA sequences (Kawamoto et al. 2007; Kawamoto 2010) provided a general overview of genetic structure of Japanese macaques; however, remaining questions including a fine-scale demographic trajectories over time and genomic diversity at the nucleotide level have not thoroughly investigated.

The Japanese macaques have also been used as experimental animals, particularly by Japanese researchers for physiological and behavioral studies. In the field of neuroscience, Japanese macaques are particularly valued and employed for their intelligence, docility, and adeptness with their hands. The Japanese government hosts the National BioResource Project (NBRP), funded by the Ministry of Education, Culture, Sports, Science and Technology, providing a repository of Japanese macaques for research purposes (Isa et al. 2009). The NBRP for Japanese macaques has strategically introduced macaques from various regions across Japan, establishing breeding populations that preserve genetic diversity. Research into these populations has uncovered phenotypic variations unique to local groups of Japanese macaques. Over 10 years of daily breeding observations have highlighted the distinct characteristics inherent to each local group within the NBRP colonies. For instance, the Minoh population exhibits a tendency toward obesity, the Hagachizaki population shows a remarkable lack of fear toward humans, and the Okazaki population is notable for being free from Simian T-cell leukemia virus infection. While these populations from diverse geographical regions of Japanese archipelago have been maintained separately, showing unique phenotypic traits, the extent of their genetic differentiation at a genome-wide level has yet to be fully understood. A deeper understanding of their genetic backgrounds could significantly enhance the utility of Japanese macaques as valuable resources for scientific research.

Compared with other species in the *mulatta* group, the genomic information of Japanese macaques is quite

limited, and only a few individual genomes have been fully sequenced (Osada et al. 2021). Here, we performed whole-genome sequencing of 64 Japanese macaque individuals originated from 5 geographic locations and reared in NBRP and clarify their genomic background, particularly focusing on how many neutral and potentially functional variants were carried in the populations. We also sequenced four trios from different populations and performed haplotype phasing using the information. This large-scale dataset and phased haplotype panels would become important genetic resources for future experimental and genetic studies using Japanese macaques. We found the level of genetic differentiation among populations is quite high and highly differentiated variants caused the functional loss of hundreds of genes in each population, which potentially contribute to characteristic phenotypes in each population. The population history of Japanese macaques was also inferred from the whole-genome dataset, showing a highly complex history.

Results

Genome Sequencing of Japanese Macaques From Five Populations

We sequenced the genomes of 64 Japanese macaques from 5 NBRP colonies—Jigokudani, Hagachizaki, Minoh, Tsubaki, and Okazaki—with their geographical origins presented in Fig. 1 with details in [supplementary table S1, Supplementary Material](#) online. Although reared in captivity, all sequenced samples originated as founders from wild populations. Additionally, genomic data for one Koshima individual, seven Chinese rhesus macaques (RMC) (*M. mulatta*), two Indian rhesus macaques (RMI), one Taiwanese macaque (*M. cyclopis*), and one Thai cynomolgus macaque (CMT) (*M. fascicularis*) were retrieved from the public databases ([supplementary table S2, Supplementary Material](#) online) (Xue et al. 2016; Liu et al. 2018; Osada et al. 2021). Using the rhesus macaque reference genome (Mmul_10), we mapped short reads to identify single nucleotide



Fig. 1. Locations of Japanese macaque samples analyzed in this study. The locations include the following: Jigokudani (Yamanouchi Town, Nagano Prefecture); Hagachizaki (Minami-Izu Town, Shizuoka Prefecture); Minoh (Minoh City, Osaka Prefecture); Tsubaki (Shirahama Town, Wakayama Prefecture); Okazaki (Shodo-shima Town, Kagawa Prefecture); and Koshima (Kushima City, Miyazaki Prefecture).

variants (SNVs) and short indels. After filtering, a total of 20,888,045 SNVs and 5,581,784 short indels were detected on the autosomes of Japanese macaque samples. Of these, 4,055,738 SNVs were fixed among all Japanese macaque individuals. On the X chromosome, we identified 543,735 SNVs and 197,404 indels.

For each Japanese macaque individual, we detected 4,000 to 10,000 heterozygous SNVs among 1,930,114,258 autosomal sites that passed our filtering criteria. We evaluated individual nucleotide diversity (heterozygosity per base pair), representing the nucleotide difference per site between two autosomes for each individual. The values were ~0.08% to 0.10%, with the notable exception of the Okazaki population, which exhibited extremely lower heterozygosity at around 0.04%. Compared with the Japanese macaques, other macaque species within the *mulatta* group showed considerably higher heterozygosity: 0.14% in the Taiwanese macaque and 0.16% to 0.24% in rhesus macaques (Fig. 2). Grouping all unrelated Japanese individuals and calculating the nucleotide diversity among all Japanese macaques yield a value of 0.13%, which still remains lower than those observed in other macaque samples.

We also assessed the number and length of runs of homozygosity (ROH) in each genome. These two values showed a negative correlation with individual nucleotide diversity. Notably, individuals in the Okazaki population exhibited a strikingly higher number and longer ROHs compared with Japanese macaques from other populations (Fig. 2; supplementary table S3, Supplementary Material online).

Functional Impact of Variants in Japanese Macaques

The low genetic diversity in Japanese macaques suggests they underwent significant population bottlenecks, likely during their migration to the Japanese archipelago or habitat reductions in glacial periods. We focused on functional mutations universally present in the autosomes of Japanese macaque populations. In total, the Japanese macaques exhibit 12,907 fixed nonsynonymous substitutions against the rhesus macaque reference genome, some of which may be shared with other macaque species. In order to focus on Japanese macaque-specific derived variants with functional importance, we only considered the sites where all 11 outgroup individuals, *M. fascicularis*, *M. mulatta*, and *M. cyclopis* (supplementary table S2, Supplementary Material online), unambiguously have the reference allele as the homozygous state. We identified 40 SNVs and 53 short indels resulting in loss-of-function (LoF), caused by frameshifts, premature stop codons, and splicing disruptions, fixed and specific to the Japanese macaque samples.

On average, each Japanese macaque carries 39.6 and 84.0 species-specific homozygous LoF variants segregating within the species due to SNVs and indels, respectively, in addition to the fixed LoF substitutions. Including heterozygous variants, Japanese macaque individuals on average have 828 genes with species-specific LoF variants, compared with the rhesus macaque reference genome. Although individuals in the Okazaki population had a higher number of homozygous LoF variants, the total number of genes affected by LoF was significantly lower in this

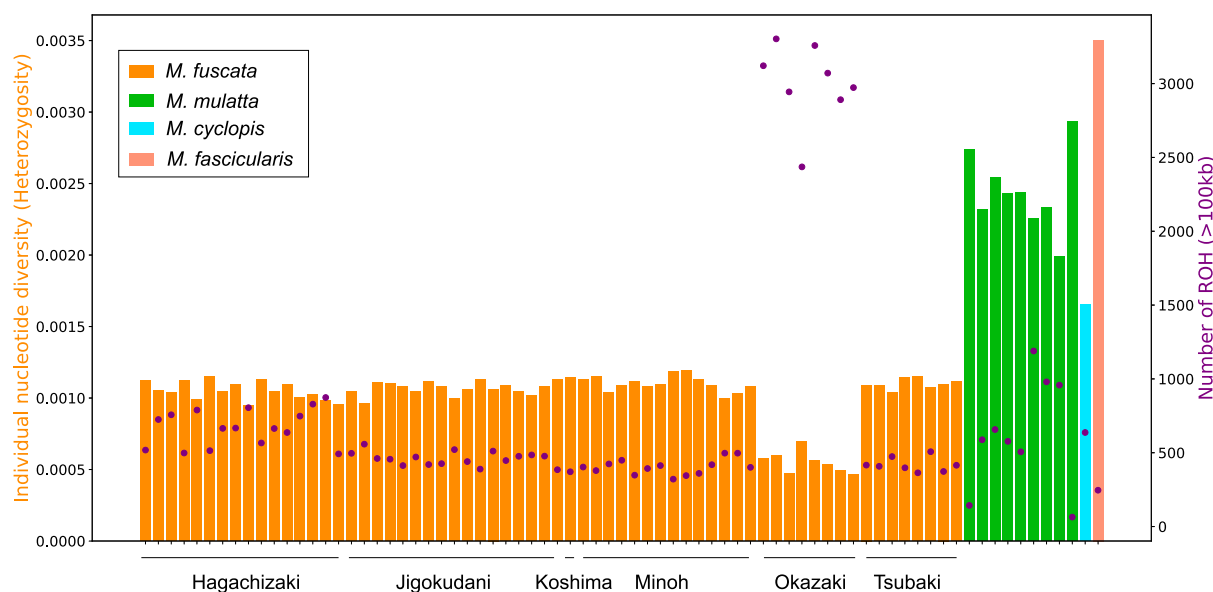


Fig. 2. Individual nucleotide diversity (heterozygosity) and the number of ROH segments >100 kb. Bar heights represent individual nucleotide diversity (left y axis), while circles indicate the number of ROH segments (right y axis).

population compared with the others ($P = 5.52 \times 10^{-6}$, two-sided Mann–Whitney U test).

Highly Differentiated Variants Among Japanese Macaque Populations

To elucidate the genetic underpinnings of population differentiation, we assessed the F_{ST} between populations. F_{ST} values on autosomes between pairwise populations were notably high, spanning from 0.112 to 0.416 across autosomes, with the Okazaki population displaying the highest F_{ST} when compared with others (Table 1). Notably, F_{ST} values on the X chromosomes consistently exceeded those on autosomes, ranging from 0.206 to 0.580. We further evaluated the Q statistics, serving as a relative impact of genetic drift on autosomes versus X chromosomes (Keinan et al. 2009). Under conditions of equal population genetics parameters, such as effective population size and migration rate for both sexes, this value is anticipated to be 0.75. Excluding the highly differentiated Okazaki population, for which Q statistics computation was impractical, Q values were between 0.36 and 0.50.

We subsequently concentrated on highly differentiated variants leading to LoF, which is likely to contribute population-specific phenotypes. To identify population-specific variants, we calculated population-specific F_{ST} within Japanese macaques for each biallelic SNV and short indel. We also focused on Japanese macaque-specific derived SNVs, following the approach used in the LoF analysis. The analysis compared allele frequencies in each target population with those in the background populations of Japanese macaques, excluding individuals within the third degree of kinship. Variants falling within the top 1% of F_{ST} values were operationally designated as population-specific SNVs although they are not necessarily exclusively found in the target population. From these, we selected population-specific LoF variants for each population. We discovered that each population possesses population-specific LoF variants, affecting 34 to 48 genes. We performed functional enrichment analysis using DAVID for these genes. The genes associated with population-specific LoF variants in the Minoh and Tsubaki populations were initially found to be significantly enriched for functions related to Cytochrome Oxidase P450 genes. However, we found that this excess was due to the misassignment of

ENSEMBL Gene IDs to gene symbols in the DAVID database. A comprehensive list of population-specific LoF genes is available in [supplementary table S4, Supplementary Material](#) online.

To verify the existence of population-specific LoF variants that might be associated with unique phenotypic traits in each population, we chose three genes harboring population-specific LoF variants for closer examination through Sanger sequencing. These genes included *TECR*, associated with neuronal disorders in humans (Çalışkan et al. 2011), found in the Hagachizaki population, and *ISG20L2* and *CTSW*, both related to immune functions, identified in the Okazaki population. For each of these three genes, we were able to confirm that the LoF variants occur at high frequencies within their respective populations. Possible phenotypic effects of these mutations are discussed in the “Discussion” section.

Population Structure of Japanese Macaques

We inferred population history of Japanese macaques using various genetic data. The whole mitochondrial genome sequences were initially assembled from the short-read data, and phylogenetic tree was reconstructed. The genealogy of mitochondrial genomes, including outgroup species, is shown in Fig. 3. In the mitochondrial tree, among Japanese macaques, the Koshima individual split first, and then Jigokudani–Hagachizaki and Minoh–Tsubaki–Okazaki split. Although the Okazaki population is highly differentiated from the other populations in the analysis of genome-wide F_{ST} , the mitochondrial genomes of Okazaki population is highly similar to those of Minoh population.

A Neighbor-Net graph constructed from the nuclear genome data is presented in Fig. 4. Although each population formed a single cluster, the pattern was incongruent with the mitochondrial tree. In the nuclear genome network, the Okazaki population was not genetically very close to the Minoh population. In addition, the Koshima individual, which had highly divergent mitochondrial haplotype, showed complex admixture pattern with the other populations.

We also performed principal component analysis (PCA) and admixture analysis for the Japanese macaque populations. In the PCA plot shown in [supplementary fig. S1](#),

Table 1 Pairwise F_{ST} between populations

	Hagachizaki	Jigokudani	Minoh	Tsubaki	Okazaki
Hagachizaki	...	0.112 (0.001)	0.200 (0.001)	0.209 (0.001)	0.416 (0.003)
Jigokudani	0.206 (0.011)	...	0.198 (0.001)	0.209 (0.002)	0.415 (0.003)
Minoh	0.371 (0.013)	0.369 (0.014)	...	0.137 (0.002)	0.336 (0.003)
Tsubaki	0.369 (0.011)	0.370 (0.009)	0.252 (0.011)	...	0.368 (0.003)
Okazaki	0.580 (0.015)	0.580 (0.015)	0.447 (0.020)	0.500 (0.012)	...

Upper right and lower left triangles show the F_{ST} values on autosomes and X chromosomes, respectively. Standard errors are presented in the parentheses.

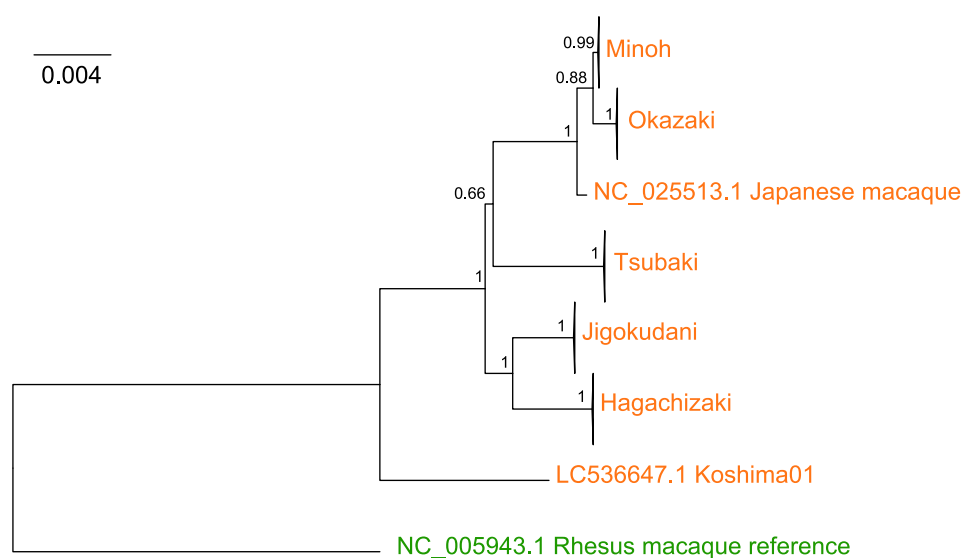


Fig. 3. Phylogenetic tree of mitochondrial genomes. Bootstrap values are shown upon the nodes. Samples in populations are collapsed because all samples in the same population were tightly clustered.

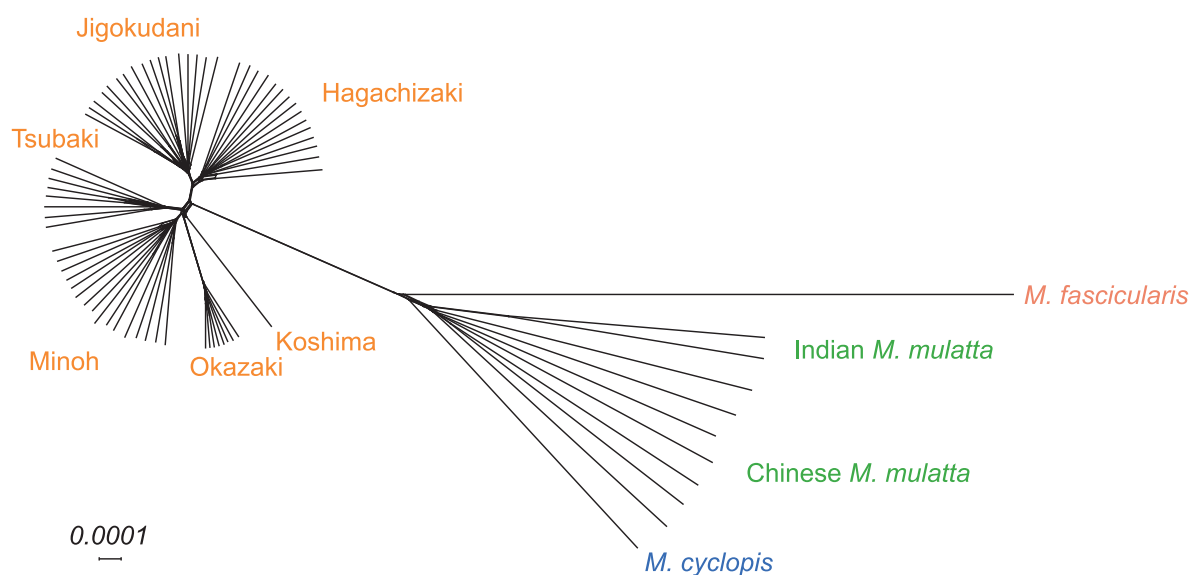


Fig. 4. The Neighbor-Net graph, constructed from nuclear genome data. The genetic distances were calculated for every pair of samples, and the network was subsequently generated using SplitTree software.

Supplementary Material online, the PC1 scores correspond to the difference between the north-east and south-west groups. PC2 scores split the Okazaki population from the Minoh–Tsubaki–Koshima group. The PC3 scores isolate the Koshima individual. The scores of PC4 and PC5 represent the genetic differentiation between Minoh and Tsubaki and between Jigokudani and Hagachizaki, respectively. The admixture plot from $K=2$ to 5 is shown in supplementary fig. S2, Supplementary Material online, where the cross-validation errors were the smallest when

$K=3$ (supplementary fig. S3, Supplementary Material online). The assignment of ancestry was generally in good agreement with the results of Neighbor-Net graph and PCA plot.

Demographic Modeling

We inferred past population size changes of Japanese macaques using a coalescent framework implemented in MSMC2 software for each population (Schiffels and

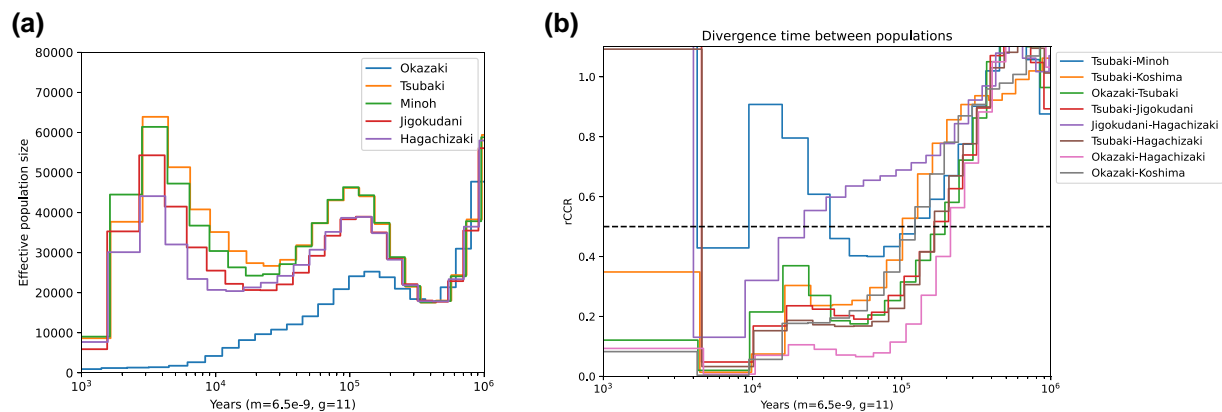


Fig. 5. Coalescent rates over time in Japanese macaques. a) The graph illustrates the changes in effective population size over time for five Japanese macaque populations on the y axis, analyzed through eight phased haplotypes (equivalent to four individuals) per population. b) The graph presents the relative cross coalescent rates rCCRs among these populations on the y axis. In both panels, the x axis indicates time from the present, displayed on a log scale.

Wang 2020). The trajectories of population size change for five populations, assuming the mutation rate of 6.5×10^{-9} per generation and the generation time of 11 years (Itoigawa et al. 1992; Koyama, et al. 1992; Wang et al. 2020), are shown in Fig. 5a. The results suggested that all five populations experienced strong population bottleneck around 400 to 500 kya. After the bottleneck, all populations were inferred to have increased in size, reaching their peak around 100 to 150 kya. The population size of Okazaki appeared to have been decreasing until very recent, but those of other four populations increased again and became maximum around 3 to 4 kya.

The relative cross coalescent rates (rCCRs) are also estimated and presented in Fig. 5b. The pattern shows complex pattern of their divergence. Assuming that rCCR of 0.5 is the divergence time of populations, many population pairs, except for between Jigokudani and Hagachizaki populations, diverged around 150 to 200 kya. The rCCR between the Tsubaki and Minoh populations decreased around the time but started to increase after the dip of rCCR, suggesting the scenario in which these populations have partially merged after the initial divergence. The rCCR between the Jigokudani and Hagachizaki implied that they diverged around 20 to 30 kya, suggesting that they have not diverged until LGM. The rCCR between some pairs increased again after 3 kya. In any population pairs, the rCCR did not monotonically decrease, which appears to indicate that the populations were split and merged multiple times after the initial colonization to the Japanese archipelago.

Genetic Relatedness Among Japanese Macaque Populations

The Neighbor-Net graph and population size trajectories imply complex population history of Japanese macaques. There is a general agreement of the division between

Jigokudani–Hagachizaki and Minoh–Tsubaki groups. However, if we calculated f_4 statistics and evaluated the relationship of quartets inferred by the Neighbor-Net graph, the value of $f_4(\text{Jigokudani, Hagachizaki; Minoh, Tsubaki})$ was significantly deviated from 0 toward the positive direction (Z-score = 3.688). All other quartet topologies showed a significant deviation in f_4 statistics. We next estimated the f_4 statistics using the RMI as outgroup to test the symmetric sister relationships between Jigokudani and Hagachizaki, $f_4(\text{RMI, Minoh/Tsubaki; Jigokudani, Hagachizaki})$ and between Minoh and Tsubaki, $f_4(\text{RMI, Jigokudani/Hagachizaki; Minoh, Tsubaki})$. The f_4 statistics showed bias in all configurations, but greater magnitude of gene flow was suggested for between Jigokudani and Minoh populations and between Hagachizaki and Tsubaki populations (supplementary table S5, Supplementary Material online).

The placement of Okazaki and Kushima populations is unclear from the Neighbor-Net graph. The outgroup f_3 statistics using the RMI at outgroup showed that Kushima individual is most closely related to Okazaki population, and the Okazaki population is most closely related to the Minoh population (supplementary table S6, Supplementary Material online), which is consistent with the pattern observed in the Neighbor-Net graph (Fig. 3).

We also calculated the f_4 statistics to assess the relatedness of Japanese macaques to RMC. Using *M. fascicularis* (CMT) as an outgroup, we computed $f_4(\text{CMT, X; RMC, RMI})$, where X represents different Japanese macaque populations. All Japanese macaque populations exhibited strongly negative f_4 values (supplementary table S7, Supplementary Material online), suggesting ancient gene flow between RMC and Japanese macaque lineages (Osada et al. 2021). We then tested the heterogeneity in the relatedness to RMC among Japanese macaque populations by calculating $f_4(\text{RMI, RMC; Y, Z})$, where Y and Z represent all possible pairs of Japanese macaque populations. (supplementary table S8,

Supplementary Material online). The results showed that Koshima individual is more distantly related to RMC compared with the other Japanese macaques.

Discussion

In this study, we determined high-coverage whole-genome sequences of 64 Japanese macaque individuals derived from 5 population sources reared in NBRP. These samples include four trios, and we performed haplotype phasing using the information. The information provides valuable insight into genetic diversity of Japanese macaques.

Low Genetic Diversity of Japanese Macaques

The entire samples sequenced in this study are wild-derived animals: i.e. founders of colonies maintained in NBRP. Although they have been fed by humans for a few decades and some samples were identified as closely related, we showed that genome-wide individual nucleotide diversity was almost constant within populations, indicating that they were not strongly inbred individuals. The genome-wide individual nucleotide diversity is around 0.1%, showing that Japanese macaques harbor similar genetic diversity with the human population and the values are much smaller than other species in the *mulatta* group macaques (Fig. 2).

The nucleotide diversity in the Okazaki population was around a half of that in the other Japanese macaques. The population size trajectory of Okazaki population presented in Fig. 5a shows that the decline of population size of the Okazaki population has probably continued since 100 to 150 kya, indicating that the low genetic diversity of the Okazaki population is not simply due to recent inbreeding or population bottleneck. Notably, the nucleotide diversity of the Okazaki population is one of the smallest values among wild primates (Osada 2015). Individuals in the Okazaki population also possessed extensive numbers and sizes of ROH, probably owing to recent population size reduction. However, we found that the number of genes affected by LoF variants was smaller in the Okazaki population than in the other populations. This is because individuals in the Okazaki population have fewer heterozygous LoF variants on average, suggesting that strongly deleterious LoF variants have already been purged to some extent in this population.

The divergence time of Japanese macaque lineages from the most recent common ancestors with rhesus macaques was estimated to be around 1.0 to 2.3 Ma using a whole-genome sequence data (Kuderna et al. 2023; Zhang et al. 2023; Zhou et al. 2024). All Japanese macaque populations shared the signature of strong population bottleneck around 400 to 500 kya. Since the oldest fossil of Japanese macaques was found in a layer dated around 430 to 630 kya (Aimi 2002), the bottleneck may

correspond to the founding event of the colonization of Japanese macaques to the Japanese archipelago. The strong population bottleneck and long history of partial isolation would have increased the genetic load in Japanese macaques. Each Japanese macaque individual carries ~120 species-specific unfixed LoF variants as homozygotes. Given that Icelandic humans, which have also experienced strong population bottleneck during migration, harbor on average 21 homozygous LoF variants including SNVs and indels (Gudbjartsson et al. 2015), the number of LoF variants carried by Japanese macaques are substantial, propounding significant impact on their phenotypes.

Population Differentiation Among Japanese Macaques

Population differentiation among the Japanese macaque populations, as measured by F_{ST} , was relatively high, particularly given that these macaques are restricted to the Japanese archipelago. This elevated F_{ST} is partly attributed to the small effective population size of each group and a history of isolation. The least differentiated pair, Hagachizaki and Jigokudani, showed an F_{ST} value of 0.112, comparable with the differentiation observed among African, European, and Asian human populations (Bhatia et al. 2013). The high level of differentiation observed in the Okazaki population is similar to that reported in golden snub-nosed monkeys (*Rhinopithecus roxellana*), an endangered species living in fragmented inland habitats (Zhou et al. 2016).

The Q statistic, representing the relative impact of genetic drift on autosomes vs X chromosomes (Keinan et al. 2009), was below 0.75 in all pairwise comparisons, except for those where the statistic could not be calculated due to high F_{ST} . This pattern suggests a scenario in which males migrated more frequently than females between populations, consistent with observations of female philopatry in Japanese macaques (Sprague et al. 1998).

Functional Differentiation in Proteins

Although many genes were specifically lost in a population, we investigated more deeply a part of candidate genes that would be the basis of population-specific phenotypes. We focused on behavior-related genes in the Hagachizaki population and immune system-related genes in the Okazaki population, as these traits are characteristic features of the colonies. It should be noted that these genes were arbitrarily selected for validation of our short-read sequencing, and many other candidate genes were also identified (supplementary table S3, Supplementary Material online). The Hagachizaki-specific LoF mutation in *TECR* (ENSMUT00000055868.2) is a substitution from C to T, changing the codon for glutamine to stop codon. The premature stop codon reduces the length of intact protein from 374 to 262 aa long. *TECR*, a membrane protein

involved in fatty acid elongation, is associated with mental retardation in humans (Çalışkan et al. 2011). This gene would be an intriguing candidate for characterizing the phenotypes of the Hagachizaki population, given their lack of fear toward humans. However, according to the UniProt database, the C-terminal region of protein (A0A1D5QQQ1) does not form rigid structure, and the mutation only affects the amino acid sequences of one transcript variant among three possible variants. In addition, the stop codon–gained transcript isoform was only annotated in the Ensembl database and not in the GenBank database, of *M. mulatta*, and it may result from an error in the database annotation. The mutation is almost exclusively found in the Hagachizaki population. Among 15 individuals we sequenced, 3 and 11 individuals were homozygous and heterozygous for the mutation, respectively. Only one individual in the Hagachizaki population was a complete wild type. On the other hand, only one individual in the Minoh population carries the mutation as heterozygous state in our dataset.

The frameshift-causing deletions in the *CTSW* (ENSMMUG00000008649) and *ISG20L2* (ENSMMUG00000000037) genes were found exclusively in the Okazaki population. Within this population, only one individual was heterozygous for the deletions, while seven were homozygous for each gene. The *CTSW* gene is encoding Cathepsin W, a cysteine proteinase that is selectively expressed in CD8⁺ T cells and natural killer cells, which are crucial for immune responses. The mutation specific to the Okazaki population leads to a frameshift at Q40X. Given the absence of an alternative start codon immediately following the frameshift site, it is probable that the Cathepsin W protein is nonexistent in Okazaki individuals, potentially influencing the immune characteristics of this group. Another candidate gene *ISG20L2* encodes an exoribonuclease that is activated by interferon and plays a significant role in immunity. Research by Rodríguez-Galán et al. (2023) has demonstrated that the lack of this gene can amplify the activation of T cells in cultured human cells. The deletion found in the Okazaki population affects the C-terminal region of the protein, removing sites essential for catalytic activity and likely leading to phenotypic differences.

Population History of Japanese Macaques

The pattern of differentiation among six populations including Koshima individual implied a highly complex population history. The initial divergence among populations probably occurred around 150 to 200 kya between north-east and south-west groups. The estimated divergence timing within the Japanese archipelago is quite deep, considering that the split of RMI and RMC was estimated to have occurred around a similar time frame (Hernandez

et al. 2007; Zhou et al. 2024), which overlaps with the penultimate glacial period. Furthermore, the trajectories of rCCR suggested that the process of divergence was not simple, and probably there have been multiple split and merge events following to their initial divergence. The gene flow between populations in the nuclear genomes was also inferred by Ito et al. (2021). In many species pairs, there were increases of rCCR ~10 to 30 kya, when all population increased their within-population coalescent rates (a signature of population size decrease). If the shrink of population sizes is caused by habitat fragmentation, rCCR should decrease but the rates rather increased in our dataset, particularly between Tsubaki and Mihoh populations. The pattern therefore suggests that their habitats during the shrinkage of native range might have been interconnected or overlapped. The timing may correspond to the LGM, when the habitats of Japanese macaques might have been restricted to a few refugia along the southern coastal areas (Ito et al. 2021). Altogether, they may have expanded their range and start genetically diverging during interglacial periods with the isolation-by-distance mode and contracted the range and partly merge with the other populations during the glacial periods.

Clear genetic distinctions were identified between the Jigokudani–Hagachizaki and Minoh–Tsubaki–Okazaki populations. This differentiation echoes findings by Ito et al. (2021), which reported significant genetic divergence of populations from Shimokita—the northernmost point of Honshu—, Yamagata, and Gunma, compared with other groups. However, the Jigokudani and Hagachizaki populations are geographically located near the boundary of the north-east and south-west division. To ascertain whether the Jigokudani and Hagachizaki populations are part of the north-east group defined by Ito et al. (2021), we integrated our dataset with the ddRAD-seq data and constructed a Neighbor-Net graph, presented in [supplementary fig. S4, Supplementary Material](#) online. The integrated analysis confirmed the classification of the Jigokudani and Hagachizaki populations with the Shimokita, Yamagata, and Gunma populations, hence assigning them to the north-east group. Moreover, the Okazaki population aligned with the western Japanese Kansai area cluster, which is in agreement with the patterns observed in our whole-genome sequence analysis.

Conclusion

The research delivers high-coverage whole-genome sequences for 64 Japanese macaques, shedding light on the genetic underpinnings of the species, with significant implications for both evolutionary biology and biomedical research. Our findings reveal well-differentiated local populations of Japanese macaques, each with a complex and dynamic evolutionary history involving multiple splits and

merges over the past 500,000 years. The prolonged presence in the Japanese archipelago, coupled with their small effective population sizes, has resulted in each group harboring numerous unique genetic variants. Some of these variants have considerable effects on individual phenotypes and may underlie observed differences between populations. The genomic insights from this study pave the way for future research on these uniquely fascinating non-human primates.

Materials and Methods

Sample Preparation and Genome Sequencing

Whole blood cells for genomic DNA were obtained from 5 populations of macaques: the Okazaki (8 individuals), Jigokudani (17 individuals), Hagachizaki (15 individuals), Minoh (16 individuals), and Tsubaki (8 individuals) populations (Fig. 1; [supplementary table S1, Supplementary Material](#) online). They were housed at the Center for the Evolutionary Origins of Human Behavior (EHUB), Kyoto University, Inuyama, Aichi, Japan, in accordance with Guidelines for Care and Use of Nonhuman Primates (Ver. 3, 2010, Primate Research Institute, Kyoto University and Guidelines for the Captive Maintenance, Storage, and Utilization of Japanese Macaques 2018, NBRP).

Blood collection was conducted at the EHUB in accordance with the guidelines of the Laboratory Biosafety Manual, World Health Organization. Genomic DNA was isolated from 2 mL of peripheral blood with EDTA using a Qiagen Genomic DNA purification kit (Qiagen K. K., Tokyo, Japan). The isolated DNA samples were kept at -30°C until use. The protocol was approved by the Animal Welfare and Animal Care Committee, the Center for the EHUB, Kyoto University (permission no. 2022-104, 2023-148, 2024-019).

Genome sequencing was performed using the DNBSEQ-G400 (MGI). Mate-pair (150 bp \times 2) libraries were generated using the macaque genomic DNA. Paired-end libraries of 350 bp insert sizes were prepared. Library preparations and all DNBSEQ runs were performed as per the standard manufacturer's protocols. Demultiplexing was performed by the DNBSEQ built-in software, and the raw data of each individual were subsequently retrieved by means of specific barcodes. All sequence reads have been deposited to the DDBJ database of the National Institute of Genetics, accession number DRA016173.

Variant Calling and Filtering

In addition to the newly sequenced data, we also downloaded whole-genome sequence data for *M. fascicularis*, *M. mulatta*, *M. cyclopis*, and *M. fuscata* from public databases ([supplementary table S2, Supplementary Material](#) online) (Xue et al. 2016; Liu et al. 2018; Osada et al. 2021).

Paired-end reads were aligned to the reference genome sequence (Mmul_10) using the Parabricks 4.0 fq2bam pipeline (Alser et al. 2020), which also sorted the reads and marked polymerase chain reaction duplicates. Initial genotyping was conducted with GATK HaplotypeCaller as implemented in Parabricks 4.0 (Van der Auwera and O'Connor 2020), followed by joint calling using GATK 4.4 GenomicsDBImport and GenotypeGVCFs. Hard filtering criteria for SNVs and indels were set as follows: for SNVs, "QD < 2.0, QUAL < 30.0, SQR > 3.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, ExcessHet > 30.0"; for indels, "QD < 2.0, QUAL < 30.0, FS > 200.0, ReadPosRankSum < -20.0, and ExcessHet > 30.0." SNVs and indels were annotated using SNPEff 5.1 with the Mmul_10.105 annotation (Cingolani et al. 2012).

We further applied mappability and accessibility filters on the SNVs. The mappability score for each genomic position was calculated using GenMap software (Pockrandt et al. 2020). We specifically selected sites with a (30, 2)-mappability score of 1. This indicates that a 30-mer sequence starting from the filter-passed site is unique within the genome, even when allowing for two mismatches. To assess genome accessibility, we evaluated the read coverage at each site across all individuals. We determined the mode of the coverage distribution, and sites falling within the range of half to twice the mode value were classified as accessible, corresponding to 17.5- to 67.5-fold coverages per individual genome. Autosomes encompassed 1,930,114,258 sites in mappable and accessible regions. We also genotyped non-pseudoautosomal region (PAR) of X chromosomes. The non-PAR was assigned to chrX:2,364,962-151,635,290. Male and females were separately genotyped, and the variants were merged after the genotyping. The hardfiltering and mappability filtering were applied with the same criteria for autosomes. For the accessibility filter, we adjusted the coverage range calculated using autosomal data, assuming that the X chromosomes of males and females are haploid and diploid, respectively. We used BCFtools to identify regions of ROH (Narasimhan et al. 2016). Regions with Phred quality scores below 30 or shorter than 100 kb were filtered out.

Haplotype Phasing

We first phased haplotypes with four trio groups, focusing exclusively on autosomal SNVs. Each trio sample was phased using WhatsHap 1.7, utilizing trio-based information (Martin et al. 2016). These phased haplotypes from the trios were then employed as a reference panel to help the statistical phasing of all Japanese macaque individuals. The statistical phasing was executed using SHAPEIT 5.1.1, conforming to the standard procedures outlined on the SHAPEIT5 website (Hofmeister et al. 2023). This included steps such as chunking, ligating, and phasing common

and rare variants. Prior to being processed with SHAPEIT, each non trio sample underwent initial phasing using WhatsHap 1.7. For the non-PAR of X chromosomal data, we only used SHAPEIT for haplotype phasing.

Analysis of Population Structures

We utilized only biallelic SNVs for studying genetic structures. The filtered biallelic autosomal SNVs in the VCF file were converted into the PLINK 1.9 file format for population genetics analyses (Purcell et al. 2007). After format conversion, individuals with a kinship equal or higher than the third degree were excluded using PLINK 2.0's "--king-cutoff 0.0884" option (Chang et al. 2015), resulting in the removal of 20 individuals (marked in [supplementary table S1, Supplementary Material](#) online). SNVs under strong linkage disequilibrium (LD) were pruned using PLINK 1.9 with "--indep-pairwise 50 5 0.5." Subsequent PCA was also performed with PLINK 1.9. We used the LD-pruned dataset to run ADMIXTURE (Alexander et al. 2009). The genetic distance between individuals was estimated by calculating the genotype distance, where heterozygous sites for both individuals were assigned a value of 0.5. These data were then used with SplitTree4 for reconstructing a neighbor-joining tree and a Neighbor-Net graph (Huson and Bryant 2006). Admixtools was employed to estimate f_3 and f_4 statistics using the non-LD-pruned SNV dataset. (Patterson et al. 2012). F_{ST} for each SNV was estimated using Hudson's method implemented in PLINK 2.0. Genome-wide F_{ST} was calculated by averaging the numerators and denominators, followed by taking the ratio (Bhatia et al. 2013). Confidence intervals were estimated using a jackknife block size of 25,000 SNVs, roughly corresponding to 5 cM in our dataset. For the analysis of LoF SNVs with high F_{ST} , we considered variants that were successfully called in all unrelated Japanese macaque samples. Functional enrichment analysis of highly differentiated LoF mutations was performed using DAVID (Jiao et al. 2012). Enrichment in the three gene ontology categories (biological process, cellular component, and molecular function) was assessed using ENSEMBLE Gene IDs as queries, with a false discovery rate threshold of 0.05.

Demographic Inference

The inference of demography was performed using MSMC2 (Schiffels and Wang 2020). In order to estimate effective population sizes of each population of Japanese macaque, for each population eight phased autosomal haplotypes from four unrelated individuals were used ([supplementary table S9, Supplementary Material](#) online). The population of Koshima was not analyzed because it contains only single individual. Mutation rate of 6.5×10^{-9} was used (Wang et al. 2020), assuming the generation time of 11 years (Itoigawa et al. 1992; Koyama et al. 1992).

The Sequencing of Variants Within the *TECR*, *CTSW*, and *ISG20L2* Regions

The sequencing was based on the Sanger technique and the BigDye Terminator V3.1 (Life Technologies, Carlsbad, CA, USA) was used. The primer sequences used for sequencing of *TECR*, *CTSW*, and *ISG20L2* regions are listed in [supplementary table S10, Supplementary Material](#) online. The sequencing protocol was as follows: 30 cycles of 10 s at 98 °C, 30 s at 55 °C, and 60 s at 72 °C, and finally the sequences were maintained at 4 °C. The DNA sequence was determined using the ABI3730x Genetic Analyzer (Thermo Fisher Scientific-Applied Biosystems, Foster City, CA, USA). The sequences obtained were read and aligned with the MEGA X (Knyaz et al. 2018).

Analysis of ddRAD-seq Data

The ddRAD-seq data from 88 Japanese macaques, as reported by Ito et al. (2021), were retrieved from the NCBI sequence read archive database. A comprehensive list of the samples analyzed in this study can be found in [supplementary table S11, Supplementary Material](#) online. Note that these combined data were exclusively used for generating [supplementary fig. S4, Supplementary Material](#) online. We aligned the short reads to the rhesus macaque Mmul_10 reference genome using the Parabricks 4.0 software, consistent with the procedures applied to whole-genome sequencing data. SNVs were genotyped employing the Stacks software pipeline (Catchen et al. 2013), and the resulting variant data were integrated with the whole-genome sequence data utilizing the BCFtools program. Prior to the integration, the orientation of reference and alternative alleles, as outputted by Stacks, was normalized using BCFtools. We included only those autosomal sites where more than 80% of the samples were successfully genotyped, assuring that these sites were variant sites in both datasets. The combined dataset was used to construct a Neighbor-Net graph, applying the method consistent with that used for the whole-genome sequence data analysis.

Supplementary Material

[Supplementary material](#) is available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by the Genome Information Upgrading Program of the National BioResource Project of MEXT. Blood samples were provided by the NBRP Japanese macaques, which is part of the NBRP of the MEXT. We are grateful to Miho Hakukawa, Miyuki Ido, and Mayumi Morimoto for their invaluable assistance

with the collection of blood samples. We used ChatGPT-4 software for sentence-by-sentence proofreading of the manuscript and declare that our original intention has not been modified by the use of artificial intelligence.

Data Availability

Raw sequencing reads are available in the DDBJ database of the National Institute of Genetics, accession number DRA016173.

Literature Cited

- Aimi M. The oldest fossil macaque from Japan (in Japanese). *Primate Res.* 2002;18(2):239–245. <https://doi.org/10.2354/psj.18.239>.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;19(9):1655–1664. <https://doi.org/10.1101/gr.094052.109>.
- Alser M, Bingöl Z, Cali DS, Kim J, Ghose S, Alkan C, Mutlu O. Accelerating genome analysis: a primer on an ongoing journey. *IEEE Micro.* 2020;40(5):65–75. <https://doi.org/10.1109/MM.2020.3013728>.
- Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P, Chen Y, Felkel S, Hallast P, Kamm J, et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science.* 2020;367(6484):eaay5012. <https://doi.org/10.1126/science.aay5012>.
- Bhatia G, Patterson N, Sankararaman S, Price AL. Estimating and interpreting F_{ST} : the impact of rare variants. *Genome Res.* 2013;23(9):1514–1521. <https://doi.org/10.1101/gr.154831.113>.
- Çalışkan M, Chong JX, Uricchio L, Anderson R, Chen P, Sougnéz C, Garimella K, Gabriel SB, dePristo MA, Shakir K, et al. Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the *TECR* gene on chromosome 19p13. *Hum Mol Genet.* 2011;20(7):1285–1289. <https://doi.org/10.1093/hmg/ddq569>.
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol.* 2013;22(11):3124–3140. <https://doi.org/10.1111/mec.12354>.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience.* 2015;4(1):7. <https://doi.org/10.1186/s13742-015-0047-8>.
- Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* 2012;6(2):80–92. <https://doi.org/10.4161/fly.19695>.
- Cooper A, Turney C, Hughen KA, Brook BW, McDonald HG, Bradshaw CJ. Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. *Science.* 2015;349(6248):602–606. <https://doi.org/10.1126/science.aac4315>.
- Feng Y, Comes HP, Chen J, Zhu S, Lu R, Zhang X, Li P, Qiu J, Olsen KM, Qiu Y. Genome sequences and population genomics provide insights into the demographic history, inbreeding, and mutation load of two ‘living fossil’ tree species of Dipteronia. *Plant J.* 2024;117(1):177–192. <https://doi.org/10.1111/tj.16486>.
- Gudbjartsson DF, Helgason H, Gudjonsson SA, Zink F, Oddson A, Gylfason A, Besenbacher S, Magnusson G, Halldorsson BV, Hjartarson E, et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat Genet.* 2015;47(5):435–444. <https://doi.org/10.1038/ng.3247>.
- Haffer J. Speciation in amazonian forest birds. *Science.* 1969;165(3889):131–137. <https://doi.org/10.1126/science.165.3889.131>.
- Hernandez RD, Hubisz MJ, Wheeler DA, Smith DG, Ferguson B, Rogers J, Nazareth L, Indap A, Bourquin T, McPherson J, et al. Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques. *Science.* 2007;316(5822):240–243. <https://doi.org/10.1126/science.1140462>.
- Hofmeister RJ, Ribeiro DM, Rubinacci S, Delaneau O. Accurate rare variant phasing of whole-genome and whole-exome sequencing data in the UK Biobank. *Nat Genet.* 2023;55(7):1243–1249. <https://doi.org/10.1038/s41588-023-01415-w>.
- Hubisz MJ, Williams AL, Siepel A. Mapping gene flow between ancient hominins through demography-aware inference of the ancestral recombination graph. *PLoS Genet.* 2020;16(8):e1008895. <https://doi.org/10.1371/journal.pgen.1008895>.
- Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* 2006;23(2):254–267. <https://doi.org/10.1093/molbev/msj030>.
- Isa T, Yamane I, Hamai M, Inagaki H. Japanese macaques as laboratory animals. *Exp Anim.* 2009;58(5):451–457. <https://doi.org/10.1538/expanim.58.451>.
- Ito T, Hayakawa T, Suzuki-Hashido N, Hamada Y, Kurihara Y, Hanya G, Kaneko A, Natsume T, Aisu S, Honda T, et al. Phylogeographic history of Japanese macaques. *J Biogeogr.* 2021;48(6):1420–1431. <https://doi.org/10.1111/jbi.14087>.
- Itoigawa N, Tanaka T, Ukai N, Fujii H, Kurokawa T, Koyama T, Ando A, Watanabe Y, Imakawa S. Demography and reproductive parameters of a free-ranging group of Japanese macaques (*Macaca fuscata*) at Katsuyama. *Primates.* 1992;33(1):49–68. <https://doi.org/10.1007/BF02382762>.
- Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics.* 2012;28(13):1805–1806. <https://doi.org/10.1093/bioinformatics/bts251>.
- Kawamoto Y. Modes of differentiation in Japanese macaques: perspectives from population genetics. In: Nakagawa N, Nakamichi M, Sugiura H, editors. *The Japanese macaques*. Tokyo: Springer Japan; 2010. p. 53–76.
- Kawamoto Y, Shotake T, Nozawa K, Kawamoto S, Tomari K, Kawai S, Shirai K, Morimitsu Y, Takagi N, Akaza H, et al. Postglacial population expansion of Japanese macaques (*Macaca fuscata*) inferred from mitochondrial DNA phylogeography. *Primates.* 2007;48(1):27–40. <https://doi.org/10.1007/s10329-006-0013-2>.
- Kawamoto Y, Tomari K-i, Kawai S, Kawamoto S. Genetics of the Shimokita macaque population suggest an ancient bottleneck. *Primates.* 2008;49(1):32–40. <https://doi.org/10.1007/s10329-007-0057-y>.
- Keinan A, Mullikin JC, Patterson N, Reich D. Accelerated genetic drift on chromosome X during the human dispersal out of Africa. *Nat Genet.* 2009;41(1):66–70. <https://doi.org/10.1038/ng.303>.
- Knyaz C, Stecher G, Li M, Kumar S, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018;35(6):1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Koyama N, Takahata Y, Huffman M, Norikoshi K, Suzuki H. Reproductive parameters of female Japanese macaques: thirty years data from the Arashiyama troops, Japan. *Primates.* 1992;33(1):33–47. <https://doi.org/10.1007/BF02382761>.
- Kuderna LFK, Gao H, Janiak MC, Kuhlwlilm M, Orkin JD, Bataillon T, Manu S, Valenzuela A, Bergman J, Rousselle M, et al. A global catalog of whole-genome diversity from 233 primate species. *Science.* 2023;380(6648):906–913. <https://doi.org/10.1126/science.abn7829>.
- Liu Z, Tan X, Orozco-terWengel P, Zhou X, Zhang L, Tian S, Yan Z, Xu H, Ren B, Zhang P, et al. Population genomics of wild Chinese rhesus

- macaques reveals a dynamic demographic history and local adaptation, with implications for biomedical research. *GigaScience*. 2018;7(9):giy106. <https://doi.org/10.1093/gigascience/giy106>.
- Lorenzana GP, Figueiró HV, Kaelin CB, Barsh GS, Johnson J, Karlsson E, Morato RG, Sana DA, Cullen L, May JA Jr, et al. Whole-genome sequences shed light on the demographic history and contemporary genetic erosion of free-ranging jaguar (*Panthera onca*) populations. *J Genet Genomics*. 2022;49(1):77–80. <https://doi.org/10.1016/j.jgg.2021.10.006>.
- Lozier JD, Strange JP, Heraghty SD. Whole genome demographic models indicate divergent effective population size histories shape contemporary genetic diversity gradients in a montane bumble bee. *Ecol Evol*. 2023;13(2):e9778. <https://doi.org/10.1002/ece3.9778>.
- Martin M, Patterson M, Garg S, Fischer SO, Pisanti N, Klau GW, Schöenhuth A, Marschall T. WhatsHap: fast and accurate read-based phasing. *bioRxiv* 085050. <https://doi.org/10.1101/085050>, 14 November 2016, preprint: not peer reviewed.
- Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. BCFtools/ROH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*. 2016;32(11):1749–1751. <https://doi.org/10.1093/bioinformatics/btw044>.
- Osada N. Genetic diversity in humans and non-human primates and its evolutionary consequences. *Genes Genet Syst*. 2015;90(3):133–145. <https://doi.org/10.1266/ggs.90.133>.
- Osada N, Matsudaira K, Hamada Y, Malaivijitnond S. Testing sex-biased admixture origin of macaque species using autosomal and X-chromosomal genomic sequences. *Genome Biol Evol*. 2021;13(1):evaa209. <https://doi.org/10.1093/gbe/evaa209>.
- Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. Ancient admixture in human history. *Genetics*. 2012;192(3):1065–1093. <https://doi.org/10.1534/genetics.112.145037>.
- Pockrandt C, Alzamel M, Iliopoulos CS, Reinert K. GenMap: ultra-fast computation of genome mappability. *Bioinformatics*. 2020;36(12):3687–3692. <https://doi.org/10.1093/bioinformatics/btaa222>.
- Provan J, Bennett KD. Phylogeographic insights into cryptic glacial refugia. *Trends Ecol Evol*. 2008;23(10):564–571. <https://doi.org/10.1016/j.tree.2008.06.010>.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575. doi: <https://doi.org/10.1086/519795>.
- Ragsdale AP, Weaver TD, Atkinson EG, Hoal EG, Möller M, Henn BM, Gravel S. A weakly structured stem for human origins in Africa. *Nature*. 2023;617(7962):755–763. <https://doi.org/10.1038/s41586-023-06055-y>.
- Rodríguez-Galán A, Dosil SG, Hrková A, Fernández-Messina L, Feketová Z, Pokorná J, Fernández-Delgado I, Camafeita E, Gómez MJ, Ramírez-Huesca M, et al. ISG20L2: an RNA nuclease regulating T cell activation. *Cell Mol Life Sci*. 2023;80(9):273. <https://doi.org/10.1007/s00018-023-04925-2>.
- Schiffels S, Wang K. MSMC and MSMC2: the multiple sequentially Markovian coalescent. In: Duthiel JY, editor. *Statistical population genomics*. New York (NY): Springer US; 2020. p. 147–166.
- Sprague DS, Suzuki S, Takahashi H, Sato S. Male life history in natural populations of Japanese macaques: migration, dominance rank, and troop participation of males in two habitats. *Primates*. 1998;39(3):351–363. <https://doi.org/10.1007/Bf02573083>.
- Van der Auwera GA, O'Connor BD. *Genomics in the cloud: using Docker, GATK, and WDL in Terra*. California (CA): O'Reilly Media; 2020.
- Wang RJ, Thomas GWC, Raveendran M, Harris RA, Doddapaneni H, Muzny DM, Capitanio JP, Radivojac P, Rogers J, Hahn MW. Paternal age in rhesus macaques is positively associated with germline mutation accumulation but not with measures of offspring sociability. *Genome Res*. 2020;30(6):826–834. <https://doi.org/10.1101/gr.255174.119>.
- Willis KJ, Bennett KD, Walker D, Hewitt GM. Genetic consequences of climatic oscillations in the quaternary. *Philos Trans R Soc Lond B Biol Sci*. 2004;359(1442):183–195. doi: [10.1098/rstb.2003.1388](https://doi.org/10.1098/rstb.2003.1388).
- Xue C, Raveendran M, Harris RA, Fawcett GL, Liu X, White S, Dahdouli M, Rio Deiros D, Below JE, Salerno W, et al. The population genomics of rhesus macaques (*Macaca mulatta*) based on whole genome sequences. *Genome Res*. 2016;26(12):1651–1662. <https://doi.org/10.1101/gr.204255.116>.
- Zhang BL, Chen W, Wang Z, Pang W, Luo MT, Wang S, Shao Y, He WQ, Deng Y, Zhou L, et al. Comparative genomics reveals the hybrid origin of a macaque group. *Sci Adv*. 2023;9(22):eadd3580. <https://doi.org/10.1126/sciadv.add3580>.
- Zhou X, Meng X, Liu Z, Chang J, Wang B, Li M, Wengel PO, Tian S, Wen C, Wang Z, et al. Population genomics reveals low genetic diversity and adaptation to hypoxia in snub-nosed monkeys. *Mol Biol Evol*. 2016;33(10):2670–2681. <https://doi.org/10.1093/molbev/msw150>.
- Zhou Y, Tian J, Jiang H, Han M, Wang Y, Lu J. Phylogeography and demographic history of macaques, *fascicularis* species group, in East Asia: inferred from multiple genomic markers. *Mol Phylogenet Evol*. 2024;194:108042. <https://doi.org/10.1016/j.ympev.2024.108042>.
- Zinner D, Fickenscher GH, Roos C. Family Cercopithecidae (Old World monkeys). In: Mittermeier RA, Rylands AB, Wilson DE, editors. *Handbook of the mammals of the world*. Barcelona: Lynx Edicions; 2013. p. 550–627.

Associate editor: Aida Andres