Doctoral Dissertation

Genetic variation and diversity in wild small mammals: an analysis on a mutation in coat color-determining gene of Norway rat (*Rattus norvegicus*) and phylogeography in Japanese moles

Takeru Tsunoi

Course of Bioengineering and Bioinformatics,

Graduate School of Information Science and Technology, Hokkaido University

Acknowledgements

I would like to express my sincere gratitude to my advisors, Associate Professor Naoki Osada and Professor Emeritus Hitoshi Suzuki for their guidance and support. I would like to thank Gohta Kinoshita, Jun J. Sato, and Yasuhiro Go for their experimental guidance. I also thank Kimi Araki and the laboratory members for their help in mice breeding, conducting experiments, and providing samples. I also thank Koki Noju and Reiko Mitsuhashi for their assistance in analysis and providing data, Shoji Tatsumoto for his cooperation in genomic assembly and construction, and Kimiyuki Tsuchiya, Masashi Harada, Alexey P. Kryukov, Han Sang-Hoon, Masaru Notsu, Takeshi Eto, Shin-ichiro Kawada, Shuji Yachimori, Shinji Yano, Tetsuji Hosoda, and Hajime Fukaya for their help with sample collection. I would also like to thank Hiroe Ishikawa, Chika Usui, and Kyoko Noguchi for their experimental assistance, and Masaru Kato for his efforts in assigning tissue and DNA museum codes. I would also like to express my gratitude to my laboratory members and my family for their support and encouragement.

Table of Contents

Acknowledgements 2

Table of Contents 3

Abstract 6

List of Abbreviations 9

General Introduction 12

PART I: An *Mc1r* mutation responsible for a yellowish color variant of Norway rat (*Rattus norvegicus*) and insights into mammalian *MC1R* variants 18

Introduction 18

Materials and Methods 19

Study animals 19

DNA amplification and sequencing 19

Coat color measurements 20

Results 20

Microscopic observations 20

Gene sequencing 21

Coat color measurements 21

Discussion 21

Phenotypic change due to a severe loss-of-function mutation in *Mc1r* 21

Pigment production and body color pattern formation in *Mc1r*-deficient mutants 23

Figures and Legends 25

References 29

PART II: Evolutionary history of the relict populations of the lesser Japanese mole (*Mogera imaizumii*) inferred from mitochondrial cytochrome *b* gene sequences 36

Introduction 36

Materials and Methods 38

Materials 38

Determination of mitochondrial *Cytb* sequences 38

Phylogenetic analyses 39

Results 39

Discussion 41

Interspecific phylogenetic history of moles of the Japanese Archipelago 41

Impact of Quaternary climate changes and interspecies competitive exclusion 42

Conclusions 43

Figures and Legends 45

References 59

PART III: Insights into the population structure of the lesser Japanese mole (*Mogera imaizumii*) and the large Japanese mole (*M. wogura*) revealed by reduced representation genome analysis 63

Introduction 63

Materials and Methods 65

Whole genome sequencing and assembling the draft genome 65

MIG-seq library preparation and SNP data generation 66

Analysis of population structure 67

Results 68

Genomic data assembly 68

Genetic structure of *M. imaizumii* 68
Genetic structure of *M. wogura* 69
Discussion 70
Population structures of *M. imaizumii* 70
Population structure of *M. wogura* 71
Common genetic boundary in the Kinki area observed in both *Mogera imaizumii* and *M. wogura* 72
Appendix 73
Figures and Legends 80
References 92
PART IV: Whole-genome sequencing analysis of *Mogera imaizumii* and *M. wogura* 97
Introduction 97
Materials and Methods 97
Reference genome construction 97
Whole genome sequencing 99
Estimation of population divergence times and admixture between populations 101
Inference of past demography 101
Analyses on the natural selection on genes 101
Results 102
Discussion 103
References 132

Abstract

Animals harbor a wide range of phenotypic and genotypic polymorphisms. They are shaped under the influence of biotic/abiotic factors and offer significant insights into the genetic mechanisms of the animals' evolution and adaptation. Investigating the genetic and evolutionary background of this biodiversity is not only one of the most important topics in evolutionary biology, but also a crucial step toward understanding and improving the interactions between humans and their surrounding natural environment. In wild populations, where natural selection, genetic drift, and gene flow interact in complex ways, this is a topic of great significance. This thesis looks into the topic by focusing on two distinct, yet complementary sets of studies: 1) examination of the genetic basis of a coat color variant occurrence in a wild Norway rat (*Rattus norvegicus*), and 2) exploration of the phylogeographic patterns of a subterranean species, the lesser Japanese mole (*Mogera imaizumii*) and the large Japanese mole (*Mogera wogura*).

In PART I, a coat color-mutant *Rattus norvegicus* sample collected on Sado Island, Japan, was investigated in search of the genetic basis of this mutation. This *R. norvegicus* was easily distinguished from wild-type rats, by its striking yellowish coat color. We sequenced the sequence of the coat color-determining gene *Mc1r* and identified a single-nucleotide deletion (c.337del) as the mutation responsible for this phenotypic characteristic. The color of the coat was quantitatively evaluated and was compared with those of wild rats and genome-edited mice (*Mus musculus*). It was shown that mice with similar frameshift-inducing mutations (c.97del, c.95_96del) do not always result in a yellowish coat color as this naturally occurring mutation found on Sado, indicating a complex and unknown mechanisms working behind the determination of coat color. We went through previous studies that refer to color-diluting mutations in *MC1R* of mammals and pointed out that deletions tend to occur at extracellular regions of *MC1R*, while substitutions tend to occur at transmembrane and intracellular regions. The molecular basis of this trend is not resolved, but since deletions lead to frameshifts and therefore easily result in corruption in the protein structure, we suspect that this might be a consequence of an unknown mechanism protecting the intracellular regions from highly-damageable deletions.

In PART II, we conducted a phylogeographic study on the isolated populations of the lesser Japanese mole, *Mogera imaizumii*. *Mogera imaizumii* mainly inhabits eastern

Japan, while western Japan is mainly inhabited by its close relative, the large Japanese mole, *M. wogura*. However, there are several patches of *M. imaizumii* populations in western Japan, in western Honshu (Mt. Hiba), Shikoku (Mts. Ishizuchi and Tsurugi), and two offshore islands (Dogo and Shodoshima). *Mogera imaizumii* is thought to have once dominated the whole Archipelago but retreated eastwards subsequent to the invasion of *M. wogura* from the continent via Tsushima Strait. The isolated populations, or "relicts", are thought to have been left behind from the rest of *M. imaizumii*. Our study focuses on uncovering the historical background of how these relict populations were formed, by analyzing their mitochondrial *Cytb* sequences. We analyzed 122 *Cytb* sequences of *M. imaizumii*, including 12 samples from the relict populations, and 157 sequences of *M. wogura*, including 14 samples from the continental populations. Our analysis showed that the *Cytb* sequences of these relict populations belonged to a distinctive haplogroup (mt*Mim*-IV), and the divergence of these populations occurred approximately 0.7 mya (million years ago) from the three major lineages identified in previous studies. *Mogera wogura* is said to have been present in Japan since at least 1.3 mya, thereby we hypothesize that the genetic structures of *M. imaizumii* and *M. wogura* in western Japan may have been formed by the repeated expansion and contraction in response to the recurrent environmental fluctuations during the 100,000-year Quaternary glacial cycles.

In PART III, we performed a reduced representation genome analysis on both *M. imaizumii* and *M. wogura*, further looking into their population structures. The MIG-seq method was chosen for reduced representation analysis, and the MIG-seq reads were mapped onto a draft genome of *M. wogura*, newly created in this study using 10x Genomics linked reads. From the analysis of population structure using fastSTRUCTURE, *M. imaizumii* and *M. wogura* were separated into three (nc*Mim*-1 to 3) and five (nc*Mwo*-1 to 5) genetically distinctive clusters, respectively. The clusters of *M. imaizumii* were well-separated with high F_{ST} values ranging from 0.376 to 0.478. The clusters of *M. wogura* were also well-separated (F_{ST}: 0.291–0.757), except for samples from Chugoku region, assigned to both nc*Mwo*-2 and nc*Mwo*-3 genetic clusters. The pattern indicates genetic introgression due to secondary contact between the two clusters, and this result was consistent with other analyses such as Principal Component Analysis (PCA) and Neighbor-Net analysis. The Kinki region, where both species shared a similar population boundary, may be associated with current and past geographic barriers. We identify that the Kinki region served as an

important site for the diversification of moles, where multiple factors, such as topographic barriers, interspecific interactions, and/or isolation related to vegetation may have shaped their population genetic structures.

In PART IV, we aim to unveil a more detailed demographic history of the genetic clusters defined in the previous PART III. While whole-genome-based approaches are expected to provide us with more detailed and robust data on the historical background of populations. Whole-genome sequencing was conducted on six *M. imaizumii* and ten *M. wogura* individuals. Adding to the draft genome created in PART III, Omni-c was performed to combine the reads to create a more robust reference sequence. The high F_{ST} values supported that each genetic cluster defined in PART III was genetically well-differentiated. NeighborNet analysis and Neighbor-Joining phylogenetic tree did not fully resolve the order of divergence, thereby indicating that there was a significant level of gene flow between the clusters, probably right after their divergence. Demographic inference by PSMC showed that *M. imaizumii* and *M. wogura* show a different trend in population expansion/contraction in response to climatic fluctuations. The results support the hypothesis that *M. imaizumii* is more cold-adapted than *M. wogura*, and that their population expanded during the colder periods.

From a single SNP that generates a phenotypical variant to a whole-genome-scale set of SNPs that infer the demographic history of a certain population over a long scale of time, this set of studies demonstrates how significant genetic variants can be for wild animal populations. It also provides new insights into the evolution and adaptation of wild animal populations, highlighting the potential and significance of various methods with different levels of resolution and robustness. Our findings not only advance our knowledge of evolutionary biology but also serve as a foundation for other related studies by providing a genomic basis of certain species, contributing to the proceedings of ecological, environmental, and conservation biology.

List of Abbreviations

α-MSH	α-melanocyte-stimulating hormone
ASIP	agouti signaling protein
BWA-MEM	Burrows-Wheeler aligner-maximum exact matches
cAMP	cyclic adenosine 3, 5'-monophosphate
CIE	International Commission on Illumination
CO1	cytochrome c oxidase subunit 1
CV	cross-validation
Cytb	cytochrome *b*
ddRAD-seq	Double Digested Restriction site-Associated DNA sequencing
DNA	deoxyribonucleic acid
DSRE	double SRE
eChugoku	eastern Chugoku
HMW DNA	high molecular weight DNA
HPD	highest posterior density
ISSR	inter-simple sequence repeat
JS	Japan Sea
LGM	last glacial maximum
MC1R	melanocortin 1 receptor
M. etigo	*Mogera etigo*
MIG-seq	multiplexed ISSR genotyping by sequencing
M. imaizumii	*Mogera imaizumii*
MJ	median joining

ML	maximum likelihood
M. musculus	*Mus musculus*
mtDNA	mitochondrial DNA
M. tokudae	*Mogera tokudae*
M. wogura	*Mogera wogura*
M. smithii	*Myodes smithii*
mya	million years ago
NCBI	National Center for Biotechnology Information
nKinki	northern Kinki
NJ	neighbor joining
PC	Principal Component
PCA	Principal Component Analysis
PCR	polymerase chain reaction
PFGE	pulse-field gel electrophoresis
PGM	penultimate glacial maximum
PO	Pacific Ocean
PSMC	pairwise sequentially Markovian coalescent
RNA	ribonucleic acid
RNA-seq	RNA sequencing
R. norvegicus	*Rattus norvegicus*
sKinki	southern Kinki
SNP	single nucleotide polymorphism
SRE	Short Read Eliminator

wChugoku western Chugoku

General Introduction

Investigating the genetic and evolutionary mechanisms behind phenotypic traits and population differentiation is a central goal in biology, as it addresses fundamental questions about the origins and maintenance of biodiversity. Phenotypic traits, such as coloration, and population-level processes, such as genetic structuring across landscapes, offer critical insights into how organisms adapt to their environments and how evolutionary forces shape biodiversity. These topics are particularly relevant in wild populations, where natural selection, genetic drift, and gene flow interact in complex ways. Understanding these mechanisms not only helps us to elucidate the evolutionary history of a species but also to predict their responses to future environmental changes. Such research mediates between molecular genetics and evolutionary biology, fostering an integrative perspective that is essential for addressing future challenges. Furthermore, these studies provide a framework for conservation strategies by identifying evolutionary significant units and adaptive traits critical for species survival. These studies advance our ability to understand, preserve, and sustainably interact with the natural world. This thesis addresses these themes by focusing on two studies. PART I examines the genetic basis of a visible phenotypic trait. PART II to PART IV explores the phylogeographic patterns of subterranean species.

Firstly, PART I investigates a naturally occurring coat color mutation in a wild Norway rat (*Rattus norvegicus*). Coat color is a key trait with significant ecological and evolutionary implications, influencing camouflage, inter/intra-species communication, and physiological regulation (Caro, 2005; Hubbard et al. 2010; Cuthill et al. 2017; Caro and Mallarino, 2020). The MC1R-ASIP system is a key mechanism in determining the coat color of mammals, regulating the type of melanin produced by the pigment-producing melanocytes. This system switches the type of melanin produced during hair growth, between black eumelanin and yellow pheomelanin, generating a banded pattern (agouti pattern) on a single hair (Dry 1928; Galbraith et al. 1979; Mills and Patterson 2009). Such microscopic patterning creates a complex and sophisticated visual variation in mammals. Since this MC1R-ASIP system is the core mechanism of coat color determination, mutations in the related genes can cause the coat color to change drastically. The Mc1r gene sequence of this individual was determined, and a single-nucleotide deletion was identified as the mutation responsible for this phenotype. The lightness of the coat color was quantitatively evaluated and compared to those of genome-edited mice (*Mus musculus*), showing that

similar mutations do not necessarily lead to identical coat color and that their determination system is not fully resolved. Previous studies that refer to coat color-diluting mutations were searched, to see if there is a correlation between the type of mutation and the place they occur. This study sheds light on the molecular basis of coat color variation in natural populations. These findings contribute to our understanding of phenotypic evolution in small wild mammals and also highlight the genetic mechanisms through which small-scale mutations can have observable effects on fitness and adaptation.

Secondary, we focus on the phylogeography of two wild mole species. Understanding the factors that influence current patterns of biodiversity is essential for evolutionary, ecological, and conservation biology. Physical barriers such as mountains or oceans are abiotic factors that contribute to genetic diversity, whereas competitive exclusion is a biotic factor that can maintain historical phylogeographic division (Waters 2011). The 100,000-year glacial cycles of the middle to late Quaternary influenced the phylogeographic/genetic differentiation occurring during glacial periods (McCormack et al. 2008; He and Jiang 2014). Such historical environmental fluctuations often lead to population expansions or contractions, which can result in shifts in species distributions, with glacial cycles being one of the most significant drivers (Haffer 1969; Bennett and Provan 2008).

While the genetic diversity and population structure of many mammal species have been extensively studied (Stewert and Cooper 2008), this topic has been less explored in strictly subterranean mammals, such as moles. Although these animals live in environments largely insulated from short-term and seasonal climatic changes (Feuda et al. 2015), they are likely influenced by changes in vegetation. The availability of food resources and plant cover are crucial factors that impact population dynamics (Nevo 1979). Consequently, vegetation changes driven by past climatic shifts may have played a role in shaping the population dynamics of subterranean species. Therefore, long-term climatic changes, such as glacial cycles, must have highly influenced the distributions of the underground insectivores. Despite having low dispersal ability, some mole species have succeeded in widely expanding their distribution range, as seen in moles in the Japanese Archipelago. Their nationwide distribution makes this Far Eastern set of islands an ideal site to look into the relationships between the moles and environmental fluctuations.

The Japanese Archipelago is home to diverse mammalian fauna, each with its own migration and dispersal history (Sato 2017). The complex topographic features of the Archipelago, along with climatic oscillations during the Pleistocene, have played a crucial role in shaping its current biodiversity. The key to understanding the evolutionary history of terrestrial animals in the Japanese Archipelago is the dramatic changes of distribution in response to climate fluctuation during the Quaternary, including the glacial and interglacial cycles, and the repeated formation of land bridges between Japan's four main islands and the Asian continent, as well as the smaller surrounding islands (Millien-Parra and Jaeger 1999). The expansion/contraction in population size and formation/submergence of land bridges cause the populations to merge/segregate. The repeated occurrence of such historical population dynamics is engraved into genomic features, contributing to the formation of genetic population structures. These factors have contributed to the biogeographic boundaries observed in terrestrial species. For example, a significant boundary divides the Archipelago into eastern and western sections, as evidenced by the distinct karyotypes of Japanese wood mice (Tsuchiya 1974) and greater Japanese shrew moles (Harada et al. 2001). This boundary may also have marked the separation between the two main mole species, the lesser Japanese mole (*Mogera imaizumii*) in the east and the large Japanese mole (*Mogera wogura*) in the west. Although these underground dwellers are emblematic of the high biodiversity of the Archipelago, the biotic and abiotic factors that shaped their evolution and dispersal remain understudied.

From PART II to PART IV, we investigate the phylogeography of Japanese moles using different genetic markers with different levels of resolution. In PART II, we focus on the isolated patches, or relict populations, of *M. imaizumii*. *Mogera imaizumii* mainly inhabits eastern Japan, but these relicts are distributed in western Japan, surrounded by their close relative, *M. wogura*. Phylogenetic analyses on 281 *Cytb* sequences of *M. imaizumii* and *M. wogura*, including samples of all known relict populations of *M. imaizumii* were performed. The goal is to reveal the historical background of when and how these relict populations diverged, leading to their isolation. In PART III, a reduced genome representation analysis (MIG-seq) on a vast number of sequences of the two species was performed, to comprehensively undercover the genetic structure of these species. In PART IV, a whole-genome approach was taken to look further deep into the demographic history of the populations.

These studies contribute to the broader fields of evolutionary biology, molecular ecology, and conservation science. It also underscores the relevance of wild populations as natural laboratories for understanding the genetic and environmental forces shaping life on Earth.

References

Bennett KD and Provan J. 2008. What do we mean by 'refugia'? Quaternary Science Reviews 27: 2449–2455.

Caro T. 2005. The adaptive significance of coloration in mammals. BioScience 55: 125–136.

Caro T and Mallarino R. 2020. Coloration in mammals. Trends in Ecology & Evolution 35: 357–366.

Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, Hill GE, Jablonski NG, Jiggins CD, Kelber A, Mapppes J, Marshall J, Merrill R, Osorio D, Prum R, Roberts NW, Roulin A, Rowland HM, Sherratt TN, Skelhorn J, Speed MP, Stevens M, Stoddard MC, Stuart-Fox D, Talas L, Tibbetts E, and Caro T. 2017. The biology of color. Science 357: eaan0221.

Dry FW. 1928. The agouti coloration of the mouse (*Mus musculus*) and the rat (*Mus norvegicus*). Journal of Genetics 20: 131–144.

Feuda R, Bannikova AA, Zemlemerova ED, Di Febbraro M, Loy A, Hutterer R, Aloise G, Zykov AE, Annesi F, and Colangelo P. 2015. Tracing the evolutionary history of the mole, *Talpa europaea*, through mitochondrial DNA phylogeography and species distribution modelling. Biological Journal of the Linnean Society 114: 495–512.

Galbraith DB, Wolff GL, and Brewer NL. 1979. Tissue microenvironment and the genetic control of hair pigment patterns in mice. Developmental Genetics 1: 167–179.

Haffer J. 1969. Speciation in Amazonian forest birds: most species probably originated in forest refuges during dry climatic periods. Science 165: 131–137.

Harada M, Ando A, Tsuchiya K, and Koyasu K. 2001. Geographical variations in chromosomes of the greater Japanese shrew-mole, *Urotrichus talpoides* (Mammalia: insectivora). Zoological Science 18: 433–442.

He K and Jiang X. 2014. Sky islands of southwest China. I: an overview of phylogeographic patterns. Chinese Science Bulletin 59: 585–597.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, and Safran RJ. 2010. Vertebrate pigmentation: from underlying genes to adaptive function. Trends in Genetics 26: 231–239.

McCormack JE, Bowen BS, and Smith TB. 2008. Integrating paleoecology and genetics of bird populations in two sky island archipelagos. BMC Biology 6: 1–12.

Millien-Parra V and Jaeger JJ. 1999. Island biogeography of the Japanese terrestrial mammal assemblages: an example of a relict fauna. Journal of Biogeography 26: 959–972.

Mills MG and Patterson LB. 2009. Not just black and white: pigment pattern development and evolution in vertebrates. Seminars in Cell & Developmental Biology 20: 72–81.

Nevo E. 1979. Adaptive convergence and divergence of subterranean mammals. Annual Review of Ecology and Systematics 10: 269–308.

Sato JJ. 2017. A review of the processes of mammalian faunal assembly in Japan: insights from molecular phylogenetics. In (Motokawa M and Kajihara H eds.) Species Diversity of Animals in Japan, pp. 49–60.

Stewert JR and Cooper A. 2008. Ice Age refugia and Quaternary extinctions: An issue of Quaternary evolutionary palaeoecology. Quaternary Science Reviews 27: 2443–2448.

Tsuchiya K. 1974. Cytological and biochemical studies of *Apodemus speciosus* group in Japan. The Journal of the Mammalogical Society of Japan 6: 67–87 (in Japanese with English summary).

Waters JM. 2011. Competitive exclusion: phylogeography's 'elephant in the room'? Molecular Ecology 20: 4388–4394.

PART I: An *Mc1r* mutation responsible for a yellowish color variant of Norway rat (*Rattus norvegicus*) and insights into mammalian *MC1R* variants

Introduction

Body pigmentation is an important factor for adaptation in mammals. They can function as concealment, as communication signals, or as physiological regulators (e.g. sunscreen, thermoregulation, Caro, 2005; Hubbard et al. 2010; Cuthill et al. 2017; Caro and Mallarino, 2020). A number of genes are reported to be associated with pigmentation or pattern formation in mammals, but the central genes that take part in the main pigment-determining system, the MC1R-ASIP system, are the melanocortin-1 receptor gene (*MC1R*) and the agouti-signaling protein gene (*ASIP*). *MC1R* consists of a single exon carrying a 954-bp amino acid coding region. The gene encodes a G protein-coupled receptor localized in the melanocyte membrane. When the agonist, alpha-melanocyte-stimulating hormone (α-MSH), binds to MC1R, the coupled adenyl cyclase is activated, leading to the production of cAMP, a second messenger within the cell. This path ultimately results in the production of black-brown eumelanin. By contrast, the perception of an antagonist (or inverse agonist), ASIP, induces the production of red-yellow pheomelanin in the melanocyte. These alternations between signaling proteins during hair growth result in a striped pattern known as the agouti pattern (Dry 1928; Galbraith et al. 1979; Mills and Patterson 2009). The MC1R-ASIP system is highly conserved among vertebrates (Bennett and Lamoreux 2003), and amino acid substitutions in *MC1R* cause inter-/intraspecific color polymorphisms in several mammalian species, such as laboratory mice (Robbins et al. 1993), domestic dogs (Newton et al. 2000; Schmutz and Berryere 2007a), domestic cats (Peterschmitt et al. 2009), horses (Marklund et al. 1996; Rieder et al. 2001), jaguars and jaguarundis (Eizirik et al. 2003) and martens (Ishida et al. 2013). In some cases, a single-nucleotide polymorphism is enough to cause striking diversity, from complete black (melanism) to white or bright yellow, representing both ends of the color spectrum (e.g. Hoekstra et al. 2006; Kambe et al. 2011; Ishida et al. 2013; Suzuki 2013). However, the genetic changes underlying the variation in coat color seen in natural populations (e.g. Łopucki and Mróz 2010; Tsuchihashi et al. 2011; Tomozawa et al. 2014; van der Geer 2019) remain to be investigated.

Here, we report a case of a Norway rat (*Rattus norvegicus*) captured on Sado

Island, Japan, with a yellowish color variant. Because *Mc1r* was a strong candidate gene for the phenotype due to its specificity (e.g. Suzuki 2013), we sequenced the gene to find the mutation responsible for this phenotypic variant. The animal's coat color was also measured using a spectrophotometer and compared with those of wild-type *R. norvegicus* and genome-edited laboratory mice (*Mus musculus*). Finally, we went through previous studies that report color-diminishing mutations in *MC1R* of mammals and discuss the consequences of an *MC1R*-deficient mutation with regard to coat color phenotype in mammals.

Materials and Methods

Study animals

A *R. norvegicus* individual with a yellowish coat color was captured on Sado Island, Japan, on December 12, 2015, for coat color measurement and genetic analysis. For the representatives of the wild-type *R. norvegicus*, three specimens stored in Hokkaido University (KT3884, HS5895, KT4093) were used for coat color measurements. The knockout mice (*Mus musculus*) with modified *Mc1r* generated from C3H/HeJ (Suzuki et al. 2020) were also used. These mice were homozygous for a 1-bp deletion (c.97del) or a 2-bp deletion (c.95_96del) in the *Mc1r* open reading frame that were both expected to cause frameshifts. Thus, these mice were characterized as *Mc1r*-deficient mutants.

DNA amplification and sequencing

DNA was extracted using a QIAmp DNA Mini kit (QIAGEN, Hilden, Germany) and the whole *Mc1r* coding region was sequenced. PCR was performed using AmpliTaq Gold 360 Master Mix (Applied Biosystems, Foster City, CA, USA) with the same primers and cycle conditions used to analyze *M. musculus* samples by Shimada et al. (2009). Two sets of primers were used to amplify the N-terminal and C-terminal halves of the *Mc1r* sequence: 5'*Mc1r* (-52, 5'-GCTCATACCACCTGGAGCTGCAGCC-3') and 3'*Mc1r* (+504, 5'-AAGAGGGTGCTGGAGACGATGCTGACC-3') for N-terminal half, and 5'*Mc1r* (+131, 5'-ATCCCAGATGGCCTCTTCCT-3') and 3'*Mc1r* (+1025, 5'-CCCTTAGACAAATGGAGATCAGG-3') for the C-terminal half. Base numbers in the primer names refer to the nucleotide positions relative to the start codon (+1) in the mouse gene (Moutjoy et al. 1992). Samples were then sequenced with a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) using an ABI 3130*xl* Genetic Analyzer (Applied

Biosystems). The sequences obtained were aligned using MEGA X software (Kumar et al. 2018) and searched manually to find the mutation responsible for the coat color variation. Nucleotide sequence data obtained are available in the DDBJ/EMBL/GenBank databases under the accession number LC579562.

Coat color measurements

A spectrophotometer (CM-700d; Konica Minolta Japan, Tokyo, Japan) was used to quantitatively evaluate coat color. Coat color was measured at 10 locations each on the dorsal and ventral sides of each specimen using a 3-mm diameter window. Any datapoints that were beyond 1.5x the interquartile range were considered outliers and excluded from further analysis. A one-sample *t*-test was performed to compare the color scores between specimens. The spectrophotometer evaluates color within a CIELab color space. The CIELab color space, defined by the International Commission on Illumination (CIE) in 1976, is a three-dimensional color space with three axes, L*, a*, and b*, which are transformations of R, G, and B values. An L* value is a measure of lightness (black to white, [0 – 100]), an a* value represents redness or greenness (green to red, [-60 – +60]), and a b* value is measure of blueness or yellowness (blue to yellow, [-60 – +60]) (CIELab color space, 2003, http://docs-hoffmann.de/cielab03022003.pdf). These units are suited for this kind of study because they can be related to visual perception (Joiner, 2004).

Results

Microscopic observations

The *R. norvegicus* individual captured on Sado Island had yellowish hair on its dorsal side and white hair on its ventral side (Fig. 1). In wild-type rats, pigment granules are observed throughout the whole hair, except for the basal end near the hair root. By contrast, pigment granules were observed only in the basal half of the hairs of our yellowish rat, and none seemed to be present in the other half. The tip was especially colorless, in contrast to wild-type *R. norvegicus*, the tip of which did not transmit light under an optical microscope (Fig. 2AB). In the genome-edited mice (Fig. 2EFGH), hair tips were transparent, although the transparency was not as apparent in the hair tips of the mice as in those of the rat (Fig. 2B). Ventral hairs appeared to lack pigmentation.

Gene sequencing

We sequenced the predicted whole 954-bp *Mc1r* coding region of the Sado individual and compared it with a reference rat sequence downloaded from GenBank (accession number: AB306978.1). Three homozygous nucleotide variants were detected: two synonymous nucleotide substitutions (c.C792T and c.G879C) and a homozygous 1-bp deletion at site 337 (c.337del, Fig. 3). Translation of the mutated sequence indicated that the deletion caused a frameshift starting from the amino acid at site 113 (Val113Cys) and produced a premature stop codon at site 140 (Ile140ter).

Coat color measurements

L* scores for both the dorsal and ventral sides of the Sado *Mc1r* mutant were higher than those of wild-type *R. norvegicus* (Fig. 4). L* scores for the dorsal and ventral sides of the rat with the homozygous deletion were compared with those of genome-edited mice homozygous for *Mc1r*-deficient alleles (Suzuki et al. 2020) and were found to be significantly higher ($P < 1.0 \times 10^{-8}$, Fig. 4).

Discussion

Phenotypic change due to a severe loss-of-function mutation in *Mc1r*

The 1-bp deletion (c.337del) found in the *Mc1r* gene during sequence analysis is expected to induce a frameshift starting at codon 113, situated in the region encoding the end of the second extracellular domain. This would introduce a premature stop codon at position 140, potentially resulting in a significant loss-of-function or null mutation. Similar mutations have been observed in other mammalian species (Everts et al. 2000; Newton et al. 2000; García-Borrón et al. 2005; Sánchez-Más et al. 2005), suggesting that this 1-bp deletion in *Mc1r* may be responsible for the yellowish coat color of the rat from Sado Island.

The deletion observed in this study was located within the gene region encoding an extracellular domain, aligning with patterns of deletion mutations in *MC1R* noted in other mammals. Notably, six loss-of-function mutations caused by indels are situated within regions that code extracellular domains (Fig. 3). In contrast, most mutations (14 of 15) that result in amino acid substitutions and contribute to lighter coat color are found in cell membrane-spanning and intracellular domains (Peters et al. 2016), with only one substitution (p.Cys35Phe) affecting an extracellular region that encodes the N-terminal

extracellular domain. While the reason for this distribution is unknown, it is hypothesized that key amino acids essential for *MC1R* function, such as those determining the protein structure and G protein interactions, are conserved and tend to occur in transmembrane and intracellular domains (Peters et al. 2016). This suggests that amino acid changes in extracellular regions might rarely impact *MC1R* function. Notably, six of the 13 mutations linked to lighter hair color involve the gain or loss of a cysteine residue (Fig. 3). The reason for the predominance of insertion-deletion mutations in extracellular regions is also unclear. However, certain structural features of these regions may predispose them to indels. Indels in *MC1R* are often associated with specific sequence structures, such as 6-bp duplications and inverted repeats within duplicated sequences, as seen in in *MC1R* variants linked to melanism (Eizirik et al. 2003: Hosoda et al. 2005) and somatic reversion (Kijas et al. 2001).

While both the Sado rat and the genome-edited mice possess *Mc1r*-deficient mutations, the degree of coat lightness differs between these two murine species (Fig. 4). Loss-of-function mutations in *MC1R* are typically associated with a halt in eumelanin production, resulting in a shift toward pheomelanin production. However, *MC1R*-deficient mutations produce a range of coat color variations (Fig. 3), from red to nearly white, across different mammalian species, even within a species, as seen in dogs (Newton et al., 2000; Schmutz and Berryere, 2007b). Dark red or chestnut coat color has been observed in species such as horses (Marklund et al. 1996; Rieder et al. 2001; Andersson 2003), pigs (Andersson 2003), and dogs (Irish Setters; Newton et al. 2000). In humans, mutations in *MC1R* commonly lead to red hair, with multiple possible mutation sites linked to this trait (e.g. four potential sites for amino acid substitution; Wong and Rees 2005). Conversely, nearly-white and faintly pheomelanic (cream) coat colors have been noted in Labrador Retrievers and Golden Retrievers (Schmutz and Berryere, 2007b), coyotes (Brockerville et al. 2013), Antarctic fur seals (Peters et al. 2016), black bears (Ritland et al. 2001) and Alaskan and Siberian Huskies (Dürig et al. 2018), suggesting a regulatory mechanism that limits pheomelanin production in the absence of *MC1R* (Dürig et al. 2018). Notably, *Mc1r*-deficient mice still retain some eumelanin in their hair (Lamoreux et al. 2001; Suzuki et al. 2020). These findings imply that loci outside the MC1R-ASIP system may play a role in supporting eumelanin production in mammals.

Pigment production and body color pattern formation in Mc1r-deficient mutants

From our present understanding of MC1R-ASIP system and its role in the formation of an agouti-like pattern, *MC1R*-deficient mutants are expected to be fully pheomelanic. However, microscopic observations of the Mc1r-deficient rat showed that its body hair did contain pheomelanic regions, but only in the basal portion, with highly depigmented tips, which contrasts with this expectation. In contrast, hairs from the genome-edited mice displayed eumelanic subapical regions along with colorless tips. The banded pattern in individual hairs is especially pronounced on the dorsal side of *Mc1r*-null mice, which have dark-tipped hairs (Bennett and Lamoreux 2003). These findings suggest that pigment switching in the hair cycle is regulated by more than just the MC1R-ASIP system.

Bennett and Lamoreux (2003) reported that mice with *Mc1r* loss-of-function mutations ($Mc1r^e/Mc1r^e$) are mostly yellow but display slightly darker pigmentation at the hair tips, suggesting that melanocytes lacking functional *Mc1r* may still produce eumelanin. Conklin and Bourne (1993) proposed that ASIP could act as a competitive antagonist by binding receptors other than MC1R. However, Abdel-Malek et al. (2001) provided conflicting evidence, observing that melanocytes in $Mc1r^e/Mc1r^e$ mice do not alter cAMP production levels in response to ASIP. Certain signaling proteins can stimulate cAMP production, potentially through direct action on melanocytes (Schallreuter et al. 2008; Enshell-Seijffers et al. 2010; Kondo and Hearing 2011). Additionally, eumelanin production is retained in proopiomelanocortin-deficient (i.e. α-MSH-deficient) mice, suggesting that eumelanin synthesis may be supported by alternative, non-melanocortin pathways or *Mc1R*-independent mechanisms (Slominski et al. 2005).

As previously mentioned, the MC1R-ASIP system may not be the only mechanism responsible for the generation of pigment patterns, as illustrated by the rat in our study, which displayed a weak countershading pattern. The contrast between the dorsal and ventral coat colors was clearly visible, with corresponding significant differences in L* scores ($P < 1.0 \times 10^{-9}$; data not shown). While the ventral-specific promoter in *ASIP* plays a crucial role in shaping the countershading pattern (Nadeau et al. 2008; Hubbard et al. 2010), it cannot fully explain the distinct body pattern seen in $Mc1r^e/Mc1r^e$ rats and mice. This is because double null mutations in *Mc1r* and *Asip* result in similar coat color patterns in *Mc1r*-null individuals (Suzuki et al. 2020). Therefore, other gene(s) are likely involved in determining

the body color pattern. Gonçalves et al. (2012) demonstrated that tuco-tucos (Rodentia, Ctenomyidae) express *MC1R* in a gradient from their dorsal to ventral sides. Based on this, we hypothesize that in the absence of functional *MC1R*, there may be a gradient of expression in other genes encoding signaling proteins that regulate melanin production. This suggests that the countershading pattern likely evolved prior to the development of the MC1R-ASIP system, specifically around the time of vertebrate divergence (Cortés et al. 2014). This idea aligns with the widespread occurrence of countershading across diverse taxa, including non-vertebrates (Rowland, 2009).

Figures and Legends

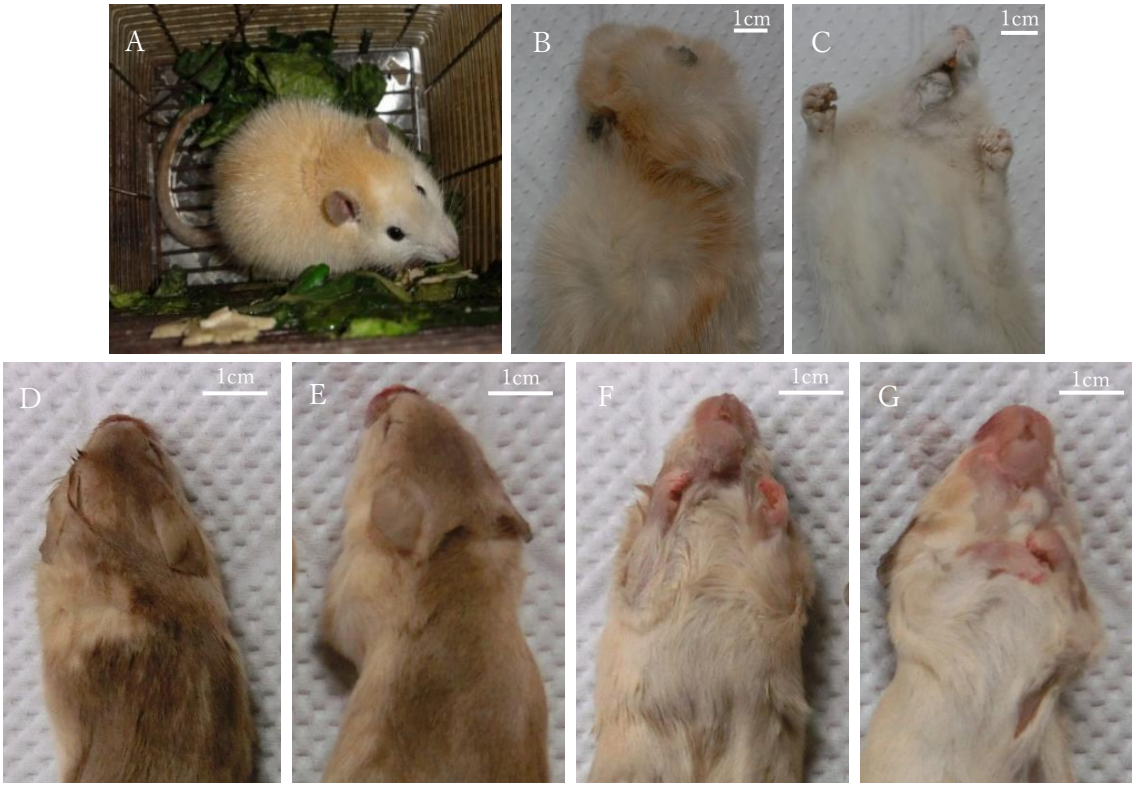


Figure 1. The yellowish Norway rat (*Rattus norvegicus*) from Sado Island, Japan. Its eyes are black, and its coat color is mostly yellowish on the dorsal (B) and white on the ventral side (C). The dorsum (D, E), and the venter (F, G) of genome-edited mice (*Mus musculus*) (1-bp deletion: D, F; 2-bp deletion: E, G).

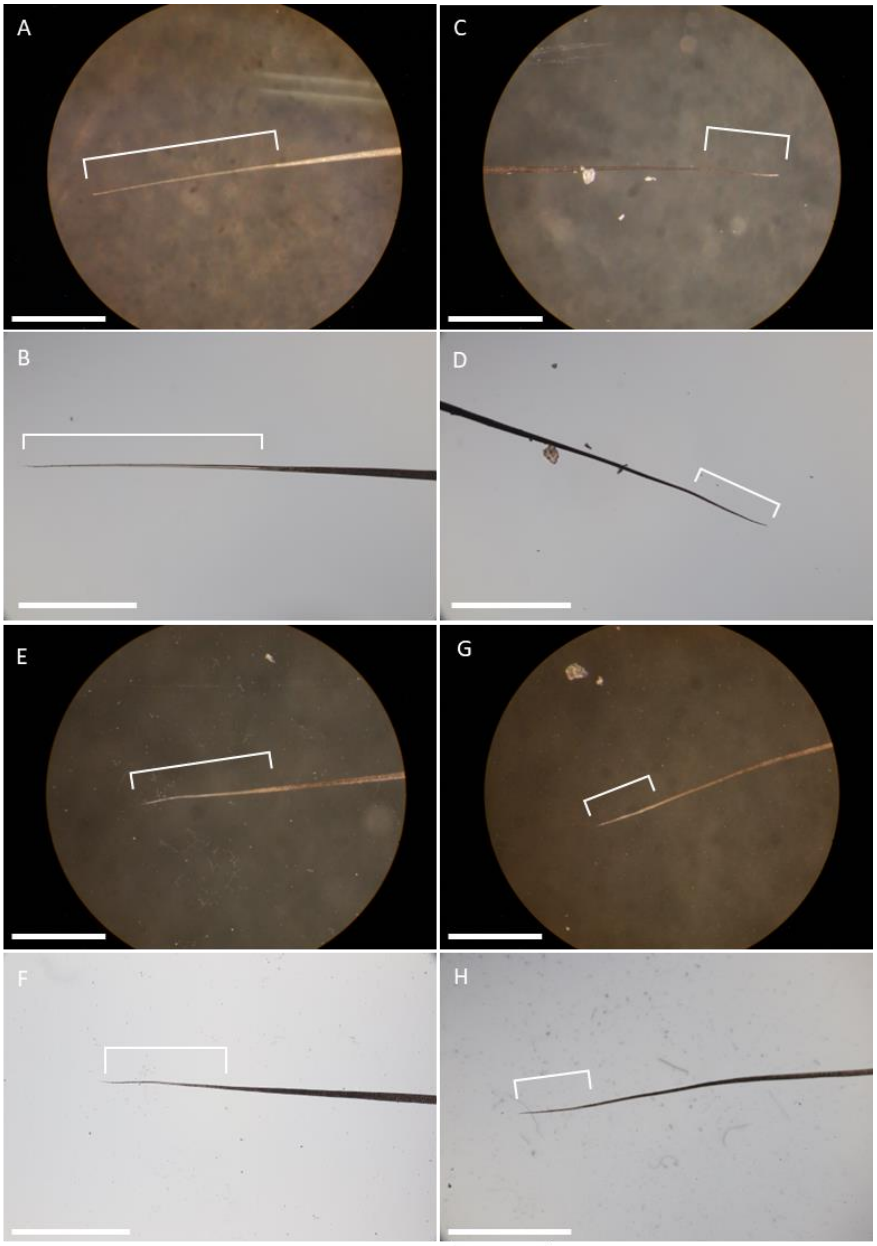
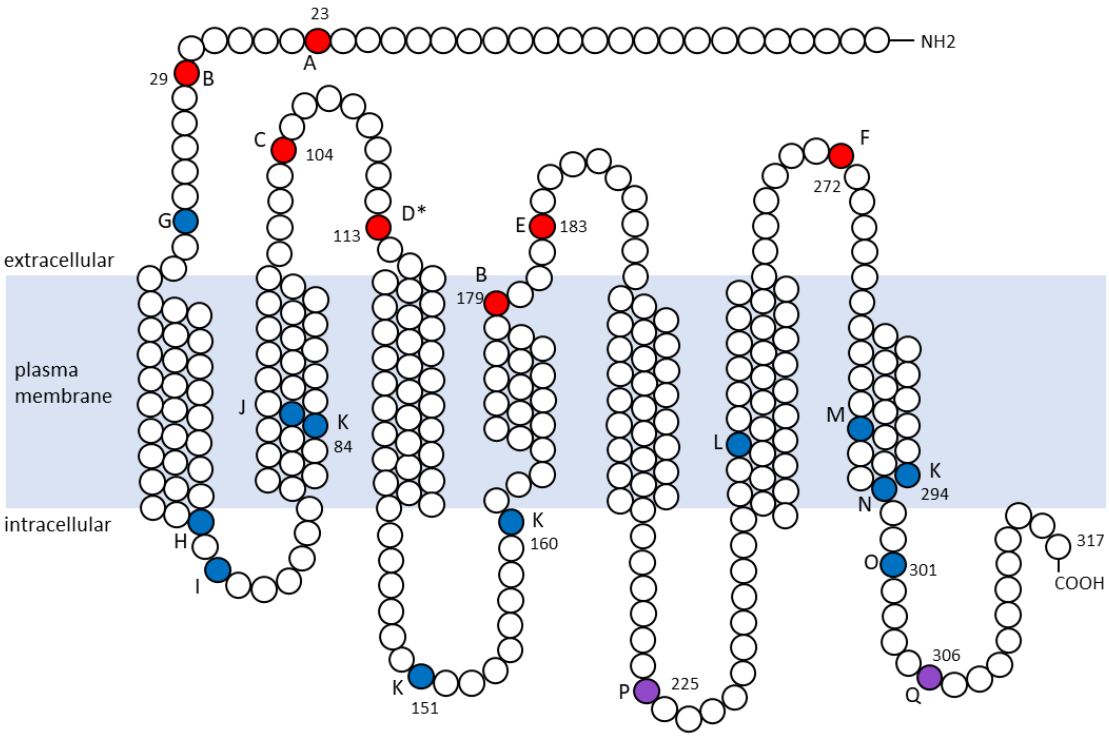


Figure 2. Tips of hairs of the yellowish rat (A, B), wild rat (C, D), and genome-edited mice (1-bp deletion: E, F; 2-bp deletion: G, H). The tips are indicated by square brackets. Under a stereo microscope, the tips are lighter in color for all specimens (A, C, E, G), but the tip of the wild rat hair does not transmit light of the optical microscope (D). This contrasts with the transparent tips of the yellowish rat (B), as well as those of genome-edited mice (F–H), although the transparency is less obvious in the mouse hairs than it is in the yellowish rat hair. Scale bars, 0.5 mm.



Frameshifts

A: Red, pigs (c.67_68insCC, Kijas et al., 2001)
B: Red hair, humans (c.86_87insA, c.537_538insC, Beaumont et al., 2008)
C: Red, cattle (c.310del, Klungland et al., 1995; Joerg et al., 1996; Rouzaud et al., 2000)
D*: Yellowish, rat (c.337del, this study)
E: Yellow, mice (c.549del, Robbins et al., 1993)
F: White, Alaskan and Siberian huskies (c.816_817del, Dürig et al., 2018)

Amino acid substitutions

G: Yellow, sable (p.Cys35Phe, Ishida et al., 2013)
H: Light, beach mice (p.Arg65Cys, Hoekstra et al., 2005)
I: Light, mammoth and sheep (p.Arg67Cys, Römpler et al., 2006, Fontanesi et al., 2010)
J: Chestnut, domestic horses (p.Ser83Phe, Marklund et al., 1996)
K: Red hair, humans (p.Asp84Glu, p.Arg151Cys, p.Arg160Trp, p.Asp294His, Kennedy et al., 2001; Smith et al., 1998), amber cats (p.Asp84Asn, Peterschmitt et al., 2009)
L: Red, pigs (p.Ala243Thr, Andersson, 2003)
M: Cream, Antarctic fur seals (p.Ser291Phe, Peters et al., 2016)
N: Cream/white, Kermode bears (p.Tyr298Cys, Ritland et al., 2001)
O: White, Arabian camels (p.Arg301Cys, Almathen et al., 2018)

Truncation by a premature stop codon

P: Red, goats (p.Gln225ter, Fontanesi et al., 2009)
Q: Red/yellow and cream, domestic dogs and coyotes (p.Arg306ter, Newton et al., 2000; Everts et al., 2000; Schmutz and Berryere, 2007ab; Brockerville et al., 2013)

Figure 3. Schematic illustration of MC1R and the mutation sites considered to be responsible for lighter coloration in mammals, based on information from this and other studies.

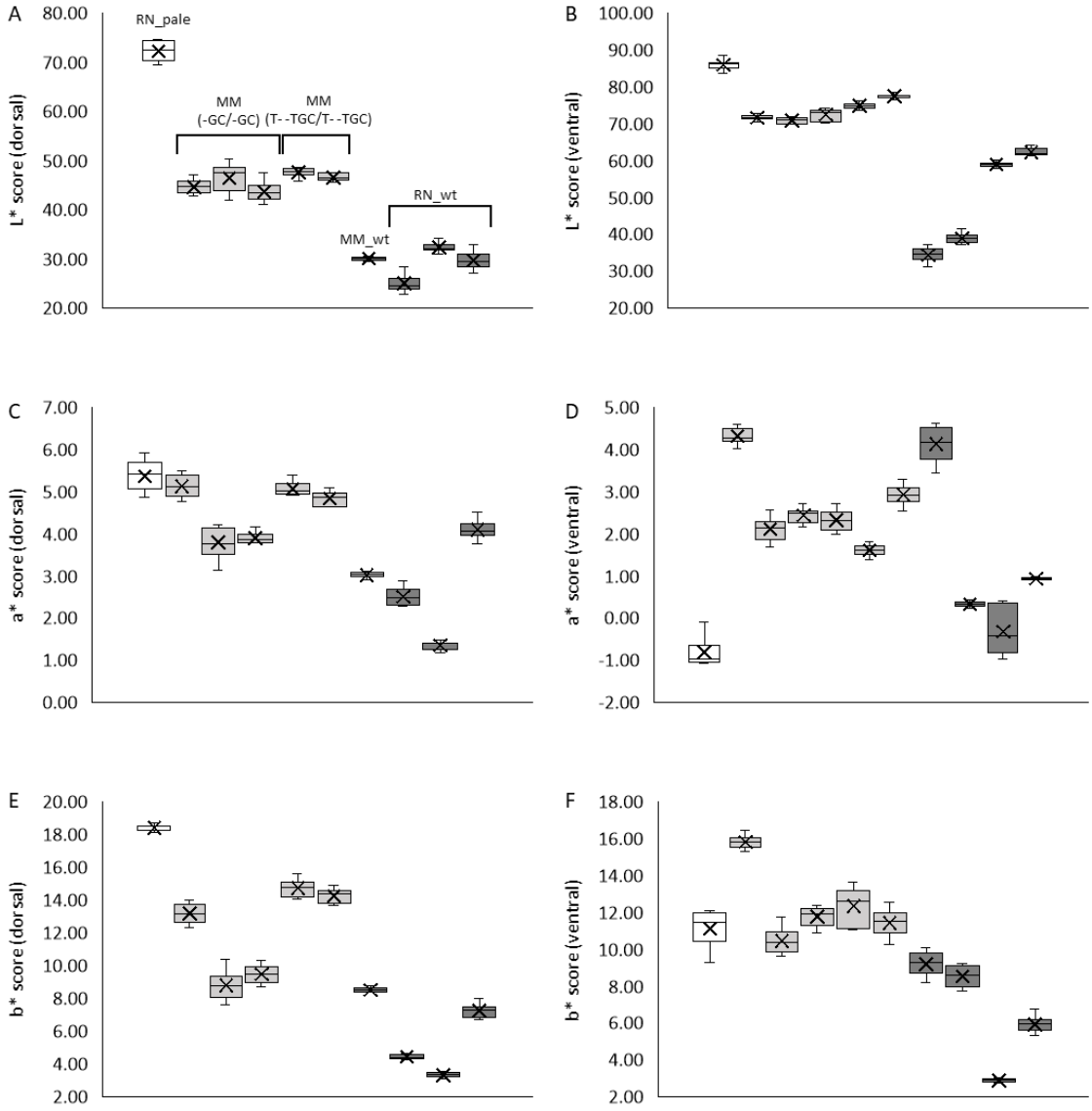


Figure 4. Box plots of L*, a*, and b* scores for the dorsal (A, C, E) and ventral (B, D, F) sides of the yellowish rat from Sado (white) and five genome-edited mice (gray), and a wild-type mouse and three wild-type rats (black). In the genome-edited mice, 1 or 2 bp was deleted at site 97T (MM, -GC/-GC) or at sites 95G and 96G (MM, T--TGC/T--TGC). Both dorsal and ventral L* scores of the yellowish rat (RN_pale) were substantially higher than those of the genome-edited mice and the wild-type rats (RN-wt). The genome-edited mice also had significantly higher L* scores compared to the wild-type control (MM_wt).

References

Abdel-Malek ZA, Scott MC, Furumura M, Lamoreux ML, Ollmann M, and Barsh GS. 2001. The melanocortin 1 receptor is the principal mediator of the effects of agouti signaling protein on mammalian melanocytes. Journal of Cell Science 114: 1019–1024.

Almathen F, Elbir H, Bahbahani H, Mwacharo J, and Hanotte O. 2018. Polymorphisms in *MC1R* and *ASIP* genes are associated with coat color variation in the Arabian camel. Journal of Heredity 109: 700–706.

Andersson L. 2003. Melanocortin receptor variants with phenotypic effects in horse, pig, and chicken. Annals of the New York Academy of Sciences 994: 313–318.

Beaumont KA, Shekar SN, Cook AL, Duffy DL, and Sturm RA. 2008. Red hair is the null phenotype of MC1R. Human Mutation 29: E88–E94.

Bennett DC and Lamoreux ML. 2003. The color loci of mice – a genetic century. Pigment Cell Research 16: 333–344.

Brockerville RM, McGrath MJ, Pilgrim BL, and Marshall HD. 2013. Sequence analysis of three pigmentation genes in the Newfoundland population of *Canis latrans* links the Golden Retriever *Mc1r* variant to white coat color in coyotes. Mammlian Genome 24: 134–141.

Caro T. 2005. The adaptive significance of coloration in mammals. BioScience 55: 1235–136.

Caro T and Mallarino R. 2020. Coloration in mammals. Trends in Ecology & Evolution 35: 357–366.

Conklin BR and Bourne HR. 1993. Mouse coat colour reconsidered. Naure 364: 110.

Cortés R, Navarro S, Agulleiro MJ, Guillot R, García-Herranz V, Sánchez E, and Cerdá-Reverter JM. 2014. Evolution of the melanocortin system. General and Comparative Endocrinology 209: 3–10.

Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, Hill GE, Jablonski NG, Jiggins CD, Kelber A, et al. 2017. The biology of color. Science 357: eaan0221.

Dry FW. 1928. The agouti coloration of the mouse (*Mus musculus*) and the rat (*Mus*

norvegicus). Journal of Genetics 20: 131–144.

Dürig N, Letko A, Lepori V, Hadji Rasouliha S, Loechel R, Kehl A, Hytönen MK, Lohi H, Mauri N, Diedtrich J, et al. 2018. Two *MC1R* loss-of-function alleles in cream-coloured Australian Cattle Dogs and white Huskies. Animal Genetics 49: 284–290.

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, and O'Brien SJ. 2003. Molecular genetics and evolution of melanism in the cat family. Current Biology 13: 448–453.

Enshell-Seijffers D, Lindon C, Wu E, Taketo MM, and Morgan BA. 2010. β-Catenin activity in the dermal papilla of the hair follicle regulates pigment-type switching. Proceedings of the National Academy of Sciences 107: 21564–21569.

Everts RE, Rothuizen J, and van Oost BA. 2000. Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (*MC1R*) in Labrador and Golden retrievers with yellow coat colour. Animal Genetics 31: 194–199.

Fontanesi L, Beretti F, Riggio V, Dall'Olio S, Calascibetta D, Russo V, and Portolano B. 2010. Sequence characterization of the melanocortin 1 receptor (*MC1R*) gene in sheep with different coat colours and identification of the putative *e* allele at the ovine *Extension* locus. Small Ruminant Research 91: 200–207.

Fontanesi L, Beretti F, Riggio V, Dall'Olio S, González EG, Finocchiaro R, Davoli R, Russo V, and Portolano B. 2009. Missense and nonsense mutations in melanocortin 1 receptor (*MC1R*) gene of different goat breeds: association with unexpected evidences. BMC Genetics 10: 47.

Galbraith DB, Wolff GL, and Brewer NL. 1979. Tissue microenvironment and the genetic control of hair pigment patterns in mice. Developmental Genetics 1: 167–179.

García-Borrón JC, Sánchez-Laorden BL, and Jiménez-Cervantes C. 2005. Melanocortin-1 receptor structure and functional regulation. Pigment Cell Research 18: 393–410.

Gonçalves GL, Hoekstra HE, and Freitas TRO. 2012. Striking coat colour variation in tuco-tucos (Rodentia: Ctenomyidae): a role for the melanocortin-1 receptor? Biological Journal of the Linnean Society 105: 665–680.

Hoekstra HE. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity 97: 222–234.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, and Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313: 101–104.

Hoekstra HE Krenz JG, and Nachman MW. 2005. Local adaptation in the rock pocket mouse (*Chaetodipus intermedius*): natural selection and phylogenetic history of populations. Heredity 94: 217–228.

Hosoda T, Sato JJ, Shimada T, Campbell KL, and Suzuki H. 2005. Independent nonframeshift deletions in the MC1R gene are not associated with melanistic coat coloration in three mustelid lineages. Journal of Heredity 96: 607–613.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, and Safran RJ. 2010. Vertebrate pigmentation: from underlying genes to adaptive function. Trends in Genetics. 26: 231–239.

Ishida K, Sato JJ, Kinoshita G, Hosoda T, Kryukov AP, and Suzuki H. 2013. Evolutionary history of the sable (*Martes zibellina brachyura*) on Hoddaido inferred from mitochondrial *Cytb* and nuclear *Mc1r* and *Tcf25* gene sequences. Acta Theriologica. 58: 13–24.

Joerg H, Fries HR, Meijerink E, and Stranzinger GF. 1996. Red coat color in Hostein cattle is associated with a deletion in the *MSHR* gene. Mammalian Genome 7: 317–318.

Joiner A. 2004. Tooth colour: a review of the literature. Journal of Dentistry 32: 3–12.

Kambe Y, Tanikawa T, Matsumoto Y, Tomozawa M, Aplin KP, and Suzuki H. 2011. Origin of agouti-melanistic polymorphism in wild Black Rats (*Rattus rattus*) inferred from *Mc1r* gene sequences. Zoological Science 28: 560–567.

Kennedy C, ter Huurne J, Berkhout M, Gruis N, Bastiaens M, Bergman W, Willemze R, and Bavinck JNB. 2001. Melanocortin 1 receptor (*MC1R*) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. Journal of Investigative Dermatology 117: 294–300.

Kijas JMH, Moller M, Plastow G, and Andersson L. 2001. A frameshift mutation in *MC1R* and a high frequency of somatic reversions cause black spotting in pigs. Genetics 158: 779–785.

Klungland H, Våge DI, Gomez-Raya L, Adalsteinsson S, and Lien S. 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. Mammalian Genome 6: 636–639.

Kondo T and Hearing VJ. 2011. Update on the regulation of mammalian melanocyte function and skin pigmentation. Expert Review of Dermatology 6: 97–108.

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549.

Lamoreux ML, Wakamatsu K, and Ito S. 2001. Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. Pigment Cell Research 14: 23–31.

Łopucki R and Mróz I. 2010. Cases of colouration anomalies in small mammals of Poland, and reasons for their incidence. Annales UMCS, Biologia 65: 67–76.

Mauklund L, Moller M, Sandberg K, and Andersson L. 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses. Mammlian Genome 7: 895–899.

Mills MG and Patterson LB. 2009. Not just black and white: pigment pattern development and evolution in vertebrates. Seminars in Cell & Developmental Biology 20: 72–81.

Mountjoy KG, Robbins LS, Mortrud MT, and Cone RD. 1992. The cloning of a family of genes that encode the melanocortin receptors. Science 257: 1248–1251.

Nadeau NJ, Minvielle F, Ito S, Inoue-Murayama M, Gourichon D, Follett SA, Burke T, and Mundy NI. 2008. Characterization of Japanese quail *yellow* as a genomic deletion upstream of the avian homolog of the *ASIP* (*agouti*) gene. Genetics 178: 777–786.

Newton JM, Wilkie AL, He L, Jordan SA, Metallinos DL, Holmes NG, Jankson IJ, and

Barsh GS. 2000. Melanocortin 1 receptor variation in the domestic dog. Mammalian Genome 11: 24–30.

Peters L, Humble E, Kröcker N, Fuchs B, Forcada J, and Hoffman JI. 2016. Born blonde: a recessive loss-of-function mutation in the melanocortin 1 receptor is associated with cream coat coloration in Antarctic fur seals. Ecology and Evolution 6: 5705–5717.

Petershmitt M, Grain F, Arnaud B, Deléage G, and Lambert V. 2009. Mutation in the melanocortin 1 receptor is associated with amber colour in the Norwegian Forest Cat. Animal Genetics 40: 547–552.

Rieder S, Taourit S, Mariat D, Langlois B, and Guérin G. 2001. Mutations in the agouti (*ASIP*), the extension (*MC1R*), and the brown (*TYRP1*) loci and their assoctiation to coat color phenotypes in horses (*Equus caballus*). Mammlian Genome 12: 450–455.

Ritland K, Newton C, and Marshall HD. 2001. Inheritance and population structure of the white-phased "Kermode" black bear. Current Biology 11: 1468–1472.

Robbins LS, Nadeau JH, Johnson Kr, Kelly MA, Roselli-Rehfuss L, Baack E, Mountjoy KG, and Cone RD. 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSYH receptor fuction. Cell 72: 827–834.

Römpler H, Rohland N, Lalueza-Fox C, Willerslev E, Kuznetsova T, Rabeder G, Bertranpetit J, Schöneberg T, and Hofreiter M. 2006. Nuclear gene indicates coat-color polymorphism in mammoths. Science 313: 62.

Rouzaud F, Martin J, Gallet PF, Delourme D, Goulemot-Leger V, Amigues Y, Ménissier F, Levéziel H, Julien R, and Oulmouden A. 2000. A first genotyping assay of French cattle breeds based on a new allele of the *extension* gene encoding the melanocortin-1 receptor (Mc1r). Genetics Selection Evolution 32: 511–520.

Rowland HM. 2009. From Abbott Thayer to the present day: what have we learned about the function of countershading? Philosophical Transactions of the Royal Society B: Biological Sciences 364: 519–527.

Sánchez-Más J, Sánchez-Laorden BL, Guillo LA, Jiménez-Cervantes C, and Garcia-Borrón JC. 2005. The melanocortin-1 receptor carboxyl terminal pentapeptide is essential for MC1R function and expression on the cell surface. Peptides 26: 1848–1857.

Schallreuter KU, Kothari S, Chavan B, and Spencer JD. 2008. Regulation of melanogenesis – controversies and new concepts. Experimental Dermatology 17: 395–404.

Schmutz SM and Berryere TG. 2007a. Genes affecting coat colour and pattern in domestic dogs: a review. Animal Genetics 38: 539–549.

Schmutz, SM and Berryere TG. 2007b. The genetics of cream coat color in dogs. Journal of Heredity 98: 544–548.

Shimada T, Sato JJ, Aplin KP, and Suzuki H. 2009. Comparative analysis of evolutionary modes in *Mc1r* coat color genet in wild mice and mustelids. Genes & Genetic Systems 84: 225–231.

Slominski A, Plonka PM, Pisarchik A, Smart JL, Tolle V, Wortsman J, and Low MJ. 2005. Preservation of eumelanin hair pigmentation in proopiomelanocortin-deficient mice on a nonagouti(*a/a*) genetic background. Endocrinology 146: 1245–1253.

Smith R, Healy E, Siddiqui S, Flanagan N, Steijlen PM, Rosdahl I, Jacques JP, Rogers S, Turner R, Jackson IJ, et al. 1998. Melanocortin 1 receptor variants in a nIrish population. Journal of Investigative Dermatology 111] 119–122.

Suzuki H. 2013. Evolutionary and phylogeographic views on *Mc1r* and *Asip* variation in mammals. Genes & Genetic Systems 88: 155–164.

Tomozawa M, Nunome M, Suzuki H, and Ono H. 2014. Effect of founding events on coat colour polymorphism of *Apodemus speciosus* (Rodentia: Muridae) on the Izu islands. Biological Journal of the Linnean Society 113: 522–535.

Tsuchihashi A, Tamate H, and Yokohata Y. 2011. Frequent occurrence of partial albinism in lesser Japanese moles (*Mogera imaizumii*) on Kinkazan Island, Miyagi Prefecture, northeastern Japan. Mammal Study 36: 141–146.

van der Geer AAE. 2019. Effect of isolation on coat colour polymorphism of Polynesian rats in Island Southeast Asia and the Pacific. PeerJ 7: e6894.

Wong TH and Rees JL. 2005. The relation between melanocortin 1 receptor (MC1R) variation and the generation of phenotypic diversity in the cutaneous response to ultraviolet radiation. Peptides 26: 1965–1971.

PART II: Evolutionary history of the relict populations of the lesser Japanese mole (*Mogera imaizumii*) inferred from mitochondrial cytochrome *b* gene sequences

Introduction

The climatic changes during the Quaternary including the 100,000-year glacial cycles have shown to have had a large impact on Japanese mammals, as shown in the case of the large Japanese wood mouse, *Apodemus speciosus* (Suzuki et al. 2015), and large mole species of genus *Mogera* (Iwasa et al. 2006; Kirihara et al. 2013; Nakamoto et al. 2021). One important aspect of impact on Japanese mammals is that distributions of species or groups of closely related species expand and contract reciprocally between populations at the northern and southern extremes of the Archipelago in a manner consistent with Quaternary climate changes. This trend has been reported in large mammals as shown in a study on Asian black bears in Japan (Wu et al. 2015). In small mammals such as voles or moles, northern (or eastern) species can have fragmented populations within the ranges of southern (or western) species (Kirihara et al. 2013; Honda et al. 2019). For example, the northern species *Myodes andersoni* has a fragmented population in the Kii Peninsula within the ranges of the southern species *M. smithii* (Honda et al. 2019). This fragmented population is thought to reflect the repeated expansion and contraction of the northern species, highlighting the impact of the Quaternary climate changes.

The temperate zone of the Japanese Archipelago (i.e., Honshu, Shikoku, and Kyushu islands) contains several mole and shrew-mole species (family Talpidae), most of which are endemic, following repeated ancient migrations from the continent to Japan via the Tsushima Strait (Shinohara et al. 2005; Kirihara et al. 2013; Sato 2017). Large mole species of the genus *Mogera* have a subterranean lifestyle, which allows them to expand into both plains and mountains; they feed mainly on earthworms and other invertebrates (Abe 1968; Yokohata 2005) and evolved under the influence of Quaternary environmental changes (Iwasa et al. 2006; Kirihara et al. 2013; Nakamoto et al. 2021). The lesser Japanese mole (*Mogera imaizumii*) inhabits eastern Japan (northern Honshu and relicts in western Honshu and Shikoku; Fig. 5), and the large Japanese mole (*M. wogura*) inhabits western Japan (southern Honshu, Shikoku, Kyushu, and other islands including the Oki islands, Shodoshima Island, Yakushima Island, Tanegashima Island, and the Tsushima Islands). In addition, *M. tokudae* is endemic to Sado Island, and *M. etigo* is endemic to the Echigo plain

(Abe et al. 2008; Ohdachi et al. 2009). *Mogera wogura* is also found on the Asian continent (Korean Peninsula, northeastern China, and Russian Primorsky Territory), and these continental populations are sometimes considered as a separate species, *M. robusta* (Zemlemerova et al. 2019). *Mogera wogura* is generally larger than *M. imaizumii* and commonly excludes *M. imaizumii* through competition under soft and fertile soil conditions (Abe 1974, 2001, 2010; Moribe and Yokohata 2011). A previous study suggested that *M. imaizumii*, which has traits that are likely cold-adapted, moved south during glacial periods of the Quaternary cycle, and *M. wogura*, which has a larger body size and therefore dominant over *M. imaizumii*, moved north during interglacial periods, in response to environmental fluctuation (Kirihara et al. 2013). In addition to its Kyoto and Kii Peninsula populations, *M. imaizumii* has relict populations on Mt. Hiba in Chugoku, Mts. Ishizuchi and Tsurugi in Shikoku, Shodoshima Island, the Nobi Plain in Tokai, and the Suzuka Mountains in Kinki (Kushihashi 1982; Abe et al. 2008; Ohdachi et al. 2009). A new population of *M. imaizumii* was recently discovered on Dogo Island (Tsunoi et al. 2023) among the Oki Islands in Chugoku (Fig. 5). The five isolated relict populations of *M. imaizumii* in western Japan, including the two offshore islands inhabited by *M. wogura*, provide opportunities for investigating the factors contributing to habitat segregation among the relict populations.

Phylogeographic variation in *M. imaizumii* and *M. wogura* has been investigated with sequences of the cytochrome *b* gene (*Cytb*) in mitochondrial DNA (mtDNA; Tsuchiya et al. 2000; Iwasa et al. 2006; Nakamoto et al. 2021). *Mogera imaizumii* has three distinct phylogroups generally corresponding to the administrative units of Japan: *Mim*-I (Tohoku), *Mim*-II (Kanto), and *Mim*-III (Hokuriku and Kinki). *Mogera wogura* has three phylogroups in the Japanese Archipelago — as well as one phylogroup (*Mwo*-IV) in the neighboring continental area (Korean Peninsula and Russian Far East). In Kinki, *M. imaizumii* has relict populations in Kyoto, Nara, and Wakayama Prefectures (*Mim*-III), which show substantial genetic differentiation from *Mim*-I and *Mim*-II (Kirihara et al. 2013; Nakamoto et al. 2021). Okamoto (1999) examined mitochondrial cytochrome c oxidase subunit 1 gene (*CO1*) sequences in relict *M. imaizumii* populations of Mt. Hiba on western Honshu, Mts. Ishizuchi and Tsurugi in Shikoku, and an offshore island (Shodoshima). The relict mtDNA haplotypes formed a distinct cluster separated from other *M. imaizumii* clusters, although no sequences were deposited in the nucleotide database. Recently, Tsunoi et al. (2023)

discovered a fragmented *M. imaizumii* population on Dogo Island and demonstrated that its lineage (*Mim*-IV) was distinct from the three known linages. Previous studies based on *Cytb* sequences did not examine all relict populations (Tsuchiya et al. 2000; Kirihara et al. 2013; Nakamoto et al. 202), and the phylogeographical history of the relict populations was not discussed in Okamoto (1999). Thus, the phylogenetic relationships of all known *M. imaizumii* relict populations remain poorly understood, and the factors leading to the generation and persistence of these fragmented populations have not been studied in detail.

In this study, we examined *M. imaizumii* samples collected from Shodoshima and Mts. Hiba, Tsurugi, and Ishizuchi and determined their *Cytb* sequences to analyze their phylogenetic relationships and historical dynamics. We performed molecular phylogenetic analyses of these sequences together with those reported from Dogo Island (Tsunoi et al. 2023) to elucidate the evolutionary history of isolated relict populations of *M. imaizumii* in western Japan. We further investigated how the evolutionary history of *M. imaizumii* was influenced by late Quaternary environmental changes and interactions with the closely related species *M. wogura*, which has been possibly competitive at the boundary between two species distributions.

Materials and Methods

Materials

We collected 43 specimens of *M. imaizumii* from 20 localities, with a special emphasis on isolated populations in western Honshu (Mt. Hiba, Shodoshima Island, and Mts. Ishizuchi and Tsurugi; Fig. 5). Thirteen specimens were newly collected for *M. wogura*. Genomic DNA was extracted with a QIAmp DNA Mini Kit (Qiagen, Hilden, Germany) from tissue samples (liver, muscle, or bone) preserved in 99.5% ethanol solution. Nucleotide sequence data newly obtained in this study are available in the DDBJ/EBNA/GenBank databases under the accession numbers LC777529–LC777584, LC807168, and LC833857.

Determination of mitochondrial Cytb sequences

Mitochondrial *Cytb* sequences (1,140 bp) were determined with polymerase chain reation (PCR) and Sanger sequencing. PCR was performed with the universal primer pair L14724 and H15915 (Kocher et al. 1989). The thermocycling parameters for the PCR were 94°C for 2 min; 30 cycles at 98°C for 10 s, 50°C for 30 s, and 60°C for 2 min; followed by

72°C for 7 min. The PCR mixtures (20 μL) consisted of 1 μL DNA solution, 10 μL 2x buffer, 2 μL 2mM dNTPs, 0.6 μL of each primeer (concentration: 10 pmol), 5.4 deionized water, and 0.4 μL KOD FX Neo polymerase (Toyobo, Osaka, Japan).

Sequencing was performed with a BigDye Terminator Cycle Sequencing Kit v.3.1 (Applied Biosystems, Waltham, MA, USA) and the sequence primers L14724 and H15915. The labeled PCR products were sequenced with an ABI3130 automated sequencer (Applied Biosystems). The sequences were aligned in MEGA X v10.0 (Kumar et al. 2018).

Phylogenetic analyses

Including the newly determined sequences (n = 56) and those obtained from the GenBank, ENA, and DDBJ nucleotide databases, we used 122 and 157 *Cytb* sequences (1140 bp) of *M. imaizumii* and *M. wogura*, respectively, in our phylogenetic analyses. Sequences of *M. tokudae* and *M. etigo* were obtained from the databases and used ads outgourps for phylogenetic inference.

A maximum likelihood (ML) phylogenetic tree was constructed in MEGA v10.0, with the substitution models TN93 + G + I. The best-fitting model was determined with Akaike's information criterion in MEGA. Bootstrap scores were calculated using 500 replicates. Divergence time estimation was performed with BEAST v2.7.4 (Bouckaert et al. 2014). We used an evolutionary rate of 0.03 substitutions/site/million years for the TN93 + G + I substitution model based on rates for mole species determined in the previous studies (Kirihara et al. 2013; Nakamoto et al. 2021).

Based on the results of ML phylogenetic tree, a haplotype network was generated for the relict (mt*Mim*-IV) haplogroup. The aligned fasta file was converted to nexus format using DnaSP version 6.12.03 (Rozas et al. 2017) and the haplotype network was generated by PopART (Bandelt et al. 1999).

Maternal demographic inference of each haplogroup were estimated by Bayesian Skyline Plot using BEAST version 2.7.4 (Bouckaert et al. 2014). BEAST was run with HKY substitution model and with mean clock rate of 0.11 (Nakamoto et al. 2021), chain length of 50,000,000 and pre burnin of 10%.

Results

Sample collection and *Cytb* sequence determination were conducted mainly for *M.*

imaizumii in the Chugoku and Shikoku regions to analyze their phylogenetic relationships and historical dynamics. The ML phylogenetic tree constructed from sequences of geographically isolated populations of *M. imaizumii* from western Japan (Dogo Island, Mt. Hiba, Shodoshima Island, and Mts. Ishizuchi and Tsurugi), together with those obtained from the databases, indicated four major monophyletic clades (mt*Mim*-I–IV; Fig. 5B). All haplotypes from Chugoku and shikoku were integrated into a single clade, mt*Mim*-IV, that was previously assigned to the lineage of Dogo Island haplotypes (Tsunoi et al. 2023). The remaining three lineages represented specific geographic regions, as reported previously (Kirihara et al. 2013; Nakamoto et al. 2021; Fig. 5B). The mt*Mim*-IV clade exhibited three distinct sublineages: Shodoshima–Shikoku (mt*Mim*-Iva), Mt. Hiba (mt*Mim*-IVb, and Oki–Dogo (mt*Mim*-IVc) (Fig. 5B). The ML tree placed the mt*Mim*-II lineage (Kanto) in the basal position. The divergence patterns of mt*Mim*-I (Tohoku), -III (Kinki), and -IV were not clearly resolved, which suggests that they diverged within a short period of evolutionary time. In *M. wogura*, the divergence of mt*Mwo*-III (Kyushu) and mt*Mwo*-IV (continental) was followed by that of mt*Mwo*-I (Tokai and Kinki) and mt*Mwo*-II (Chugoku and Shikoku) (Fig. 5B).

The BEAST analyses produced a different topology for the phylogenetic relationships from that estimated by the ML analysis. The analyses estimated the basal divergence of mt*Mim*-II at approximately 0.74 million years ago (mya, 95% highest posterior density [HPD], 0.59–0.90), that of mt*Mim*-I vs. mt*Mim*-IV and mt*Mim*-III at 0.56 mya (95% HPD, 0.44–0.68), and that of mt*Mim*-IV vs. mt*Mim*-III at 0.55 mya (95% HPD, 0.43–0.68; Fig. 6). In mt*Mim*-IV, sublineage divergence times were estimated to be 0.18 mya (95% HPD, 0.123–0.26) and 0.2 mya (95% HPD, 0.13–0.27) for the early divergence of mt*Mim*-Iva (Dogo Island) and that of mt*Mim*-IVc (Shodoshima and Shikoku Island) and mt*Mim*-IVb (Kyoto) and mt*Mim*-IIIc (Kii Peninsula) lineages from mt*Mim*-IIIa lineage (Hokuriku) was estimated to have occurred at approximately 0.26 mya (95% HPD, 0.18–0.34).

The major phylogroups of *M. wogura* were estimated to have diverged 0.91 mya (95% HPD, 0.74–1.09), diverging to mt*Mwo*-III/mt*Mwo*-IV and mt*Mwo*-I/mt*Mwo*-II. The haplotypes of *M. wogura* from Shodoshima appear to have been integrated into the Kinki and Tokai mt*Mwo*-I phylogroup rather than mt*Mwo*-II.

A haplotype network for mt*Mim*-IV is shown in Fig. 7. It is shown that individuals from Shikoku can be separated into two subgroups. This does not correspond to the two main mountain systems (Mts. Ishizuchi and Tsurugi).

The results of Bayesian Skyline Plot are shown in Fig. 8. While *M.imaizumii* showed a sign of expansion starting in the late LGP, the expansion of *M. wogura* started after the termination of LGP.

Discussion

Interspecific phylogenetic history of moles of the Japanese Archipelago

Previous studies of mole evolution using mtDNA sequences identified three major phylogroups (mt*Mim*-I–III), including the Kii Peninsula population (Tsuchiya et al. 2000; Kirihara et al. 2013; Nakamoto et al. 2021). In this study, we found that all *M. imaizumii* haplotypes from Chugoku and Shikoku were members of a fourth lineage (mt*Mim*-IV) of *M. imaizumii* from Dogo in the Oki Islands, where this species was only recently discovered (Tsunoi et al. 2023; Fig. 5). By contrast, Kinki populations (Kii Peninsula and Kyoto) had the third major clade (mt*Mim*-III), which is represented by haplotypes from Hokuriku.

Phylogenetic analyses showed that among Chugoku and Shikoku populations, mt*Mim*-IV diverged from the northwestern Honshu lineages (mt*Mim*-I and -III) approximately 0.55 mya (Fig. 6). This finding suggests that the Chugoku-Shikoku lineage has persisted in western Japan as early as the mid-Pleistocene.

Phylogenetic relationships among distinct populations of western Japan showed three sublineages: the Oki Islands (mt*Mim*-IVa), Mt. Hiba (mt*Mim*-IVb), and Shodoshima and Shikoku (Mts. Ishizuchi and Tsurugi; mt*Mim*-IVc, Fig. 5B). These sublineages are estimated to have diverged0.18–0.2 mya. The ancient divergence of isolated *M. imaizumii* populations in western Japan appears to have not been associated with the last glacial maximum (LGM) but rather with the post-penultimate glacial maximum (PGM), which has not been discussed in previous studies on climatic effects on granivorous and herbivorous small mammals in Japan (Oshida et al. 2009; Suzuki et al. 2015; Hanazaki et al. 2017; Honda et al. 2019). Our results clearly indicate a similarity between haplotypes from Shodoshima and Shikoku, which suggests their post-LGM divergence. In summary, our results imply that the observed mtDNA diversity among *M. imaizumii* populations in western Japan could be

influenced by a series of expansion and contraction events, possibly in response to glacial and interglacial periods during the 100,000-year cycle of Quaternary environmental changes.

The *Cytb* sequences between Ootoyo-cho (Fig. 5A, locality 56) and Kumakogen-cho (Fig. 5A, locality 55) in Shikoku differed by a single nucleotide (Fig. 7). This suggests that the genetic differentiation may not be significant between the two populations of these localities, despite previous beliefs that *M. imaizumii* have two isolated habitats on Mts. Ishizuchi and Tsurugi in Shikoku (Ohdachi et al. 2009). Instead, the haplotype network (Fig. 7) shows that there is an alternative genetic structure that separates populations of Tokushima Prefecture (Fig. 5A, localities 53, 54) from the rest. The phylogenetic trees also showed that these individuals are more closer to the ones from Shodoshima. Recent distribution surveys suggest that *M. imaizumii* is found in the mountains between Mts. Ishizuchi and Tsurugi (Tanioka 2020). Based on these facts, a biological border certainly existed in Shikoku to separate the *M. imaizumii* in two. However, this border does not correspond to the geographical border that separates the two mountains. It is therefore presumable that the Mt. Ishizuchi populations have expanded its way to Mt. Tsurugi.

Impact of Quaternary climate changes and interspecies competitive exclusion

Our results are consistent with the established view that interaction between *M. imaizumii* and *M. wogura* influenced their evolution. *Mogera imaizumii* may have adapted to cold environments and expanded its territory during glacial periods (Kirihara et al. 2013; Nakamoto et al. 2021), whereas *M. wogura* is thought to have experienced rapid demographic expansion during periods of abrupt warming. For example, Nakamoto et al. (2021) reported expansion signals of post-LGM and post-PGM interglacial periods in *Cytb* sequence data for *M. wogura* from western Japan. Thus, it is reasonable to hypothesize that both species underwent changes in their territory and ecological dominance throughout the 100,000-year Quaternary glacial cycle (Kirihara et al. 2013; Nakamoto et al. 2021).

Our BEAST phylogenetic analyses showed that western Japanese *M. wogura* diverged as early as 0.91 mya (Fig. 6). The mt*Mim*-IV lineage was present in western Japan for at least 0.55 million years; therefore, *M. imaizumii* may have played a role in shaping the three main lineages observed in *M. wogura* distribution in western Japan. The mechanisms underlying the creation and maintenance of the distribution boundary between mt*Mwo*-II

and mt*Mwo*-III mtDNA lineages of *M. wogura* in Chugoku, at the westernmost tip of Honshu, have not been elucidated to date. mt*Mim*-IV lineage may have existed near the boundary between mt*Mwo*-II and mt*Mwo*-III (e.g., Fig. 5, locality 50). It can be inferred that the population of mt*Mim*-IV lineage had intermittently interrupted the range expansions and gene flows of mt*Mwo*-II and mt*Mwo*-III during the middle Quaternary period, which would have led to the current phylogeographic patterns of mt*Mwo*-II and mt*Mwo*-III. Likely, Kirihara et al. (2013) have speculated that the range expansions of *M. imaizumii* may have contributed to the genetic distinctiveness of mt*Mwo*-I. We also consider the presence of the mt*Mim*-IV populations (e.g., localities 39–43) during glacial periods as a potential factor in the segregation of mt*Mwo*-I and mt*Mwo*-II. Previous studies have also speculated that the presence of an inland sea (Kirihara et al. 2013) and a difference in ecological preference for mountains and plains (Nakamoto et al. 2021) may have formed a boundary between mt*Mwo*-I and mt*Mwo*-II, and these multiple factors may have worked complexly.

The Bayesian Skyline Plot shows that both *M. imaizumii* and *M. wogura* populations started to expand from the end of LGP and through the Holocene. However, it is indicated that *M. imaizumii* started its expansion earlier than *M. wogura*. The reason of this trend is unknown, but it could be related to the hypothesis that *M. imaizumii* is cold-adapted, gaining the chance to expand during the cold LGP. As BSP is not capable of estimating the effective population sizes far back in time and requires a certain amount of sample sizes, this alone is not sufficient to discuss how mole populations fluctuated. To estimate the overall demography of these species, combining other methods, such as PSMC, is required.

Conclusions

This study is the first to provide the genetic structure of *M. imaizumii*, with an emphasis on isolated populations in western Japan. Our phylogenetic analyses showed that mtDNA sequences from relict populations were integrated into a fourth major lineage of *M. imaizumii* (mt*Mim*-IV). Divergences within mt*Mim*-IV were estimated at 200,000 years ago, which suggests that in the Japanese Archipelago, cold-adapted *M. imaizumii* underwent genetic exchange between sites during glaciation two and three epochs ago and that the origin of these western Japanese lineages dates to 550,000 years ago. Our results suggest

that the 100,000-year Quaternary glacial cycle caused mole populations to shrink and expand repeatedly. The factors that determine the distributions of *M. imaizumii* and *M. wogura* are only partly uncovered, but we speculate that they had been in a state of mutual competitive exclusion during the long period of the Quaternary, and *M. imaizumii* may have been involved in the generation of the three main lineages of *M. wogura* in western Japan. Mole species in the Japanese Archipelago appear to have been affected by a variety of factors, including Quaternary environmental changes, geographic barriers, interspecific competition, and regional habitat diversity. In conclusion, both biotic and abiotic factors are likely responsible for creating and maintaining the within-species genetic diversity in *M. imaizumii* and *M. wogura* in the Japanese Archipelago.

Figures and Legends

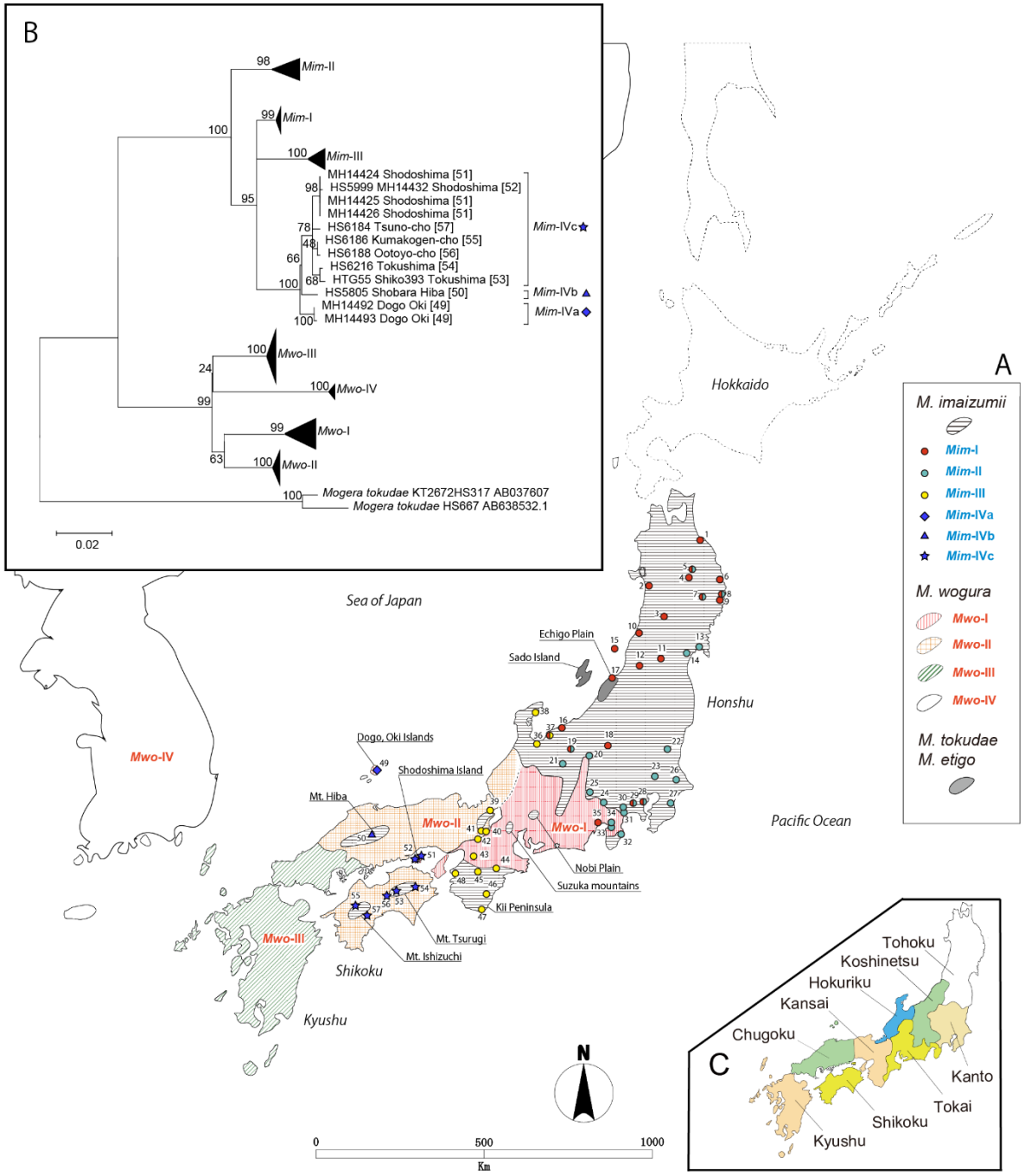


Figure 5. (A) Map of the Japanese Archipelago indicating sampling sites for *Mogera imaizumii* and *M. wogura* used in this study. (B) Maximum Likelihood (ML) tree for *M. imaizumii*, *M. wogura*, *M. tokudae*, and *M. etigo* based on *Cytb* sequences. Evolutionary history was inferred with Tamura-Nei model (Tamura and Nei 1993). Bootstrap scores for 500 replicates are shown in percentages at nodes. (C) Administrative units of Japan.

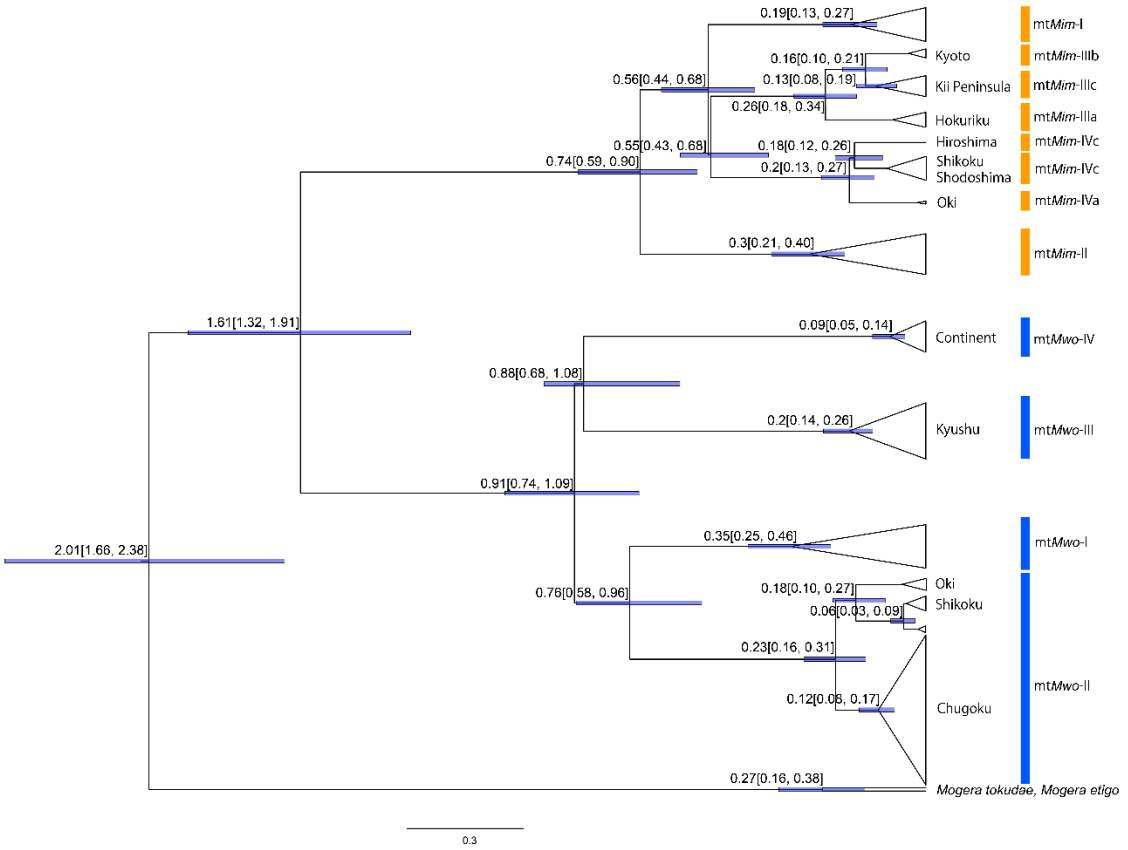


Figure 6. Chronogram of *M. imaizumii*, *M. wogura*, *M. tokudae*, and *M. etigo* based on *Cytb* sequences obtained using the BEAST software. Estimated divergence times and 95% highest posterior density (HPD) interval are indicated at nodes. Blue horizontal bars indicate 95% HPD intervals.

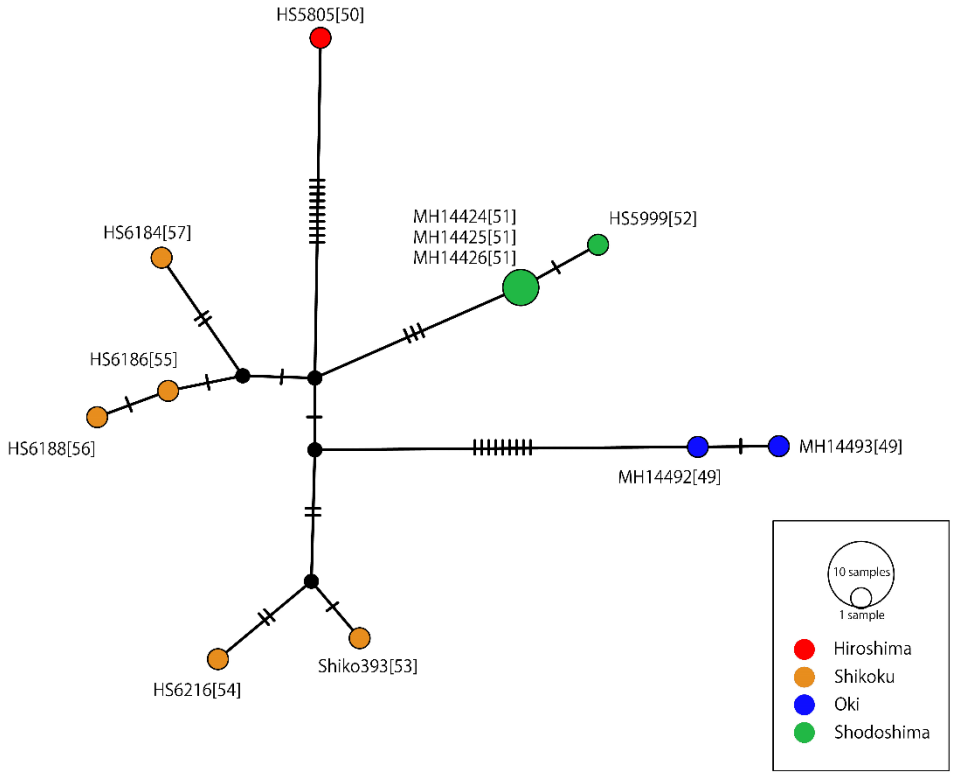


Fig. 7. A haplotype network of mt*Mim*-IV haplogroup. The network was generated using Median Joining (MJ) method. The numbers in brackets represent the sampling locations (see Fig. 5).

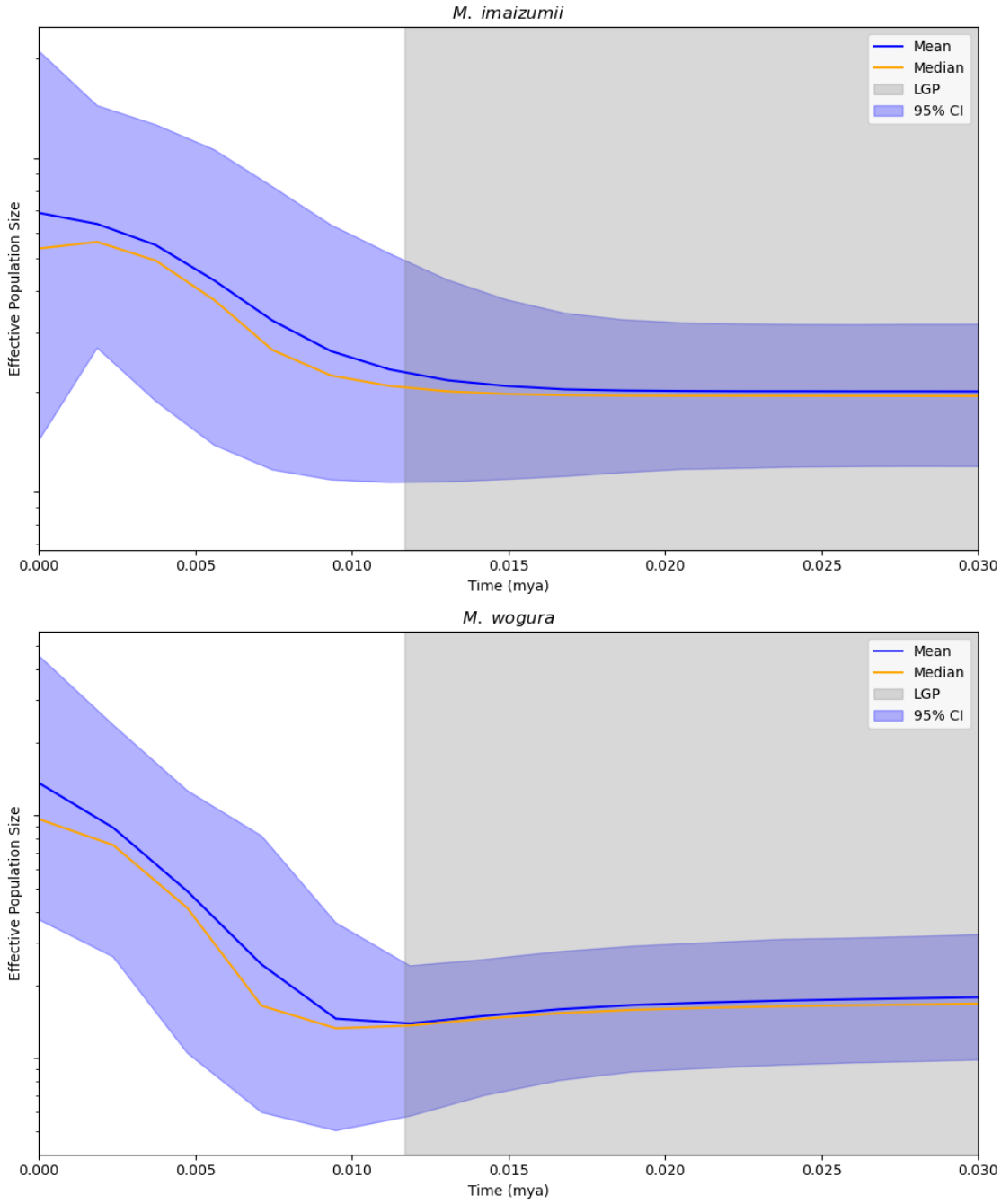


Fig. 8. Bayesian Skyline Plots of *M. imaizumii* and *M. wogura*. The span of Last Glacial Period (LGP) is highlighted in gray.

Appendix I. List of *Mogera* samples and *Cytb* sequences used in this study.

Species	**Locality No.**	**Location**		**Specimen code**	***Cytb* haplogroup**
Mogera imaizumii	1	Aomori	Hachinohe	KT2793/HS499; AB037611*	*Mim*-Ia
Mogera imaizumii	2	Akita	Akita	MH14257/HS5786; LC554619****	*Mim*-Ia
Mogera imaizumii	3	Akita	Yuzawa	MH14258/HS5787; LC554620****	*Mim*-Ia
Mogera imaizumii	3	Akita	Yuzawa	MH14259/HS5788; LC554621****	*Mim*-Ia
Mogera imaizumii	3	Akita	Yuzawa	MH14263/HS5792; LC554622****	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG319; AB245936**	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG320; AB245958**	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG321; AB245937**	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG322; AB245938**	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG323; AB245939**	*Mim*-Ia
Mogera imaizumii	4	Iwate	Morioka	HEG350; AB245940**	*Mim*-Ia
Mogera imaizumii	5	Iwate	Himekami	HS5986	*Mim*-Ia
Mogera imaizumii	5	Iwate	Himekami	HS5988	*Mim*-IIa
Mogera imaizumii	6	Iwate	Miyako	HS1777; AB245960**	*Mim*-Ia
Mogera imaizumii	7	Iwate	Tono	HEG346; AB245952**	*Mim*-IIa
Mogera imaizumii	7	Iwate	Tono	HS2315; AB245961**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG170; AB245941**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG172/HS1498; AB245942**	*Mim*-IIa
Mogera imaizumii	8	Iwate	Otsuchi	HEG352; AB245943**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG353; AB245944**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG354; AB245945**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG197; AB245948**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG200; AB245949**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG180/HS1514; AB245946**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG344; AB245947**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG347; AB245951**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG325; AB245962**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG326; AB245955**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG327; AB245954**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HEG348; AB245950**	*Mim*-Ia
Mogera imaizumii	8	Iwate	Otsuchi	HS1508**	*Mim*-I
Mogera imaizumii	9	Iwate	Kamaishi	HEG181/HS1515; AB245956**	*Mim*-Ia

Mogera imaizumii	9	Iwate	Kamaishi	HEG328; AB245953**	*Mim*-Ia
Mogera imaizumii	10	Yamagata	Tsuruoka	MH14264/HS5793; LC554623****	*Mim*-Ia
Mogera imaizumii	10	Yamagata	Tsuruoka	MH14265/HS5794; LC554624****	*Mim*-Ia
Mogera imaizumii	11	Yamagata	Kamio	MH14266/HS5795; LC554625****	*Mim*-Ia
Mogera imaizumii	11	Yamagata	Kamio	MH14267/HS5796; LC554626****	*Mim*-Ia
Mogera imaizumii	11	Yamagata	Kamio	MH14268/HS5797; LC554627****	*Mim*-Ia
Mogera imaizumii	11	Yamagata	Kamio	MH14269/HS5798; LC554628****	*Mim*-Ib
Mogera imaizumii	12	Yamagata	Ogunimachi	MH14270/HS5799; LC554629****	*Mim*-Ia
Mogera imaizumii	13	Miyagi	Ishinomaki	HEG175; AB245957**	*Mim*-IIa
Mogera imaizumii	13	Miyagi	Ishinomaki	HEG343; AB245963**	*Mim*-IIa
Mogera imaizumii	14	Miyagi	Sendai	HS582; AB037613*	*Mim*-IIa
Mogera imaizumii	15	Niigata	Awashima	HS1521; AB638494***	*Mim*-Ia
Mogera imaizumii	16	Niigata	Itoigawa	MH14303/HS5810	*Mim*-Ib
Mogera imaizumii	16	Niigata	Itoigawa	MH14304/HS5811	*Mim*-Ib
Mogera imaizumii	16	Niigata	Itoigawa	MH14305/HS5812	*Mim*-Ib
Mogera imaizumii	16	Niigata	Itoigawa	MH14302/HS5809	*Mim*-Ib
Mogera imaizumii	17	Niigata	Ikarashi	KT2795/HS364; AB270529**	*Mim*-Ib
Mogera imaizumii	17	Niigata	Ikarashi	HS470*	*Mim*-Ib
Mogera imaizumii	18	Gunma	Minakami	HS3057; LC554630****	*Mim*-Ib
Mogera imaizumii	18	Gunma	Minakami	KT3331/HS3077; LC554631****	*Mim*-Ib
Mogera imaizumii	19	Nagano	Chikuma	HS4491; LC554631****	*Mim*-Ib
Mogera imaizumii	19	Nagano	Chikuma	KT4014/HS4572; LC554632****	*Mim*-IIb
Mogera imaizumii	20	Nagano	Karuizawa	HS1778; AB245959**	*Mim*-IIb
Mogera imaizumii	21	Nagano	Matsumoto	MH14479/HS6091	*Mim*-IIb
Mogera imaizumii	21	Nagano	Matsumoto	MH14481/HS6093	*Mim*-IIb
Mogera imaizumii	22	Ibaraki	Hitachiohmiya	HS4665	*Mim*-IIb
Mogera imaizumii	23	Ibaraki	Tsukubamirai	KT4469/HS5630; LC554633****	*Mim*-IIb
Mogera imaizumii	23	Ibaraki	Tsukubamirai	KT4470/HS5631; LC554634****	*Mim*-IIb
Mogera imaizumii	23	Ibaraki	Tsukubamirai	KT4471/HS5632; LC554635****	*Mim*-IIb
Mogera imaizumii	24	Yamanashi	Aokigaharajukai	KT3468/HS3187; LC554637****	*Mim*-IIc
Mogera imaizumii	25	Yamanashi	Hokuto	HS5603	*Mim*-IIb
Mogera imaizumii	26	Chiba	Katori	MH14343/HS5864	*Mim*-IIb
Mogera imaizumii	26	Chiba	Katori	MH14344/HS5865	*Mim*-IIb
Mogera imaizumii	26	Chiba	Katori	MH14345/HS5866	*Mim*-IIb
Mogera imaizumii	26	Chiba	Katori	MH14346/HS5867	*Mim*-IIb

Mogera imaizumii	27	Chiba	Shirako	MH14341/HS5872	*Mim*-IIb
Mogera imaizumii	27	Chiba	Shirako	MH14342/HS5873	*Mim*-IIb
Mogera imaizumii	28	Kanagawa	Yokohama	KT2697/HS365; AB037616*	*Mim*-IIb
Mogera imaizumii	28	Kanagawa	Yokohama	HS472; LC554638****	*Mim*-Ib
Mogera imaizumii	29	Kanagawa	Atsugi	HS3896; AB638496***	*Mim*-II
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1444/HS6004	*Mim*-IIb
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1450/HS6007	*Mim*-IIb
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1001/HS6012	*Mim*-IIb
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1292/HS6067	*Mim*-IIb
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1274/HS6064	*Mim*-IIb
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1314/HS6073	*Mim*-IIb08
Mogera imaizumii	29	Kanagawa	Atsugi	TUA1315/HS6074	*Mim*-IIb
Mogera imaizumii	30	Kanagawa	Yamakita	TUA1084/HS6009	*Mim*-IIb
Mogera imaizumii	30	Kanagawa	Yamakita	TUA1066/HS6010	*Mim*-IIb
Mogera imaizumii	30	Kanagawa	Yamakita	KT4314/HS5901	*Mim*-IIb
Mogera imaizumii	30	Kanagawa	Yamakita	TUA1167/HS6047	*Mim*-IIb
Mogera imaizumii	30	Kanagawa	Yamakita	TUA1198/HS6049	*Mim*-IIb
Mogera imaizumii	31	Kanagawa	Hakone	HS3092; LC554639****	*Mim*-IIc
Mogera imaizumii	32	Shizuoka	Ito	HS1022*	*Mim*-IIc
Mogera imaizumii	33	Shizuoka	Mishima	KT3469/HS3184; LC554640****	*Mim*-IIc
Mogera imaizumii	33	Shizuoka	Mishima	TUA1110/HS6030	*Mim*-IIc
Mogera imaizumii	34	Shizuoka	Susono	TUA1205/HS6040	*Mim*-IIc
Mogera imaizumii	35	Shizuoka	Shizuoka	SIK:0775; HG737872*****	*Mim*-Ib
Mogera imaizumii	36	Toyama	Toyama	HS5456; LC554641****	*Mim*-IIIa
Mogera imaizumii	36	Toyama	Toyama	HS5457; LC554642****	*Mim*-IIIa
Mogera imaizumii	36	Toyama	Toyama	HS5458; LC554643****	*Mim*-IIIa
Mogera imaizumii	36	Toyama	Toyama	HS5462; LC554644****	*Mim*-IIIa
Mogera imaizumii	37	Toyama	Unazuki	MH14299/HS5806	*Mim*-IIIa
Mogera imaizumii	37	Toyama	Unazuki	MH14300/HS5807	*Mim*-Ib
Mogera imaizumii	38	Ishikawa	Noto	HS3366; AB638495***	*Mim*-IIIa
Mogera imaizumii	39	Shiga	Takashima	HS3367; AB638497***	*Mim*-IIIb
Mogera imaizumii	39	Shiga	Takashima	HS3368; LC554207***	*Mim*-IIIb
Mogera imaizumii	40	Kyoto	Mt. Hiei	M13285/HS4736; LC554645****	*Mim*-IIIb
Mogera imaizumii	41	Kyoto	Iwakurahasemachi	HS463; AB037619*	*Mim*-IIIb
Mogera imaizumii	42	Kyoto	Nagaokakyo	HS1843; LC554646****	*Mim*-IIIb

Mogera imaizumii	43	Osaka	Kawachinagano	MH14237/HS5757; LC554647****	*Mim*-IIIc
Mogera imaizumii	43	Osaka	Kawachinagano	MH14253/HS5783; LC554648****	*Mim*-IIIc
Mogera imaizumii	44	Nara	Mt. Wasamata	MH9967/HS4620; LC554649****	*Mim*-IIIc
Mogera imaizumii	45	Nara	Nosegawa	MH14387	*Mim*-IIIc
Mogera imaizumii	46	Wakayama	Kumanogawa-cho	KT2707/HS368; LC554208***	*Mim*-IIIc
Mogera imaizumii	47	Wakayama	Kozagawa	MH14218/HS5758; LC554650****	*Mim*-IIIc
Mogera imaizumii	47	Wakayama	Kozagawa	MH14219/HS5759; LC554651****	*Mim*-IIIc
Mogera imaizumii	48	Wakayama	Wakayama	KT2807/HS462; LC554652****	*Mim*-IIIc
Mogera imaizumii	49	Shimane	Dogo, Oki Is.	MH14492******	*Mim*-IVa
Mogera imaizumii	49	Shimane	Dogo, Oki Is.	MH14493******	*Mim*-IVa
Mogera imaizumii	50	Hiroshima	Hiba	MH14329/HS5805	*Mim*-IVb
Mogera imaizumii	51	Kagawa	Uchinomi	MH14424/HS5991	*Mim*-IVc
Mogera imaizumii	51	Kagawa	Uchinomi	MH14425/HS5992	*Mim*-IVc
Mogera imaizumii	51	Kagawa	Uchinomi	MH14426/HS5993	*Mim*-IVc
Mogera imaizumii	52	Kagawa	Tonosho	MH14432/HS5999	*Mim*-IVc
Mogera imaizumii	53	Tokushima	Tonosho	Shiko393	*Mim*-IVc
Mogera imaizumii	54	Tokushima	Kamiyama-cho	HS6216	*Mim*-IVc
Mogera imaizumii	55	Ehime	Kumakogen-cho	HS6186	*Mim*-IVc
Mogera imaizumii	56	Kochi	Otoyo-cho	HS6188	*Mim*-IVc
Mogera imaizumii	57	Kochi	Tsuno-cho	HS6184	*Mim*-IVc
Mogera wogura	31	Kanagawa	Hakone	KT3491/HS3096; AB638498***	*Mwo*-Ib
Mogera wogura	31	Kanagawa	Hakone	KT3490/HS6080	*Mwo*-Ib
Mogera wogura	34	Shizuoka	Susono	TUA1078/HS6037	*Mwo*-Ib
Mogera wogura	33	Shizuoka	Mishima	KT3201/HS1023; AB037623*	*Mwo*-Ib
Mogera wogura	58	Shizuoka	Fujinomiya	TUA1519/HS3897; AB638499***	*Mwo*-Ib
Mogera wogura	59	Aichi	Okazaki	HS4429; AB638500***	*Mwo*-Ia
Mogera wogura	60	Aichi	Kasugai	HS581; AB037625*	*Mwo*-Ia
Mogera wogura	61	Ishikawa	Mattoh	TUA1332/HS6003	*Mwo*-IIa
Mogera wogura	61	Ishikawa	Mattoh	HS1416***	*Mwo*-II
Mogera wogura	62	Fukui	Natashomura	NH11483/HS4776; LC554653****	*Mwo*-IIa
Mogera wogura	63	Shiga	Biwa Lake	MH11604/HS4884; LC554654****	*Mwo*-IIa
Mogera wogura	64	Shiga	South coast of Biwa Lake	MH11720/HS4925; LC554655****	*Mwo*-Ia-1
Mogera wogura	64	Shiga	South coast of Biwa Lake	MH11721/HS4926; LC554656****	*Mwo*-Ia-1
Mogera wogura	64	Shiga	South coast of Biwa Lake	MH11722/HS4927; LC554657****	*Mwo*-Ia-1
Mogera wogura	65	Shiga	Kusatsu	MH13056/HS5625; LC554658****	*Mwo*-Ia-1

Mogera wogura	65	Shiga	Kusatsu	MH43057/HS5626; LC554659****	*Mwo*-Ia-1
Mogera wogura	66	Mie	Ogaito	MH11723/HS4932; LC554660****	*Mwo*-Ia-1
Mogera wogura	67	Mie	Otowa	MH11726/HS4935; LC554661****	*Mwo*-Ia-1
Mogera wogura	68	Mie	Onogi	MH11717/HS4929; LC554662****	*Mwo*-Ia-1
Mogera wogura	68	Mie	Onogi	MH11712/HS4930; LC554663****	*Mwo*-Ia-1
Mogera wogura	68	Mie	Onogi	MH11713/HS4931; LC554664****	*Mwo*-Ia
Mogera wogura	69	Nara	Nara	MH11530/HS4898; LC554665****	*Mwo*-Ia
Mogera wogura	70	Nara	Yoshino	MH14355/HS5889	*Mwo*-Ia
Mogera wogura	71	Nara	Koryo	MH11463/HS4764; LC554668****	*Mwo*-Ia-2
Mogera wogura	72	Nara	Mt. Abe	MH11517/HS4894; LC554669****	*Mwo*-Ia-1
Mogera wogura	72	Nara	Mt. Abe	MH11518/HS4895; LC554670****	*Mwo*-Ia-1
Mogera wogura	43	Osaka	Kawachinagano	KT2767/HS314; AB638503***	*Mwo*-Ia
Mogera wogura	43	Osaka	Kawachinagano	MH7937/HS4623; LC554671****	*Mwo*-Ia
Mogera wogura	73	Kyoto	Kameoka	MH11474/HS4766; LC554672****	*Mwo*-IIa
Mogera wogura	73	Kyoto	Kameoka	MH11476/HS4767; LC554673****	*Mwo*-IIa
Mogera wogura	74	Kyoto	Ashiu	HS683; AB638501***	*Mwo*-IIa
Mogera wogura	74	Kyoto	Ashiu	HS810; AB638502***	*Mwo*-IIa
Mogera wogura	75	Kyoto	Wachi	MH11482/HS4774; LC554674****	*Mwo*-IIa
Mogera wogura	76	Kyoto	Fukuchiyama	MH11477/HS4768; LC554675****	*Mwo*-IIa
Mogera wogura	77	Hyogo	Toyooka	MH11298/HS4690; LC554676****	*Mwo*-IIa
Mogera wogura	78	Hyogo	Tanba	MH11478/HS4769; LC554677****	*Mwo*-IIa
Mogera wogura	78	Hyogo	Tanba	MH11479/HS4770; LC554678****	*Mwo*-IIa
Mogera wogura	79	Hyogo	Sasayama	MH11480/HS4771; LC554679****	*Mwo*-IIa
Mogera wogura	79	Hyogo	Sasayama	MH11481/HS4772; LC554680****	*Mwo*-IIa
Mogera wogura	80	Hyogo	Takarazuka	MH11271/HS4642; LC554681****	*Mwo*-Ia
Mogera wogura	81	Hyogo	Arimaonsen	MH11269/HS4641; LC554682****	*Mwo*-IIa
Mogera wogura	82	Hyogo	Kobe	HS4664; LC554683****	*Mwo*-IIa
Mogera wogura	83	Hyogo	Aioi	HS4582; LC554684****	*Mwo*-IIa
Mogera wogura	84	Hyogo	Awaji Is.	MH11279/HS4657; LC554685****	*Mwo*-Ia
Mogera wogura	85	Tottori	Tohaku	MH14180/HS5629; LC554686****	*Mwo*-IIa
Mogera wogura	86	Tottori	Yonago	HA56115/HS4432; AB638505***	*Mwo*-IIa
Mogera wogura	87	Okayama	Maniwa	MH14176/HS5627; LC554687****	*Mwo*-IIa
Mogera wogura	87	Okayama	Maniwa	MH414177/HS5628; LC554688****	*Mwo*-IIa
Mogera wogura	88	Okayama	Akaiwa	HS4585; LC554689****	*Mwo*-IIa
Mogera wogura	88	Okayama	Akaiwa	HS4586; LC554690****	*Mwo*-IIa

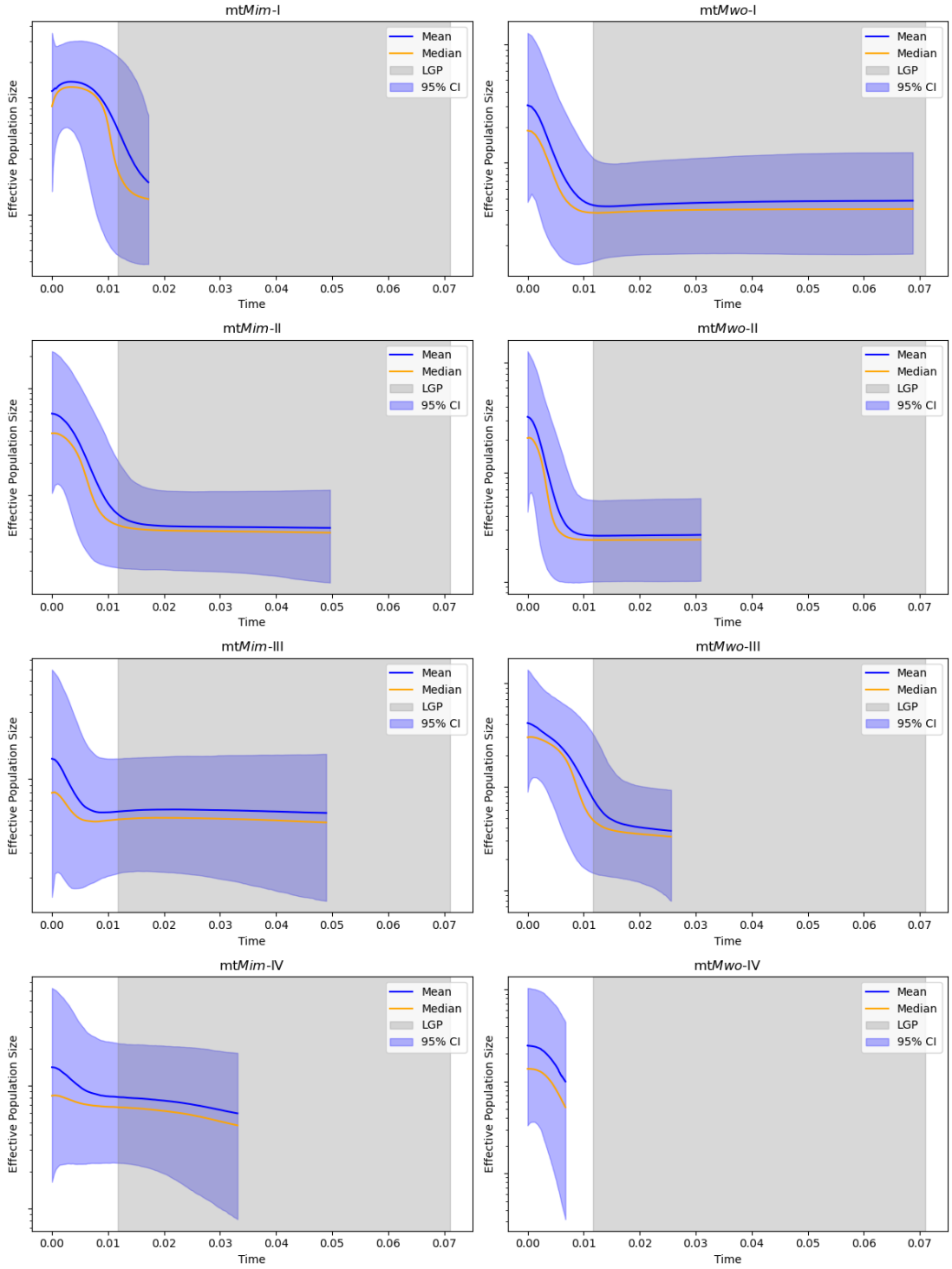
Mogera wogura	49	Shimane	Dogo, Oki Is.	HA56101/HS4433; AB638504***	*Mwo*-IIb
Mogera wogura	49	Shimane	Dogo, Oki Is.	MH14485	*Mwo*-IIb
Mogera wogura	49	Shimane	Dogo, Oki Is.	MH14486	*Mwo*-IIb
Mogera wogura	49	Shimane	Dogo, Oki Is.	MH14487	*Mwo*-IIb
Mogera wogura	49	Shimane	Dogo, Oki Is.	MH14488	*Mwo*-IIb
Mogera wogura	49	Shimane	Dogo, Oki Is.	MH14499	*Mwo*-IIb
Mogera wogura	89	Shimane	Matsukawa	MH13052/HS5153; LC554691****	*Mwo*-IIa
Mogera wogura	89	Shimane	Matsukawa	MH13053/HS5154; LC554692****	*Mwo*-IIa
Mogera wogura	90	Shimane	Ninomiya	MH13054/HS5155; LC554693****	*Mwo*-IIb
Mogera wogura	90	Shimane	Ninomiya	MH13055/HS5156; LC554694****	*Mwo*-IIb
Mogera wogura	91	Shimane	Nagaya	MH13407/HS5342; LC554695****	*Mwo*-IIa
Mogera wogura	91	Shimane	Nagaya	MH13408/HS5343; LC554696****	*Mwo*-IIa
Mogera wogura	92	Shimane	Kanagi	MH13605/HS5362; LC554697****	*Mwo*-IIa
Mogera wogura	92	Shimane	Kanagi	MH13606/HS5363; LC554698****	*Mwo*-IIb
Mogera wogura	93	Shimane	Kochi	MH13603/HS5360; LC554699****	*Mwo*-IIa
Mogera wogura	93	Shimane	Kochi	MH13604/HS5361; LC554700****	*Mwo*-IIa
Mogera wogura	94	Shimane	Sufu	MH14024/HS5597; LC554701****	*Mwo*-IIIa
Mogera wogura	95	Shimane	Uchida	MH13691/HS5423; LC554702****	*Mwo*-IIIb
Mogera wogura	96	Shimane	Masuda:Umetsuki	MH13689/HS5421; LC554703****	*Mwo*-IIIb
Mogera wogura	96	Shimane	Masuda:Umetsuki	MH13690/HS5422; LC554704****	*Mwo*-IIIe
Mogera wogura	97	Hiroshima	Hiwa	HA55944/HS4430; AB638506***	*Mwo*-IIa
Mogera wogura	98	Hiroshima	Kimita	MH14017/HS5479; LC554705****	*Mwo*-IIa
Mogera wogura	98	Hiroshima	Kimita	MH14018/HS5480; LC554706****	*Mwo*-IIa
Mogera wogura	99	Hiroshima	Mirasaka	MH14023/HS5485; LC554707****	*Mwo*-IIa
Mogera wogura	100	Hiroshima	Kisa	MH14019/HS5481; LC554708****	*Mwo*-IIa
Mogera wogura	100	Hiroshima	Kisa	MH14020/HS5482; LC554709****	*Mwo*-IIa
Mogera wogura	100	Hiroshima	Kisa	MH14021/HS5483; LC554710****	*Mwo*-IIa
Mogera wogura	100	Hiroshima	Kisa	MH14022/HS5484; LC554711****	*Mwo*-IIa
Mogera wogura	101	Hiroshima	Uzuto	MH14043/HS5604; LC554712****	*Mwo*-IIa
Mogera wogura	101	Hiroshima	Uzuto	MH14044/HS5605; LC554713****	*Mwo*-IIa
Mogera wogura	102	Hiroshima	Fuchu	MH13042/HS5143; LC554714****	*Mwo*-IIa
Mogera wogura	102	Hiroshima	Fuchu	MH13043/HS5144; LC554715****	*Mwo*-IIa
Mogera wogura	102	Hiroshima	Fuchu	MH13044/HS5145; LC554716****	*Mwo*-IIa
Mogera wogura	103	Hiroshima	Fukuyama	HS5425; LC554717****	*Mwo*-IIa
Mogera wogura	104	Hiroshima	Mihara	HS5323; LC554718****	*Mwo*-IIIb

Mogera wogura	104	Hiroshima	Mihara	HS5324; LC554736****	*Mwo*-IIIe
Mogera wogura	105	Hiroshima	Kui	MH14045/HS5606; LC554719****	*Mwo*-IIa
Mogera wogura	105	Hiroshima	Kui	MH14046/HS5607; LC554720****	*Mwo*-IIIb
Mogera wogura	106	Hiroshima	Miwa	MH43614/HS5369; LC554722****	*Mwo*-IIIe
Mogera wogura	107	Hiroshima	Koda	MH13613/HS5368; LC554723****	*Mwo*-IIIe
Mogera wogura	108	Hiroshima	Yoshida	MH13612/HS5367; LC554724****	*Mwo*-IIId
Mogera wogura	109	Hiroshima	Midori	MH13045/HS5146; LC554725****	*Mwo*-IIa
Mogera wogura	109	Hiroshima	Midori	MH13046/HS5147; LC554726****	*Mwo*-IIa
Mogera wogura	110	Hiroshima	Yachiyo	MH13047/HS5148; LC554727****	*Mwo*-IIId
Mogera wogura	111	Hiroshima	Oasa	MH13409/HS5344; LC554728****	*Mwo*-IIa
Mogera wogura	111	Hiroshima	Oasa	MH13410/HS5345; LC554729****	*Mwo*-IIb
Mogera wogura	112	Hiroshima	Totani	MH13608/HS5365; LC554730****	*Mwo*-IIIc
Mogera wogura	112	Hiroshima	Totani	MH13609/HS5366; LC554731****	*Mwo*-IIIc
Mogera wogura	113	Hiroshima	Tahara	MH13607/HS5364; LC554732****	*Mwo*-IIIb
Mogera wogura	114	Hiroshima	Togochi	HA55940/HS4449; AB638507***	*Mwo*-IIIb
Mogera wogura	114	Hiroshima	Togochi	HA55943/HS4450; AB638508***	*Mwo*-IIIe
Mogera wogura	115	Hiroshima	Shiraki	MH13048/HS5149; LC554733****	*Mwo*-IIId
Mogera wogura	115	Hiroshima	Shiraki	MH13049/HS5150; LC554734****	*Mwo*-IIId
Mogera wogura	116	Hiroshima	Higashihiroshima	HS5322; LC554735****	*Mwo*-IIa
Mogera wogura	116	Hiroshima	Higashihiroshima	HS5424; LC554705****	*Mwo*-II
Mogera wogura	117	Yamaguchi	Mizugasako	MH12401/HS5157; LC554737****	*Mwo*-IIIb
Mogera wogura	118	Yamaguchi	Waki	MH11524/HS4899; LC554738****	*Mwo*-IIIb
Mogera wogura	118	Yamaguchi	Waki	MH11525/HS4900; LC554739****	*Mwo*-IIIb
Mogera wogura	119	Yamaguchi	Nagato	MH11084/HS4606; LC554740****	*Mwo*-IIIc
Mogera wogura	119	Yamaguchi	Nagato	MH11528/HS4896; LC554741****	*Mwo*-IIIb
Mogera wogura	119	Yamaguchi	Nagato	MH11529/HS4897; LC554742****	*Mwo*-IIIc
Mogera wogura	119	Yamaguchi	Nagato	MH14350/HS5887	*Mwo*-IIIb
Mogera wogura	119	Yamaguchi	Nagato	MH14351/HS5888	*Mwo*-IIIb
Mogera wogura	120	Yamaguchi	Shimonoseki	HS4448; AB638509***	*Mwo*-IIIb
Mogera wogura	120	Yamaguchi	Shimonoseki	MH11526/HS4901; LC554743****	*Mwo*-IIIb
Mogera wogura	120	Yamaguchi	Shimonoseki	MH11527/HS4902; LC554744****	*Mwo*-IIIc
Mogera wogura	120	Yamaguchi	Shimonoseki	MH14047/HS5608; LC554745****	*Mwo*-IIIb
Mogera wogura	121	Kagawa	Mannou	MH11782/HS4951; LC554746****	*Mwo*-IIc
Mogera wogura	51	Kagawa	Uchinomi	MH14427/HS5994	*Mwo*-Ia
Mogera wogura	51	Kagawa	Uchinomi	MH14428/HS5995	*Mwo*-Ia

Mogera wogura	122	Tokushima	Mima	MH5921/HS4605; LC554747****	*Mwo*-IIc
Mogera wogura	123	Kochi	Kagamiogachi	MH11778/HS4947; LC554748****	*Mwo*-IIc
Mogera wogura	124	Kochi	Shimantonakamura	MH11780/HS4949; LC554749****	*Mwo*-IIc
Mogera wogura	125	Ehime	Iyo	HS5500; LC554750****	*Mwo*-IIc
Mogera wogura	126	Nagasaki	Tsushima Is.	KT2711/HS366; AB638510***	*Mwo*-IIIe
Mogera wogura	126	Nagasaki	Tsushima Is.	KT3473/HS2820; AB638511***	*Mwo*-IIIe
Mogera wogura	126	Nagasaki	Tsushima Is.	HA55911/HS4434; AB638512***	*Mwo*-IIIe
Mogera wogura	127	Fukuoka	near Fukuoka airport	HS5501; LC554751****	*Mwo*-IIIb
Mogera wogura	128	Fukuoka	Ukiha	HA55915/HS4437; AB638513***	*Mwo*-IIIa
Mogera wogura	128	Fukuoka	Ukiha	HA55916/HS4438; AB638514***	*Mwo*-IIIa
Mogera wogura	129	Fukuoka	Kurume	HA55919/HS4439; AB638515***	*Mwo*-IIIa
Mogera wogura	130	Kumamoto	Yatsushiro	HA55922/HS4441; AB638519***	*Mwo*-IIIa
Mogera wogura	131	Kumamoto	Amakusa	YK05/HS4428; AB638518***	*Mwo*-IIIa
Mogera wogura	132	Kumamoto	Ashikita	HA55926/HS4442; AB638517***	*Mwo*-IIIa
Mogera wogura	133	Kumamoto	Menda	HS414; AB638516***	*Mwo*-IIIa
Mogera wogura	134	Kumamoto	Hitoyoshi	HA55932/HS4446; AB638520***	*Mwo*-IIIa
Mogera wogura	134	Kumamoto	Hitoyoshi	HS4447; AB638521***	*Mwo*-IIIa
Mogera wogura	135	Miyazaki	Kiyotake	KT3276/HS3066; AB638522***	*Mwo*-IIIa
Mogera wogura	136	Kagoshima	Makizono	KT2699/HS369; LC554209***	*Mwo*-IIIa
Mogera wogura	137	Kagoshima	Uchinoura	HS415; LC554210***	*Mwo*-IIIa
Mogera wogura	138	Kagoshima	Tanegashima Is	HA55930/HS4445; AB638523***	*Mwo*-IIIa
Mogera wogura	139	Kagoshima	Yakushima Is.	HA55926/HS4443; AB638524***	*Mwo*-IIIa
Mogera wogura	139	Kagoshima	Yakushima Is.	HA55928/HS4444; AB638525***	*Mwo*-IIIa
M. robusta	140	Russia	Krasnoarmeysky	AK884; AB638530***	*Mwo*-IV
M. robusta	141	Russia	Ussuriisk	AK820/HS2004; AB638528***	*Mwo*-IV
M. robusta	141	Russia	Ussuriisk	AK821/HS2005; AB638529***	*Mwo*-IV
M. robusta	142	Russia	Shkotovo	AK002/HS2009****	*Mwo*-IV
M. robusta	143	Russia	Vladivostok	AK001/HS2008; LC554752****	*Mwo*-IV
M. robusta	144	Russia	Kedrovaya	AK002/HS1170; AB037646*	*Mwo*-IV
M. robusta	144	Russia	Kedrovaya	AK609/HS2001; LC554753****	*Mwo*-IV
M. robusta	145	Russia	Hasan	AK707/HS1392; AB638527***	*Mwo*-IV
M. robusta	146	China	Liaoning	LN111102; HG737874*****	*Mwo*-IV
M. robusta	146	China	Liaoning	LN111103; HG737875*****	*Mwo*-IV
M. robusta	146	China	Liaoning	LN111104; HG737873*****	*Mwo*-IV
M. robusta	147	Korea	Seoraksan	HS930; AB037641*	*Mwo*-IV

M. robusta	148	Korea	Pusan	KT2755/HS373; AB638526***	*Mwo*-IV
M. robusta	149	Korea	Namphae	HS929; AB037640*	*Mwo*-IV
M. tokudae	150	Niigata	Sado Is.	KT2672/HS317 AB037607	
M. etigo	151	Niigata	Niigata	HS667 AB638532.1	

*Tsuchiya et al. 2000, **Iwasa et al. 2006, ***Kirihara et al. 2013, ****Nakamoto et al. 2021, *****He et al. 2014, *******Tsunoi et al. 2023



Supplementary Data 1. Baysian Skyline Plots for all *Cytb* haplogroups of *M. imaizumii* and *M. wogura*.

References

Abe H. 1968. Classification and biology of Japanese insectivore (Mammalia): II. Biological aspects. Journal of the Faculty of Agriculture, Hokkaido University 55: 429–458.

Abe H. 1974. Change of the boundary-line of two moles' distributions in a period of 14 years. The Journal of the Mammalogical Society of Japan 6: 13–23 (in Japanese with English summary).

Abe H. 2001. Isolated relic populations and their keeping mechanisms in moles. Mammalian Science 41: 35–52 (in Japanese with English summary).

Abe H. 2010. The northeastern from in the distribution of *Mogera wogura* in the central Honshu, Japan in 2009, especially the past 50 years' change in Nagano Prefecture. Mammalian Science 50: 13–23 (in Japanese with English summary).

Abe H. 2008. Insectivora. In (Abe H, Ishii N, Itoo T, Kaneko Y, Maeda K, Miura S, and Yoneda M, eds.) A Guide to the Mammals of Japan, Second revised edition, pp. 21–24. Tokai University Press, Kanagawa.

Bandelt H, Forster P, and Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37–48.

Bouckaert R, Heled J, Kuhnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, and Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology 10: e1003537.

Hanazaki K, Tomozawa M, Suzuki Y, Kinoshita G, Yamamoto M, Irino T, and Suzuki H. 2017. Estimation of evolutionary rates of mitochondrial DNA in two Japanese wood mouse species based on calibrations with Quaternary environmental changes. Zoological Science 34: 201–210.

Honda A, Murakami S, Harada M, Tsuchiya K, Kinoshita G, and Suzuki H. 2019. Late Pleistocene climate change and population dynamics of Japanese *Myodes* voles inferred from mitochondrial cytochrome *b* sequences. Journal of Mammalogy 100: 1156–1168.

Iwasa MA, Kawakubo C, Tsuchiya K, and Suzuki H. 2006. Intraspecific differentiation in the lesser Japanese mole in eastern Honshu, Japan, indicated by nuclear and mitochondrial gene analyses. Zoological Science 23: 955–961.

Kawada S and Yokohata Y. 2009. *Mogera imaizumii* (Kuroda, 1957). In (Ohdachi SD, Ishibashi Y, Iwasa MA, and Saitoh T eds.) The Wild Mammals of Japan, pp. 34–35.

Kawada S and Yokohata Y. 2009. *Mogera wogura* (Temminck, 1842). In (Ohdachi SD, Ishibashi Y, Iwasa MA, and Saitoh T eds.) The Wild Mammals of Japan, pp. 36–37.

Kirihara T, Shinohara A, Tsuchiya K, Harada M, Kryukov AP, and Suzuki H. 2013. Spatial and temporal aspects of the occurrence of *Mogera* species in the Japanese Islands inferred from mitochondrial and nuclear gene sequences. Zoological Science 30: 267–281.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, and Wilson AC. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceeding of Natural Academy of Sciences, USA 86: 6196–6200.

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549.

Kushihashi H. 1982. Occurrence of two species in *Mogera* from the Is. Of shodo-shima, Kagawa Prefecture, Japan. Kagawa Seibutsu 10: 43–51 (in Japanese).

Millien-Parra V and Jaeger JJ. 1999. Island biogeography of the Japanese terrestrial mammal assemblages: an example of a relict fauna. Journal of Biogeography 26: 959–972.

Mitsuhashi R, Tsunoi T, Harada M, Yachimori S, Yano S, and Suzuki H. 2025. Evolutionary history of geographically isolated populations of the lesser Japanese mole (*Mogera imaizumii*) inferred from mitochondrial cytochrome b gene sequences. Mammal Study 50

Moribe J and Yokohata Y. 2011. Mixed distribution of *Mogera imaizumii* and *Mogera wogura* in the Tedori Alluvial Fan in Ishikawa prefecture, Japan. Mammal Study 36: 135–139.

Nakamoto A, Harada M, Mitsuhashi R, Tsuchiya K, Kryukov AP, Shinohara A, and Suzuki H. 2021. Influence of Quaternary environmental changes on mole populations inferred from mitochondrial sequences and evolutionary rate estimation. Mammal Study 7:1–11.

Okamoto M. 1999. Phylogeny of Japanese moles inferred from mitochondrial CO1 gene sequences. In (Yokohata Y and Nakamura S, eds.) Recent Advances in the Biology of Japanese Insectivora, pp. 21–27. Hiba Society of Natural History, Shobara.

Oshida T, Masuda R, and Ikeda K. 2009. Phylogeography of the Japanese giant flying squirrel, *Petaurista leucogenys* (Rodentia: Sciurida): implication of glacial refugia in an arboreal small mammal in the Japanese Islands. Biological Journal of the Linnean Society 98: 47–60.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, and Sánchez-Garcia A. 2017. DnaSP v6: DNA Sequence Polymorphism Analysis of Large Datasets. Molecular Biology and Evolution 34: 3299–3302.

Sato JJ. 2017. A review of the processes of mammalian faunal assembly in Japan: insights from molecular phylogenetics. In (Motokawa M and Kajihara H eds.) Species Diversity of Animals in Japan, pp. 49–60.

Shinohara A, Campbell KL, and Suzuki H. 2005. An evolutionary view on the Japanese talpids based on nucleotide sequences. Mammal Study 30: S19–S24.

Suzuki Y, Tomozawa M, Koizumi Y, Tsuchiya K, and Suzuki H. 2015. Estimating the molecular evolutionary rates of mitochondrial genes referring to Quaternary ice age events with inferred population expansions and dispersals in Japanese *Apodemus*. BMC Evolutionary Biology 15: 1–17.

Tamura K and Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10: 512–526.

Tanioka H. 2020. A collection of lesser Japanese mole (*Mogera imaizumii*) from Kami-shi, Kochi Prefecture and recent records in the surrounding area. Shikoku Natural History Research 13: 24–33 (in Japanese with English summary).

Tsuchiya K, Suzuki H, Shinohara A, Harada M, Wakana S, Sakaizumi M, Han SH, Lin LK, and Kryukov AP. 2000. Molecular phylogeny of east Asian moles inferred from the sequence variation of the mitochondrial cytochrome *b* gene. Genes & Genetic Systems 75: 17–24.

Tsunoi T, Harada M, Notsu M, Mitsuhashi R, and Suzuki H. 2023. First record for the lesser Japanese mole *Mogera imaizumii* (Soricomorpha, Talpidae) from Dogo, Oki Islands, Japan. Mammalian Science 63: 62–68 (in Japanese with English summary).

Wu J, Kohno N, Mano S, Fukumoto Y, Tanabe H, Hasegawa M, and Yonezawa T. 2015. Phylogeographic and demographic analysis of the Asian black bear (*Ursus thibetanus*) based on mitochondrial DNA. PLoS ONE 10: e0136398. DOI: 10.1371/journal.pone.0136398.

Yokohata Y. 2005. A brief review of the biology on moles in Japan. Mammal Study 30: S25–S30.

Zemlemerova E, Abramov A, Kryukov A, Lebedev V, Min MS, Lee SF, and Bannikova A. 2019. Genetic and morphologic diversity of the moles (Talpomorpha, Talpidae, *Mogera*) from the continental Far East. Journal of Zoological Systematics and evolutionary Research 57: 662–678.

PART III: Insights into the population structure of the lesser Japanese mole (*Mogera imaizumii*) and the large Japanese mole (*M. wogura*) revealed by reduced representation genome analysis

Introduction

Mogera imaizumii and *M. wogura* have inhabited the Japanese Archipelago for approximately 2.4 and 1.3 (or 1.6) million years, respectively (Kirihara et al. 2013; Sato 2017). These timelines encompass periods of substantial environmental shifts, such as the temperature and sea-level fluctuations related to the Quaternary 100,000-year glacial cycle. Their broad habitat range makes these two mole species suitable for analyzing how topographic and climatic changes have shaped the population structure of subterranean species in Japan. Prior studies using mitochondrial and limited nuclear gene data suggest that both species underwent severe population bottlenecks during the Last Glacial Maximum (LGM) and the Penultimate Glacial Maximum, both marked by extremely cold temperatures (Nakamoto et al. 2021). This evolutionary history makes them ideal subjects for exploring factors that influence population differentiation in subterranean animals. Based on mitochondrial and nuclear gene data, Kirihara et al. (2013) proposed that warming during the Quaternary drove an eastward shift in the distribution border of the moles, leaving isolated populations of *M. imaizumii* in small areas of western Japan, including the mountainous areas of the Kinki area (including the Kii Peninsula), the Chugoku area, the Shikoku area, Shodoshima Island, and small islands, such as Oki-Dogo Island (Tsunoi et al. 2023) (Fig. 9). However, this hypothesis remains untested on a large molecular scale, leaving uncertainty about whether the eastward migration of *M. imaizumii* and subsequent replacement by *M. wogura* resulted in a fragmented distribution of *M. imaizumii* in western Japan.

Previous studies have indicated that both *M. imaizumii* and *M. wogura* consist of several genetically distinct subpopulations (Okamoto 1997; Tsuchiya et al. 2000). Iwasa et al. (2006) identified two lineages of *M. imaizumii* in eastern Honshu, generally distributed along the western and eastern coasts, with some gene flow occurring between these regions. These lineages appear to be separated mainly by high mountain ranges, particularly the Ou and Kitakami ranges. Phylogeographic studies of *M. imaizumii* and *M. wogura* across Japan have been advanced using a broad sampling of mitochondrial DNA (mtDNA) and nuclear

gene sequences (Okamoto 1997; Tsuchiya et al. 2000; Kirihara et al. 2013; Zemlemerova et al. 2019; Nakamoto et al. 2021). Analyses of the mitochondrial cytochrome *b* (*Cytb*) gene have revealed that both *M. imaizumii* and *M. wogura* are composed of four geographically distinct mitochondrial lineages (Supplementary Data S2). For *M. imaizumii*, these lineages, labeled mt*Mim*-I to IV, are distributed in specific regions: the Japan Sea side (Tohoku/Hokuriku areas), Pacific Ocean side (Kanto/Tohoku areas), Kinki area, and an isolated island in western Japan known as Oki-Dogo (Tsunoi et al. 2023). For *M. wogura*, the four lineages are mt*Mwo*-I found in the Tokai and southern Kinki (sKinki) areas; mt*Mwo*-II located in the Hokuriku, northern Kinki (nKinki), and eastern Chugoku (eChugoku) areas; mt*Mwo*-III found in western Chugoku (wChugoku) and Kyushu areas, and mt*Mwo*-IV, sometimes identified as a separate species, *M. robusta*, along the Eurasian coastline, including Russia and Korea (Nakamoto et al. 2021). Nakamoto et al. (2021) determined genetically distinguishable borders between mitochondrial lineages and identified a geographical boundary between mt*Mwo*-I and mt*Mwo*-II in the Osaka plain and in the Rokko and Hokusetsu mountains. However, no physical barrier was found between mt*Mwo*-II and mt*Mwo*-III, leaving the cause of their genetic differentiation unresolved. Additionally, previous studies struggled to clarify the fine-scale population differentiation due to limited molecular data and reliance on phenotypic data, such as morphology, which did not provide a complete picture of the population structures.

Recent advances in next-generation sequencing technology have enabled the collection of extensive data from non-model organisms, moving beyond the traditional reliance on classical genetic markers such as mitochondrial DNA, nuclear gene-coding regions, and microsatellites (e.g. Kirihara et al. 2013; Satoh et al. 2019; Nakamoto et al. 2021). Methods for reduced genome representation, such as double-digest restriction site-associated DNA sequencing (ddRAD-seq, Peterson et al. 2012), genotyping by random amplicon sequencing-direct (GRAS-Di, Hosoya et al. 2019), and multiplexed inter-simple sequence repeat genotyping by sequencing (MIG-seq, Suyama and Matsuki 2015), are increasingly used in genetic research on Japanese mammals (Ito et al. 2021; Sato and Yasuda 2022; Kinoshita et al. 2024). Among these, MIG-seq is particularly advantageous due to its cost-effectiveness and ability to process large sample sizes, yielding a high volume of single nucleotide polymorphisms (SNPs), that addresses the limited variation in earlier genetic studies. Additionally, SNP count in MIG-seq data can be enhanced using information from a

reference genome sequence (Shafer et al. 2017; Kunvar et al. 2021). Currently, no reference genome is available for the genus *Mogera*. Therefore, in this study, we developed a reference genome for *M. wogura* to increase the SNP yield in MIG-seq analysis. Using this newly assembled genome sequence, we explored the phylogeographic structure of *M. imaizumii* and *M. wogura*, focusing on 1) analyzing the discordance between mitochondrial and nuclear phylogeographic patterns, 2) examining how geographic barriers influence gene flow and introgression among populations, and 3) identifying signs of recent shifts in species boundaries between *M. imaizumii* and *M. wogura*.

Materials and Methods

Whole genome sequencing and assembling the draft genome

An *M. wogura* sample collected in Nara, Japan, was used for whole genome sequencing and assembly. The samples were preserved at minus 80°C. The brain, liver, lung, and muscle tissues were used for high-molecular weight (HMW) DNA extraction. HMW DNA was extracted using the Nanobind Tissue Big DNA Kit (Circulomics Inc. #900-001-01) following the manufacturer's protocol, and fragmented DNA was removed using the Short Read Eliminator (SRE) XL Kit (Circulomics Inc #SS-100-111-01). The supernatant from the first wash was also centrifuged and washed twice (double SRE; DSRE), followed by elution with 20 μL elution buffer. Ten samples of the extracted HMW DNA with no SRE, SRE and washed supernatants of SRE were subjected to pulse-field gel electrophoresis (PFGE) using a CHEF-MAPPER system (Bio-Rad). Based on the PFGE band patterns and DNA integrity number value of the TapeStation (Agilent Technologies), the highest HMW DNA derived from the brain sample was selected for further analysis. A linked-read library was constructed from 1.25 ng of >50 kb HMW DNA using the 10x genomics chromium system and Chromium™ Genome Reagent Kits v2 (10x Genomics) following the manufacturer's protocol. The final constructed library was quantified using a TapeStation, Bioanalyzer (Agilent Technologies), and Qubit Fluorometer (Thermo Fisher Scientific), and sequenced on an Illumina HiSeq X platform (Illumina, San Diego, USA) with 150 paired-end reads. Sequencing reads were assembled using Supernova version 2.1.1 (Weisenfeld et al. 2017) with default settings. The sequenced reads were deposited in the DNA Data Bank of Japan Center (DDBJ) and are available under accession numbers BAABMH010000001–BAABMH010020570.

MIG-seq library preparation and SNP data generation

DNA was extracted from tissue samples (liver, muscle, or bone) preserved in 99.8% ethanol using a QIAmp DNA Mini Kit (QIAGEN, Hilden, Germany). DNA extracted from 178 samples (*M. imaizumii*: 89, *M. wogura*: 89) collected between 1995 and 2022 were used for MIG-seq analyses. The samples used in this study are shown in Appendices I and II. The DNA samples are deposited at the Botanic Garden & Museum, Field Science Center for Northern Biosphere, Hokkaido University, Sapporo (HUNHM; appendices I, II). The sampling sites covered the main habitats of *M. imaizumii* and *M. wogura*, including the coastal areas of the Asian Continent, where *M. wogura* is sometimes referred to as *M. robusta*. Library construction for MIG-seq was performed following the protocol of Suyama and Matsuki (2015), but with the annealing temperature of the first PCR set to 38°C. The PCR products were run using the MultiNA system (Shimadzu Corporation, Kyoto, Japan) after the first and second PCR to determine whether the DNA was successively amplified. The second PCR products were then pooled and fragments of 300–800 bp were selected using SPRIselect (Beckman Coulter). The constructed library was sequenced with 76 bp paired-end reads using the MiSeq system (Illumina) with the MiSeq Reagent Kit v3 (150 cycles) and 2% PhiX spike-in. Raw FASTQ data from MIG-seq were deposited in DDBJ and are available under accession numbers PRJDB17819 (DRR545064–DRR545159, DRR545214–DRR545312).

Raw data obtained from MiSeq system were trimmed using Trimmomatic version 0.39 (Bolger et al. 2014). The first 5 bp and the last 1 bp of the reads were trimmed from the raw data using the following parameters: SLIDINGWINDOW, 4:15; CROP, 75; HEADCROP, 5; and MINLEN, 70. The trimmed reads were then mapped onto the *M. wogura* reference genome using the Burrows-Wheeler aligner-maximum exact matches (BWA-MEM) program version 0.7.17 (Li 2013), converted to BAM format by the SAMtools program version 1.7 (Li et al. 2009; Danecek et al. 2021), and assigned to read-groups using the AddOrReplaceReadGroups function of Picard (Broad Institute, https://broadinstitute.github.io/picard/). Contig generation and SNP detection were performed using *gstacks* analysis implemented in Stacks 2.62 (Catchen et al. 2011; Catchen et al. 2013). Contigs that were not shared by 80% of each population and alleles with allele

frequencies < 0.01 or heterozygosity > 0.5 were filtered out by *populations*, also implemented in Stacks. Upon conducting *populations*, the individuals were assigned to each population based on their sampling sites and mtDNA lineages defined in Nakamoto et al. (2021). Subsequent filtering of samples and SNP sites was performed using TASSEL version 5.2.79 (Bradbury et al. 2007). This process filtered out samples that retained < 80% and < 70% of the SNPs for *M. wogura* and *M. imaizumii*, respectively, as well as SNP sites where minor alleles were observed in only one sample. The final all-SNP dataset included 86 samples and 2988 variable sites for *M. imaizumii* and 86 samples and 3550 variable sites for *M. wogura*. Different SNP datasets were simultaneously created for both species by randomly extracting a single SNP per locus by adding a --write-random-snp flag by *populations*. Filtering of samples and SNP sites was also conducted for these datasets under the same conditions, resulting in 86 samples and 1005 variable sites for *M. imaizumii*, and 86 samples and 998 variable sites for *M. wogura* (single SNP per locus dataset). These single SNP per locus datasets were created for analysis on population structure, since fastSTRUCTURE does not consider linkage disequilibrium between genetic markers (Raj et al. 2014). Conversion of SNP datasets in VCF format into STRUCTURE, FASTA, PHYLIP, and NEXUS formats was conducted using PGDSpider 2.1.1.5 (Lischere and Excoffier 2012).

Analysis of population structure

Population assignment analysis was performed using fastSTRUCTURE (Raj et al. 2014) with K values (number of clusters) of 2–10 over 20 independent runs for each K value using the single SNP pre locus dataset. The most likely number of clusters was evaluated using the chooseK.py function implemented in the fastSTRUCTURE. Principal component analysis (PCA) was performed using the PLINK program 1.90b6.21 version (Chang et al. 2015; Purcell and Chang; www.cog-genomics.org/plink/1.9) with the following options: --pca --header 20 --double-id --allow-extra-chr. PCA results were visualized using the R program 4.1.3 version. The phylogenetic relationships within species were visualized using Neighbor-Net analysis implemented in SplitsTree4 4.18.2 version (Huson and Bryant, 2006). PCA and Neighbor-Net analyses were performed on the all-SNP dataset. The interspecies fixation index (F_{ST}) and intraspecies pairwise F_{ST} values were calculated using the SMARTPCA algorithm implemented in the EIGENSOFT package (Patterson et al. 2006; Price et al. 2006). Average heterozygosity was calculated for each species and genetic cluster

by outputting all sites to a vcf file using *populations* implemented in Stacks 2.67 version (Catchen et al. 2011; Catchen et al. 2013) with --vcf-all option. Maximum Likelihood (ML) tree was constructed with RAxML-NG 1.2.2 version (Kozlov et al. 2019). The evolutionary model was selected by ModelTest-NG 0.1.7 version (Flouri et al. 2014; Darriba et al. 2020). Invariant sites were removed using the raxml_ascbias.py script obtained from https://github.com/btmartin721/raxml_ascbias. Under the TVM+G4 model, ML tree was constructed with 1000 bootstrap replicates.

Results

Genomic data assembly

The constructed daft genome assembly was 1.85 Gb in size with 58.48 x raw coverage and the estimated genome size was 2.09 Gb. It comprised 20,579 scaffolds, with 1.63 K scaffolds longer than 10 Kb. The N50 contig size was 109.18 Kb and the N50 scaffold size was 41.0-2 Mb. We identified 11,144 (91.1%) complete single-copy BUSCOs and 183 (1.5%) complete duplicated BUSCOs using BUSCO version 5.3.2 (Manni et al. 2021). By mapping MIG-seq reads to the draft genome, we analyzed MIG-seq data from 87 *M. imaizumii* and 86 *M. wogura* samples. In total, we obtained 2988 and 3550 SNPs in *M. imaizumii* and *M.wogura*, respectively. The mean effective per-sample coverage at SNP sites was 6.0x for *M. imaizumii* (ranging from 1.8x to 19.1x) and 6.8x for *M. wogura* (ranging from 2.4x to 17.8x). The average rates of mapped reads were 97.454% for *M. imaizumii* and 97.93% for *M. wogura*, suggesting that the reference genome of *M. wogura*, despite being slightly less efficient, is suitable for analyzing *M. imaizumii*. The two species shared only 381 SNPs, and the F_{ST} value between *M. imaizumii* and *M. wogura* was 0.868, indicating that they were genetically well-differentiated, although they were sister species. We first constructed a Neighbor-Net graph and confirmed that the two species were genetically well separated and differentiated without any sign of hybridization and introgression (Supplementary Data S3). The ML tree also showed that the two species are separated with a high bootstrap score (bootstrap score rate: 100, Supplementary Data S9). Therefore, we present the results of the genetic differentiation among the populations of each species in the subsequent subsections.

Genetic structure of M. imaizumii

PCA based on the SNPs of *M. imaizumii* showed that PC1 and PC2 segregated the samples into approximately four to six groups (Fig. 10A). PC1 separated the populations

from the sKinki area (Osaka, Nara, Wakayama, and Mie) and isolated populations (Shodoshima, Oki, and Hiroshima) from other samples. Individuals from the nKinki area (Shiga, Kyoto, Hyogo) were grouped with other individuals from the Japan Sea coast (JS; Aomori, Iwate, Akita, Yamagata, Niigata, Nagano, Toyama, and Ishikawa) and differentiated from those from the Pacific Ocean coast (PO; Miyagi, Fukushima, Ibaraki, Chiba, Gunma, Kanagawa, Yamanashi, Shizuoka) on PC2. PC3 and PC4 further separated the populations from Oki, Shodoshima, and Hiroshima from the sKinki group (Supplementary Data S4). PC5 distinguished individuals from JS and nKinki by grouping the Ishikawa population together with the nKinki population, whereas samples from Toyama were interspersed between JS and nKinki (Supplementary Data S4). The Neighbor-Net graph presented in Fig. 11 is in good agreement with the PCA results, in which the individuals were separated into three to six major clades. The results of fastSTRUCTURE analyses showed that the value of K that maximized the marginal likelihood was three for *M. imaizumii*. The populations of JS/nKinki, PO, and sKinki (Hiroshima, Shodoshima, and Oki) were mainly composed of nc*Mim*-1, nc*Mim*-2, and nc*Mim*-3, respectively (Fig. 12A). Although most samples were assigned to a single genetic cluster, an individual from Hiroshima showed signs of admixed ancestry between nc*Mim*-1 and nc*Mim*-3 (Fig. 12A). The sKinki populations were similar to the three isolated populations (Oki, Shodoshima, and Hiroshima) in terms of nc*Mim*-3, indicating that the nKinki/sKinki border is also a border between nc*Mim*-1/nc*Mim*-3 genetic groups. Pairwise F_{ST} values were high, ranging from 0.376 to 0.487 (Supplementary Data S5). The average heterozygosity is summarized in Supplementary Data S7.

Genetic structure of M. wogura

For *M. wogura*, PC1 and PC2 separated the populations from the Tokai/sKinki, Hokuriku/nKinki/Chugoku/Shikoku/Kyushu, and continental regions into distinct clusters (Fig. 10B). PC3 reflected the genetic variations in individuals from the Chugoku and Kyushu areas (Supplementary Data S4), PC4 separated Shikoku individuals from the others, and PC5 separated Tsushima, Yakushima, and Tanegashima Island individuals from the other samples within the Kyushu area. The Neighbor-Net graph shows a complex network separated into five major clusters. Several individuals from Shimane and Hiroshima showed signs of hybridization of introgression between Hokuriku/nKinki/eChugoku and Kyushu/wChugoku clusters (Fig. 11). The fastSTRUCTURE analysis showed that the value

of *K* that maximized the marginal likelihood was five and identified the five most probable genetic clusters: nc*Mwo*-1, nc*Mwo*-2, nc*Mwo*-3, nc*Mwo*-4, and nc*Mwo*-5, mainly observed in the Tokai/sKinki, Hokuriku/nKinki/eChugoku, wChugoku/Kyushu, continental, and Shikoku populations, respectively (Fig. 12B). This result aligns well with the mtDNA haplotype groups defined in a previous study (Nakamoto et al. 2021); however, the individuals from the Shikoku region were assigned to the fifth distinct genetic cluster, nc*Mwo*-5 (Fig. 12B). Consistent with the Neighbor-Net results, some individuals from Shimane and Hiroshima showed mixed ancestry of nc*Mwo*-2 and nc*Mwo*-3 (Fig. 12B). These admixed individuals were collected from the distribution border of nc*Mwo*-2 and nc*Mwo*-3, most of which are host of the mt*Mwo*-III mtDNA haplotypes (Nakamoto et al. 2021). The pairwise F_{ST} values were very high between continental and Japanese populations, ranging from 0.597 to 0.701. The lowest F_{ST} was observed between nc*Mwo*-2 and nc*Mwo*-3, consistent with the fact that these two populations showed signs of introgression. F_{ST} values are shown in Supplementary Data S6. The average heterozygosity of each genetic cluster is summarized in Supplementary Data S7. The heterozygosity for nc*Mwo*-4 was considerably low compared to other genetic clusters.

Discussion

Population structures of M. imaizumii

The results from PCA, Neighbor-Net, and fastSTRUCTURE consistently divided *M. imaizumii* into three genetically distinct clusters, with pairwise F_{ST} values also indicating strong differentiation between these genetic groups (Supplementary Data S5). Iwasa et al. (2006) demonstrated that *M. imaizumii* in eastern Japan comprises two genetically distinct groups based on mitochondrial *Cytb* sequences and restriction fragment length polymorphism analysis of the nuclear 28S ribosomal RNA gene spacer. Our findings from PCA (Fig. 10), fastSTRUCTURE (Fig. 12), and the high F_{ST} values (Supplementary Data S5), from the MIG-seq analyses of nuclear DNA reveal a high degree of differentiation between the two *M. imaizumii* groups (nc*Mim*-1/nc*Mim*-2), supporting the hypothesis that the mountain chain in eastern Japan, which forms the boundary between the nc*Mim*-1 and nc*Mim*-2 distributions, acts as a barrier (Iwasa et al. 2006). Iwasa et al. (2006) proposed a secondary contact model, suggesting that these two lineages dispersed along distinct routes (along the coastlines of the JS and PO) and converged in Iwate Prefecture, where gene

introgression was observed. However, we found no strong gene introgression between nc*Mim*-1 and nc*Mim*-2. Additionally, isolated populations from Oki, Hiroshima, and Shodoshima were also genetically distinct from each other, likely due to genetic drift following isolation, as suggested by their relatively low heterozygosity (Oki: 0.00050, Hiroshima: 0.00152, Shodoshima: 0.00030), Among these isolated population samples, one individual from the Chugoku Mountains (Mt. Kenashi, Hiroshima) displayed signs of mixed ancestry (K = 3; Fig. 12B), suggesting possible genetic contributions from both nc*Mim*-1 and nc*Mim*-3. It is noteworthy that the patchy distribution of *M. imaizumii* in Mt. Kenashi, Hiroshima is now entirely surrounded by the range of *M. wogura*. This suggests that *M. imaizumii* populations may have expanded westward during the cold periods of the Quaternary, as this species, now largely distributed in northern Japan, is better adapted to cold environments than *M. wogura* (Kirihara et al. 2013; Nakamoto et al. 2021). During these colder periods, we hypothesize that the distribution of nc*Mim*-1 shifted westward from the Kinki region to the Chugoku region, mixing with an *M. imaizumii* population related to nc*Mim*-3. Later, as temperatures rose during the post-glacial periods and *M. imaizumii* populations began migrating eastward, nc*Mim*-1 individuals may have remained in isolated areas, persisting as admixed populations. However, it is important to note that fastSTRUCTURE requires an adequate sample size to accurately interpret population structure (Raj et al. 2014), and since only one *M. imaizumii* individual was available from the Chugoku Mountains, this mixed ancestry could be an artifact. Further sampling and analysis of *M. imaizumii* in this isolated region are required to test this hypothesis.

Population structure of M. wogura

From PCA, Neighbor-Net, and fastSTRUCTURE analyses, *M. wogura* was categorized into five genetic clusters (nc*Mwo*-1 to 5), and a sign of introgression was detected in the Chugoku area. In this study, we identified seemingly genetically admixed individuals at the boundary between these two populations, spanning Shimane and Hiroshima prefectures (Supplementary Data S8B). This contrasts with the boundary of the nKinki/sKinki populations, where no clear sign of admixture was observed (Supplementary Data S8A), except that an individual was observed to hold the mt*Mwo*-I haplotype and was mainly assigned to nc*Mwo*-2 genetic cluster (HS4642). It is notable that there seem to be no clear abiotic barriers, such as mountains or rivers that segregate eChugoku/wChugoku populations. The boundary between these two populations was previously identified based

on the distribution of the mt*Mwo*-II and mt*Mwo*-III mitochondrial haplotypes. However, no signature of introgression was observed in the mitochondrial DNA, owing to a lack of recombination (Nakamoto et al. 2021). Since most of these genetically admixed individuals contained mtDNA of the m*Mwo*-III haplotype, this introgression would be unidirectional, either nc*Mwo*-3 to nc*Mwo*-2, or vice versa. Although it remains unclear whether sex-biased dispersal occurs in moles (Satoh et al. 2019), if male-biased gene flow is present, as is the case for most mammals (Greenwood 1980), the observed pattern suggests male-biased gene flow from nc*Mwo*-2- to nc*Mwo*-3-dominated populations. This type of asymmetric introgression is common in secondary contact zones of differentiated populations (Chan and Levin 2005; Currat et al. 2008; Toews and Brelsford 2012; Johnson et al. 2015). Therefore, it is presumed that this boundary is a case in which two populations make secondary contact after their isolation. A large river (Gono River) runs east of the boundary (Supplementary Data S8B), and this may have acted as a barrier that did not completely, but significantly, hinder the migration of moles. Nakamoto et al. (2021) showed that *M. wogura* populations in the eChugoku/wChugoku region exhibited signs of recent rapid population expansion, which could have allowed them to cross this barrier, possibly making a detour around the river source and admixing with each other.

Common genetic boundary in the Kinki area observed in both Mogera imaizumii and M. wogura

The nc*Mwo*-1/nc*Mwo*-2 border is also observant in the nc*Mim*-1/nc*Mim*-3 border of *M. imaizumii*. This nKinki/sKinki border could have played an important role in the segregation of mole populations, as no clear sign of gene introgression between genetic clusters are seen. This distinguishing feature of this region can be related to its geological history. The Osaka Plain is known to have been underwater multiple times during the Quaternary, and this possibly acted as a barrier to divide the mole populations. The Jomon transgression, a sea-level rise spanning from the early Pleistocene to the middle Holocene, has been proposed by Nakamoto et al. (2021). The emergence of the second Seto Inland Sea (current Osaka Plain), an inland water body that formed approximately 3 mya and persisted until at least the middle Pleistocene (Sugiyama 1991; Yoshida 1992), is also a strong candidate that could have contributed to the formation of this border. The Osaka Plain is also known to have undergone the Rokko movements, a set of tectonic movements started in the middle Pleistocene (Huzita 1990), and this movement shaped many of the current rises

and falls of mountains and basins. These dynamic geographic histories are very likely to have influenced the distribution of moles. Kirihara et al. (2013) also proposed that the genetic distinctiveness of the M. wogura population from Tokai/Kinki (in this case, ncMwo-1) could be the result of the Hokuriku/nKinki *M. imaizumii* population expanding its range southwards, serving as a barrier to prevent *M. wogura* migration. The change in vegetation may also have played a role in population segregation. The current Kinki region is mostly covered in broad-leaved forests, with temperate mixed deciduous forests generally corresponding to the border of the nKinki/sKinki region (Dobson 1994). This vegetation shifts to a boreal coniferous forest in the LGM and could have acted as a barrier to prevent interactions between mole populations.

Appendix

Appendix II: List of *Mogera imaizumii* samples used in this study.

Species	Locality No.	Collection Locality		*Specimen code	**HUNHM code	DDBJ Accession No.
Mogera imaizumii	1	Aomori	Hachinohe	HS499/KT2793	79062	DRR545082
Mogera imaizumii	2	Iwate	Himekami	HS5986	84549	DRR545220
Mogera imaizumii	2	Iwate	Himekami	HS5988	84551	DRR545221
Mogera imaizumii	3	Iwate	Miyako	HS1777	80340	DRR545085/DRR545216
Mogera imaizumii	4	Iwate	Tono	HS2315	80878	DRR545087
Mogera imaizumii	5	Iwate	Otsuchi	HS1498	80061	DRR545086/DRR545217
Mogera imaizumii	5	Iwate	Otsuchi	HS1514	80077	DRR545218
Mogera imaizumii	5	Iwate	Otsuchi	HS1508	80071	DRR545072
Mogera imaizumii	6	Iwate	Kamaishi	HS1515	80078	DRR545219
Mogera imaizumii	7	Akita	Akita	HS5786/MH14257	84349	DRR545083
Mogera imaizumii	8	Akita	Yuzawa	HS5787/MH14258	84350	DRR545084
Mogera imaizumii	8	Akita	Yuzawa	HS5788/MH14259	84351	DRR545214
Mogera imaizumii	8	Akita	Yuzawa	HS5792/MH14263	84355	DRR545215
Mogera imaizumii	9	Yamagata	Tsuruoka	HS5793/MH14264	84356	DRR545073
Mogera imaizumii	9	Yamagata	Tsuruoka	HS5794/MH14265	84357	DRR545088
Mogera imaizumii	10	Yamagata	Kamio	HS5795/MH14266	84358	DRR545089
Mogera imaizumii	10	Yamagata	Kamio	HS5796/MH14267	84359	DRR545222

Mogera imaizumii	10	Yamagata	Kamio	HS5797/MH14268	84360	DRR545223
Mogera imaizumii	10	Yamagata	Kamio	HS5798	84361	DRR545242
Mogera imaizumii	11	Yamagata	Ogunimachi	HS5799/MH14270	84362	DRR545090
Mogera imaizumii	12	Miyagi	Kinkazan	HS1497	80060	DRR545224
Mogera imaizumii	13	Miyagi	Sendai	HS582	79145	DRR545091
Mogera imaizumii	14	Fukushima	Kooriyama	HS575	79138	DRR545225
Mogera imaizumii	15	Niigata	Awashima	HS1521	80075	DRR545064
Mogera imaizumii	16	Niigata	Ikarashi	HS470	79033	DRR545092
Mogera imaizumii	16	Niigata	Ikarashi	HS364/KT2795	78927	DRR545226
Mogera imaizumii	17	Niigata	Itoigawa	HS5809/MH14302	84372	DRR545227
Mogera imaizumii	17	Niigata	Itoigawa	HS5810/MH14343	84373	DRR545243
Mogera imaizumii	17	Niigata	Itoigawa	HS5811/MH14344	84374	DRR545244
Mogera imaizumii	18	Toyama	Unazuki	HS5806	84369	DRR545234
Mogera imaizumii	18	Toyama	Unazuki	HS5807	84370	DRR545235
Mogera imaizumii	18	Toyama	Unazuki	HS5808	84371	DRR545236
Mogera imaizumii	19	Toyama	Toyama	HS5457	84020	DRR545098
Mogera imaizumii	19	Toyama	Toyama	HS5458	84021	DRR545099
Mogera imaizumii	20	Toyama	Fuchu	HS5456	84019	DRR545233
Mogera imaizumii	21	Ishikawa	Noto	HS3366	81929	DRR545067/DRR545237
Mogera imaizumii	22	Ibaraki	Hitachioomiya	HS4665	83228	DRR545229
Mogera imaizumii	23	Ibaraki	Tsukubamirai	HS5632/KT4471	84195	DRR545066
Mogera imaizumii	23	Ibaraki	Tsukubamirai	HS5630/KT4469	84193	DRR545077
Mogera imaizumii	23	Ibaraki	Tsukubamirai	HS5631/KT4470	84194	DRR545228
Mogera imaizumii	24	Gunma	Minakami	HS3057	81620	DRR545074
Mogera imaizumii	24	Gunma	Minakami	HS3077/KT3331	81640	DRR545075
Mogera imaizumii	25	Nagano	Chikuma	HS4491	83054	DRR545078
Mogera imaizumii	25	Nagano	Chikuma	HS4572/KT4014	83135	DRR545065
Mogera imaizumii	26	Nagano	Karuizawa	HS1778	80341	DRR545076
Mogera imaizumii	27	Nagano	Matsumoto	HS6090/MH14478	84653	DRR545309
Mogera imaizumii	27	Nagano	Matsumoto	HS6091/MH14479	84654	DRR545310
Mogera imaizumii	27	Nagano	Matsumoto	HS6092/MH14480	84655	DRR545311
Mogera imaizumii	27	Nagano	Matsumoto	HS6093/MH14481	84656	DRR545312
Mogera imaizumii	28	Chiba	Katori	HS5864/MH14343	84427	DRR545248
Mogera imaizumii	28	Chiba	Katori	HS5866/MH14345	84429	DRR545249

Mogera imaizumii	29	Chiba	Shirako	HS5872/MH14341	84435	DRR545247
Mogera imaizumii	29	Chiba	Shirako	HS5873	84436	DRR545103
Mogera imaizumii	30	Yamanashi	Ichikawamisato	HS4545	83108	DRR545230
Mogera imaizumii	31	Yamanashi	Aokigaharajukai	HS3187/KT3468	81750	DRR545079
Mogera imaizumii	32	Kanagawa	Atsugi	HS3896	82459	DRR545095
Mogera imaizumii	32	Kanagawa	Atsugi	HS6007	84570	DRR545257
Mogera imaizumii	32	Kanagawa	Atsugi	HS6001	84564	DRR545258
Mogera imaizumii	33	Kanagawa	Yamakita	HS5901/KT4314	84464	DRR545231
Mogera imaizumii	33	Kanagawa	Yamakita	HS6013	84576	DRR545255
Mogera imaizumii	34	Kanagawa	Matsuda	HS6011	84574	DRR545256
Mogera imaizumii	35	Kanagawa	Yokohama	HS472	79035	DRR545093
Mogera imaizumii	35	Kanagawa	Yokohama	HS365/KT2697	78928	DRR545094
Mogera imaizumii	36	Kanagawa	Hakone	HS3092	81655	DRR545096
Mogera imaizumii	37	Shizuoka	Gotenba	HS6029	84592	DRR545260
Mogera imaizumii	38	Shizuoka	Susono	HS6028/KT3630	84591	DRR545259
Mogera imaizumii	39	Shizuoka	Mishima	HS3184/KT3469	81747	DRR545097
Mogera imaizumii	39	Shizuoka	Mishima	HS6030	84593	DRR545261
Mogera imaizumii	40	Shizuoka	Ito	HS1022	79585	DRR545232
Mogera imaizumii	41	Shiga	Kutsuki	HS3367	81930	DRR545238
Mogera imaizumii	41	Shiga	Kutsuki	HS3368	81931	DRR545068/DRR545239
Mogera imaizumii	42	Kyoto	Mt. Hiei	HS4736	83299	DRR545069
Mogera imaizumii	43	Kyoto	Nagaokakyo	HS1843	80406	DRR545070
Mogera imaizumii	44	Kyoto	Iwakuranagatani	HS463	79026	DRR545100/DRR545240
Mogera imaizumii	45	Osaka	Kawachinagano	HS5757/MH14237	84320	DRR545101
Mogera imaizumii	45	Osaka	Kawachinagano	HS5783/MH14253	84346	DRR545241
Mogera imaizumii	46	Nara	Mt. Wasamata	HS4620/MH9967	83183	DRR545071
Mogera imaizumii	47	Nara	Nosegawa	HS6436/MH14387	84999	DRR545081
Mogera imaizumii	48	Wakayama	Wakayama	HS462/KT2807	79025	DRR545246
Mogera imaizumii	49	Wakayama	Kumanogawa	HS368/KT2707	78931	DRR545080
Mogera imaizumii	50	Wakayama	Kozagawa	HS5758/MH14218	84321	DRR545102
Mogera imaizumii	50	Wakayama	Kozagawa	HS5759/MH14219	84322	DRR545245
Mogera imaizumii	51	Shimane	Uzuki	HS6444/MH14492	85007	DRR545306
Mogera imaizumii	51	Shimane	Uzuki	HS6445/MH14493	85008	DRR545307
Mogera imaizumii	52	Hiroshima	Mt. Kenashi	HS5805	84368	DRR545104/DRR545250

Mogera imaizumii	53	Kagawa	Yoshida	HS5991/MH14424	84554	DRR545251
Mogera imaizumii	53	Kagawa	Yoshida	HS5992/MH14425	84555	DRR545252
Mogera imaizumii	53	Kagawa	Yoshida	HS5993/MH14426	84556	DRR545253
Mogera imaizumii	54	Kagawa	Tonosho	HS5999/MH14432	84562	DRR545254

*Specimen codes refer to DNA in the personal collection of HS, MH, and KT (Kimiyuki Tsuchiya).

**Numbers refer to DNA deposited at the Botanic Garden & Museum, Field Science Center for Northern Biosphere, Hokkaido University, Sapporo.

Appendix III: List of *Mogera wogura* samples used in this study.

Species	Locality No.	Collection Locality		*Specimen code	**HUNHM code	DDBJ Accession No.
Mogera wogura	36	Kanagawa	Hakone	HS3096/KT3491	81659	DRR545106/DRR545295
Mogera wogura	39	Shizuoka	Mishima	HS1023	79586	DRR545263
Mogera wogura	55	Shizuoka	Fujinomiya	HS3897	82460	DRR545124
Mogera wogura	56	Aichi	Kasugai	HS581	79144	DRR545117/DRR545264
Mogera wogura	57	Ishikawa	Matoou	HS1416/KT3095	79979	DRR545121
Mogera wogura	58	Fukui	Natashomura	HS4776	83339	DRR545272
Mogera wogura	58	Shiga	Takashima	HS4884/MH11604	83447	DRR545108
Mogera wogura	60	Shiga	Otsu	HS4888/MH11608	83451	DRR545131
Mogera wogura	60	Shiga	Otsu	HS4925/MH11720	83488	DRR545285
Mogera wogura	61	Kyoto	Kameoka	HS4766/MH11474	83329	DRR545129
Mogera wogura	62	Kyoto	Fukuchiyama	HS4768/MH11477	83331	DRR545273
Mogera wogura	63	Osaka	Takatsuki	HS4607/MH11242	83170	DRR545107/DRR545294
Mogera wogura	64	Osaka	Ibaraki	HS4654/M H 11273	83217	DRR545276
Mogera wogura	45	Osaka	Kawachinagano	HS314	78877	DRR545269
Mogera wogura	65	Mie	Iga	HS4929/MH11717	83492	DRR545136
Mogera wogura	66	Nara	Nara	HS4898/MH11530	83461	DRR545132
Mogera wogura	67	Nara	Mt. Katsuragi	HS4659/MH8004	83222	DRR545275
Mogera wogura	68	Wakayama	Kanaya	HS1073	79636	DRR545119
Mogera wogura	69	Hyogo	Tanba	HS4769/MH11478	83332	DRR545130
Mogera wogura	70	Hyogo	Sasayama	HS4771/MH11480	83334	DRR545274
Mogera wogura	71	Hyogo	Takarazuka	HS4642/MH11271	83205	DRR545277
Mogera wogura	72	Hyogo	Kobe	HS4664	83227	DRR545128
Mogera wogura	73	Hyogo	Awajishima	HS4657/MH11279	83220	DRR545127
Mogera wogura	74	Tottori	Misasa	HS5629/MH14180	84192	DRR545150
Mogera wogura	75	Okayama	Akaiwa	HS4585	83148	DRR545125
Mogera wogura	76	Shimane	Harada	HS6439/MH14485	85002	DRR545302
Mogera wogura	77	Shimane	Yamada	HS6440/MH14486	85003	DRR545303
Mogera wogura	77	Shimane	Yamada	HS6442/MH14487	85005	DRR545304
Mogera wogura	77	Shimane	Yamada	HS6441/MH14488	85004	DRR545305
Mogera wogura	78	Shimane	Kamo	HS6443/MH14499	85006	DRR545308

Mogera wogura	79	Shimane	Oki	HS4433	82996	DRR545151
Mogera wogura	80	Shimane	Dozen	HS6446	85009	DRR545298
Mogera wogura	81	Shimane	Matsukawa	HS5153/MH13052	83716	DRR545139
Mogera wogura	82	Shimane	Nagatani	HS5342/MH13407	83905	DRR545287
Mogera wogura	83	Shimane	Ninomiya	HS5155/MH13054	83718	DRR545140
Mogera wogura	83	Shimane	Ninomiya	HS5156/MH13055	83719	DRR545278
Mogera wogura	84	Shimane	Kochi	HS5360/MH13603	83923	DRR545288
Mogera wogura	84	Shimane	Kochi	HS5361/MH13604	83924	DRR545289
Mogera wogura	85	Shimane	Kanagi	HS5363/MH13606	83926	DRR545290
Mogera wogura	85	Shimane	Kanagi	HS5362/MH13605	83925	DRR545145
Mogera wogura	86	Shimane	Sufu	HS5597/MH14024	84160	DRR545149
Mogera wogura	87	Hiroshima	Mt. Hiba	HS5947/MH14384	84510	DRR545152
Mogera wogura	88	Hiroshima	Miyoshi	HS5948/MH14386	84511	DRR545153
Mogera wogura	89	Hiroshima	Fuchu	HS5143/MH13042	83706	DRR545138
Mogera wogura	90	Hiroshima	Mihara	HS5323/MH13347	83886	DRR545142
Mogera wogura	91	Hiroshima	Miwa	HS5369/MH13614	83932	DRR545292
Mogera wogura	92	Hiroshima	Koda	HS5368/MH13613	83931	DRR545147
Mogera wogura	93	Hiroshima	Yoshida	HS5367/MH13612	83930	DRR545146
Mogera wogura	94	Hiroshima	Yachiyo	HS5148/MH13047	83711	DRR545281
Mogera wogura	95	Hiroshima	Shiraki	HS5150/MH13049	83713	DRR545279
Mogera wogura	95	Hiroshima	Shiraki	HS5149/MH13048	83712	DRR545280
Mogera wogura	96	Hiroshima	Totani	HS5365/MH13608	83928	DRR545109
Mogera wogura	97	Hiroshima	Oasa	HS5344/MH13409	83907	DRR545144
Mogera wogura	98	Hiroshima	Higashiyawatabara	HS6438/MH14399	85001	DRR545159
Mogera wogura	99	Hiroshima	Higashihiroshima	HS5322/MH13346	83885	DRR545141
Mogera wogura	99	Hiroshima	Higashihiroshima	HS5324/MH11349	83887	DRR545143
Mogera wogura	100	Hiroshima	Tahara	HS5364/MH13607	83927	DRR545291
Mogera wogura	101	Hiroshima	Matsubara	HS6437/MH14394	85000	DRR545158
Mogera wogura	102	Yamaguchi	Waki	HS4899/MH11524	83462	DRR545133
Mogera wogura	103	Yamaguchi	Abu	HS5157/MH12401	83720	DRR545286
Mogera wogura	104	Yamaguchi	Nagato	HS4606/MH11084	83169	DRR545262
Mogera wogura	104	Yamaguchi	Nagato	HS4896/MH11528	83459	DRR545270
Mogera wogura	105	Yamaguchi	Misumi	HS4882/MH11529	83445	DRR545271
Mogera wogura	106	Yamaguchi	Shimonoseki	HS5608/MH14047	84171	DRR545301

Mogera wogura	106	Yamaguchi	Shimonoseki	HS4901/MH11526	83464	DRR545134
Mogera wogura	106	Yamaguchi	Shimonoseki	HS4902/MH11527	83465	DRR545135
Mogera wogura	107	Kagawa	Yasuda	HS5995/MH14428	84558	DRR545299
Mogera wogura	107	Kagawa	Yasuda	HS5994/MH14427	84557	DRR545300
Mogera wogura	108	Kagawa	Mannou	HS4953/MH11784	83516	DRR545282
Mogera wogura	108	Kagawa	Mannou	HS4952/MH11783	83515	DRR545283
Mogera wogura	108	Kagawa	Mannou	HS4951/MH11782	83514	DRR545284
Mogera wogura	109	Tokushima	Mima	HS4605/MH-5921	83168	DRR545126
Mogera wogura	110	Kochi	Shimantonakamura	HS4949/MH11780	83512	DRR545137
Mogera wogura	111	Ehime	Iyo	HS5500	84063	DRR545148
Mogera wogura	112	Fukuoka	Fukuoka	HS5501	84064	DRR545293
Mogera wogura	113	Nagasaki	Tsushima	HS409	78972	DRR545267
Mogera wogura	113	Nagasaki	Tsushima	HS366/KT2711	78929	DRR545112
Mogera wogura	113	Nagasaki	Tsushima	HS2820/KT3473	81383	DRR545123
Mogera wogura	114	Miyazaki	Kiyotake	HS3066/KT3276	81629	DRR545105/DRR545296
Mogera wogura	115	Kumamoto	Menda	HS414	78977	DRR545266
Mogera wogura	115	Kumamoto	Menda	HS404	78967	DRR545115
Mogera wogura	116	Kagoshima	Makizono	HS369/KT2699	78932	DRR545268
Mogera wogura	117	Kagoshima	Tanegashima	HS467	79030	DRR545116/DRR545265
Mogera wogura	118	Kagoshima	Yakushima	HS4444	83007	DRR545297
Mogera wogura	118	Kagoshima	Yakushima	HS371/KT2731	78934	DRR545113
Mogera robusta	119	Russia	Ussuriisk	HS2004	80567	DRR545122
Mogera robusta	120	Russia	Hasan	HS1392	79955	DRR545120
Mogera robusta	121	Korea	Seoraksan	HS930	79493	DRR545118
Mogera robusta	122	Korea	Pusan	HS373/KT2755	78936	DRR545114

*Specimen codes refer to DNA in the personal collection of HS, MH, and KT (Kimiyuki Tsuchiya).

**Numbers refer to DNA deposited at the Botanic Garden & Museum, Field Science Center for Northern Biosphere, Hokkaido University, Sapporo.

Figures and Legends

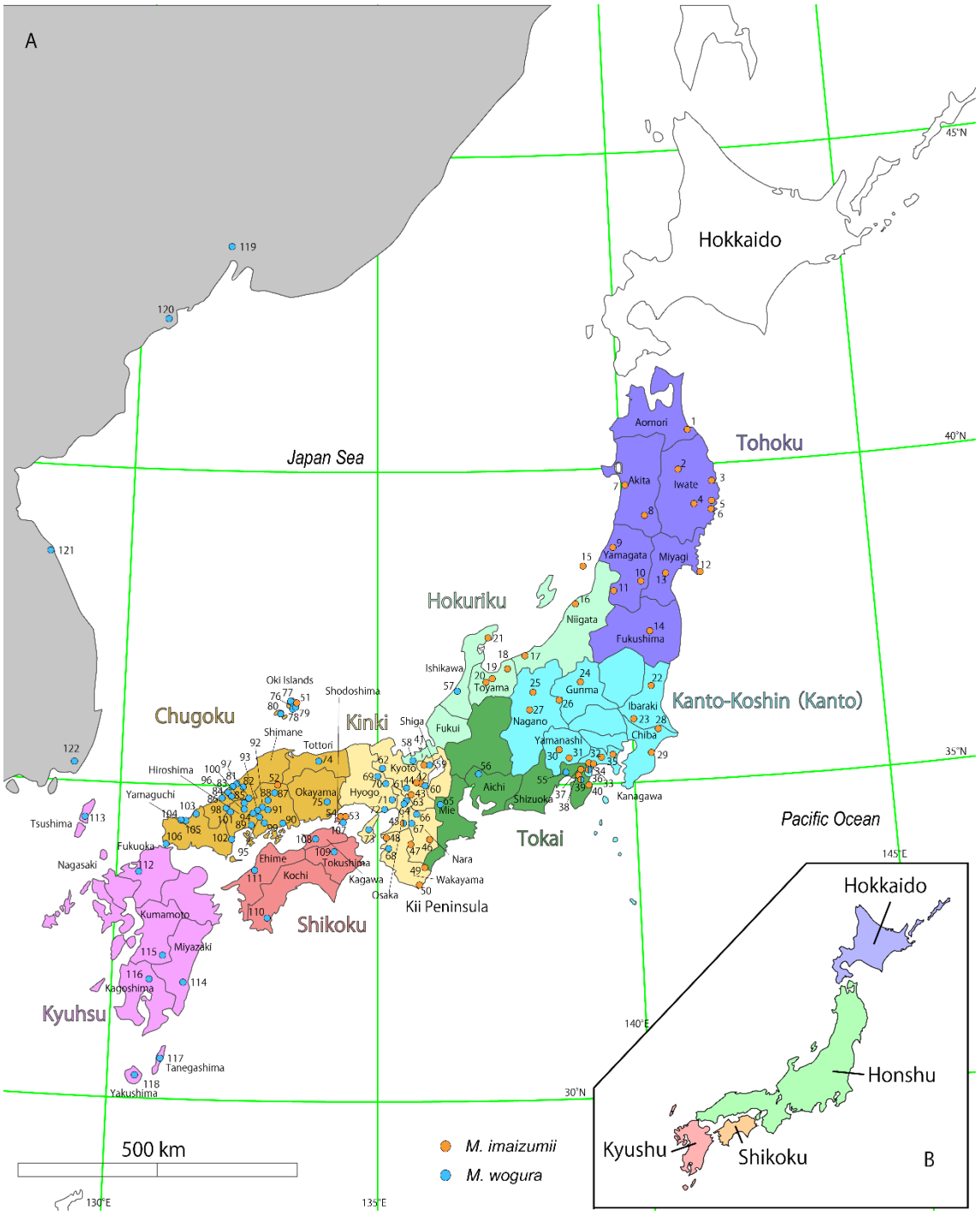


Figure 9. (A) Geographic locations of the sampling sites of *Mogera imaizumii* and *M. wogura*. (B) Four main islands of the Japan Archipelago.

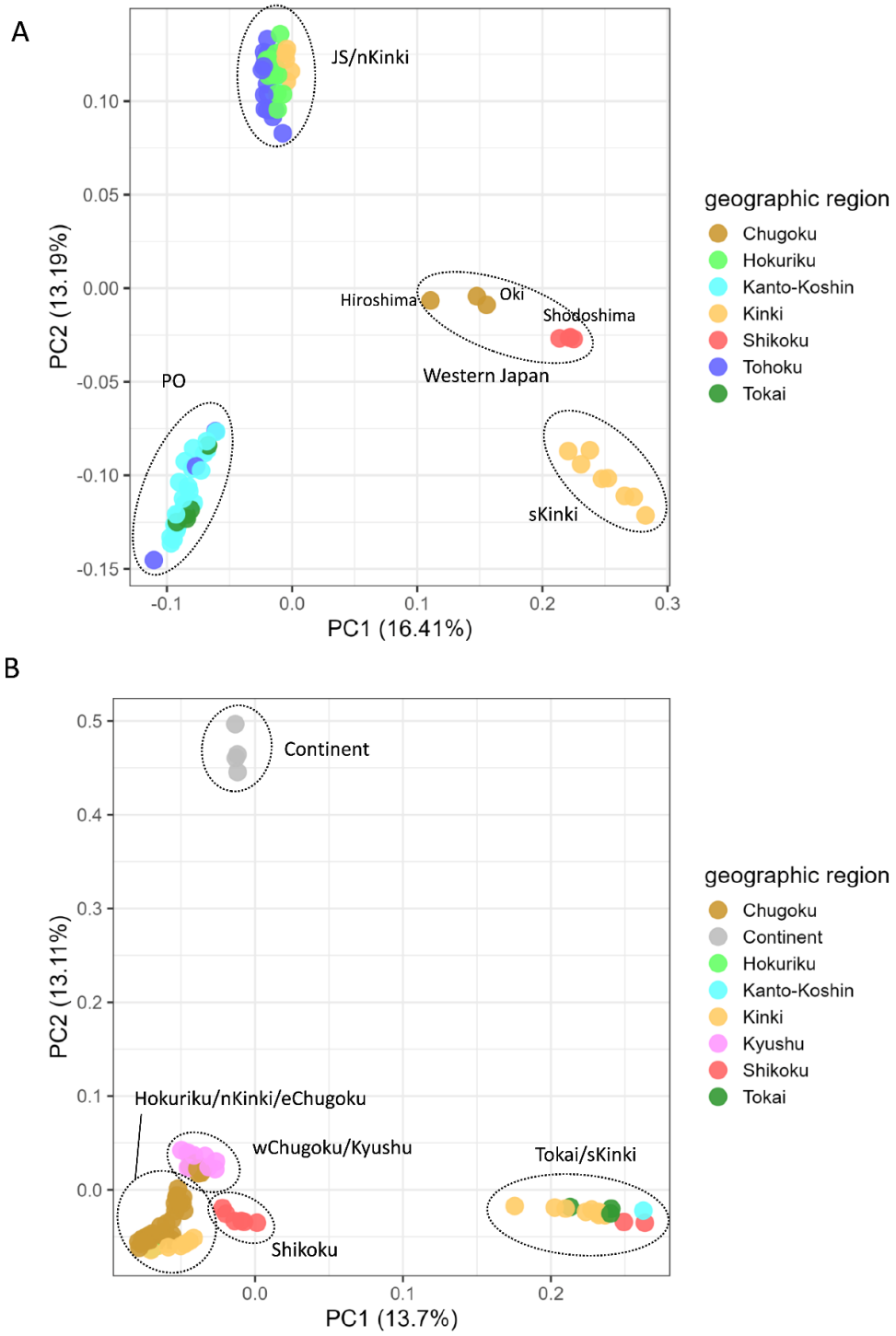


Figure 10. The results of the PCA analysis of *Mogera imaizumii* (A) and *M. wogura* (B).

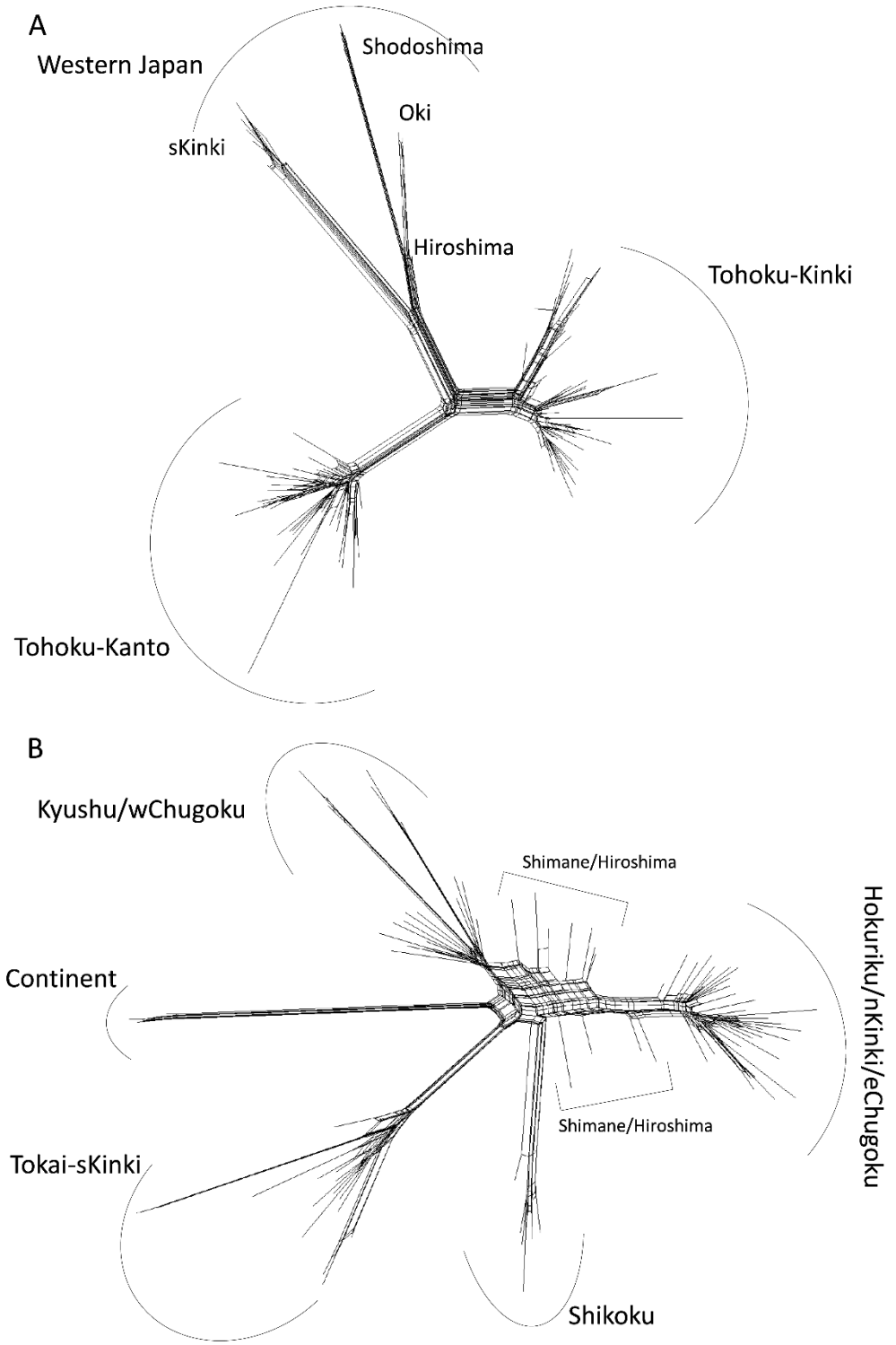


Figure 11. The results of the NeighborNet analysis for *Mogera imaizumii* and *M. wogura*.

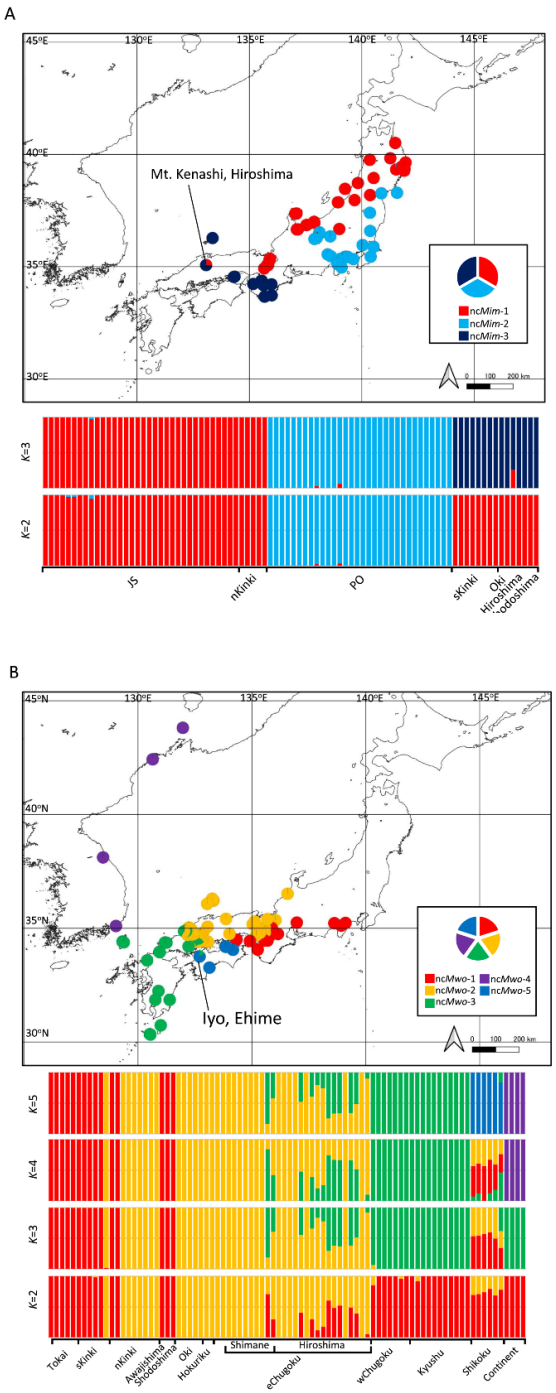
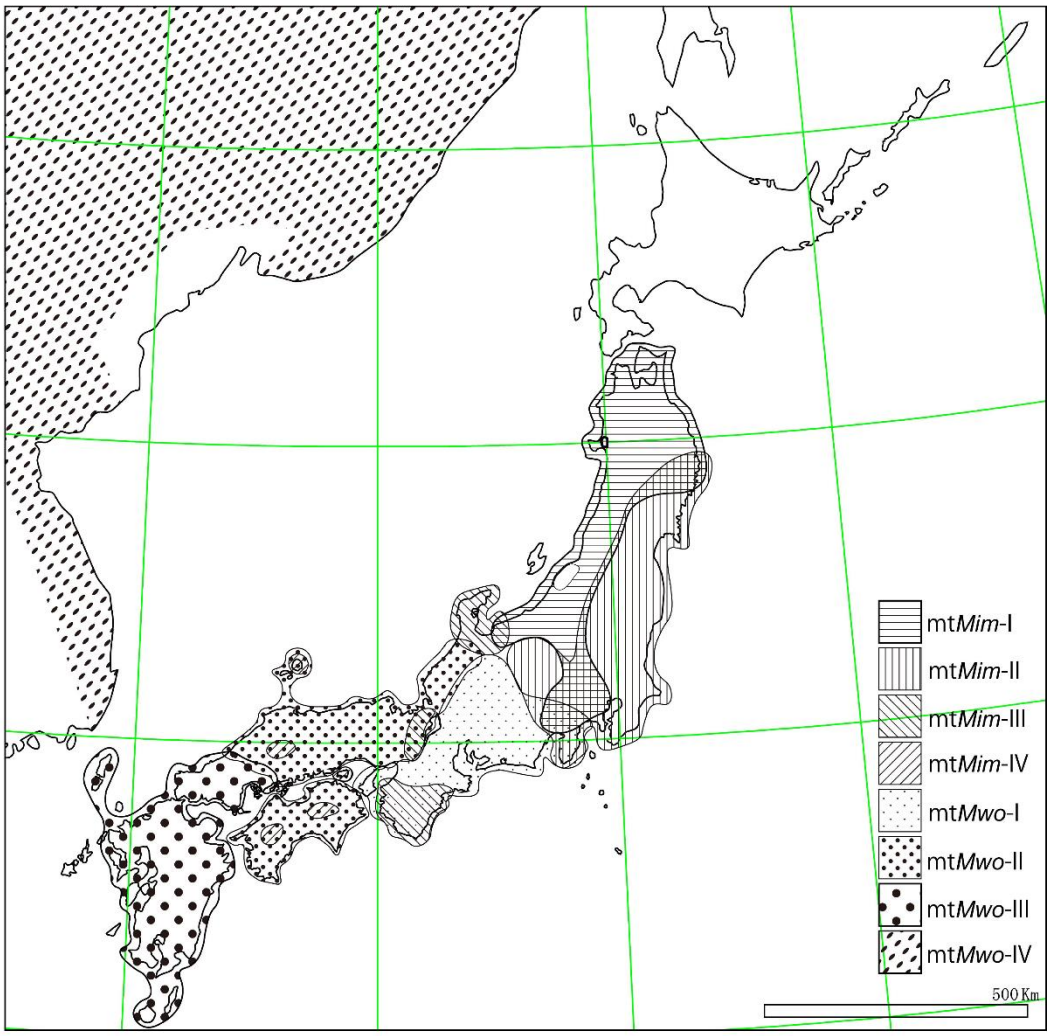


Figure 12. The results of fastSTRUCTURE analysis on *Mogera imaizumii* and *M. wogura*. Each pie chart or line indicate an individual. The pie charts are assigned to the sampling sites.

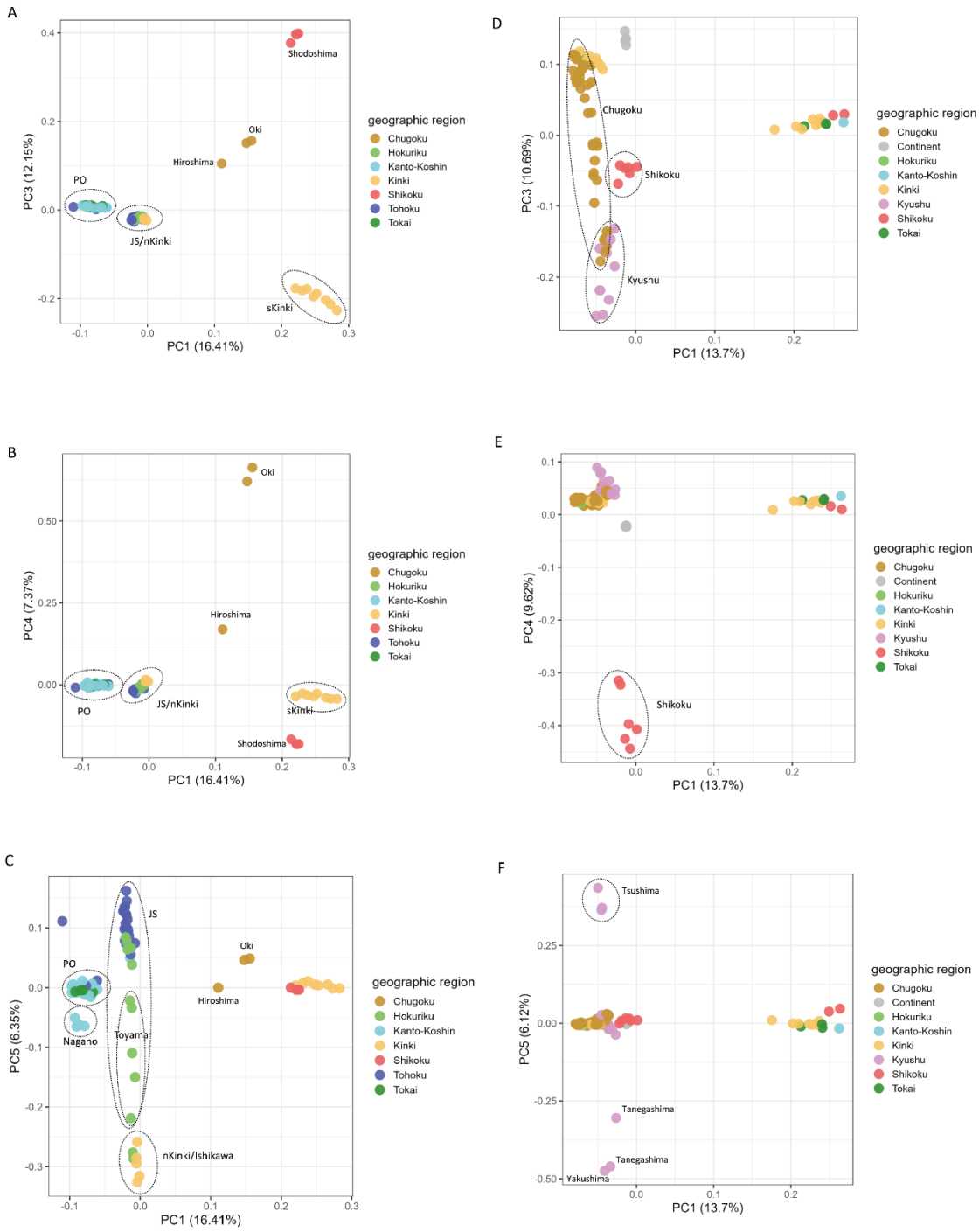


Supplementary Data S2.

M. wogura

M. imaizumii

Supplementary Data S3.



Supplementary Data S4.

Supplementary Data S5: Pairwise F_{ST} values for *Mogera imaizumii*

	nc*Mim*-1	nc*Mim*-2	nc*Mim*-3
nc*Mim*-1	0	0.376	0.458
nc*Mim*-2	-	0	0.487
nc*Mim*-3	-	-	0

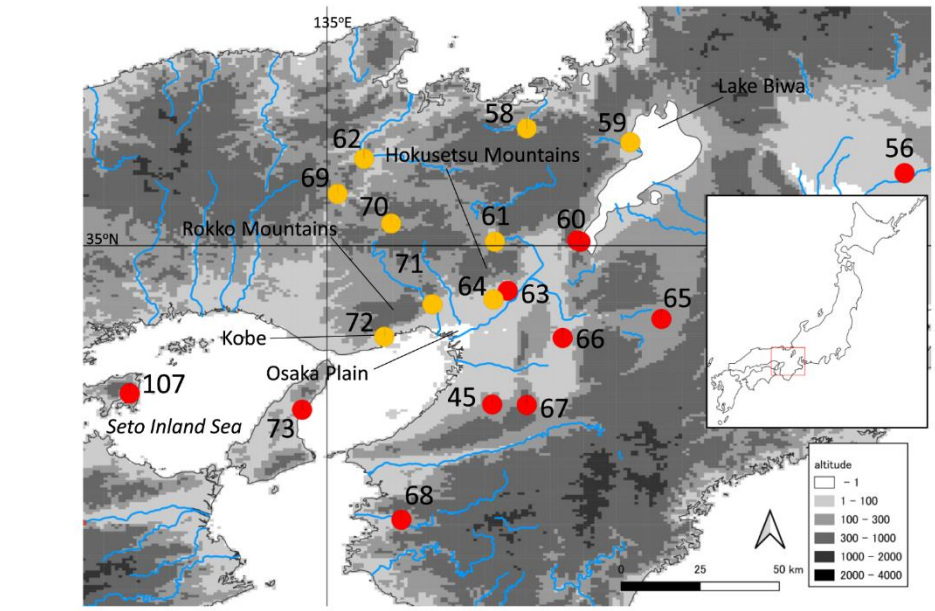
Supplementary Data S6: Pairwise F_{ST} values for *Mogera wogura*

	nc*Mwo*-1	nc*Mwo*-2	nc*Mwo*-3	nc*Mwo*-4	nc*Mwo*-5
nc*Mwo*-1	0	0.446	0.362	0.701	0.502
nc*Mwo*-2	-	0	0.291	0.700	0.460
nc*Mwo*-3	-	-	0	0.597	0.390
nc*Mwo*-4	-	-	-	0	0.757
nc*Mwo*-5	-	-	-	-	0

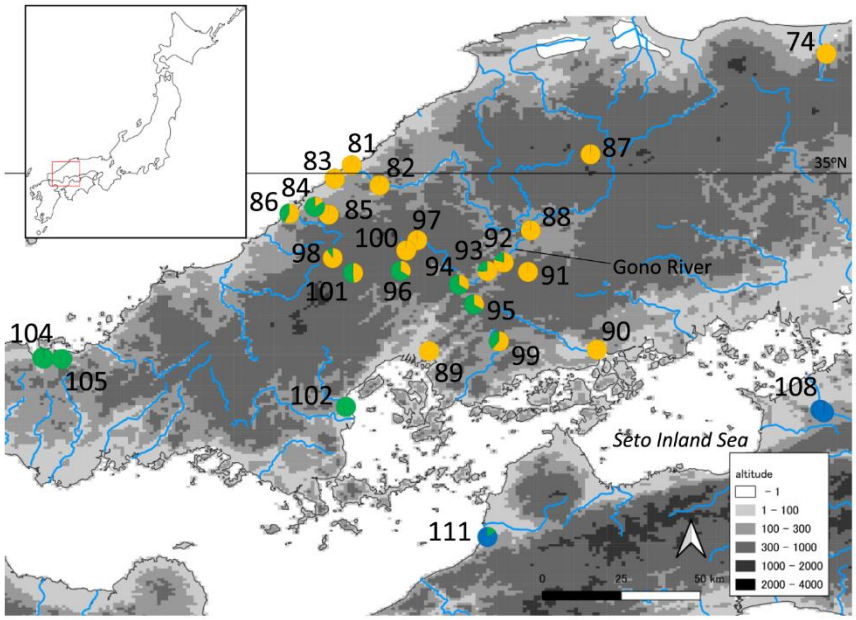
Supplementary Data 7: Average heterozygosity for each genetic cluster.

	average heterozygosity
M. imaizumii	0.001053088
nc*Mim*-1	0.000990768
nc*Mim*-2	0.001337198
nc*Mim*-3	0.000613173
M. wogura	0.000974172
nc*Mwo*-1	0.000846439
nc*Mwo*-2	0.000875681
nc*Mwo*-3	0.00129332
nc*Mwo*-4	0.000235526
nc*Mwo*-5	0.001116706

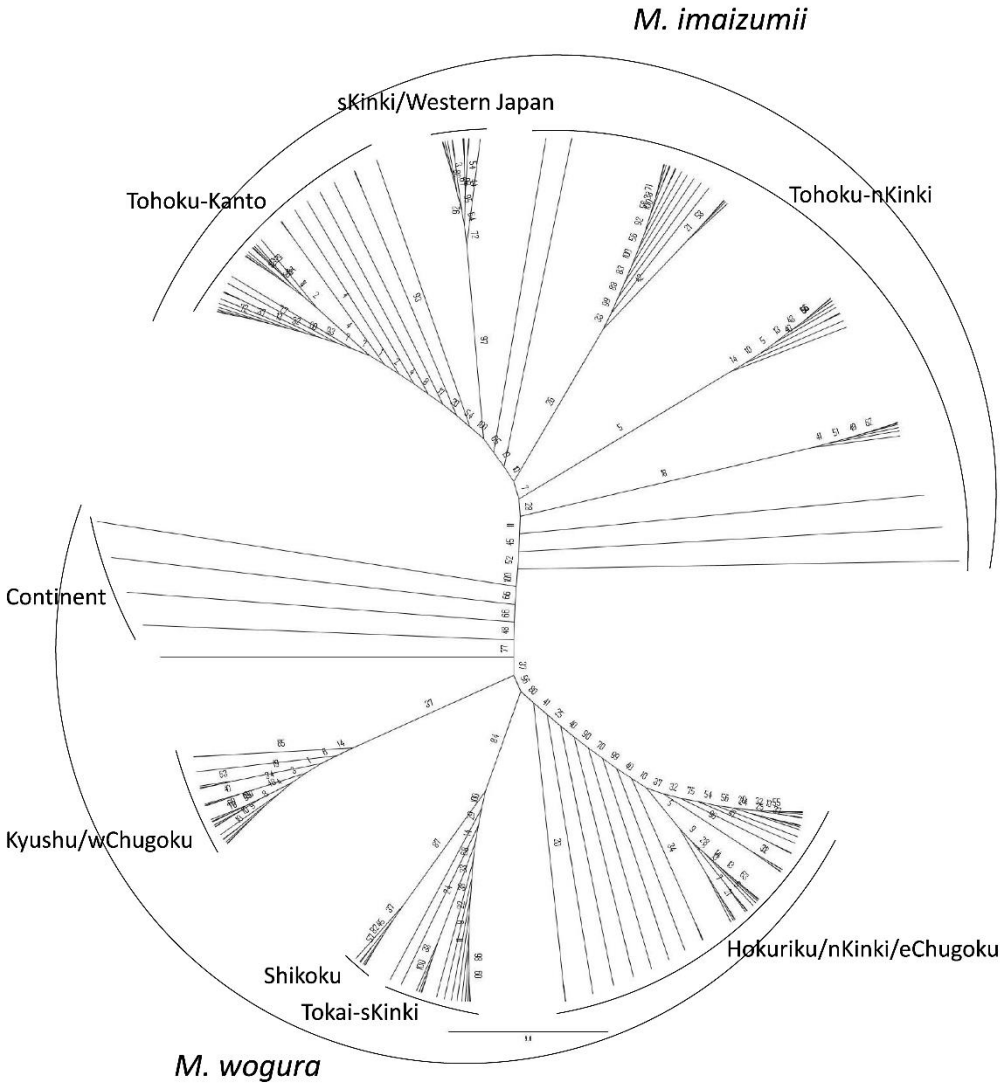
A



B



Supplementary Data S8.



Supplementary Data S9.

References

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y and Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633–2635.

Catchen J, Hohenlohe PA, Bassham S, Amores A, and Cresko WA. 2013. Stacks: an analysis tool set for population genomics. Molecular ecology 22: 3124–3140.

Catchen JM, Amores A, Hohenlohe P, Cresko W, and Postlethwait JH. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. G3: Genes| genomes| genetics 1: 171–182.

Chan KMA and Levin SA. 2007. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59: 720–729.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, and Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4: s13742–015–0047–8.

Currat M, Ruedi M, Petit RJ, and Excoffier L. 2008. The hidden side of invasions: massive introgression by local genes. Evolution 62: 1908–1920.

Darriba D, Psada D, Kozlov AM, Stamatakis AM, and Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Molecular Biology and Evolution 37: 291–294.

Dobson M. Patterns of distribution in Japanese land mammals. Mammal Review 24: 91–111.

Flouri T, Izquierdo-Carrasco F, Darriba D, Aberer AJ, Nguyen LT, Minh BQ, von Haeseler A, and Stamatakis A. 2014. The phylogenetic likelihood library. Systematic Biology 64: 356–362.

Greenwood PJ. 1980. Mating systems, philopatry and dispersal in birds and mammals. Animal behaviour 28: 1140–1162.

Hosoya S, Hirase S, Kikuchi K, Nanjo K, Nakamura Y, Kohno H, and Sano M. 2019. Random PCR-based genotyping by sequencing technology GRAS-Di (genotyping by

random amplicon sequencing, direct) reveals genetic structure of mangrove fishes. Molecular Ecology Resources 19: 1153–1163.

Huson DH and Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23: 254–267.

Huzita K. 1990. Manchidani unconformity and Rokko Movements -middle Pleistocene fault-block movement and sea-level rise in Kinki, Japan. The Quaternary Research 29: 337–349.

Ito T, Hayakawa T, Suzuki-Hashido nm, Hamada Y, Kurihara Y, Hanya G, Kaneko A, Natsume T, Aisu S, Honda T, Yachimori S, Anezaki T, Omi T, Hayama S, Tanaka M, Wakamori H, and Imai H. 2021. Phylogeographic history of Japanese macaques. Journal of Biogeography 48: 1420–1431.

Iwasa MA, Kawakubo C, Tsuchiya K, and Suzuki H. 2006. Intraspecific differentiation in the lesser Japanese mole in eastern Honshu, Japan, indicated by nuclear and mitochondrial gene analyses. Zoological Science 23: 955–961.

Johnson BB, White TA, Phillips CA, and Zamudio KR. 2015. Asymmetric introgression in a spotted salamander hybrid zone. Journal of Heredity 106: 608–617.

Kinoshita G, Sato T, Murakami S, Monakhov V, Kryukov AP, Frisman LV, Tsunamoto Y, Suyama Y, Murakami T, Suzuki H, and Sato JJ. 2024. Ice age land bridges to continental islands: repeated migration of the forest-dwelling sable in northeastern Asia. Journal of Biogeography 51: 924–939.

Kirihara T, Shinohara A, Tsuchiya K, Harada M, Kryukov AP, and Suzuki H. 2013. Spatial and temporal aspects of occurrence of Mogera species in the Japanese islands inferred from mitochondrial and nuclear gene sequences. Zoological Science 30: 267–281.

Kozlov AM, Darriba D, Flouri T, Morel B, and Stamatakis A. 2019. RAxML-NG: a fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35: 4453–4455.

Kunvar S, Czarnomska S, Pertoldi C, and Tokarska M. 2021. In search of species-specific SNPs in a non-model animal (European bison (*Bison bonasus*))-comparison of de novo

and reference-based integrated pipeline of STACKS using genotyping-by-sequencing (GBS) data. Animals 11: 2226.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv: 1303.3997.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Lischer HE and Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 28: 298–299.

Manni M, Berkeley MR, Seppey M, Simao FA, and Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eucaryotic, prokaryotic, and viral genomes. Molecular Biology and Evolution 38: 4647–4654.

Nakamoto A, Harada M, Mitsuhashi R, Tsuchiya K, Kryukov AP, Shinohara A, and Suzuki H. 2021. Influence of Quaternary environmental changes on mole populations inferred from mitochondrial sequences and evolutionary rate estimation. Zoological Letters 7: 1–11.

Okamoto M. 1999. Phylogeny of Japanese moles inferred from mitochondrial CO1 gene sequences. In (Yokohata Y and Nakamura S, eds.) Recent Advances in the Biology of Japanese Insectivora, pp. 21–27. Hiba Society of Natural History, Shobara.

Patterson N, Price AL, and Reich D. 2006. Population structure and eigenanalysis. PLoS genetics 2: e190.

Peterson BK, Weber JN, Kay EH, Fisher HS, and Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS One 7: e37135.

Price AL, Patterson NJ, Plenge RM, Weinblatt Me, Shadick NA, and Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 38: 904–909.

Raj A, Stephens M, and Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 197: 573–589.

Sato JJ. 2017. A review of the processes of mammalian faunal assembly in Japan: insights from molecular phylogenetics. In (Motokawa M and Kajihara H eds.) Species Diversity of Animals in Japan, pp. 49–60.

Sato JJ and Yasuda K. 2022. Ancient rivers shaped the current genetic diversity of the wood mouse (*Apodemus speciosus*) on the islands of the Seto Inland Sea, Japan. Zoological Letters 8: 9.

Satoh SS, Murata C, Yoshida K, Shibata K, and Tamate HB. 2019. Population genetic study of the lesser Japanese mole *Mogera imaizumii* using novel microsatellite markers with special reference to sex-biased dispersal. Mammal Study 44: 91–97.

Shafer AB, Peart CR, Tusso S, Maayan I, Brelsford A, Wheat CW, and Wolf JB. 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Methods in Ecology and Evolution 8: 907–917.

Sugiyama Y. 1991. Right-lateral strike-slip basins in the Second Paleo-Seto Inland Sea, -a model of basin development associated with the migration of active domain of a large-scale strike-slip fault. Structural Geology (The Journal of the Tectonic Research Group of Japan) 36: 99–108.

Suyama Y and Matsuki Y. 2015. MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. Scientific reports 5: 16963.

Toews DPL and Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. Molecular Ecology 21: 3907–3930.

Tsuchiya K, Suzuki H, Shinohara A, Harada M, Wakana S, Sakaizumi M, Han SH, Lin LK, and Kryukov AP. 2000. Molecular phylogeny of East Asian moles inferred from the sequence variation of the mitochondrial cytochrome *b* gene. Genes & Genetic Systems 75: 17–24.

Tsunoi T, Harada M, Notsu M, Mitsuhashi R, and Suzuki H. 2023. First record for the lesser Japanese mole *Mogera imaizumii* (Soricomorpha, Talpidae) from Dogo, Oki Islands, Japan. Mammalian Science 63: 63–68.

Weisenfeld NI, Kumar V, Shah P, Church DM, and Jaffe DB. 2017. Direct determination of diploid genome sequences. Genome research 27: 757–767.

Yoshida F. 1992. Geologic development of the Setouchi Geologic Province since Early Miocene -with special reference to the First and Second Setouchi Inland Sea times. Bulletin of the Geological Survey of Japan 43: 43–67.

Zemlemerova E, Abramov A, Kryukov A, Lebedev V, Min MS, Lee SJ, and Bannikova A. 2019. Genetic and morphologic diversity of the moles (Talpomorpha, Talpidae, Mogera) from the continental far east. Journal of Zoological Systematics and Evolutionary Research 57: 662–678.

PART IV: Whole-genome sequencing analysis of *Mogera imaizumii* and *M. wogura*

Introduction

Building on the insights gained from MIG-seq analysis, whole-genome sequencing provides a powerful tool to investigate genetic diversity and population structure at an unprecedented resolution. Unlike reduced representation methods which focus on a subset of the genome, whole-genome analysis enables the comprehensive assessment of genetic variation across the entire genome. This approach allows the identification of rare and novel variants, elucidates evolutionary processes in greater detail, and provides robust data for inferring historical demography, local adaptation, and patterns of gene flow.

In the context of subterranean mammals, whole-genome sequencing offers unique advantages, as their cryptic lifestyles and relatively low dispersal capabilities can lead to complex patterns of genetic differentiation that are not always apparent from traditional markers. For *M. imaizumii* and *M. wogura*, whole-genome analysis can help clarify unresolved questions, such as the historical climatic fluctuations on genome-wide diversity and the precise mechanisms underlying population differentiation.

In this study, we sequenced high-coverage whole genomes for a subset of individuals from both *M. imaizumii* and *M. wogura*, using a newly developed reference genome of *M. wogura*. By integrating whole-genome data with the findings from MIG-seq (PART III), we aim to: 1) refine our understanding of population structure and demographic history for both species, 2) detect signals of selection and adaptation across their genome. This comprehensive genomic approach not only complements the MIG-seq results but also provides a broader evolutionary context for understanding the factors shaping the diversity of subterranean mole species in Japan.

Materials and Methods

Reference genome construction

For the whole genome-based approach, another draft genome was created using the 10x Genomics linked reads data, which was generated in the previous chapter, Omni-C library data, and RNA-seq data additionally created in this chapter. Omni-C was performed using Dovetail Omni-C Kit (Dovetail Genomics #21005 8rxns) following the official protocol (Omni-C™Proximity Ligation Assay Mammalian Samples Protocol version 1.3) on

the same sample used to generate a draft genome in the previous chapter. 20 mg of heart and muscle tissues were cut and crushed with a Bead Crusher μT-01 (TAITEC). The samples were treated with 37% formaldehyde and washed twice. The washed samples were filtered through 200 μm and 50 μm filters and treated with 1X Nuclease Digest Buffer and Nuclease Enzyme Mix. The reaction was terminated with 0.5M EDTA resulting in lysates of the two samples. The lysates were quantified with Qubit Fluorometer (Thermo Fisher Scientific) and TapeStation, and the heart lysate, with a higher concentration, was chosen for further experiments. Proximity ligation was performed following the official protocol, but as the lysate concentration was lower than expected, the whole quantity of the heart lysate was used. For library preparation and index PCR, KAPA HyperPlus Library Preparation Kit was used. The library was sequenced on an Illumina HiSeq X platform (Illumina, San Diego, USA).

RNA-seq was also performed for gene annotation. The same specimen as the one used for draft genome construction was used. Brain, liver, heart, lung, stomach, kidney, gonad, muscle, and skin tissues were selected for RNA-seq. RNA extraction was performed with RNeasy Plis Universal Mini Kit 50 (QIAGEN).

10x Genomics linked reads amounting to 112.9Gb (809.7 M reads, read1 = 128 bp, read2 = 151 bp) were assembled using 10x Genomics assembly tool Supernova version 2.1.1 (Weisenfeld et al. 2017) with the following parameters: supernova --id=Mogera --fastqs=Mogera/FQ --localmem=480 --localcores=24 --maxreads=all. As the scaffolds assembled by Supernova contain redundant scaffolds, the redundancy was removed using a homology-based approach. Subsequently, the Omni-C library data (116.2 Gb, 384.7 M paired-end reads) was used to perform super-scaffolding on the non-redundant scaffolds with the programs SALSA2 version 2.3 (Ghurye et al. 2017; Ghurye et al. 2019) and 3D-DNA version 180922 (Dudchenko et al. 2017). SALSA2 was run with the default parameters and an additional parameter “-e DNASE”. 3D-DNA was run with the default parameters and an additional parameter “-i 1000”. The results obtained from SALSA2 and 3D-DNA were evaluated using BUSCO version 5.3.2 (lineage dataset = laurasiatheria_odb10, Manni et al. 2021). As a result, the assembly generated by SALSA2 had a higher rate of complete BUSCOs, higher scaffold N50, and lower scaffold L50 compared to the one generated by 3D-DNA and was selected as the draft genome for further analysis. The RNA-seq reads were mapped and assembled with Trinity version

2.14.0.

The assembled genome was repeat-masked by RepeatModeler version 2.0.3 (Flynn et al. 2020) and RepeatMasker version 4.1.4, and the gene was annotated with four patterns, i.e. using two different gene prediction tools and two different gene models. The gene prediction tools were BRAKER2 version 2.1.6 (Bruna et al. 2021) and GeMoMa version 1.9 (Keilwagen et al. 2018), and the gene models were *Talpa occidentalis* (talOcc4.1) and *Homo sapiens* (hg83). The obtained predictions were compared by gene models using GffCompare (Pertea and Pertea 2020). When a gene was included in both predictions, the gene obtained from hg83 gene model was adopted, while when a gene was included in only one of the predictions, it was adopted as it is. Gene annotation was performed by EnTAP version 0.10.8 (Hart et al. 2020). Three protein databases (refseq_vertebrate_mammalian, uniprot_sprot, uniprot_trembl_std) were used to complement the EggNog database implemented in EnTAP. The genes were subjected to homology searches in the protein databases, and those corresponding to vertebrates were retained. Other genes were subjected to homology searches in the EggNog database, and those with the "Mammals" Tax Scope were retained. The gff file was reconstructed with the annotation information (description) obtained from EnTAP. Gene symbols were added based on the NCBI data (https://ftp.ncbi.nlm.nih.gov/gnee/DATA/). Gene IDs and transcript IDs were newly constructed based on these gene symbols. The same IDs were assigned to the ones with the same gene symbols, except for the ones from different strands. The resulting number of genes was 21,901, and the number of transcripts was 22,614.

Whole genome sequencing

Six samples of *M. imaizumii* and ten samples of *M. wogura* were selected for whole genome sequencing. Two samples were selected from each genetic cluster (nc*Mim*-1–3, nc*Mwo*-1–5) defined by the MIG-seq-based analyses in PART III. Sample DNA was extracted from tissue samples (liver, muscle, or bone) preserved in 99.8% ethanol using a QIAmp DNA Mini Kit (QIAGEN, Hilden, Germany). The sequencing was outsourced to BGI Japan and performed using DNBseq platform with 150 bp pair-end sequencing reads. The quality of the sequence data was measured and visualized with FastQC version 0.11.9 (Andrews 2010) and MultiQC version 1.14 (Ewels et al. 2016). The reads were mapped onto the reference genome using BWA version 0.7.17-r1188 (Li 2013), and duplicate reads

were marked with samblaster version 0.1.24 (Faust and Hall 2014). The data were compressed using SAMtools version 1.7 (Li et al. 2009; Danecek et al. 2021), and further processing of data was performed with gatk4 version 4.0.5.1 (McKenna et al. 2010; Van der Auwera and O'Connor 2020). The SNPs and indels were called using gatk HaplotypeCaller. The reads were merged using gatk GenomicsDBImport, and genotyped with gatk GenotypeGVCFs, filtered by gatk VariantFiltration, followed by extraction of SNPs and indels with gatk SelectVariants. On average, 99.5% of *M. imaizumii* reads and 99.6% of *M. wogura* reads were mapped onto the reference genome. To calculate the mappability scores of the reference genome sequence, we ran GenMap version 1.3.0 (Pockrandt et al. 2020) with options -K 30 and -E 2, using positions with the score 1. An accessible mask file was created by excluding sites with a depth of coverage greater than twice and less than half the mode. The two mask files were intersected by bedtools version 2.27.1 (Quinlan and Hall 2010) into a single mask file. The variants were filtered with the mask file using BCFtools version 1.5 (Danecek et al. 2021) and used for further analyses. Analysis of population structure.

Population assignment analysis was performed using ADMIXTURE version 1.3.0 (Alexander et al. 2009) with K values (number of clusters) of 2–10. The most likely number of clusters (K value) was evaluated using the CV error score calculated by ADMIXTURE. The heterozygosity of each sample was calculated and so was their effective population size using generation time (approximately two years, Pacifici et al. 2014) and mutation rate of Pyrenean desman (5e-9, Escoda and Castresana 2020). Principal component analysis (PCA) was performed using the PLINK program 1.90b6.21 version (Chang et al. 2015; Purcell and Chang; www.cog-genomics.org/plink/1.9). PCA results were visualized using the R program 4.1.3 version. The phylogenetic relationships within species were visualized by constructing a Neighbor-Net and Neighbor-Joining tree using SplitsTree4 4.18.2 version (Huson and Bryant, 2006). The interspecies fixation index (F_{ST}) and intraspecies pairwise F_{ST} values were calculated using the SMARTPCA algorithm implemented in the EIGENSOFT package (Patterson et al. 2006; Price et al. 2006).

Mitochondrial sequences were assembled using GetOrganelle v1.7.7.0 (Jin et al. 2020). The mitochondrial sequences were then aligned using MEGA X (Kumar et al. 2018), and their phylogenetic relationships were visualized using RAxML-NG v. 1.1.0 (Kozlov et al. 2019) under TPM2uf+I+G4 model. The selection of the model was performed by

ModelTest-NG (Flouri et al. 2014; Darriba et al. 2020) v 0.1.7. A mitochondrial sequence of *M. wogura* downloaded from NCBI (accession number: AB099482.1) was included in this analysis.

Estimation of population divergence times and admixture between populations

To estimate the divergence times of the populations, we performed fastsimcoal28 (Excoffier et al. 2013; Excoffier et al. 2021; Excoffier et al. 2023) on both *Mogera* species. Divergence scenarios were constructed based on the results of the NJ phylogenetic tree (Fig. 20, 21). The fastsimcoal2 program was run 500 times with 10,000 simulations per parameter file with 40 loops. The best run was chosen by fsc-selectbestrun.sh (https://github.com/speciationgenomics/scripts/blob/master/fsc-selectbestrun.sh).

To test whether there are gene flows between the populations from the continent and the Japanese Archipelago, f4 statistics were calculated using the qpDstat program implemented in Admixtools. As the within-cluster F_{ST} were high even between the individuals from the same genetic cluster, the f4 statistics were calculated by considering each individual as one population. The outgroup was set to one *M. imaizumii* sample with the highest F_{ST} between the *M. wogura* sample of nc*Mwo*-4 (Supplementary Data S12).

Inference of past demography

Demographic history was estimated using PSMC version 0.6.5-r67 (Li and Durbin 2011) for all 16 samples of *M. imaizumii* and *M. wogura*. The generation time was set to two years based on Pacifici et al. (2014) and the mutation rate was set to 5×10^{-9} per site per generation using the figures of Pyrenean desman (*Galemys pyrenaicus*, Escoda and Castresana 2021). To create the input file for PSMC, we used the "mpileup" command of bcftools with the -C 50 option. The options -d and -D were set individually for each sample, as the third and double of mean coverage, respectively. PSMC was conducted with the options -N25 -t15 -r5 -p "4+25*2+4+6".

Analyses on the natural selection on genes

The coding sequences, exon sequences, and amino acid sequences were extracted from the draft genome and the gff file using gffread (Pertea and Pertea 2020). For annotation, snpEff (Cingolani et al. 2012) was run to annotate each SNP. The vcf file with the annotation information was run through a python code to count the number of SNPs

that were fixed or polymorphic in *M. imaizumii* and *M. wogura*. SNPs that were fixed in one species and absent in the other were counted as fixed SNPs. Otherwise, the SNPs were counted as polymorphic SNPs. As an exception, if a SNP was present in every individual, it was excluded from the calculation, since this can be interpreted as a polymorphism only observed in the reference genome. From the annotation information, the variant type of a SNP was evaluated, and "missense_variant" and "synonymous_variant" were counted as nonsynonymous and synonymous polymorphism, respectively. A chi-square test was performed to assess the significance of polymorphism presence, and genes with p values below 0.05 were considered significant.

Results

From the result of population assignment using ADMIXTURE, the *K* value with the smallest CV error value was two, discriminating the samples by species (Supplementary Data S10). The CV error value also dropped when $K = 8$, and this segregated the samples into eight groups with two samples each, consistent with the genetic clusters defined by the population structure analysis performed using MIG-seq data, although a sample (HS4491_ima_2) showed a sign of mixed ancestry. The calculated heterozygosity and estimated effective population size for each sample are as shown in Table 1. The F_{ST} between *M. imaizumii* and *M.wogura* was 0.733. The results of pairwise F_{ST} values calculated between genetic clusters and between individuals are shown in Tables 2 and 3. The lowest pairwise F_{ST} between individuals was 0.056 between two individuals from nc*Mim*-3 cluster, and this was the only pair with F_{ST} under 0.15 (Table 3).

The results of PCA show that PC1 separated the samples by their species, and PC2, PC3, and PC5 separated samples of *M. wogura*, while PC4 separated samples of *M. imaizumii* (Fig. 17) The samples from the same genetic cluster were very closely positioned.

The results of the Neighbor-Joining tree and the Neighbor-Net analysis are shown in Fig. 15 and 16. For *M. imaizumii*, nc*Mim*-3 was the first to diverge, followed by the divergence of nc*Mim*-1 and nc*Mim*-2. For *M. wogura*, a monophyletic group of nc*Mwo*-3 and nc*Mwo*-4 diverged, followed by nc*Mwo*-1, and subsequent divergence of nc*Mwo*-2/nc*Mwo*-5 group. The Neighbor-Net analysis result, however, is not as resolvable as the phylogenetic tree. For both *M. imaizumii* and *M. wogura*, the order of divergence is unclear. It is also shown that a nc*Mim*-2 sample (HS4491_ima_2) is genetically close to nc*Mim*-1.

The ML phylogenetic tree of mitochondrial sequences is shown in Fig. 18. While most samples were clustered to the same cluster defined in PART II, an individual of nc*Mim*-2 (HS4491_ima_2) was shown to host mtDNA of mt*Mim*-I. The phylogenetic relationship of *M. wogura* was in discordance with the results of *Cytb* analysis in previous studies, showing that mt*Mwo*-I was the first to diverge, followed by mt*Mwo*-IV. Unlike the Neighbor-Net result, mt*Mwo*-III cluster formed a monophyletic clade not with mt*Mwo*-IV, but with mt*Mwo*-II and mt*Mwo*-V.

The results of fastsimcoal2 are shown in Tables 4 and 5. The estimated divergence time of nc*Mim*-1/nc*Mim*-2 was 26,939 generations ago, and the divergence time of nc*Mim*-3 was 245,931 generations ago. Based on the generation time of approximately two years (759.1857 days, Pacifici et al. 2014), these divergence times are calculated to be approximately 54,000 and 491,000 years ago. For *M. wogura*, the basal divergence was 328,145 generations ago, or approximately 660,000 years ago based on the generation time of two years (759.1857 days, Pacifici et al. 2014). The divergence time of nc*Mwo*-3 and nc*Mwo*-4 was estimated to be around 250,000 years ago, while the divergence times between the populations of Honshu and Shikoku were more recent, approximately 79,000–110,000 years ago. From the outgroup f4 statistic test, it was shown that f4 statsitics were below zero for all sets of samples (Table 6).

The results of PSMC analysis are shown in Fig. 19. It is shown that for both *M. imaizumii* and *M. wogura*, individuals from the same genetic cluster tend to show similar demographic histories (Supplementary Data S13, S14). While *M. imaizumii* populations tend to expand once at around the Penultimate Glacial Maximum (PGM), *M. wogura* populations tend to expand twice before and after the PGM.

From the chi-square test, genes that were considered significant in terms of polymorphism presence were *TRIO* and *MAP9*. The number of the fixed nonsynonymous polymorphisms of these genes were both zero (Table 7), thus indicating that these genes were significantly conserved, in either one of the two species.

Discussion

Table 1 shows that individuals from the same genetic cluster do not always show similar heterozygosity and effective population size. This is not an unexpected result, as individuals can be sampled from geographically distant populations (Fig. 13). The

continental cluster (nc*Mwo*-4) showed relatively low heterozygosity, indicating that the continental populations experienced historical population bottlenecks. This could be a consequence of the severe environmental fluctuation during the Quaternary 100,000-year glacial cycle or could be an indication that these populations expanded from a founder group of limited size. The high F_{ST} value between the two continental individuals (0.624, Table 3) indicates that the continental populations are isolated from each other. Overall, continental moles might have rapidly expanded their distribution along the eastern coastlines of the Asian continent, followed by isolations to multiple refugia due to subsequent severe changes in the environment.

The F_{ST} values were high even within genetic clusters (Table 3). The within-cluster pair with the lowest F_{ST} value was nc*Mim*-3. This was to be expected, since the two samples from this genetic cluster were sampled from two geographically close sites (Fig. 13). The other pairs, however, showed high F_{ST} values, indicating high genetic divergence within genetic clusters. This could be explained by biological aspects of moles, characterized by low dispersal ability and highly territorial behavior. These features could have made it easier for mole populations to isolate, and harder to interact with each other, resulting in high F_{ST} values between populations.

The result of admixture analysis (Fig. 14) is consistent with the PCA results (Fig. 17), along with the MIG-seq results shown in PART III, indicating that the samples can generally be classified into eight genetic clusters. When K= 8, an individual (HS4491_ima_2) was shown to be assigned to other clusters, peculiarly to nc*Mwo*-3. This result can be related to the results of Neighbor-Net analysis (Fig. 16) and mitochondrial phylogenetic tree (Fig. 18), where this individual was shown to be placed near the nc*Mim*-1 cluster, and to be a host to mtDNA of mt*Mim*-I. These results indicate that this specific individual is a result of introgression between the two clusters. This seems to contradict with the high F_{ST} value between the two clusters, but this could be an artifact of a secondary contact between clusters, possibly an event right after their divergence. While the low dispersal ability of moles allows the mole populations to isolate easily over a long period of time, it could also make it easier for them to make a secondary contact right after their divergence. This could explain why the Neighbor-Net does not fully resolve how the genetic clusters diverged. The mole populations may have made multiple secondary contacts right after their divergence, complicating the relationships between clusters.

From the demographic inference by PSMC (Fig. 19), it can be supposed that the divergence time of the genetic clusters of *M. imaizumii* is at around 400,000–500,000 years ago, where the lines of effective population size begin to diverge from one another. This is consistent with the divergence time estimated by fastsimcoal2 (Table 4), where the first divergence of populations occurred at around 492,000 years ago. For *M. wogura*, the divergence of each cluster is supposed to be around 800,000–900,000 years ago. The continental cluster (nc*Mwo*-4) may have diverged earliest from the clusters from Japan, followed by the divergence of nc*Mwo*-1 and nc*Mwo*-3. The estimated divergence time of nc*Mwo*-3 and nc*Mwo*-4 from the populations of Honshu and Shikoku was at around 660,000 years ago (Fig. 5). This is slightly younger than the time of divergence indicated by PSMC. This might be because fastsimcoal2 underestimated the divergence time between the two clusters since our current model does not consider gene flows between populations. If there was a gene flow between nc*Mwo*-3 and nc*Mwo*-4, their divergence time would be underestimated. The negative f4 statistics suggest that the populations of the continent and the Japanese Archipelago were not segregated ever since their divergence, but there has been a certain level of introgression between the two populations. However, we detected gene flows between clusters nc*Mwo*-4 and nc*Mwo*-2/nc*Mwo*-3/nc*Mwo*-5, no gene flow has been detected between nc*Mwo*-4 and nc*Mwo*-1, indicating that the former gene flow was stronger than the latter. This leads us to hypothesize that *M. wogura* has migrated to the Japan Archipelago multiple times. After the first migration, *M. wogura* probably extended its distribution to the Tokai region, but the gene introgression caused by later migration(s) was hindered by the nc*Mwo*-1/nc*Mwo*-2 border. This can be related to the nKinki/sKinki border discussed in PART III.

The two clusters of Chugoku and Shikoku, nc*Mwo*-2 and nc*Mwo*-5, seem to have gone under similar demographic histories, thereby indicating that their divergence is not as severe as other clusters. However, fastsimcoal2 estimates their divergence times as old as 80,000 years ago, sometime after the PGM. Therefore, the abrupt warming after the PGM might have caused the sea level to rise, separating the two clusters. The similarity between the two clusters shown by PSMC could be simply because of their geological proximity. It is notable that for *M. imaizumii*, the effective population size seems to have grown during the Penultimate Glacial Maximum (130,000–190,000 ya), while for *M. wogura*, the effective population size seems to have dropped or restricted. This coincides with the hypothesis in

previous studies that *M. imaizumii* is more cold-adapted than *M. wogura*, and with our hypothesis that population of *M. imaizumii* expanded during the colder periods, while *M. wogura* contracted. There is however a doubt on whether the *M. imaizumii* populations from the eastern Japan (nc*Mim*-1 and nc*Mim*-2) merged with the relict populations in western Japan (nc*Mim*-3), since nc*Mim*-3 population growth is not as apparent as the other two, and if there was a population merging, it should have appeared as a sign of population growth in PSMC.

The post-LGP population expansion detected in PART II does not seem to be so apparent in the PSMC analysis. The reason is presumably because PSMC is not capable of estimating the demographic fluctuations of too recent a time scale.

From the results of MK test and the chi-square test, it has been shown that two genes, *TRIO* and *MAP9* were considered significant in terms of SNP presence. Since both genes had zero fixed nonsynonymous sites and high polymorphic nonsynonymous site rates (Table 6), it is presumable that the genes have gone under a negative selection, and the sequences are conserved among *Mogera* species. While the gene sequences hold many polymorphic sites, nonsynonymous variations do not lead to fixation.

Appendix IV. List of *Mogera* specimens used for whole genome sequencing.

Species	Genetic cluster	Location	Sample ID
M. imaizumii	nc*Mim*-1	Ikarashi, Niigata	HS470
M. imaizumii	nc*Mim*-1	Yuzawa, Akita	HS5788
M. imaizumii	nc*Mim*-2	Chikuma, Nagano	HS4491
M. imaizumii	nc*Mim*-2	Sendai, Miyagi	HS582
M. imaizumii	nc*Mim*-3	Mt. Wasayama, Nara	HS4620
M. imaizumii	nc*Mim*-3	Kozagawa, Kyoto	HS5758
M. wogura	nc*Mwo*-1	Hakone, Kanagawa	HS3096
M. wogura	nc*Mwo*-1	Nara, Nara	HS4898
M. wogura	nc*Mwo*-2	Akaiwa, Okayama	HS4585
M. wogura	nc*Mwo*-2	Takashima, Shiga	HS4884
M. wogura	nc*Mwo*-3	Waki, Yamaguchi	HS3066
M. wogura	nc*Mwo*-3	Kiyotake, Miyazaki	HS4899
M. wogura	nc*Mwo*-4	Hasan, Russia	HS1392
M. wogura	nc*Mwo*-4	Pusan, Korea	HS373
M. wogura	nc*Mwo*-5	Mima, Tokushima	HS4605
M. wogura	nc*Mwo*-5	Mannou, Kagawa	HS4953

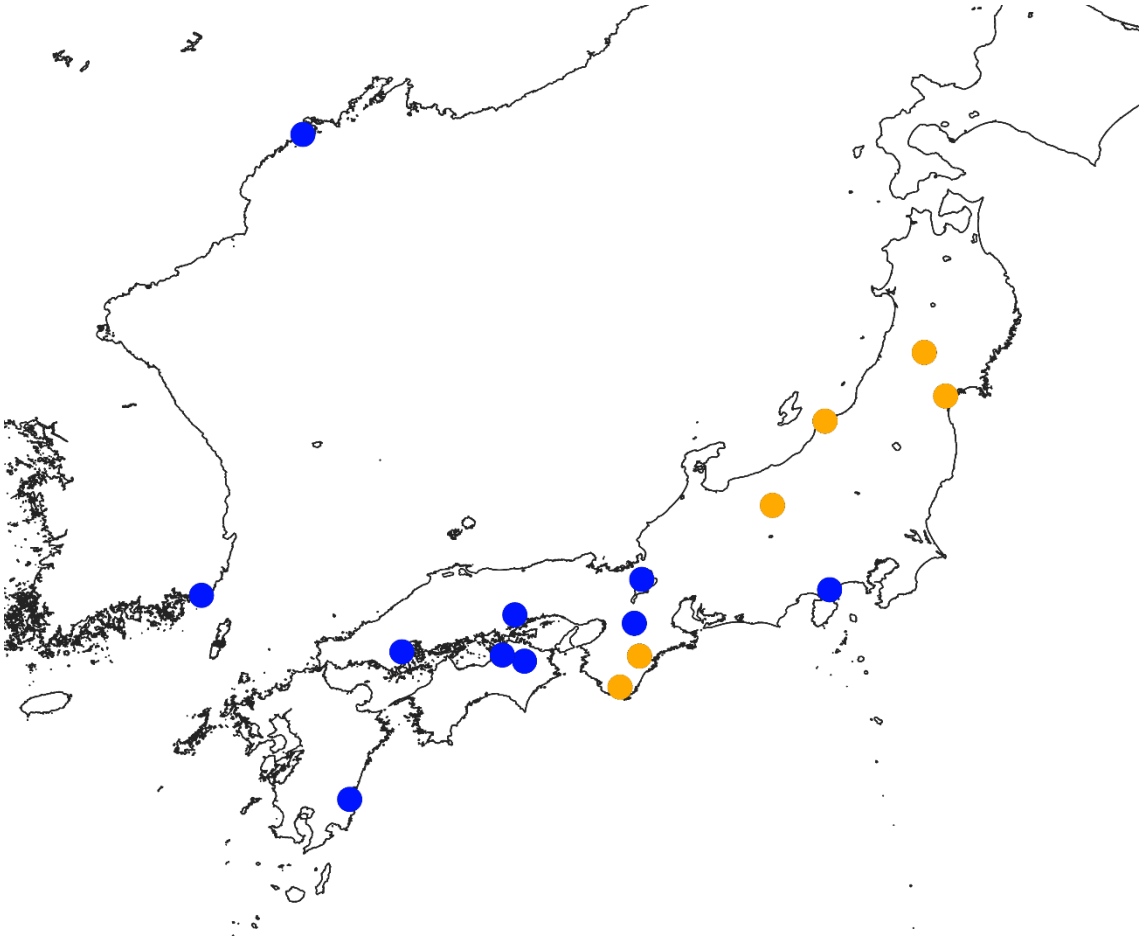


Figure 13. Geographical locations of the sampling sites where samples which whole-genome sequencing were conducted.

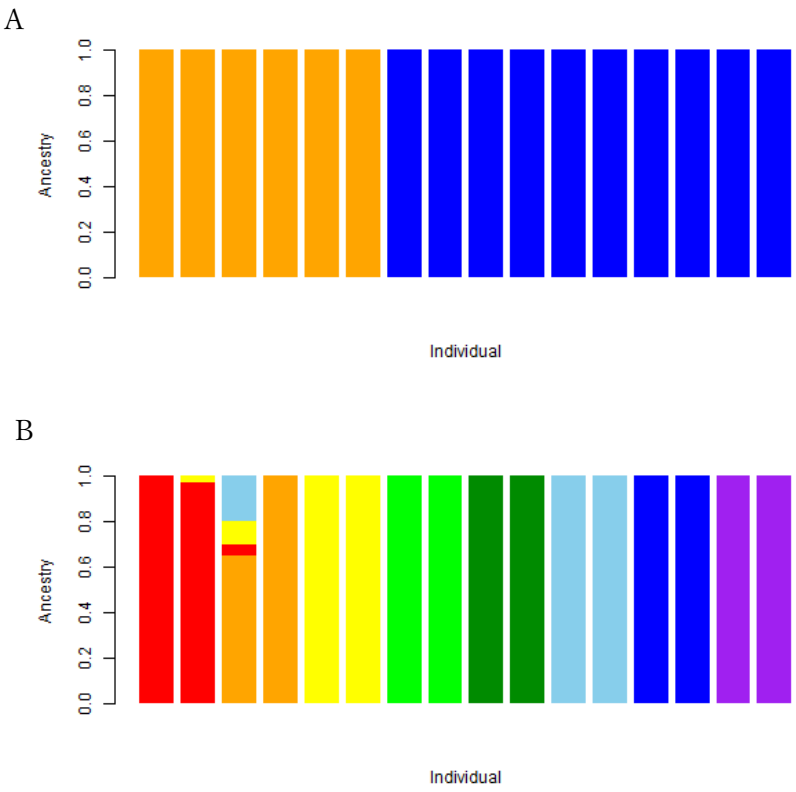


Figure 14. Population structures inferred by Admixture when *K*= 2 (A) and *K*= 8 (B).

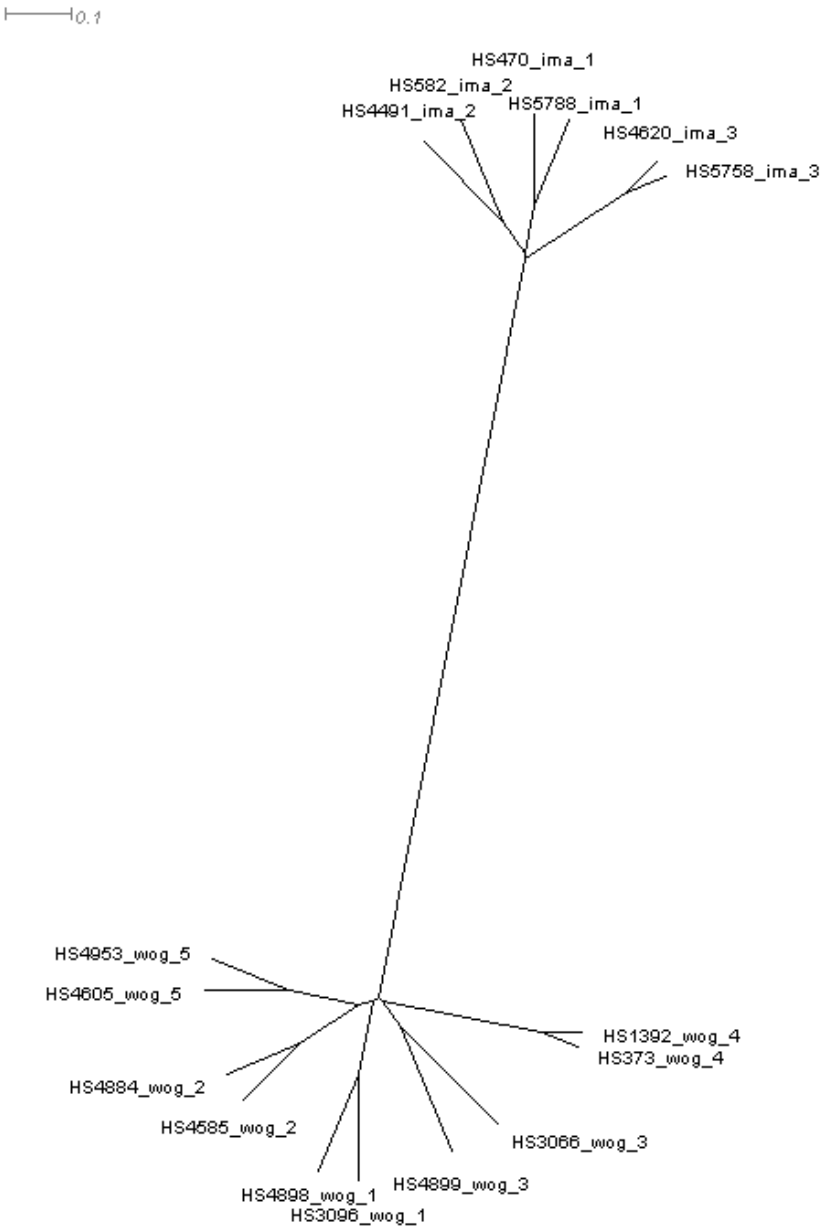


Figure 15. Neighbor-Joining tree of *M. imaizumii* and *M. wogura*.

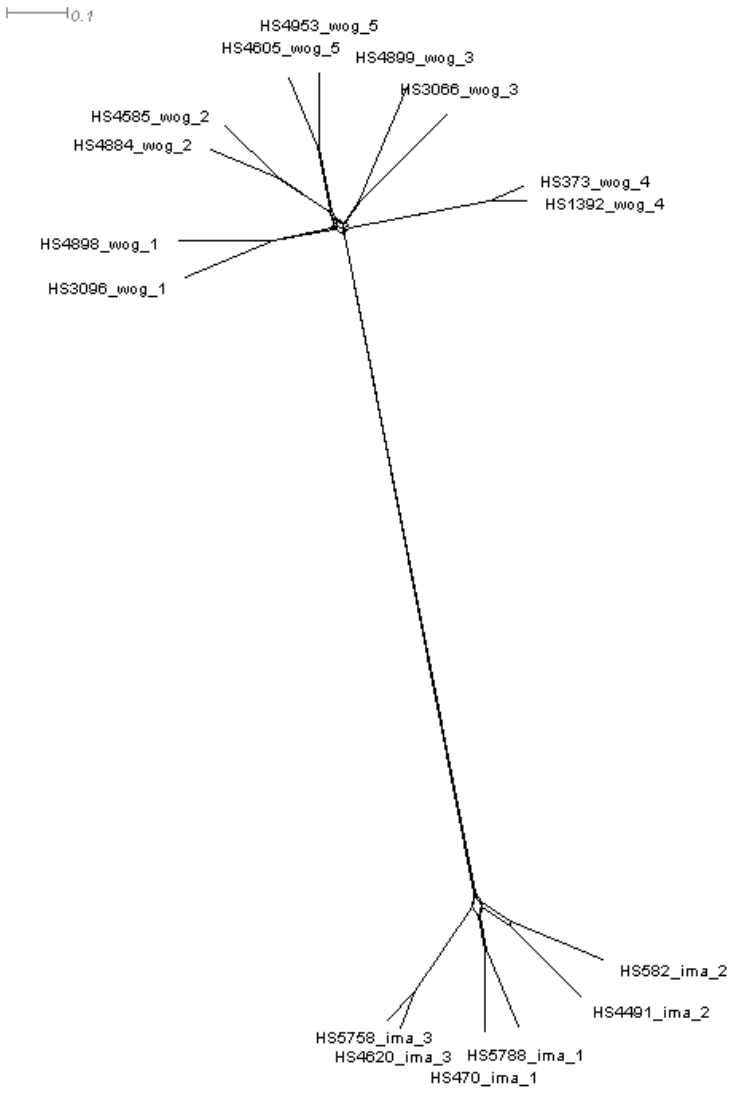


Figure 16. Neighbor-Net analysis on *M. imaizumii* and *M. wogura*.

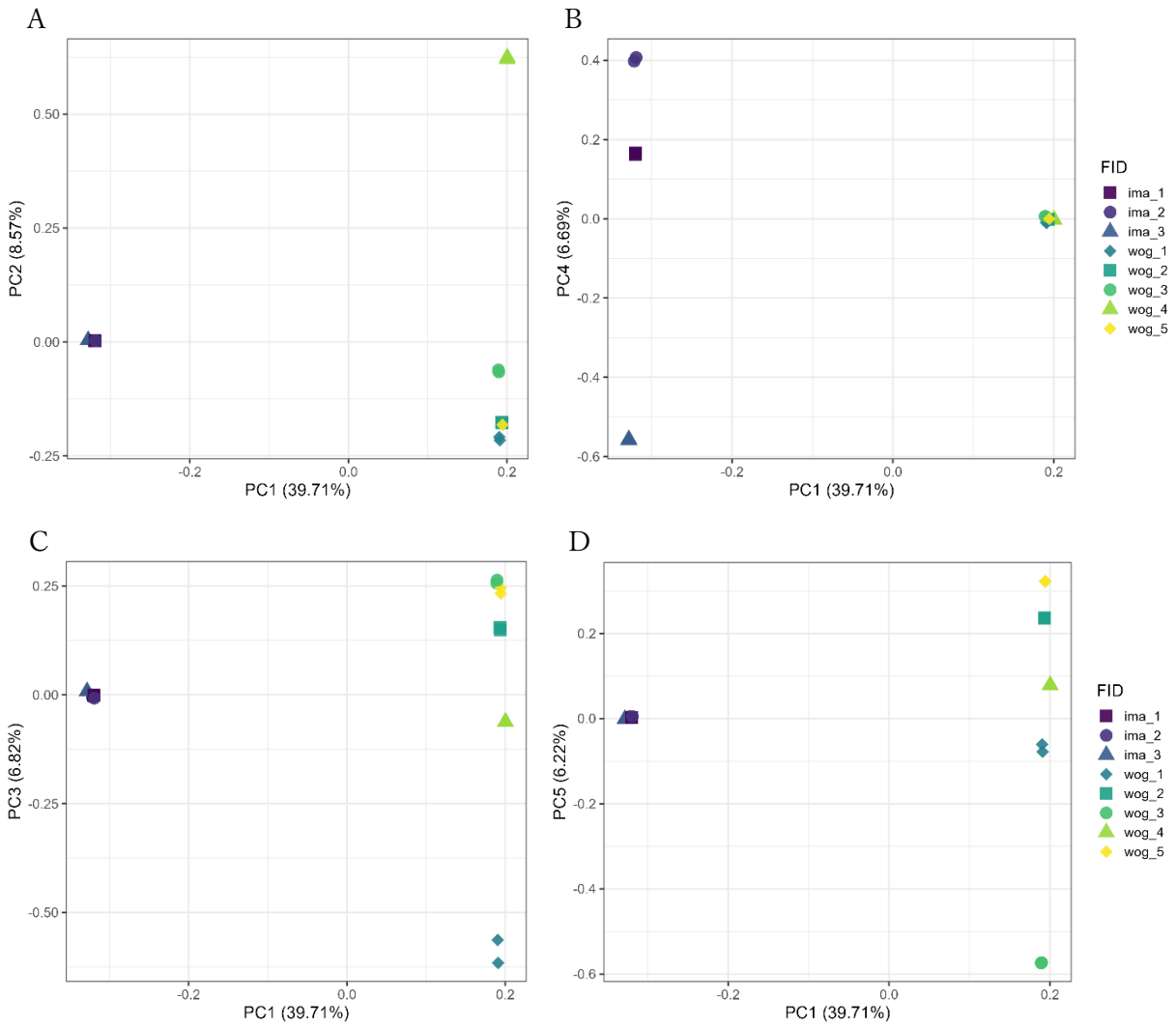


Figure 17. Results of PCA analysis. (A) PC1 vs PC2, (B) PC1 vs PC3, (C) PC1 vs PC4, (D) PC1 vs PC5. The legends (ima_1 to ima_3, wog_1 to wog_5) refer to the genetic clusters defined in PART III (nc*Mim*-1 to nc*Mim*-3, nc*Mwo*-1 to nc*Mwo*-5).

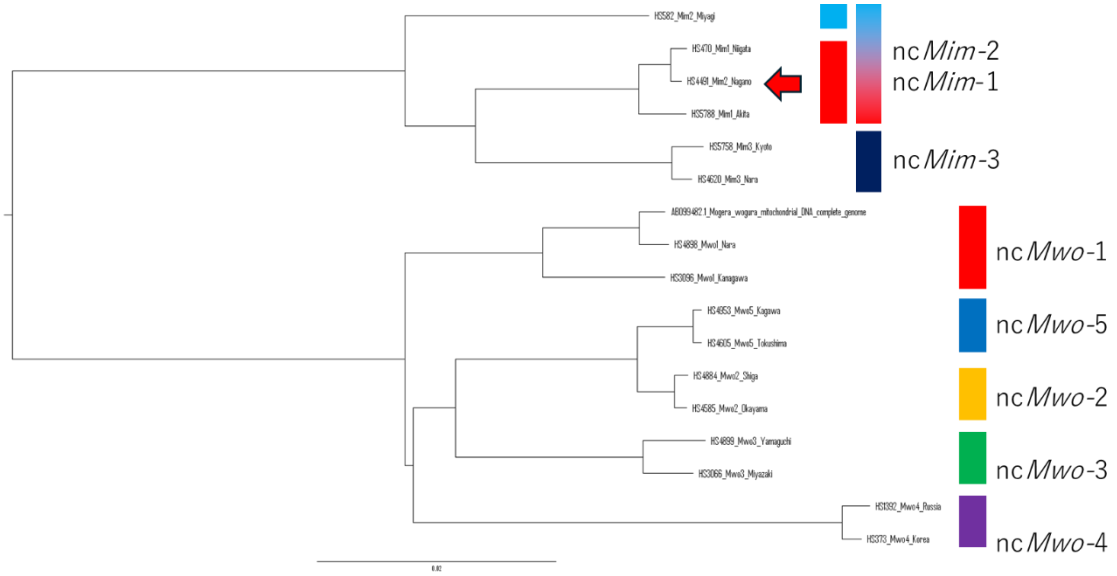


Figure 18. Maximum Likelihood phylogenetic tree of the whole mitochondrial sequences of *M. imaizumii* and *M. wogura*. Arrow indicates the individual that was host to mt*Mim*-I mitochondrial haplotype and belonged to nc*Mim*-2 nuclear genetic cluster.

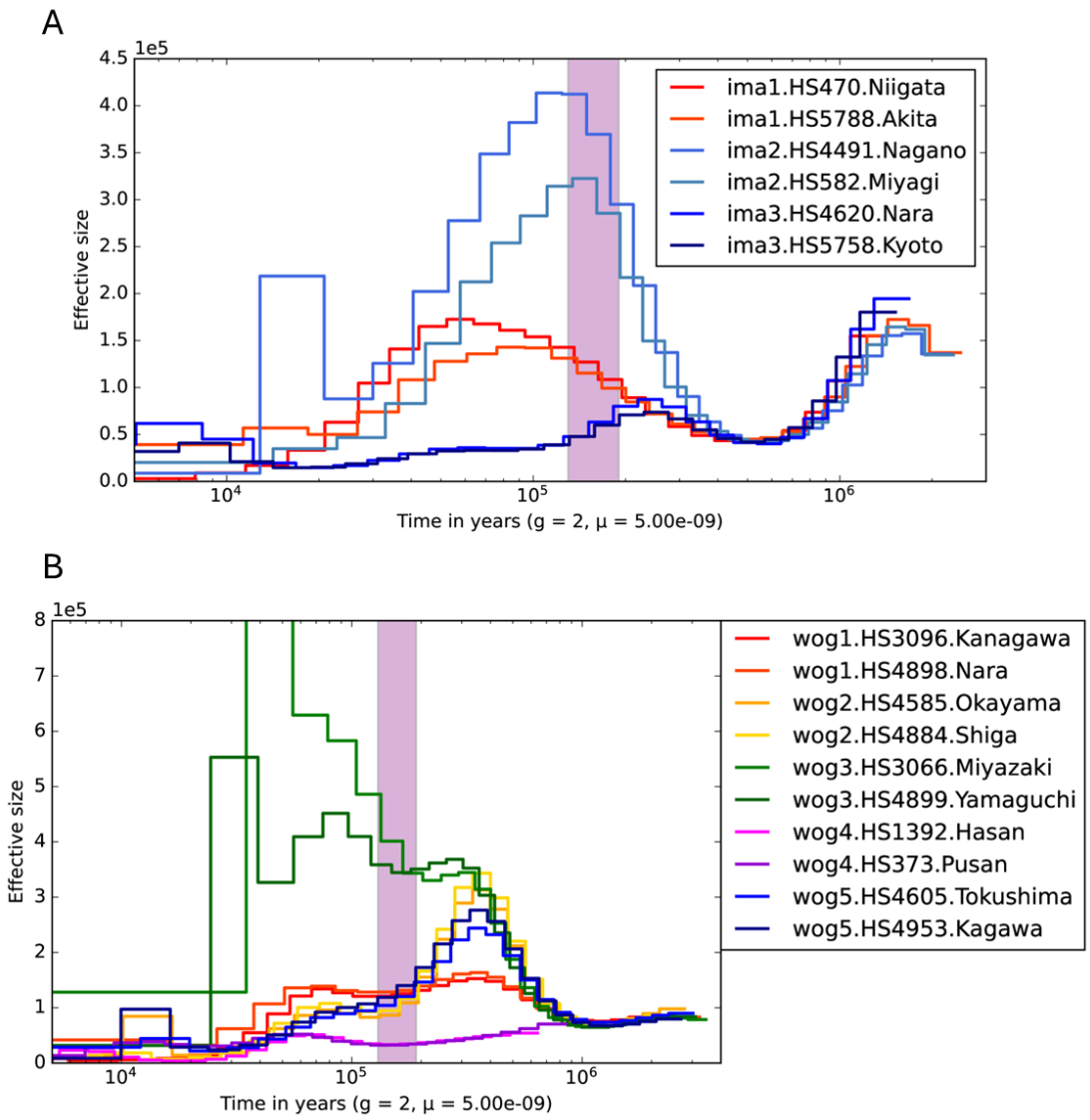


Figure 19. Demographic histories of *Mogera imaizumii* (A) and *M. wogura* (B) inferred by PSMC. The span of 130,000–190,000 years ago, representing the Penultimate Glacial Maximum (PGM) is highlighted in purple.

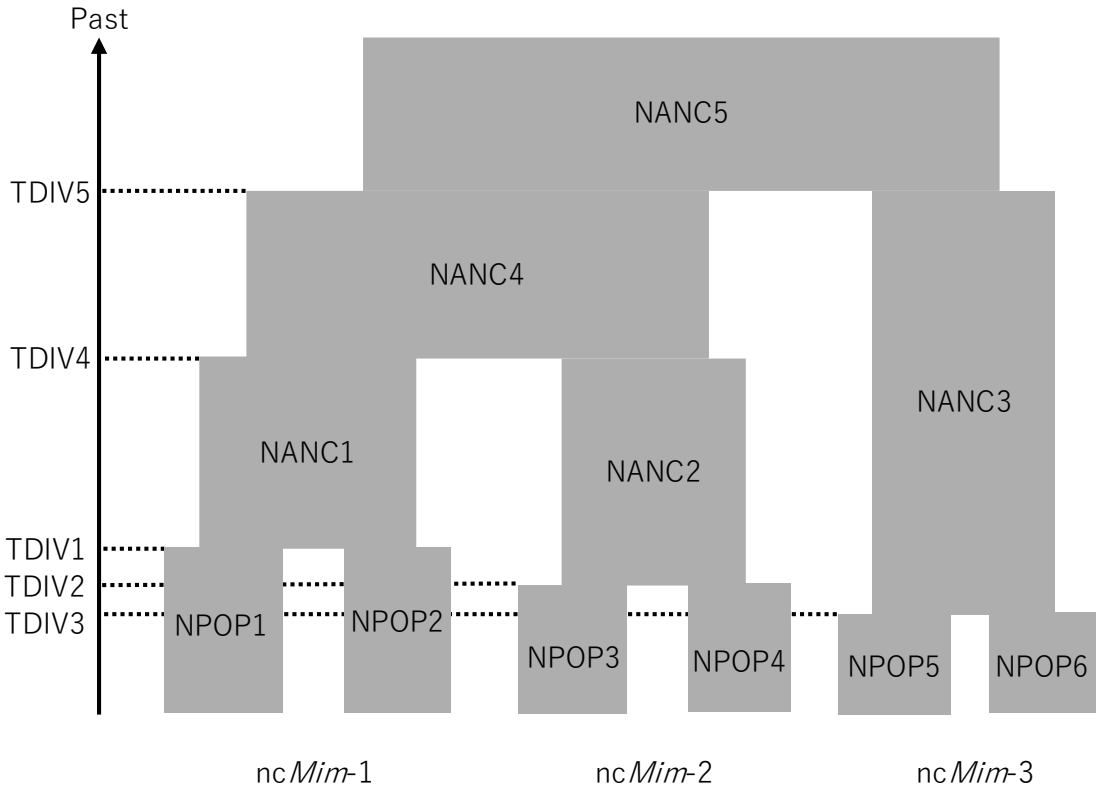


Fig. 20. Demographic model of *M. imaizumii* and parameters estimated by fastsimcoal2.

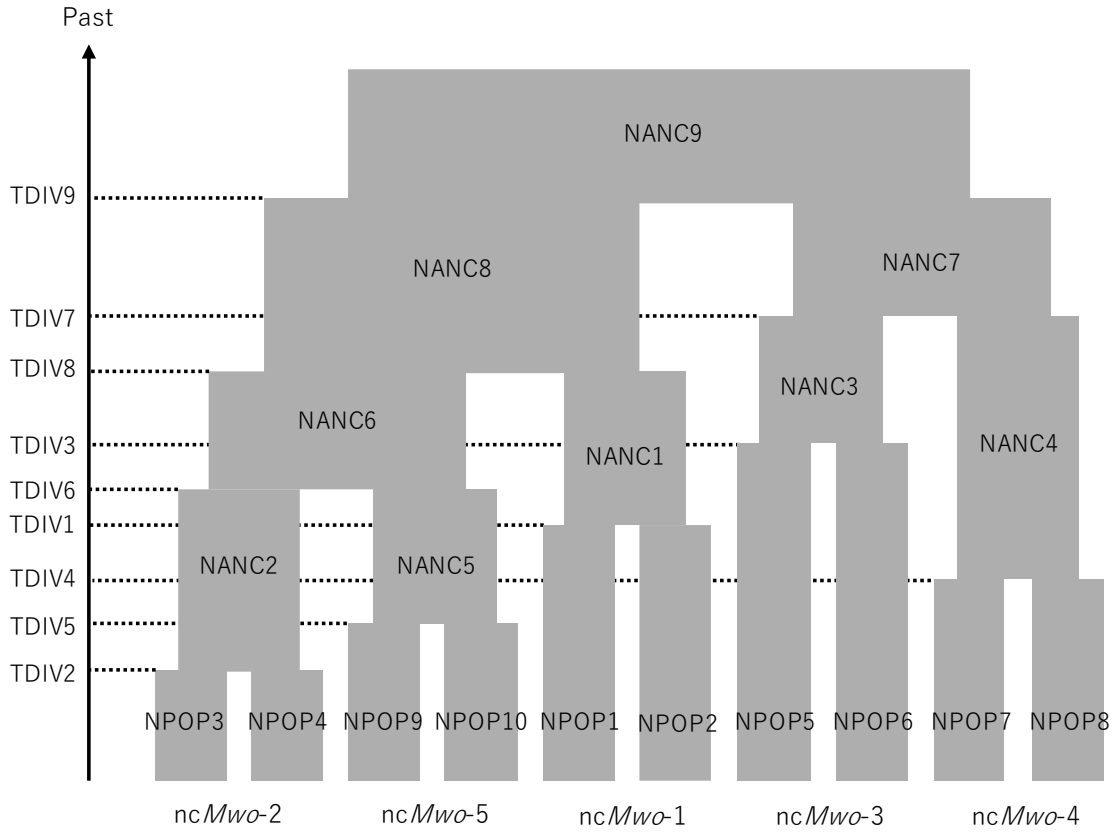


Fig. 21. Demographic model for *M. wogura* and parameters estimated by fastsimcoal2.

Sample ID	Homozygous Sites	Heterozygous Sites	Total Sites	Heterozygosity	Effective population size (mutation rate: 5e-9)
HS470_ima_1	1819699471	2885358	1845924511	0.0015631	78155
HS5788_ima_1	1820079767	5295873	1848566109	0.00286485	143242.5
HS4491_ima_2	1818717412	5260820	1847173950	0.00284804	142402
HS582_ima_2	1829259343	5710806	1858213911	0.00307328	153664
HS4620_ima_3	1819367935	3481835	1846736591	0.0018854	94270
HS5758_ima_3	1819254736	3199852	1846341186	0.00173308	86654
HS3096_wog_1	1870344891	3240498	1877354242	0.0017261	86305
HS4898_wog_1	1869060132	5629789	1878679711	0.00299667	149833.5
HS4585_wog_2	1862464020	4473327	1872854049	0.00238851	119425.5
HS4884_wog_2	1862465961	3837225	1872183236	0.0020496	102480
HS3066_wog_3	1858429772	8858003	1873566024	0.00472788	236394
HS4899_wog_3	1859096694	6591873	1872222490	0.00352088	176044
HS1392_wog_4	1859235580	1452113	1866508645	0.000777984	38899.2
HS373_wog_4	1858696841	1671878	1866390077	0.000895782	44789.1
HS4605_wog_5	1861385977	5145421	1872726219	0.00274756	137378
HS4953_wog_5	1859791992	4331854	1870204921	0.00231625	115812.5

Table 1. Heterozygosity calculated for each sample. The effective population size was calculated by dividing the heterozygosity by 4μ, where μ is the mutation rate of Pyrenean desman (Escoda and Castresana 2020).

Table 2. Pairwise F_{ST} values calculated between genetic clusters.

	ima_1	ima_2	ima_3	wog_1	wog_2	wog_3	wog_4	wog_5
ima_1	0							
ima_2	0.381	0						
ima_3	0.591	0.572	0					
wog_1	0.844	0.826	0.879	0				
wog_2	0.861	0.843	0.896	0.521	0			
wog_3	0.811	0.794	0.847	0.401	0.409	0		
wog_4	0.901	0.883	0.935	0.692	0.728	0.574	0	
wog_5	0.857	0.839	0.892	0.521	0.508	0.400	0.719	0

Table 3. F_{ST} values calculated between individuals. The table is ordered by the F_{ST} value

Sample 1	Sample 2	F_{ST}
HS4620_ima_3	HS5758_ima_3	0.056
HS4605_wog_5	HS4953_wog_5	0.193
HS3066_wog_3	HS4899_wog_3	0.194
HS4585_wog_2	HS4884_wog_2	0.275
HS4491_ima_2	HS582_ima_2	0.303
HS470_ima_1	HS5788_ima_1	0.384
HS3096_wog_1	HS4898_wog_1	0.401
HS3066_wog_3	HS4605_wog_5	0.425
HS3066_wog_3	HS4898_wog_1	0.43
HS3066_wog_3	HS4585_wog_2	0.449
HS3066_wog_3	HS4953_wog_5	0.457
HS5788_ima_1	HS582_ima_2	0.459
HS3066_wog_3	HS4884_wog_2	0.477
HS4491_ima_2	HS5788_ima_1	0.481
HS4605_wog_5	HS4899_wog_3	0.508
HS4898_wog_1	HS4899_wog_3	0.509
HS4585_wog_2	HS4899_wog_3	0.531
HS4899_wog_3	HS4953_wog_5	0.54
HS3066_wog_3	HS3096_wog_1	0.541
HS4585_wog_2	HS4605_wog_5	0.557
HS4605_wog_5	HS4898_wog_1	0.559
HS4884_wog_2	HS4899_wog_3	0.56
HS4585_wog_2	HS4898_wog_1	0.573
HS4898_wog_1	HS4953_wog_5	0.585
HS4605_wog_5	HS4884_wog_2	0.588
HS4585_wog_2	HS4953_wog_5	0.594
HS470_ima_1	HS582_ima_2	0.594
HS4884_wog_2	HS4898_wog_1	0.598
HS4620_ima_3	HS5788_ima_1	0.604
HS4491_ima_2	HS470_ima_1	0.615

HS5758_ima_3	HS5788_ima_1	0.617
HS3066_wog_3	HS373_wog_4	0.619
HS3096_wog_1	HS4899_wog_3	0.621
HS1392_wog_4	HS373_wog_4	0.624
HS4884_wog_2	HS4953_wog_5	0.624
HS4620_ima_3	HS582_ima_2	0.627
HS1392_wog_4	HS3066_wog_3	0.631
HS5758_ima_3	HS582_ima_2	0.639
HS4491_ima_2	HS4620_ima_3	0.647
HS4491_ima_2	HS5758_ima_3	0.66
HS3096_wog_1	HS4605_wog_5	0.675
HS3096_wog_1	HS4585_wog_2	0.692
HS373_wog_4	HS4899_wog_3	0.698
HS3096_wog_1	HS4953_wog_5	0.702
HS1392_wog_4	HS4899_wog_3	0.71
HS3096_wog_1	HS4884_wog_2	0.718
HS4620_ima_3	HS470_ima_1	0.731
HS470_ima_1	HS5758_ima_3	0.744
HS373_wog_4	HS4898_wog_1	0.748
HS1392_wog_4	HS4898_wog_1	0.759
HS373_wog_4	HS4605_wog_5	0.769
HS1392_wog_4	HS4605_wog_5	0.78
HS373_wog_4	HS4585_wog_2	0.792
HS373_wog_4	HS4953_wog_5	0.797
HS1392_wog_4	HS4585_wog_2	0.803
HS1392_wog_4	HS4953_wog_5	0.808
HS3066_wog_3	HS582_ima_2	0.812
HS373_wog_4	HS4884_wog_2	0.816
HS3066_wog_3	HS5788_ima_1	0.818
HS3066_wog_3	HS4491_ima_2	0.821
HS1392_wog_4	HS4884_wog_2	0.827
HS4899_wog_3	HS582_ima_2	0.839
HS4899_wog_3	HS5788_ima_1	0.844
HS4491_ima_2	HS4899_wog_3	0.847

HS3066_wog_3	HS4620_ima_3	0.849
HS3096_wog_1	HS373_wog_4	0.849
HS4898_wog_1	HS582_ima_2	0.85
HS3066_wog_3	HS470_ima_1	0.853
HS3066_wog_3	HS5758_ima_3	0.853
HS4898_wog_1	HS5788_ima_1	0.855
HS4491_ima_2	HS4898_wog_1	0.858
HS1392_wog_4	HS3096_wog_1	0.86
HS4605_wog_5	HS582_ima_2	0.86
HS4605_wog_5	HS5788_ima_1	0.865
HS4491_ima_2	HS4605_wog_5	0.868
HS4585_wog_2	HS582_ima_2	0.868
HS4953_wog_5	HS582_ima_2	0.87
HS4585_wog_2	HS5788_ima_1	0.873
HS4953_wog_5	HS5788_ima_1	0.875
HS4620_ima_3	HS4899_wog_3	0.876
HS4884_wog_2	HS582_ima_2	0.876
HS4491_ima_2	HS4585_wog_2	0.877
HS4491_ima_2	HS4953_wog_5	0.878
HS4899_wog_3	HS5758_ima_3	0.879
HS470_ima_1	HS4899_wog_3	0.88
HS4884_wog_2	HS5788_ima_1	0.882
HS4491_ima_2	HS4884_wog_2	0.885
HS3096_wog_1	HS582_ima_2	0.886
HS4620_ima_3	HS4898_wog_1	0.886
HS470_ima_1	HS4898_wog_1	0.89
HS4898_wog_1	HS5758_ima_3	0.89
HS3096_wog_1	HS5788_ima_1	0.892
HS3096_wog_1	HS4491_ima_2	0.895
HS4605_wog_5	HS4620_ima_3	0.896
HS4605_wog_5	HS5758_ima_3	0.9
HS4605_wog_5	HS470_ima_1	0.901
HS4585_wog_2	HS4620_ima_3	0.904
HS4620_ima_3	HS4953_wog_5	0.906

HS4585_wog_2	HS5758_ima_3	0.908
HS4585_wog_2	HS470_ima_1	0.909
HS470_ima_1	HS4953_wog_5	0.91
HS4953_wog_5	HS5758_ima_3	0.91
HS373_wog_4	HS582_ima_2	0.912
HS4620_ima_3	HS4884_wog_2	0.913
HS1392_wog_4	HS582_ima_2	0.915
HS4884_wog_2	HS5758_ima_3	0.916
HS373_wog_4	HS5788_ima_1	0.917
HS470_ima_1	HS4884_wog_2	0.917
HS373_wog_4	HS4491_ima_2	0.92
HS1392_wog_4	HS5788_ima_1	0.921
HS3096_wog_1	HS4620_ima_3	0.922
HS1392_wog_4	HS4491_ima_2	0.924
HS3096_wog_1	HS5758_ima_3	0.926
HS3096_wog_1	HS470_ima_1	0.927
HS373_wog_4	HS4620_ima_3	0.947
HS373_wog_4	HS5758_ima_3	0.95
HS1392_wog_4	HS4620_ima_3	0.951
HS373_wog_4	HS470_ima_1	0.952
HS1392_wog_4	HS5758_ima_3	0.954
HS1392_wog_4	HS470_ima_1	0.956

parameter	estimated value
NPOP1	11632
NPOP2	38946
NPOP3	25951
NPOP4	22744
NPOP5	13955
NPOP6	13919
NANC1	57248
NANC2	116079
NANC3	189385
NANC4	1016762
NANC5	61044
TDIV1	12511
TDIV2	11514
TDIV3	8950
TDIV4	26939
TDIV5	245931
Max Estimated Likelihood	-425176018.374
Max Observed Likelihood	-420852771.858

Table 4. Parameter values estimated by fastsimcoal2 for *Mogera imaizumii*. NPOP parameters represent the population size (number of alleles) of current populations. NANC parameters represent the number of alleles of ancestral populations. TDIV parameters represent the divergence times by the number of generations. See Fig. . for the detailed properties of the parameters.

parameters	estimated value
NPOP1	32574
NPOP2	95568
NPOP3	24714
NPOP4	17651
NPOP5	275956
NPOP6	111136
NPOP7	12054
NPOP8	15081
NPOP9	64827
NPOP10	39623
NANC1	53160
NANC2	56542
NANC3	428495
NANC4	106211
NANC5	40996
NANC6	322331
NANC7	1180672
NANC8	1718175
NANC9	25350
TDIV1	36137
TDIV2	10256
TDIV3	48697
TDIV4	27964
TDIV5	17044
TDIV6	39742
TDIV7	126311
TDIV8	56630
TDIV9	328145
Max Estimated Likelihood	-1413265739.755
Max Observed Likelihood	-1403328516.836

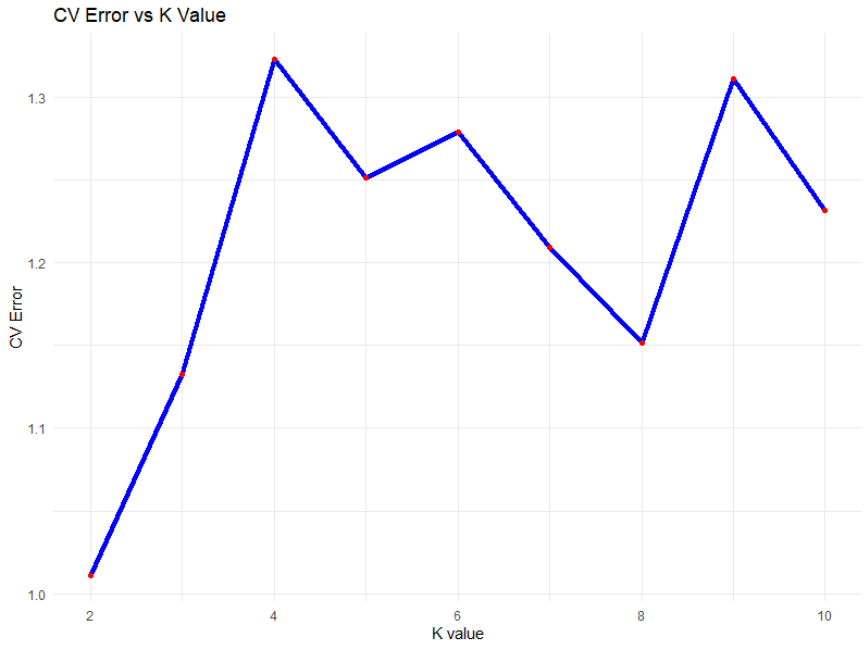
Table 5. Parameter values estimated by fastsimcoal2 for *M. wogura*. NPOP parameters represent the number of alleles in current populations. NANC parameters represent the number of alleles in ancestral populations. TDIV parameters represent divergence times in the number of generations.

Table 6. The results of f4 statistics.

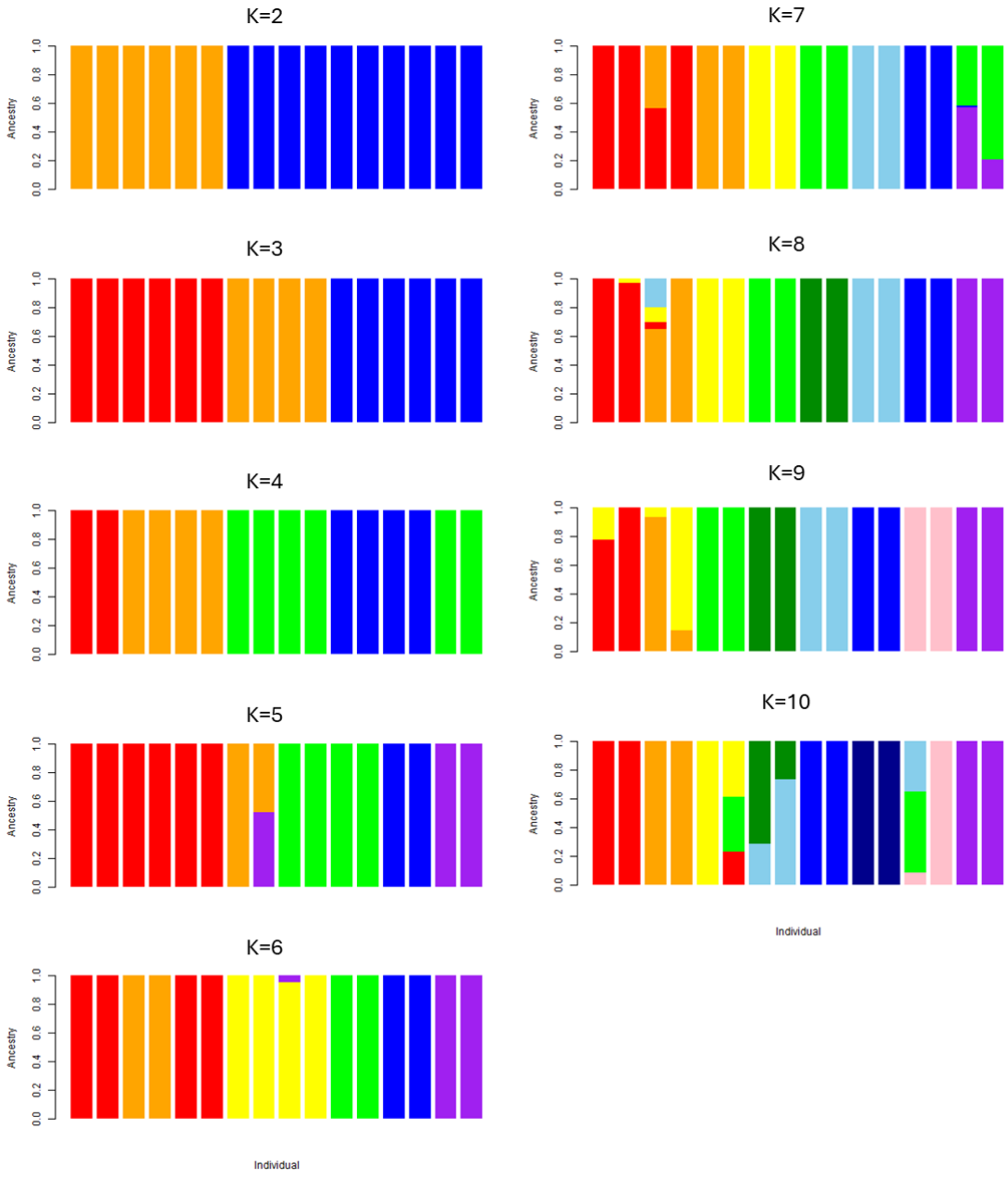
W (outgroup)	X	Y	Z	f4	Z score
HS470_ima_1	HS373_wog_4	HS4585_wog_2	HS3096_wog_1	-0.002122	-42.824
HS470_ima_1	HS373_wog_4	HS4585_wog_2	HS4898_wog_1	-0.002109	-42.741
HS470_ima_1	HS373_wog_4	HS4884_wog_2	HS3096_wog_1	-0.002104	-42.925
HS470_ima_1	HS373_wog_4	HS4884_wog_2	HS4898_wog_1	-0.002093	-42.799
HS470_ima_1	HS1392_wog_4	HS4585_wog_2	HS3096_wog_1	-0.002108	-42.664
HS470_ima_1	HS1392_wog_4	HS4585_wog_2	HS4898_wog_1	-0.002073	-41.609
HS470_ima_1	HS1392_wog_4	HS4884_wog_2	HS3096_wog_1	-0.002092	-42.674
HS470_ima_1	HS1392_wog_4	HS4884_wog_2	HS4898_wog_1	-0.002058	-41.872
HS470_ima_1	HS373_wog_4	HS3066_wog_3	HS3096_wog_1	-0.003862	-45.302
HS470_ima_1	HS373_wog_4	HS3066_wog_3	HS4898_wog_1	-0.00385	-45.079
HS470_ima_1	HS373_wog_4	HS4899_wog_3	HS3096_wog_1	-0.003859	-44.927
HS470_ima_1	HS373_wog_4	HS4899_wog_3	HS4898_wog_1	-0.003847	-44.668
HS470_ima_1	HS1392_wog_4	HS3066_wog_3	HS3096_wog_1	-0.003853	-45.438
HS470_ima_1	HS1392_wog_4	HS3066_wog_3	HS4898_wog_1	-0.003817	-44.724
HS470_ima_1	HS1392_wog_4	HS4899_wog_3	HS3096_wog_1	-0.00385	-44.884
HS470_ima_1	HS1392_wog_4	HS4899_wog_3	HS4898_wog_1	-0.003816	-44.304
HS470_ima_1	HS373_wog_4	HS1392_wog_4	HS3096_wog_1	-0.056517	-100
HS470_ima_1	HS373_wog_4	HS1392_wog_4	HS4898_wog_1	-0.056514	-100
HS470_ima_1	HS1392_wog_4	HS373_wog_4	HS3096_wog_1	-0.056513	-100
HS470_ima_1	HS1392_wog_4	HS373_wog_4	HS4898_wog_1	-0.056487	-100
HS470_ima_1	HS373_wog_4	HS4605_wog_5	HS3096_wog_1	-0.002298	-44.796
HS470_ima_1	HS373_wog_4	HS4605_wog_5	HS4898_wog_1	-0.002287	-44.341
HS470_ima_1	HS373_wog_4	HS4953_wog_5	HS3096_wog_1	-0.002206	-43.248
HS470_ima_1	HS373_wog_4	HS4953_wog_5	HS4898_wog_1	-0.002195	-42.689
HS470_ima_1	HS1392_wog_4	HS4605_wog_5	HS3096_wog_1	-0.002291	-44.089
HS470_ima_1	HS1392_wog_4	HS4605_wog_5	HS4898_wog_1	-0.002256	-43.413
HS470_ima_1	HS1392_wog_4	HS4953_wog_5	HS3096_wog_1	-0.002194	-42.807
HS470_ima_1	HS1392_wog_4	HS4953_wog_5	HS4898_wog_1	-0.002161	-41.71

gene	Dn	Ds	Pn	Ps	p-value
TRIO	0	19	29	41	0.036
MAP9	0	7	9	1	0.038

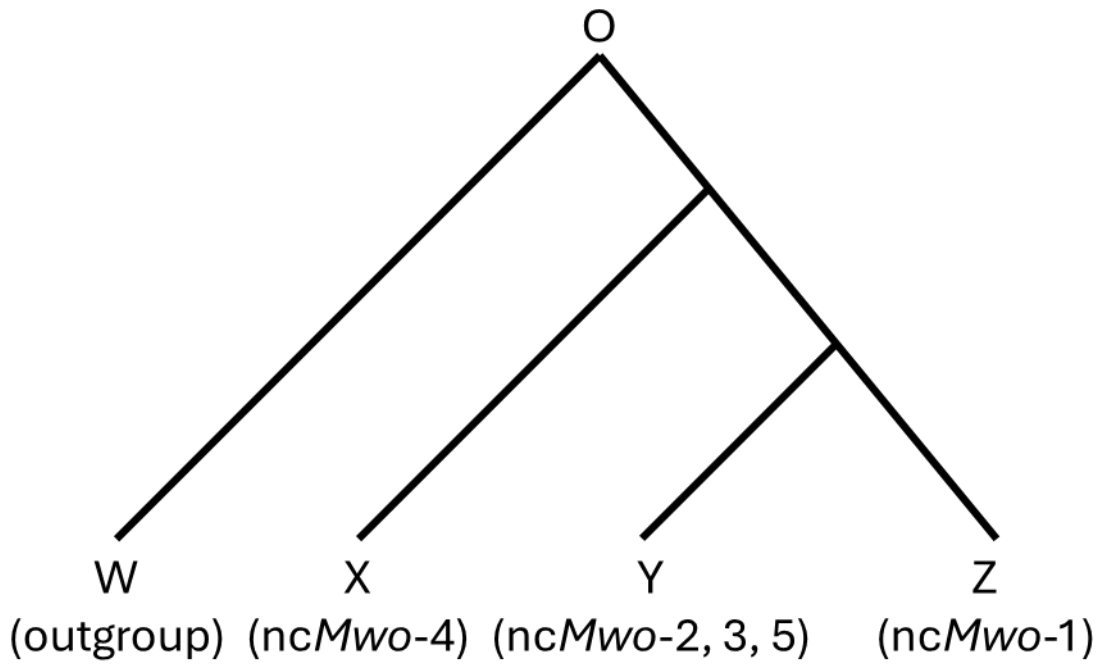
Table 7. Number of fixed nonsynonymous (Dn), fixed synonymous (Ds), polymorphic nonsynonymous (Pn), and polymorphic synonymous (Ps) sites in a gene.



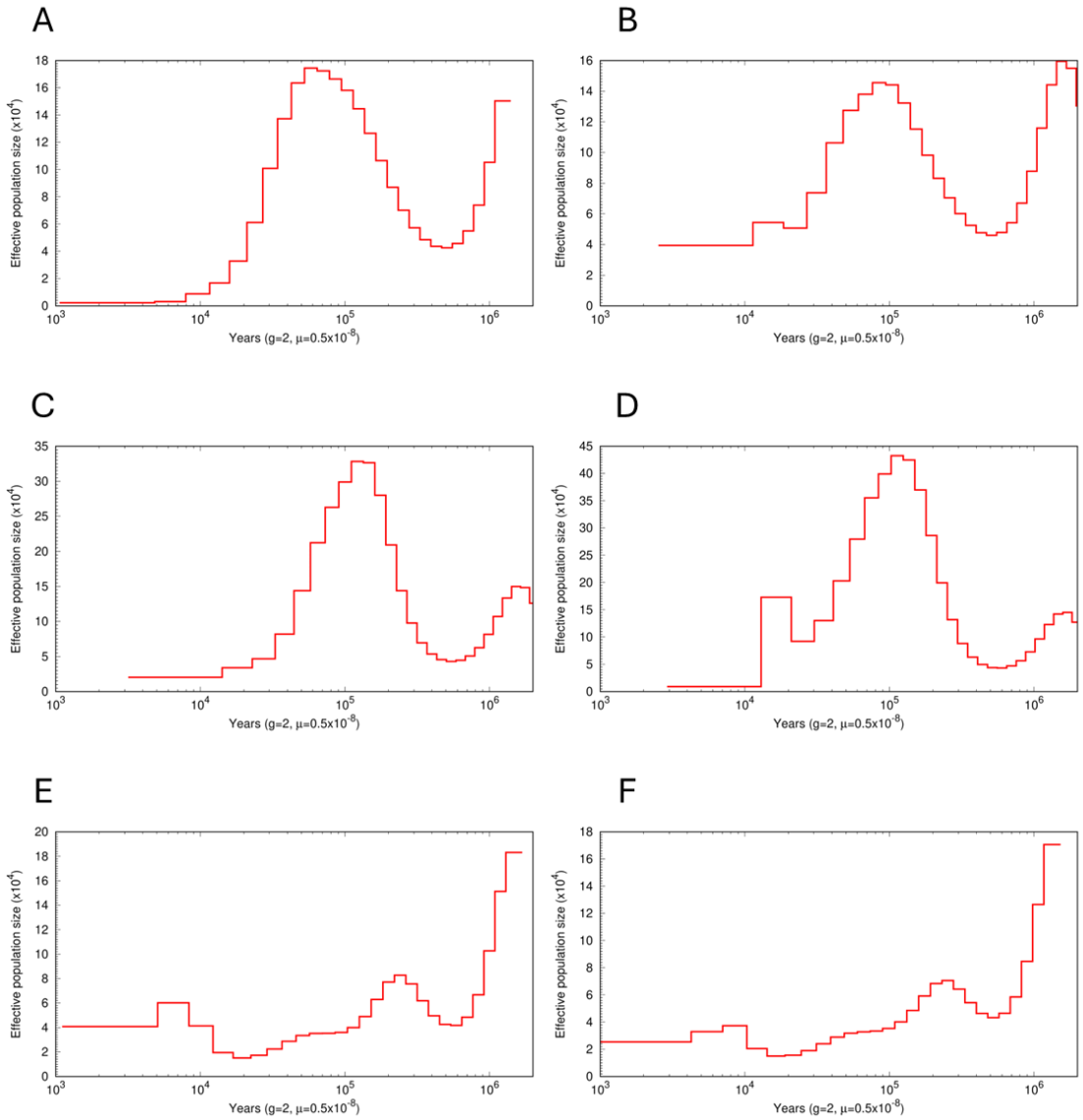
Supplementary Data S10. Fluctuation of CV error score of each *K* value in ADMIXTURE analysis.



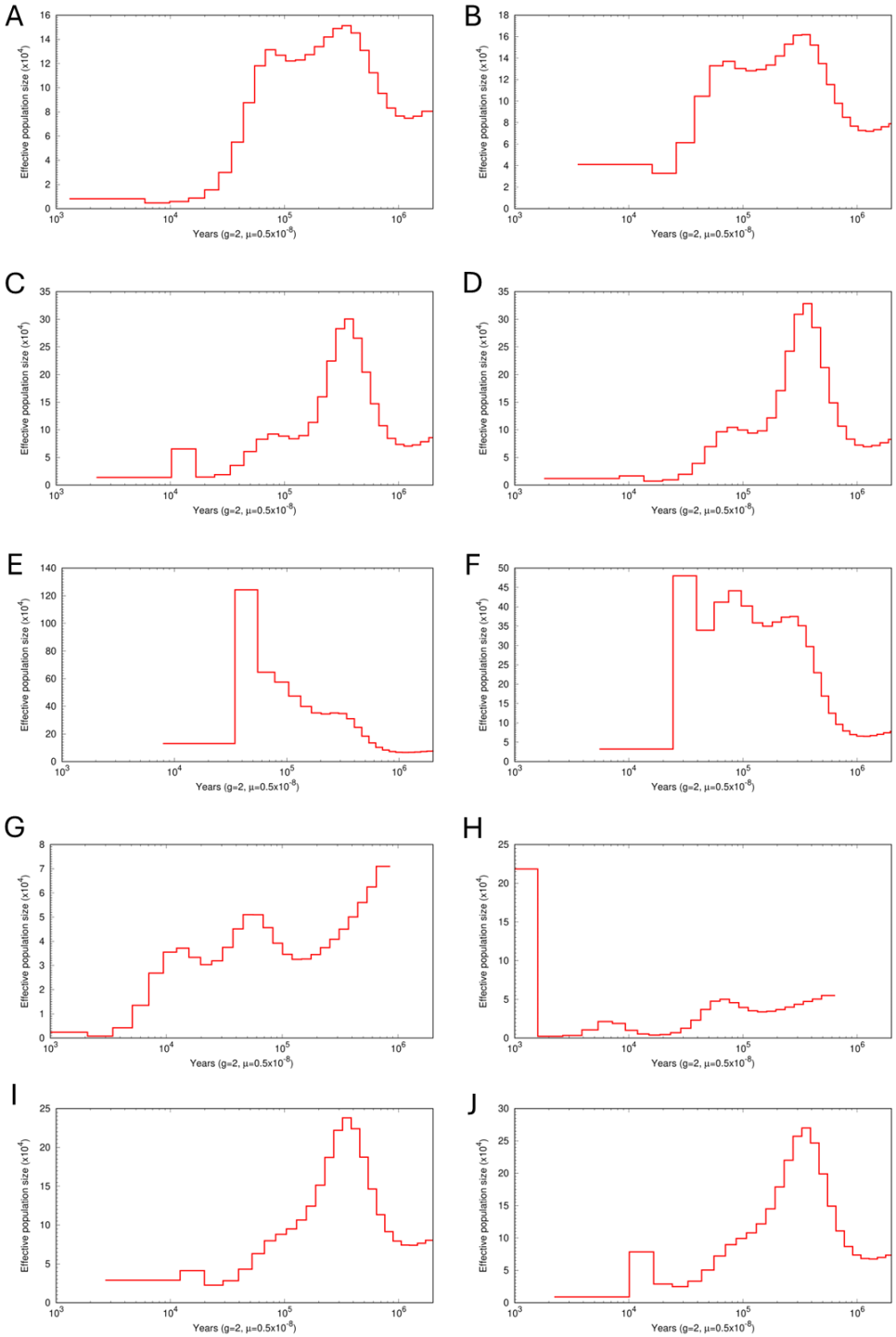
Supplementary Data S11. Population structure of *M. imaizumii* and *M. wogura* inferred by Admixture



Supplementary Data S12. A schematic depiction of f4 statistics. The f4 statistics value is zero when there are gene flows only between W/X and Y/Z.



Supplementary Data S13. The results of PSMC for *M. imaizumii* individuals. A: HS470_ima_1, B: HS5788_ima_1, C: HS4491_ima_2, D: HS582_ima_2, E: HS4620_ima_3, F: HS5758_ima_3.



Supplementary Data S14. The results of PSMC for *M. wogura* individuals. A: HS3096_wog_1, B: HS4898_wog_1, C: HS4585_wog_2, D: HS4884_wog_2, E: HS3066_wog_3, F: HS4899_wog_3, G: HS1392_wog_4, H: HS373_wog_4, I: HS4605_wog_5, J: HS4953_wog_5.

References

Allexander DH, Novembre J, and Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19: 1655–1664.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bruna T, Hoff KJ, Lomsadze A, Stanke M, and Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genomics and Bioinformatics 3.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, and Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4: s13742–015–0047–8.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, and Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso-2*; *iso-3*. Fly 6: 80–92.

Darriba D, Psada D, Kozlov AM, Stamatakis AM, and Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Molecular Biology and Evolution 37: 291–294.

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Mahol I, Lander ES, Aiden AP, and Aiden EL. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science 356: 92–95.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, and Li H. 2021. Twelve years of SAMtools and BCFtools. GigaScience 10: giab008

Escoda L and Castresana J. 2021. The genome of the Pyrenean desman and the effects of bottlenecks and inbreeding on the genomic landscape of an endangered species. Evolutionary Applications 14: 1898–1913.

Ewels P, Magnusson M, Lundin S, and Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32: 3047–3048.

Excoffier L, Dupanloup I. Huerta-Sánchez E, Sousa VC, and Foll M. 2013. Robust demographic inference from genomic and SNP data. PLoS Genetics 9: e1003905.

Excoffier L, Marchi N, Marques DA, Matthey-Doret R, Gouy A, and Sousa VC. 2021. *fastsimcoal2*: demographic inference under complex evolutionary scenarios. Bioinformatics 37: 4882–4885.

Excoffier L, Kapopulou A, and Marchi N. 2023. Demogenomic inference from spatially and temporally heterogeneous samples. Molecular Ecology Resources https://doi.org/10.1111/1755-0998.13877

Flouri T, Izquierdo-Carrasco F, Darriba D, Aberer AJ, Nguyen LT, Minh BQ, von Haeseler A, and Stamatakis A. 2014. The phylogenetic likelihood library. Systematic Biology 64: 356–362.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, and Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proceedings of the National Academy of Sciences 117: 9451–9457.

Ghurye J, Pop M, Koren S, Bickhart D, and Chin CS. 2017. Scaffolding of long read assemblies using long range contact information. BMC genomics 18: 1–11.

Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, and Koren S. 2019. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. PLoS computational biology 15: e31007273.

Hart AJ, Ginzburg S, Xu MS, Fisher CR, Rahmatpour N, Mitton JB, Paul R, and Wegrzyn JL. 2020. EnTAP: bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. Molecular Ecology Resources 20: 591–604.

Huson DH and Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23: 254–267.

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21: 241.

Keilwagen J, Hartung F, Paulini M, Twardziok SO, and Grau J. 2018. Combining RNA-

seq data and homology-based gene prediction for plants, animals and fungi. BMC bioinformatics 19: 1–12.

Kozlov AM, Darriba D, Flouri T, Morel B, and Stamatakis A. 2019. RAxML-NG: a fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35: 4453–4455.

Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv: 1303.3997.

Li H and Durbin R. 2011. Inference of human population history from individual whole-genome sequences. Nature 475: 493–496.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Manni M, Berkeley MR, Seppey M, Simao FA, and Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eucaryotic, prokaryotic, and viral genomes. Molecular Biology and Evolution 38: 4647–4654.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, and DePristo MA. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20: 1297–1303.

Pacifici M, Santini L, Marco MD, Baisero D, Francucci L, Marasini GG, Visconti P, and Rondinini C. 2014. Generation length for mammals. Nature Consevation 5: 89–94.

Patterson N, Price AL, and Reich D. 2006. Population structure and eigenanalysis. PLoS genetics 2: e190.

Pertea G and Pertea M. 2020. GFF utilities: GffRead and GffCompare. F1000Research 9.

Pockrandt C, Alzamel M, Iliopoulos CS, and Reinert K. 2020. GenMap: ultra-fast computation of genome mappability. Bioinformatics 36: 3687–3692.

Price AL, Patterson NJ, Plenge RM, Weinblatt Me, Shadick NA, and Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 38: 904–909.

Quinlan AR and Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841–842.

Van der Auwera GA and O'Connor BD. 2020. Genomics in the cloud: using Docker, GATK, and WDL in Terra (1st Edition). O'Reilly Media.

Weisenfeld NI, Kumar V, Shah P, Church DM, and Jaffe DB. 2017. Direct determination of diploid genome sequences. Genome research 27: 757–767.