



HOKKAIDO UNIVERSITY

Title	Molecular evolutionary study on the Japanese weasel (<i>Mustela itatsi</i>) and the Siberian weasel (<i>M. sibirica</i>), based on complete mitochondrial genome sequences
Author(s)	SHALABI, MOHAMMED AMIN MOHAMMED MOHAMMED
Degree Grantor	北海道大学
Degree Name	博士(理学)
Dissertation Number	甲第12428号
Issue Date	2016-09-26
DOI	https://doi.org/10.14943/doctoral.k12428
Doc URL	https://hdl.handle.net/2115/67137
Type	doctoral thesis
File Information	MOHAMMED_AMIN_MOHAMMED_MOHAMMED_SHALABI.pdf



**Molecular evolutionary study on the Japanese weasel
(*Mustela itatsi*) and the Siberian weasel (*M. sibirica*), based on
complete mitochondrial genome sequences**

(ミトコンドリアゲノム全配列に基づくニホンイタチ(*Mustela itatsi*)および
シベリアイタチ(*M. sibirica*)の分子進化学的研究)

PhD Dissertation

By

MOHAMMED AMIN MOHAMMED MOHAMMED
SHALABI

Department of Natural History Sciences
Graduate School of Science
Hokkaido University

September 2016

Contents

Acknowledgments	1
Abstract	3
General Introduction	6

Chapter I: Comparative phylogeography of the Japanese weasel (*Mustela itatsi*) and the Siberian weasel (*M. sibirica*), revealed by complete mitochondrial genome sequences

Introduction	9
Materials and Methods	14
Results	19
Discussion	22

Chapter II: Comparative sequence variations among different genes of mitochondrial genome for the Japanese weasel (*Mustela itatsi*) and the Siberian weasel (*M. sibirica*)

Introduction	34
Materials and Methods	36
Results	37
Discussion	38

Chapter III: Remarkably high variation of tandemly repeated sequences within the
mitochondrial DNA control region of the Siberian weasel (*Mustela
sibirica*) on Tsushima Island, Japan

Introduction	39
Materials and Methods	40
Results	42
Discussion	45
General Discussion	49
References	53
List of Figures	69
List of Tables	81

Acknowledgments

Acknowledgments

First of all, I would like to thank my supervisor, Professor Ryuichi Masuda, for his keen supervision and guidance throughout the PhD research. He always provided me with all kinds of advice and consultancy needed to achieve my research goals.

I am also grateful to Professor Takeo Horiguchi, Professor Kazuhiro Kogame and Professor Masaaki Takagi of Department of Natural History of Sciences for their invaluable comments on my dissertation.

My deep gratitude goes to Professor Matthew Dick who gave me invaluable advice at the beginnings of my research and later checked English for the manuscripts, and Dr. Toru Katoh for his honest and valuable advice on my research planning.

I am also grateful to my lab members; Dr. Yoshinori Nishita, my lab's alumni and all current members of Laboratory of Genetic Diversity for their continuous advice and support during my research.

I also thank the following members for providing the valuable samples: from Japan, Dr. Y. Kaneko (Tokyo University of Agriculture and Technology), Dr. K. Yamazaki (Tokyo University of Agriculture) and Mr. S. Watanabe (Seian University of Art and design), F. Sekiyama (Iwate Prefectural Museum), Mr. T. Tsujimoto (Morioka Zoo), Mr. S. Dakemoto (Takayama City, Gifu), Mr. M. Hisasue (Towada Cho, Aomori), and Dr. M. Tatara (Ministry of the Environment); from Russia, Dr. A. Abramov (Zoological Institute, Russian Academy of Sciences) and Dr. P. Kosintsev (Institute of Plant and Animal Ecology, Russian Academy of Sciences); from South Korea, Dr. S.-H. Han (National Institute of Biological Resources, Environmental Research Complex); from Taiwan, Professor L.-K. Lin (Department of Life Science, Tunghai University).

Finally, I would like to thank my parents, family and friends for their continuous support and encouragement from the beginnings and throughout my research.

Abstract

Abstract

In the present study, molecular evolution of the endemic Japanese weasel (*Mustela itatsi*) and the continental allopatric species, the Siberian weasel (*M. sibirica*) was investigated intensively using complete mitochondrial genome sequences. This dissertation on the study is divided into three chapters. In the first chapter, complete mitochondrial genome sequences for 26 individuals of *M. itatsi* and 20 individuals of *M. sibirica* were analyzed. The divergence time between *M. itatsi* and *M. sibirica* estimated from the sequence data were 2.36 million years ago (Mya), corresponding with the Early Pleistocene. This divergence time is close to that of most of other Japan-endemic/continental mammalian species pairs previously reported. *Mustela itatsi* comprised two haplotype clades that diverged an estimated 1.64 Mya, in the Middle Pleistocene: a northern (Honshu) clade comprising geographically distinct basal, northern, and eastern subclades, and a western paraphyletic group; and a southern clade comprising geographically distinct sub-clades on Kyushu, Shikoku, and adjacent small islands. The results indicate a single migration of an ancestral population from the Korean Peninsula to southern Japan across a Late Pliocene or Early Pleistocene land bridge, followed by allopatric speciation of *M. itatsi* in Japan. The southern lineage appears to have remained in place, whereas the range of the northern lineage expanded stepwise from southwestern to northern Honshu between 0.68 and 0.27 Mya. By contrast, *M. sibirica* also comprised two main clades that diverged an estimated 1.57 Mya: one containing haplotypes from continental Russia and Tsushima Island (Japan), and the other containing haplotypes from Korea, China and Taiwan. The *M. sibirica* population on Tsushima Island is likely a relict from the continental Russian population. The estimated divergence times indicated

that both *Mustela* species were the early colonists of the Japanese islands and continental Eurasia, respectively. The present study based on the complete mitochondrial genome sequences provides a higher resolution of the phylogeographic relationships between and within these closely related insular and continental species of *Mustela*, compared with previous studies using single or a few mitochondrial gene loci.

In the second chapter, the obtained complete mitochondrial (mtDNA) genome sequences from the previous chapter were used to compare sequence variations among different genes of mtDNA in *M. itatsi* and *M. sibirica*. The ratio between parsimoniously informative sites (Pi) and the length in base-pairs (bp) (Pi/length) were calculated for each gene locus. The results showed that the control region (D-loop) has the highest Pi/length ratio among gene loci. The protein-coding gene commonly used in phylogenetic studies, cytochrome *b*, did not show the second highest Pi/length in both species following the control region, but other genes like the protein coding genes CO3 for *M. itatsi* and ND4 for *M. sibirica* and ND2 for both species showed the higher values. The result indicated that a combination of control region/ ND2/ ND4/ cytochrome *b* genes could lead to better resolution of the phylogenetic relationships between mustelid species.

In the third chapter, the variation of the mtDNA control region including the C/T indel sites and the specific tandemly repeated sequences were investigated in the *M. sibirica* population native to Tsushima Island, located between the Japanese islands and the Korean Peninsula. From 31 animals examined, variants of 17 C/T insertion/deletion (indel) sites and seven different patterns of tandem repeats were detected. The tandem repeats consisting of less than 10-bp units were almost identified, although nucleotide sequences of the other parts in the mtDNA control

region were identical among all of them. The repetitious patterns of tandem repeats all shared the same starting and ending repeat units, but were different in the number of the core repeat compound units of 10-bp. Compared with repeat units of non-insular carnivores published previously, the repetitious sequences found in the present study were remarkably highly polymorphic. In addition, one nucleotide deletion at the 3' end of the last repeat unit occurred in all animals, whereas the 3' end of a previously reported unit in carnivorans was not deleted. The number variation of the compound units in the core region together with the occurrence of the particular last unit with a nucleotide deletion could have been formed by the continuous step-wise slippage. Even among the individuals sampled from the same geographic location during one year within the island, the repeat tandem repeats unit numbers were highly variable, suggesting the remarkably rapid evolution of the repeat units in the control region of *M. sibirica*. Combining the detected seven patterns of tandem repeats together with the 17 variants of C/T indel sites yielded 27 different variants among the studied 31 individuals, showing remarkably high mtDNA diversity in such a small insular population of *M. sibirica*.

General Introduction

General Introduction

Before the advent of molecular phylogenetic techniques, comparative studies between different forms of animals had been principally performed using external body measurements and skull morphology and/or anatomical differences.. Some or a combination of the previously mentioned parameters have been used to find the phylogenetic relationships between different or closely related animal taxa. Then, the researchers tried to find a sort of taxonomic or biogeographic basis for the target species. Since the advancement of phylogeography since the 1980th (Avice & Ellis, 1986), a lot of taxonomic ambiguities among different animal taxa have been clarified using the unique characteristics of maternally inherited mitochondrial DNA (mtDNA) and paternal and biparental gene loci, i.e. the case of *Canis* species complex in the study of Koepfli *et al.* (2015). At the beginnings of molecular phylogenetic study and due to some technical limitations, individual genes with high polymorphism have been studied, i.e. in the Japanese population on Hokkaido of the brown bear using the control region (D-loop) of mtDNA (Matushashi *et al.*, 1999). Then, according to the advancement of proper techniques, molecular phylogenetic studies have been developed using by the complete mtDNA sequences, for example, on the Hokkaido brown bear (Hirata *et al.*, 2013, 2014). At present, genome drafts of most of model species have been determined and deposited in DNA databases. For basic medical purposes and clinical problems, genome-wide studies of model species receive the greatest attention. Non-model species still receive less attention but recently the importance of studying the non-model species at the genome level (population genomics) have been addressed (Wayne, 2016). The purpose of those studies is to understand genetic and environmental factors, which have contributed to species evolution and phylogeographical history (Wayne, 2016).

Island biogeography has been a topic for growing interest since the beginning of accepting Darwinism. The main initiative for Darwin's hypothesis was his notes about the uniqueness of animals on the Galapagos Islands, compared with the closely related land ones (Whittaker & Fernandez-Palacios, 2007). In addition, the phylogeographic studies for continental-island mammals may provide deep insights to further understanding the evolution and speciation of island fauna, compared with their mainland counterparts (Masuda *et al.*, 2012). The Japanese islands are a good example for such kind of studies. It has both kinds of islands: (i) formation of land-bridges several times between the continent and the continental islands of Japan (Hokkaido, Honshu, Kyushu and Shikoku) and (ii) oceanic islands, which had no connections with the neighboring main-land since the first separation (Ryukyu archipelago) (Millien-Parra & Jaeger, 1999). Therefore, phylogeographic studies of mammals on the Japanese islands give deep insight to understanding the history on formation of the Japanese islands and different factors contributing to their current distribution (McKay, 2012).

The Japanese mammals on the land-bridge islands show a high degree of endemism, owing to the multiple connections and separations with the continental main land (Dobson, 1994). Among order Carnivora, family Mustelidae, which consists of weasels, otters, martens, badgers and the relatives (Koepfli *et al.*, 2008), include eight species distributed in Japan (Masuda *et al.*, 2012). The eight members of this family present an interesting model for studying the biogeographic history on the Japanese islands: (i) the sable (*Martes zibellina*) is present on Hokkaido and common to the Palearctic region; (ii) the Japanese marten (*Martes melampus*), Japanese badger (*Meles anakuma*) and Japanese weasel (*Mustela itatsi*) are endemic to the Honshu-Shikoku-Kyushu (called the Hondo); (iii) the Siberian weasel

(*Mustela sibirica*) is distributed only on Tsushima Island of Japan, and this species is widespread on the continent; (iv) the least weasel (*Mustela nivalis*) and ermine (*Mustela erminea*) occur on Hokkaido, and they also have local distribution on northern Honshu and wide distribution in the Holarctic region.

In the present study, I studied the phylogenetic relationships between one of the endemic mustelids (*Mustela itatsi* in above category ii) and its continental allopatric species (*Mustela sibirica* in above category iii). In addition, to further clarify their evolutionary history, I determined complete mitochondrial genome sequences from the samples obtained widely from their distributions to characterize the molecular features and investigate the phylogenetic relationships. In the first chapter, I elucidated the phylogenetic relationships between and within the two *Mustela* species using the highly informative complete mitochondrial genome sequences. The divergence times were calculated more precisely than the previous studies. Contents of the first chapter were accepted for publication (Shalabi *et al.*, 2016). In the second chapter, I compared the genetic variation in the complete mtDNA genome sequences between the two *Mustela* species, based on parsimoniously informative sites for each gene locus. The purpose is to figure out which genes can provide the most accurate phylogenetic relationships. In the third chapter, I examined the mitochondrial DNA (mtDNA) diversity in the insular population of *Mustela sibirica* on Tsushima Island to examine evolutionary characteristics, especially of the repetitive sequences with rapid mutation rate. The present study provides an insight to further understanding phylogeographic history and speciation of the Japanese mammals.

Chapter I

**Comparative phylogeography of the
Japanese weasel (*Mustela itatsi*) and the
Siberian weasel (*M. sibirica*), revealed by
complete mitochondrial genome sequences**

Introduction

The Japanese Archipelago provides a unique natural experiment for studies of endemism and the processes of speciation and genetic divergence. Japan contains 117 mammal species, 49 (42%) of which are endemic (Motowaka, 2009). The mammalian fauna in central Japan (the 'Hondo region', including Honshu, Shikoku, and Kyushu Islands, and adjacent smaller islands) is characterized by high species richness relative to land size, high endemism and a high degree of geographic variation (Millien-Parra & Jaeger, 1999).

The Japanese islands have a complex geological history. Some of them were connected to one another or to the adjacent continental mainland (land-bridge islands), whereas others have remained isolated throughout the Quaternary (oceanic islands) (Millien-Parra & Jaeger, 1999). Most are of the former type, and only some islands in the Ryukyu Archipelago between Kyushu and Taiwan might be of the latter type. The terrestrial mammals on the land-bridge islands are segregated into two large groups demarcated by Tsugaru Strait between Honshu and Hokkaido: (i) the northern island of Hokkaido, inhabited by high-northern Eurasian species, with limited endemism, and (ii) southern Hondo region, with a small number of Indo-Malayan elements and a high degree of endemism (Dobson & Kawamura, 1998). Based on fossil records for family Mustelidae in Japan (Kawamura *et al.*, 1989) and the native distribution (Masuda & Watanabe, 2015), *M. itatsi* (which was not distributed naturally on Hokkaido) belongs to the second group.

Because terrestrial mammals have limited capabilities for dispersal to islands (Heaney, 1984; Lawlor, 1986; Dobson, 1994), their contemporary distributions on continental islands are the result of both historical connections with the continent and climatic conditions during

periods of connection (Yalden, 1982; Dobson & Kawamura, 1998). Hokkaido remained connected to the continental for relatively long intervals during glacial periods, whereas the islands comprising Hondo were connected to the continental mainland only during some glacial maxima, when low sea-level stands exceeded the sill depth (about 130 m) of Tsushima Strait separating southern Honshu and Kyushu from the Korean Peninsula. Hondo was thus more isolated than Hokkaido (McKay, 2012). There is evidence that the Hondo islands were connected with the Korean Peninsula four times since the Middle Miocene—first in the Late Pliocene (2–3 Mya), possibly once in the Early Pleistocene (1 Mya), and twice in the Middle Pleistocene (0.5 and 0.3 Mya) (Dobson, 1994)—but not during the last glacial maximum (LGM) ca. 0.02 Mya (Park *et al.*, 2000). The Hondo islands themselves are separated from one another by small and narrow straits not exceeding 50 m in depth, so they were often connected as a single landmass in the past (McKay, 2012). Osumi Strait between Kyushu and the islands of Yukushima and Tanegashima (Oshima, 1990, 1991, 2000) could have formed 100,000–150,000 years ago. Finally, around 5000–7000 years ago, Honshu, Shikoku, and Kyushu islands were separated by the Seto Inland Sea, and Kanmon Strait between Honshu and Kyushu (Oshima, 1990, 1991, 2000).

Although the archipelago structure of Japan is conducive to allopatric speciation among animal populations, historically there has been incomplete isolation from continental populations (McKay, 2012), and phylogenetic studies have shown varied divergence times between endemic or non-endemic Japanese mammals and their continental sister species.

In the study of McKay (2012), divergence times between Japanese mammals and their allopatric continental species obtained from previous studies were re-calculated. Although the

obvious improper calibration points which were used in that study, the estimated divergences between most Japanese/continental pairs studied were in the Late Pliocene or Early Pleistocene, 1.2–2.88 million years ago (Mya) . Early Pleistocene divergences between continental Asia and the Hondo region do not coincide with the Middle Pleistocene timing of the last land bridge connecting these regions (0.3–0.5 Mya) (Dobson & Kawamura, 1998; Yasukochi *et al.*, 2009), but instead reflect the migrations of mammals from the continent to Japan during earlier episodes of land connection in the Pliocene (2.0–3.0 Mya) or early Pleistocene (1.0 Mya) (Dobson & Kawamura, 1998; McKay, 2012). There are exceptions; the estimated divergence between the Dsinezumi shrew (*Crocidura dsinezumi*) and the Ussuri white-toothed shrew (*C. lasiura*) 0.5 Mya is consistent with migration from the continent to Japan across the last land bridge (Ohdachi *et al.*, 2004; McKay, 2012). Another exception showing a much older divergence than the time of formation of the first Pliocene land bridge is the divergence 8.69 Mya between the lesser flying squirrel (*Pteromys momonga*) and the Russian flying squirrel (*P. volans*) (Oshida *et al.*, 2000; McKay, 2012).

Phylogeographic studies based on various mitochondrial DNA (mtDNA) markers have detected three general distribution patterns for mammals in the Hondo region: (i) Two geographically well-differentiated clades without a geographical contact zone between them and an Early Pleistocene divergence, i.e., the Japanese mole (*Mogera wogura*) (Tsuchiya *et al.*, 2000). (ii) Two or three distinct clades with a contact zone(s) between them and a Middle Pleistocene divergence, i.e., the Japanese hare (*Lepus brachyurus*; cytochrome *b*) (Nunome *et al.*, 2010); the sika deer (*Cervus nippon*; control region) (Nagata *et al.*, 1999), and the Japanese macaque (*Macaca fuscata*; control region) (Kawamoto *et al.*, 2007). (iii)

Populations with haplotypes dispersed between several clades, indicating recent colonization within the Japanese islands, i.e., the Japanese marten (*Martes melampus*; control region, cytochrome *b* and ND2) (Sato, Yasuda & Hosoda, 2009).

In this study, we investigated the phylogeography of a continental/endemic species pair in *Mustela*. The endemic Japanese weasel *Mustela itatsi* is naturally distributed on three of the Japanese main islands (Honshu, Shikoku, Kyushu) and adjacent southern islands (Yakushima, Tanegashima and Ohshima); introduced populations exist in the Ryukyu Archipelago and on Hokkaido Island (Masuda & Watanabe, 2015).

The Siberian weasel (*M. sibirica*) is widespread, occurring in Russia, Mongolia, Pakistan, Kashmir, Himalayan India, Nepal, Bhutan, Myanmar, Thailand, Laos, Vietnam, China, and Taiwan (Sasaki 2015). In the Japanese Archipelago, *M. sibirica* occurs naturally only on Tsushima Island located between Kyushu Island and the Korean Peninsula (Imaizumi 1970; Sasaki *et al.*, 2014), although the species has been introduced from Tsushima or the Korean Peninsula to the Hondo region within the last century (Masuda *et al.*, 2012).

Previous phylogenetic studies of these *Mustela* species based on either mtDNA or nuclear markers indicated an older divergence than the most recent isolation of the Japanese islands from the mainland. MtDNA markers included cytochrome *b* (Masuda & Yoshida, 1994a; Kurose, Abramov & Masuda, 2000), 12S rRNA (Kurose, Abramov & Masuda, 2008), and the D-loop region (Masuda *et al.*, 2012); nuclear markers included *IRBP* (Sato *et al.*, 2003) and nuclear gene loci totaling 8 kilo base-pairs (kbp) (Sato *et al.*, 2012). These studies estimated a wide range of divergence times between the two species (1.6–2.4 Mya) during the early Pleistocene. Based on 600 bp of the mtDNA control region, Masuda *et al.* (2012) found

that *M. itatsi* comprised two geographically separate clades: a Honshu clade (northern Japan) and a Shikoku-Kyushu clade (southern Japan). Some ambiguities remain for *M. sibirica*: cytochrome *b* haplotypes in the Tsushima population were phylogenetically similar to those in the Korean population (Hosoda *et al.*, 2000), whereas control region haplotypes differed from continental haplotypes (Masuda *et al.*, 2012). The divergence time of the Tsushima population from the continental population has not yet been estimated (Sato, 2013).

There is a need to infer the phylogenetic relationships and divergence times for our studied pair, mainly for two reasons. First, the uncertainty coming from the broad range of estimates of the divergence time between *M. itatsi* and *M. sibirica*, and even differences in tree topology, may have resulted from variation in substitution rates among genes. Indeed, there is evidence from previous studies of other mustelid species that these effects occur. An analysis of three concatenated mtDNA loci (control region, cytochrome *b*, and ND2) in *Martes melampus* (Sato, Yasuda & Hosoda, 2009) yielded different results from analyses of single mtDNA genes (Hosoda *et al.*, 1999, 2000; Kurose *et al.*, 1999). Analyses using cytochrome *b* (Kurose *et al.*, 2001) produced different results from those using the control region (Tashima *et al.*, 2011) for the Japanese badger *Meles anakuma*. Likewise, tree topologies were different between the control region (Kurose *et al.*, 2005) and the complete mtDNA genome (Malyarchuk, Denisova & Derenko, 2015) in studies of the ermine (*Mustela ermine*). The second reason is the growing concern for the proper choice of calibration point while estimating divergence times as explained by Ho (2007), (Ho & Duchene, 2014) and Ho *et al.*, 2008). The available coalescence-based divergence times for both *M. itatsi* and *M. sibirica* (2.88 Mya; McKay 2012 and 2.4 Mya; Sato *et al.*, 2012) may be incorrect due to

inappropriate calibration points. The former one did not consider any calibration points concerns, while the later put them into consideration, using fossil calibrations. Because the later study were mainly focused on the biogeographic history of the superfamily Musteloidea, their selected calibration points are very “external” to be accurate for within species divergence between *M. itatsi* and *M. sibirica*. There is a handful of information, which could be used for more accurate calibrations; mathematically-calculated estimated divergence times using substitution rates of the same studied populations for both species (Masuda & Yoshida, 1994; Masuda et al., 2012) and geological events (Kawamura, 1994; Oshima 1990, 1991, 2000).

To circumvent the effects of limited data, single-locus bias and inaccurate divergence times estimations, we used the whole mtDNA genome to examine the phylogenetic relationships between and within *M. itatsi* and *M. sibirica*, with the goal of more accurately estimating divergence times with a comparison between two different calibration ways and correlating evolutionary history with geological events. We used a model-based approach to analyze data set that included 12 protein coding genes, 22 tRNA loci, two rRNA loci, and the control region.

Material and Methods

Samples and DNA extraction

Muscle samples from 26 individuals of *M. itatsi* covering the native distribution were collected from Honshu, Shikoku, Kyushu and adjacent small islands (Yakushima and

Tanegashima). For *M. sibirica*, samples were collected from a total of 20 individuals from the native distribution on the continent and Tsushima Island, including four from Tsushima, one from Korea, four from Taiwan and 11 from continental Russia. Table I-1 and Figure I-1 show the sampling localities, with the sample sizes for both species.

Molecular methods

Total genomic DNA was extracted by using the DNeasy Tissue & Blood Kit (QIAGEN), following the manufacturer's protocol. To obtain complete mtDNA sequences, a set of 16 primer pairs was used to amplify mtDNA fragments 1.0–1.7 kbp long. Most of the primers were newly designed in this study (Table I-2), although one primer pair was from Knaus *et al.* (2011). The polymerase chain reaction (PCR) amplifications were conducted in 20 µl volumes containing 4.0 µl of 5X PrimeSTAR GXL DNA Buffer (Takara), 1.6 µl of dNTP mixture (2.5 mM each dNTP; Takara), 0.4 µl of PrimeSTAR GXL DNA Polymerase (1.25 U/ml, Takara), 0.2 µl each of forward and reverse primers (25 pmol/µl), 1.0–4.0 µl of DNA extract, and 9.6–12.6 µl of distilled water. Thermal cycling conditions were 30–40 cycles of 10 s at 98°C, 15 s at 50–60°C, and 2 min at 68°C. PCR products were purified with the QIAquick Purification Kit (Qiagen), following the manufacturer's protocol.

Cycle sequencing was performed by using the BigDye v3.1 Cycle Sequencing Kit (Applied Biosystems, ABI) with the sequencing primers listed in Table I-3. The PCR for sequencing was performed in 10 µl volumes containing 1.75 µl of 5X BigDye Sequencing Buffer (ABI), 1.0 µl of Ready Reaction Premix (ABI), 1.0 µl of DNA template, and 4.65 µl of distilled water. Thirty cycles of 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C were performed.

Amplified DNA fragments were purified with isopropanol, and then formamide was added. Sequences were determined on an ABI 3730 DNA Analyzer. The genomic positions of two rRNAs, 22 tRNAs, 13 coding genes, and the control region were determined by referring the complete mtDNA sequence of *M. sibirica* (Accession no. NC_020637.1). Nucleotide sequences generated in this study were deposited in the DDBJ/NCBI/EMBL databases under accession nos. AP017387–AP0173897 and AP017400–AP017421.

Phylogenetic analysis

Complete mtDNA sequences were obtained for 26 individuals of *M. itatsi* and 20 individuals of *M. sibirica*, and used for phylogenetic analysis. A previously reported complete mtDNA sequence for *M. sibirica* from China (accession no. NC_020637.1) was included in the sequence alignment. Complete mtDNA sequences for the least weasel (*Mustela nivalis*; NC_020639.1), mountain weasel (*Mustela altaica*; NC_021751.1) and ermine (*Mustela ermine*; NC_025516.1) were included to represent outgroup taxa. Sequences were aligned by using Clustal W (Thompson, Higgins & Gibson, 1994) in MEGA 6 (Tamura *et al.*, 2011). Insertions and deletions (indels), variable number tandem repeats (VNTRs) consisting of 10-bp units in the control region, and ambiguous sites were excluded from the analysis. The ND6 (NADH dehydrogenase subunit 6) sequences were also excluded because this gene is transcribed in the opposite direction from other mtDNA loci. The corrected alignment used for phylogenetic inference was 15,813 bp long and included two rRNA genes, 22 tRNA genes, 12 coding genes, and the control region (excluding VNTRs).

Phylogenetic trees were reconstructed by maximum likelihood (ML) and Bayesian

Inference (BI). In both the ML and BI analyses, a partition model was used, and different substitution models were applied to different gene partitions. The best-fit substitution model was determined for each partition by using the Akaike information criterion (AIC) for ML and the Bayesian information criterion (BIC) for BI implemented in jModelTest 4 (Tanabe, 2011). The ML tree was reconstructed by using Treefinder version March 2011 (Jobb *et al.*, 2004) and RaxMLv8.2.x (Stamatakis, 2014). Nodal support was assessed by bootstrap analyses of 1000 pseudoreplicates. BI was performed with MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003) in three simultaneous runs of 80,000,000 Markov chain Monte Carlo (MCMC) generations, with trees for estimation of the posterior probability distribution sampled every 1000 generations; the first 8,000,000 trees were discarded as burn-in.

Genetic diversity indices and sequence variation

Average haplotype diversity, average nucleotide diversity (π ; Nei, 1987), and polymorphic sites within each species and population were calculated for the complete mtDNA sequences by using DnaSP ver. 5 (Librado & Rozas, 2009). To test different scenarios of demographic expansion of *M. itatsi* and *M. sibirica*, Tajima's D (Tajima, 1989) and Fu's F_s (Fu, 1997) were computed in DnaSP.

Estimation of divergence times

Divergence times between *M. itatsi* and *M. sibirica* and between subclades (or lineages) within each species were estimated with BEAST v1.6.2 (Drummond & Rambout, 2007), using the uncorrelated lognormal model to describe a relaxed clock. BEAST xml input files

were generated by using BEAUti v1.6.2 (Drummond & Rambout, 2007), with each gene partition allowed to have its own independent base-substitution model and parameters. Each gene substitution model was selected by using the BIC implemented in Kakusan 4. The Yule process of speciation was applied to the tree prior. Posterior probability distributions of parameters, including the trees, were obtained by MCMC sampling to estimate time to the most recent common ancestor (TMRCA). Trees were sampled every 8000 generations from a total of 80,000,000 generations, with the first 40% discarded as burn-in. Five independent runs were conducted, and the most acceptable mixing and convergence to the stationary distribution were checked from the resultant output-log files by using Tracer v1.4 (Rambaut & Drummond, 2007). The maximum clade credibility tree was produced with TreeAnnotator v1.6.2 (Drummond & Rambout, 2007).

We applied two different methods for calibrating our tree; (i) in order to compare our results with the already published ones, which are already slightly inappropriate as it did not follow the recommendations of Ho (2007) and Ho *et al.* (2005, 2008) of choosing the calibration point for within species divergence. In this method we adopted a single calibration point from the previous family study, and (ii) we followed the above mentioned recommendations of applying multiple calibration points based on geological events and previously reported substitution rates. In above method (i), a single calibration point were adopted from Sato *et al.* (2012) for *Mustela itatsi*, *M. sibirica*, *M. nivalis*, and *M. altaica*. A normal distribution with a mean of 3.15 Mya and standard deviation of 0.322 was adopted to achieve the 95% posterior interval of 2.62–3.68 Mya reported in Sato *et al.* (2012); these values are consistent with a similar previous mustelid study by Koepfli *et al.* (2008).

For above method (ii), fossil records could not be used because they were not be identified to the species level until the present time, because non-appropriate use of fossil record information may bias divergence time calculations (Ho *et al.*, 2008; Herman & Searle, 2011; Mcdevitt *et al.*, 2012). Instead, we used three calibration points: root height of *Mustela itatsi*/ *M. sibirica* complex, TMCRA for clades of *M. itatsi* and TMRCA of Russia-Tsushima clade for *M. sibirica*. The root height of *M. itatsi*/ *M. sibirica* complex was given a normal distribution with lower and upper limits of 0.3 and 3.0 Mya, representing the most plausible last and the first land connection between Japanese islands and the continental land. This approach was done by Herman & Searle (2011) and Herman *et al.* (2014) on field voles colonizing different regions of Europe. The TMRCA of *M. itatsi* was given a normal distribution with lower and upper limits of 0.83-1.17 Mya, which is the estimated divergence time calculated from the previous study for the same species of Masuda *et al.* (2012) and Masuda & Yoshida (1994b), using their self-obtained substitution rates (Masuda *et al.*, 2012). We did not use straits formation information here because we thought they did not play a significant role in the species historical distribution, as explained below in the discussion. Finally, the TMRCA of Russia-Tsushima was given a normal distribution with lower and upper limits of 100,000 and 150,000 years ago.

Results

Phylogenetic analyses of *M. itatsi* and *M. sibirica*

Both the ML and BI trees yielded the same tree topology (Fig. I-2), with a clade of 19 *M. itatsi* (IT) haplotypes and a sister clade of 15 *M. sibirica* (SB) haplotypes, with high nodal

support (100/1, bootstrap value/posterior probability). *Mustela itatsi* was divided into two major clades with high nodal support (100/1): clade ITa (Honshu) and clade ITb (Shikoku, Kyushu, and adjacent small islands). Within each of these two clades, groups of haplotypes from geographically separate areas were evident; clade ITa contained subclades ITaN (northern Honshu), ITaE (eastern Honshu), and paraphyletic group ITaW (western Honshu). ITaW was divided into basal Honshu haplotypes (localities h and g) and haplotypes from Gifu, Ishikawa, and Nara in central Japan (localities e, d and f, respectively in Fig. Fig. I-1) (Fig. I-2, Table I-1). Clade ITb was divided into three subclades containing haplotypes from Kyushu (ITbK), Shikoku (ITbS), and the adjacent small islands of Yakushima and Tanegashima (ITbI).

Mustela sibirica was also divided into two major clades: clade SBa consisting of haplotypes from continental Russia and Tsushima Island, and clade SBb including haplotypes from Taiwan, China, and Korea. Clade SBa was further divided into subclades from continental Russia (SBaR) and Tsushima Island (SBaT). Despite the wide geographical range of sampling localities, all five samples from continental Russia comprised one subclade, with high nodal support (97/1). Clade SBb comprised two geographically distinct subclades, SBbT (Taiwan) and SBbC (continental China and Korea) (Fig. I-2).

Genetic diversity in *M. itatsi*

Table I-4 shows genetic diversity indices for both species and for clades within the species.

For 15,813 bp of complete mtDNA sequence, 245 sites were polymorphic among 19

haplotypes. Nucleotide diversity and haplotype diversity in *M. itatsi* were 0.00619 and 0.969,

respectively. Clade ITb was genetically more variable than clade ITa, which had more samples examined.

To test for recent expansion of *M. itatsi* populations, Tajima's *D* and Fu's *F_s* (Table 2) were statistically significant for *M. itatsi* overall but not for the Honshu (ITa) or the Kyushu-Shikoku (ITb) population. Tajima's *D* was lower and *F_s* higher for the Kyushu-Shikoku population than for the Honshu population.

Genetic diversity in *M. sibirica*

Among 20 individuals of *M. sibirica*, 213 sites were polymorphic and 14 haplotypes were identified (Table I-4). Haplotype diversity and nucleotide diversity were 0.937 and 0.00519, respectively. The haplotype diversity and nucleotide diversity were both higher for clade SBb than for SBa, despite the larger sample size of the latter. Tajima's *D* and Fu's *F_s* values were statistically significant only for *M. sibirica* overall but not for either SBa or SBb.

Estimation of divergence times

Divergence times were estimated by the BI method using (i) prior data imported from the study of Mustelidae by Sato *et al.* (2012) and (ii) multiple calibration points based on geological information and reported substitution rates from previous studies for the same species. This allowed the time to the most recent common ancestor (TMRCA) and the confidence interval (95% HPD) to be estimated between *M. itatsi* and *M. sibirica* and between clades within each species, as shown in Figure I-3 for method (i), Figure I-4 for method (ii) and Table I-5. Results of method (i) showed that the divergence time between *M. itatsi* and *M.*

sibirica was estimated to be 2.36 Mya. The TMRCA estimates for clades within *M. itatsi* and *M. sibirica* were 1.64 Mya and 1.57 Mya, respectively. The TMRCA estimates for the two main clades in *M. itatsi* and those in *M. sibirica* ranged from 0.88 Mya (Clade SBa) to 1.03 Mya (Clade ITb). Divergences within the main intraspecific clades leading to geographically localized subclades were younger; those separating the Northern Honshu (ITaN), Eastern Honshu (ITaE), and Western Honshu (ITaW-f,d,e) clades occurred 0.68–0.27 Mya. For method (ii), the divergence time between *M. itatsi* and *M. sibirica* was 1.19 Mya. The TMRCA for clades of both species were 0.93 Mya for *M. itatsi* and 0.67 Mya for *M. sibirica*. The TMRCA clades for lineages of both species follow the same age order of method (i) but with overall younger values.

The obtained substitution rates from the two calibration methods were 0.011 substitutions/site/million years (95% HPD: 0.007–0.015) using method (i) and 0.025 substitutions/site/million years (95% HPD: 0.017–0.033) using method (ii).

Discussion

Phylogenetic relationship and divergence time between *M. itatsi* and *M. sibirica*

Our results showed *M. itatsi* and *M. sibirica* each to comprise a monophyletic lineage. The single calibration point method (i) has showed an estimated divergence time of 2.36 Mya between them (95% HPD: 1.54–3.14) (Fig. I-3, Table I-5), or Early Pleistocene, which is consistent with previous studies, although our results are likely more reliable, as they are based on more data and the MCMC BEAST calculations. A previous estimate of the divergence time between the species was 2.88 Mya, based on partial cytochrome *b* sequences

from Hosoda *et al.* (2000), and Marmi, López-Giráldez & Domingo-Roura (2004) analyzed with MCMC-based software (McKay, 2012). Previous studies using equation-based calculations also reported an early Pleistocene divergence between *M. itatsi* and *M. sibirica*: Masuda *et al.* (2012) using partial D-loop sequences, Masuda & Yoshida (1994) using partial cytochrome *b* sequences, and Sato *et al.* (2003) using mitochondrial cytochrome *b* and nuclear *IRBP*. All of these previous studies and our estimate strongly indicate an early Pleistocene divergence between *M. itatsi* and *M. sibirica*. McKay (2012) re-calculated the divergence times between Japanese endemic mammals and their continental sister species, using BEAST software. Seven of eleven species pairs showed an early Pleistocene divergence, similar to that between the two species of *Mustela* in our study and that for other mustelid sister-species pairs in previous studies (Sato, Yasuda & Hosoda, 2009; Sato, 2013) (summarized in Table I-6).

Our estimated divergence between *M. itatsi* and *M. sibirica* of 2.36 Mya with a 95% HPD confidence interval of 1.54–3.14 Mya is consistent with the Late Pliocene to Early Pleistocene land connection between the Korean Peninsula and Hondo 2–3 Mya, rather than with subsequent Middle Pleistocene connections. Assuming that those estimated divergence times are accurate enough, *Mustela itatsi* was likely an “old Hondo endemic” as defined by Dobson (1994). In this category, species that originated from Pliocene or Early Pleistocene immigrants had differentiated into modern species by the time the first Middle Pleistocene land bridge formed, and have remained endemic to Hondo. All the above mentioned divergence times of our study and previous study did not consider the growing concern of the inappropriate use of a single external calibration point. An attention should be paid to the

choice of multiple various internal calibration points.

In method (ii) of the multiple calibration points, our estimated divergence time of 1.19 Mya (95% HPD: 0.87–1.6) (Fig. I-4, Table I-5)) which falls into the late early Pleistocene to Middle Pleistocene. This obtained divergence time matches with some other few recent studies for other species, which considered the calibration concern. The estimated divergence time between the Eurasian and Japanese populations of the Asiatic black bear (*Ursus thibetanus*) (1.46 Mya) using complete mtDNA sequences. One complete mtDNA genome sequence of the extinct Japanese otter (*Lutra nippon*) had a divergence time of 1.27 Mya (95% HPD: 0.98–1.59 Mya) with the continental common otter (*Lutra lutra*). This strongly suggests that there was a group of Japanese mammals, which migrated to the Japanese islands at the last early Pleistocene land connection of 1.0 Mya or at the Middle Pleistocene one, which happened 0.5 Mya. This also may cast some doubt about the “extremely early” divergence times, which had non-appropriate calibrations (Supplementary Table 3). Intensive studies for those species are necessary to validate and define migration time for “Hondo colonists”.

Masuda et al. (2012) proposed two hypotheses for the origin and present-day distribution of *M. itatsi*. One is that *M. itatsi* diverged from *M. sibirica* on the Asian continent and migrated to Japan during one or more episodes of land-bridge formation between the continent and Hondo, with the population *M. itatsi* on the continent subsequently going extinct. A corollary to this hypothesis is that the northern and southern clades on Hondo represent descendants of two independent migrations across land bridges at different times. The other hypothesis is that individuals from the continental population ancestral to *M.*

sibirica and *M. itatsi* migrated to Hondo via a land bridge, and when the land connection to continental Asia was lost, the population on Japan evolved allopatrically to become *M. itatsi*.

Parsimony favors the second hypothesis. The first hypothesis requires three or four steps: divergence of *M. itatsi* and *M. sibirica* on the continent, one or two migrations of *M. itatsi* to Japan, and extinction of *M. itatsi* on the continent. In addition, it must explain why only *M. itatsi* and not *M. sibirica* migrated to Japan; if both migrated, this requires another step for the extinction of *M. sibirica* in Japan. The second hypothesis requires only two steps: migration of an ancestral population to Japan, and allopatric divergence of that population to become *M. itatsi* after the connection between Hondo and continental Asia was lost.

Determining where the species diverged remains difficult. For the *M. itatsi*/*M. sibirica* complex, there are no fossils from the continent to investigate whether the two species diverged on the continental mainland, with continental populations of *M. itatsi* subsequently going extinct. Although there are some mustelid fossils from Japan, it is not clear whether they represent *M. itatsi* or *M. sibirica* (Kawamura, Kamei & Taruno, 1989). Surprisingly, our estimated divergence time using method (ii) is consistent with the mustelid fossils from Japan, which were excavated from middle-Pleistocene strata on Honshu (Kawamura, Kamei & Taruno, 1989) and Kyushu (Ogino *et al.*, 2009). It is previously difficult to interpret the fossil records. This is another evidence showing the power of using multiple calibration points. Except for *Ursus thibetanus*, all species pairs (including the pair we studied), for which early Pleistocene divergence times have been estimated (Supplementary table 3), are of small body size, and their morphological identification of fossils to the species level is difficult. Neither ancient DNA studies nor carbon dating have been carried out on them, but these techniques

may improve the usefulness of the fossil materials for better inference.

Phylogenetic relationships and divergence times within *M. itatsi*

We identified two main clades for *M. itatsi*, the northern Honshu clade (ITa) and the southern Kyushu-Shikoku clade (ITb) (Fig. 2), which diverged an estimated 1.64 Mya (95% HPD: 0.87–2.52) (Fig. I-3, Table I-5) at the early Pleistocene using method (i) and 0.93 Mya (95% HPD: 0.75–1.11) (Fig. I-4, Table I-5) at the late Early or Middle Pleistocene using the more accurate and reliable method (ii). The boundary between the southern and northern clades is the series of straits between Honshu and the islands of Shikoku and Kyushu, and this is somewhat different than the boundary for other mammals having similar distribution of northern and southern lineages in the Hondo region. For *Ursus thibetanus* (Yasukochi *et al.*, 2009), the sika deer (*Cervus nippon*) (Nagata *et al.*, 1999), the Japanese macaque (*Macaca fuscata*) (Kawamoto *et al.*, 2007), and the Japanese wild boar (*Sus scrofa*) (Watanobe, Ishiguro & Nakano, 2003), the contact boundary between northern and southern clades lies in the Chugoku district in western Honshu, with some haplotypes shared between the clades. All the animals mentioned above that have a contact boundary in Chugoku are of large size and large home range compared to *M. itatsi*, so they probably migrated back and forth. In our study, the genetic diversity indices for the northern and southern clades were similar and showed no evidence for founder effects (Table I-4). Tajima's *D* was higher and Fu's *F*_s was lower for the Honshu population than for the Shikoku-Kyushu clade (although neither Tajima's *D* value was significant), providing some evidence for recent expansion of the Honshu population (Table I-4).

The two hypotheses mentioned provide different explanations for the observed distribution of the northern and southern *M. itatsi* lineages. One is that clades ITa and ITb migrated separately from the Asian continent across the land bridge between the continent and southern Kyushu Island, with no contact between them; when the straits formed between Honshu and Kyushu-Shikoku, these acted as geographical barriers. However, Honshu, Shikoku, and Kyushu Islands became separated by the Seto Inland Sea and narrow straits only 5000–7000 years ago (Oshima 1990, 1991, 2000), whereas the divergence between the northern and southern clades is much older, as shown by results of the two methods used (Figs. I-3, I-4, Table I-5).

The alternative hypothesis is that there was a single migration event for the ancestral population from the continent to Kyushu and southwestern Honshu, with the northern and southern lineages subsequently diverging within Japan. In a previous study by Masuda et al. (2012), mitochondrial DNA control region sequences did not provide enough resolution to detect geographical patterns within the northern and southern lineages. Our use of complete mtDNA sequences, however, resolved geographically distinct lineages within the northern (ITa) and southern (ITb) clades despite small sample sizes. The tree topology and estimated divergence times (Figs. I-3, I-4) for clade ITa indicate a stepwise expansion of the range eastward and northward on Honshu during the Middle Pleistocene, from 0.68 Mya to 0.27 Mya or later using method (i), or even much later falling totally within the Middle to Late Pleistocene from 0.31 Mya to 0.11 Mya using method (ii). The earliest divergences occurred in the farthest southern and western part of the lineage's range (localities h and g) 1.02–0.68 and 0.49–0.38 Mya. Slightly more eastern haplotypes (localities d, e, f) diverged from

haplotypes farther north and east (clades ITaN and ITaE) 0.48 and 0.22 Mya, and the farthest north lineage (ITaN, localities a, b) diverged from the eastern lineage (ITaE, locality c) 0.27 and 0.11 Mya (Figs. I-3, I-4, Table I-5). This obvious pattern of stepwise, southwest to northeast divergence is similar to the south-to-north habitat replacement that occurred on Honshu during the Holocene. The estimated divergence times of method (ii) become much more closer to the time of the previously mentioned Holocene habitat replacement, indicating much more consistent results compared to method (i).

As with ITa, the southern clade (ITb) comprised three geographically distinct subclades (Shikoku, Kyushu and adjacent islands). Haplotypes from the most southern adjacent islands, Yakushima and Tanegashima, grouped together (ITbI), though with low nodal support; those from Shikoku formed subclade ITbS, and those from Kyushu formed subclade ITbK. The Kyushu population diverged from the others an estimated 1.03 Mya (Fig. I-3, Table I-5) and 0.52 Mya (Fig. I-3, Table I-5). Those divergence was earlier than the final formation of Osumi Strait 100,000–150,000 years ago, separating the adjacent islands of Yakushima and Tanegashima from Kyushu, but the later one of method (ii) becomes closer to Osumi Strait formation. Both divergence times are much earlier than the final formation of the Seto Inland Sea 5000–7000 years ago, separating Honshu, Shikoku and Kyushu, at which time southern Japan was already dominated by broad-leaved forests.

Assuming that *M. itatsi* arose from a single migration of an ancestral population from the continent followed by allopatric speciation, the question arises as to how the northern and southern lineages diverged in the Pleistocene and have remained distinct since then, even though the straits separating Kyushu and Shikoku from Honshu—and forming the boundary

between the lineages—were nonexistent for long periods during the Pleistocene, i.e., they were bridged by land. However, even when they were emergent during low sea-level stands, these straits nonetheless comprised valleys surrounded by lowlands, and could conceivably have been different enough in habitat from adjacent areas to constitute an impediment or barrier to migration by *M. itatsi*. A similar effect of straits-as-valleys during low sea-level stands might also explain the maintenance of geographically distinct sublineages within the southern lineage of *M. itatsi* on Kyushu, Shikoku, and smaller adjacent islands. A congruent example is the Japanese marten (*Martes melampus*) population on Tsushima Island, which is subdivided into northern and southern parts connected only by two narrow isthmuses. A population genetic study by Kamada et al. (2012) indicated that northern and southern marten populations on Tsushima were genetically distinct from each other, despite the presence of the isthmuses connecting them. Gene flow was limited because of the small home range of *M. melampus*.

Phylogenetic relationships and divergence times within *M. sibirica*

Our study detected two major haplotype clades for *M. sibirica*, SBa (continental Russia and Tsushima) and SBb (Korea, China, and Taiwan), which diverged 1.57 Mya (Fig. I-3, Table I-5), in the early Pleistocene using method (i) of calibration. Method (ii) resulted in a much later divergence time of 0.67 Mya (Fig. I-4, Table I-5) at the Middle Pleistocene. Both divergence times were only slightly more recent than, and essentially indistinguishable in timing from, that between the two major clades of *M. itatsi*. Tajima's *D* values were nearly equal (but non-significant) for clades SBa and SBb, and Fu's *F_s* values were low. Thus there

was no strong support for a recent expansion of either clade.

Both clades were subdivided into geographically southern and northern subclades. For clade SBa, the southern subclade (SBaT) was composed of Tsushima haplotypes and the northern subclade (SBaR) of continental Russian haplotypes. In clade SBb, the southern subclade (SBbT) consisted of Taiwanese haplotypes and the northern subclade (SBbC) of Korean and Chinese haplotypes. The two methods used for estimating divergence time between the continental Russian and Tsushima populations differ greatly from each other; 0.88 Mya (Fig. 3, Table 3) using the single external calibration point and 0.13 Mya (Fig. I-4, Table I-5) using the multiple calibration points. Additionally the later divergence time is more plausible and consistent with geological information. Tsushima Island last became isolated from Kyushu Island around 20,000 years ago (Nagaoka, 2001) and from the continent 100,000–150,000 years ago, although the exact timing of these events remains controversial (Oshima 1990, 1991, 2000). The close relationship between haplotypes from Tsushima and continental Russia suggests a wide distribution for this continental lineage in the past. Colonization of Tsushima Island from the continent could have occurred through an earlier land bridge than the last one connecting the island with the Korean Peninsula, with the Russian lineages subsequently disappearing from the Korean Peninsula and the Tsushima population remaining as a relict sister lineage to the Russian lineage. The leopard cat (*Prionailurus bengalensis*) has the similar distribution pattern, inhabiting Tsushima Island but not the Japanese main islands, and its mtDNA haplotypes are phylogenetically closely related to continental haplotypes, but with a much more recent divergence time (0.03 Mya) (Tamada et al., 2008).

For clade SBb, the estimated divergence between the Korean-Chinese and Taiwanese subclades was 0.94 Mya (95%HPD: 0.3–1.8) (Fig. I-3, Table I-5), toward the end of the Early Pleistocene using method (i) but it gave more recent Middle Pleistocene divergence of 0.42 Mya (0.12–0.79) using method (ii). Compared between our results and Hosoda *et al.* (2011), the effect of proper calibration is obvious. In their study, the phylogenetic relationships between continental and four Taiwanese mustelid species (*Martes flavigula*, *Melogale moschata*, *Mustela nivalis*, and *Mustela sibirica*) were examined using D-loop, ND2, and cytochrome *b* sequences. They used multiple calibration points based on fossil records. For their studied *Mustela* species, they relied on a fossil record for the *Neovison/ Mustela* complex, which is “external” to accurately estimate the divergence time between Taiwanese and the continental populations of *M. sibirica*. The single point calibration of our study and results of that study gave a very similar estimated time of divergence between Eurasian and Taiwanese *M. sibirica* (0.82 Mya) to ours (0.94 Mya). On the other hand, our properly calibrated divergence times were more recent and falling into the Middle Pleistocene (0.42 Mya). This difference probably was not only due to our use of the complete mtDNA genome, compared to their use of some mtDNA genes. Also it may not necessarily indicate an existence of multiple waves of colonization of *M. sibirica* and other Mustelidae from the continent to Taiwan during low sea-level stands associated with glacial-interglacial cycles, as suggested by Lüthi *et al.* (2008) and Hosoda *et al.* (2011). As long as there is evidence for carbon-dated well-identified fossils indicating those waves of colonization, it is probably due to bias coming from choosing the proper calibrations.

Our study gave a more detailed information about the divergence of *M. sibirica* compared

with the study of Masuda *et al.* (2012) using sequences of mtDNA control region. The differences between results are probably not due to genetic reasons, but sample size analyzed. In the study of Masuda *et al.* (2012), 25 samples were analyzed, 14 of which were from the introduced population due to the growing interest of studying the introduced populations of Japanese mammals. In the present study, we analyzed the complete mtDNA sequences for 26 individuals from the natural populations only, thus providing more clear results about the species phylogenetic history and divergence.

The divergences among clades and subclades of *M. sibirica* occurred from 1.57 Mya to 0.88 Mya, in the early Pleistocene (Fig. I-3, Table I-5) and from 0.67 to 0.13 Mya using the more proper calibration, in the Middle to Late Pleistocene. The chronology of divergences between the continental mammals and their Japanese endemic sister species or populations (Table I-6) have not been adequately studied chronologically, and so we here compare *M. sibirica*, which we studied, with three other mustelid species (sable, *Martes zibellina*; ermine, *Mustela erminea*; least weasel, *Mustela nivalis*), which, like *M. sibirica*, have a distribution on the Eurasian continent. These three mustelids also occur in northern Japan (Murakami, 2015; Masuda, 2015a, b), and the natural distributions of *Mustela erminea* and *Mustela nivalis* extend to the New World (USA and Canada). A study of *Martes zibellina* (Kinoshita *et al.*, 2015), based on mitochondrial ND2 sequences from many samples across a wide range of Eurasia, found that the lineage traces to a common ancestor 0.22 Mya. This is more recent than the comparable value of either of the 1.57 Mya or 0.67 for *M. sibirica* in our study. Likewise, based on ND2 sequences, *M. erminea* diverged within the Eurasian Continent from the common ancestor 0.3 Mya (Malyarchuk *et al.*, 2015). The D-loop sequences showed low

genetic diversity within the continental Eurasian population, although divergence times were not calculated (Kurose *et al.*, 2005). *Mustela nivalis* showed higher intraspecific genetic diversity than *M. erminea* (Kurose *et al.*, 2005), but the onset of divergence among lineages was still more recent than our estimated divergence time for *M. sibirica* (Lebarbenchon *et al.*, 2010). In the study of Martinkova *et al.* (2007), fossil calibrated divergence times indicate a recent Holocene divergence between either British or Irish stoats (*Mustela erminea*) and their continental European populations using cytochrome *b*, control region and the flanking tRNA genes between them. Mcdevitt *et al.* (2012) showed a Late Pleistocene divergence between Holarctic lineages of the least weasel (*Mustela nivalis*) using cytochrome *b* sequences for museum specimens. In summary, if looking only at the Early Pleistocene value of *M. sibirica* lineage divergence using the single calibration point, the onset of divergences for the Eurasian lineages of the three other mustelids could be in the Middle Pleistocene, more recent than the early Pleistocene divergence for *M. sibirica*. But putting the results coming from more proper calibration into account, the divergence time of *M. sibirica* lineages became more closer to the Middle Pleistocene. The broad differences among estimated divergence times from other studies may (but not necessarily) to some extent be due to the analysis of individual mtDNA genes in those studies. On the other hand, those continental species that have Japanese endemic sister species might represent basal lineages that were widespread in the Late Pliocene or Early Pleistocene. Phylogeographical studies of other continental/Japanese endemic pairs that incorporate fossil data, complete mtDNA sequences, and multiple nuclear gene loci will help to clarify divergences and their geological, geographical and climatic correlates.

Chapter II

**Comparative sequence variations among
different genes of mitochondrial genome for
the Japanese weasel (*Mustela itatsi*) and the
Siberian weasel (*M. sibirica*)**

Introduction

The mitochondrial DNA (mtDNA) consists of 37 genes: 13 protein coding genes, two ribosomal RNA (rRNA) subunits, 22 transfer RNA (tRNA) sites and a control region. One of the protein coding genes, namely the ND6 gene, lies in an opposite position in relation to other mtDNA genes, so it is not usually used for phylogeographical studies. Because the control region is thought to contain the sites of initiation of replication and transcription of mtDNA (Clyton, 1982), it is about ten times highly mutative, compared with other mtDNA genes. After mtDNA phylogeographical studies starting from 1990s, two rRNA subunits (12S and 16S rRNA genes) and a protein coding gene, cytochrome *b*, were often used to investigate the taxonomic ambiguities between different species and subspecies of mammals. One of the examples is a case of the phylogenetic study of *M. itatsi* and *M. sibirica*. Masuda & Yoshida (1994) examined the sequences of cytochrome *b* gene and reported the first genetic evidence for complete distinction between *M. itatsi* and *M. sibirica* as separate species, not as two sub-species belonging to the same species as previously thought. Another study used 12S rRNA gene (Kurose *et al.*, 2008) confirmed the same pattern between both species. On the other hand, the control region of mtDNA was widely used to study biogeography within a species and among populations of the species (Masuda *et al.*, 2012). In addition, cytochrome *b* has also been used for the same purpose as that of the control region. This gene loci shows a relatively moderate level of polymorphisms (lower than control region and higher than other protein coding genes in mtDNA. Then, combined multi-locus studies

including both the control region and cytochrome *b* were carried out for various mammalian species. The wide use of cytochrome *b* gene as a highly polymorphic gene following the control region was based on the first complete mtDNA genome draft for the first studied non model species in 1990s such as the cat (*Felis catus*) (Lopez, Ceverio & O'Brien 1996), followed by intensively studied either extinct or extant polar and brown bear populations (Delisle & Strobeck, 2002; Krause *et al.*, 2008; Lindqvist *et al.*, 2010; Hirata *et al.*, 2013, 2014). This concept was then adopted for other mammals, without carrying out complete mtDNA studies for each species or family independently.

In Japan, there are five extant mustelid species other than the two species studied here, none of which has been studied based on complete mtDNA genome sequences. There was only one mtDNA multi-locus study for a mustelid species; on the Japanese marten (*Martes melampus*), Sato, Yasuda & Hosoda (2009) examined ND2 sequences, and reported that this gene has higher parsimoniously informative sites in relation to length in base pairs $Pi/length$ than the cytochrome *b* and followed by the control region. Then, studies for other mustelid species were reported: the sable marten (*Martes zibellina*) (Kinoshita *et al.*, 2015) and the ermine (*Mustela ermine*) (Malyarchuk, Denisova & Derenko, 2015) showed the ability of ND2 gene to solve the phylogenetic relationships better than other previous studies using cytochrome *b* gene. Furthermore, Sato, Yasuda & Hosoda, (2009) selected ND2 gene together with cytochrome *b* and control region based on previously reported study and primers of a cat species study, not a mustelid. Compared to other families of the order Carnivora,

Complete mtDNA genome sequences for the species of the family Mustelidae is relatively few and not yet fully covered. Obtained complete mtDNA sequences for both *M. itatsi* and *M. sibirica* in Chapter I can then be used to figure out the best genes, which have the highest polymorphism for clarification of the molecular evolutionary features and for future use of multi-locus mtDNA studies for different genera of family Mustelidae.

Materials and Methods

Samples and DNA extraction

The samples and methods were the same as Chapter I.

Molecular methods

The molecular methods were the same as Chapter I.

Sequence variations

For the clades and lineages of mtDNA inferred from phylogenetic analysis, a total of 15,813 bp were then used for the present analysis. Parsimony informative sites (Pi) for of each of the protein coding regions, rRNA subunits and the control region were calculated using DnaSP ver5 (Librado and Rozas, 2009).

Results

Sequence variations within *M. itatsi*

Table II-1 showed parsimony informative sites (Pi) and the ratio of Pi to base-pair length of gene, Pi/length, were calculated for each of the 12 protein coding genes, two rRNA subunits and the control region. For the 26 individuals of *M. itatsi*, three protein-coding genes, CO1, ND4, ND5 (1,300-1,800 bp) showed the lower Pi/length among all protein coding genes and the control region (0.015, 0.013 and 0.013, respectively). On the other hand, smaller length protein-coding genes (1,000 bp or less) showed remarkably higher Pi/length; the highest value was 0.024 at CO3, and followed by 0.022 for ND2 gene. Cytochrome *b* gene has relatively lower Pi/length, (0.017), compared with other genes, which have the similar length to CO3, ND1 and ND2. Despite relatively shorter length (600 bp), the control region has twice larger Pi/length (0.045) than that of CO3 gene. The two rRNA subunits, 12S and 16S rRNA, showed the lowest Pi/length of 0.006 and 0.009 among all other genes.

Sequence variations within *M. sibirica*

Among 20 individuals of *M. sibirica*, ND4 has the highest Pi/length of 0.024 among all protein coding genes, even higher than that of the cytochrome *b* gene and almost equal to that of the control region. Other gene loci indicated similar values to each other. Like *M. itatsi*, two rRNA subunits showed the lowest Pi/length of 0.005 and 0.003 among all genes of mtDNA.

Discussion

Sequence variation of different gene loci of mtDNA for *M. itatsi* and *M. sibirica*

Comparison of parsimony informative sites in relation to each gene length (Pi/length) for the 12 protein coding genes, two rRNA subunits and the control region, showed useful information on which individual genes or selecting multiple genes in mtDNA should be used for future phylogenetic studies. For both species, cytochrome *b* did not have the highest Pi/length following the 600 bp of the control region, but another protein coding genes; CO3 for *M. itatsi* and ND4 for *M. sibirica*. On both species scale, ND2 gene had higher Pi/length than cytochrome *b* gene, and the similar pattern was reported in another mustelid species, *Martes melampus* (Sato *et al.*, 2009). In addition, *Mustela erminea* (Malyarchuk *et al.*, 2015) confirmed the similar pattern. These results suggest that ND2 gene can be used as a phylogenetic marker for studies of family Mustelidae together with the hypervariable control region, rather than the cytochrome *b* gene. In the case of selecting multiple genes for future phylogenetic studies of this family, a combination of control region/ ND2/ ND4/ cytochrome *b* could be used to resolve phylogenetic relationships, closely to the complete mtDNA genome sequences.

Chapter III

**Remarkably high variation of tandemly
repeated sequences within the
mitochondrial DNA control region of the
Siberian weasel (*Mustela sibirica*) on
Tsushima Island, Japan**

Introduction

In all animal taxa, the mitochondrial DNA (mtDNA) has a control region, which is an only major non-coding segment (Fumagalli *et al.*, 1996). This region is thought to contain the sites of initiation of replication and transcription of mtDNA (Clyton, 1982). The control region consists of a central conserved region (CCR) and conserved sequence blocks (CSB) (Anderson *et al.*, 1981; Walberg & Clayton, 1981). Upstream of the CCR or between CSBs, there are variable A/T-rich flanking sequences (Hoelzel *et al.*, 1994). Within the A/T rich flanking region, occurrence of short interspersed repeats was reported in primates (Hayasaka *et al.*, 1991) and cetaceans (Hoelzel *et al.*, 1991) and canivorans (Hoelzel *et al.*, 1993, 1994). The reason for the occurrence of the repetitious sequences is most probably due to turnover by DNA slippage (Holezel *et al.*, 1991, 1993). Those repetitious sequences, commonly referred as variable number tandem repeats (VNTRs), have variable positions among different species of vertebrates, named RS1, RS2, RS3, RS4 and RS5. Following the repetitious sequences, the CSB consists of three parts CSB1-3, and could to be associated with the initiation of the heavy strand replication (Chang *et al.*, 1985).

Since the advancement of molecular phylogeography, the polymorphisms of the mtDNA control region sizes have been found among more than 150 species of different animal taxa, due to the presence of VNTRs (Lund *et al.*, 1998). Among mammals, the VNTRs have been reported among individuals of the same species or within the same individual in the Japanese monkey (Hayasaka *et al.*, 1991), evening bat (Wiikinson & Chapman, 1991), rabbit (Biju-Duval *et al.*, 1990; Mignotte *et al.*, 1990), pig (Ghivizzani *et al.*, 1993), harbor seal (Aranson & Johanson, 1992), elephant seal (Hoelzel *et al.*, 1993), shrews (Stewart & Baker, 1994; Fumagalli *et al.*, 1996) and Japanese sika deer (Nagata *et al.*, 1998). Hoelzel *et al.* (1994) reported the presence of VNTRs in 21 carnivoran species belonging to eight families: the microsatellite-like repeats (2-10 base-pairs, bp, per unit and 14-103 repeats) were found in

the RS3 region of the control region, with the highest level of heteroplasmy, compared with animals reported to have VNTRs in regions other than RS3. Lunt *et al.* (1998) reviewed the published works of VNTRs of mtDNA, and listed the utility and common problems associated with their application in molecular ecology. After the recent advances in the auto-sequencer's technology for the last decade, sequence determination of VNTRs by direct nucleotide sequencing has become technically possible, if the target sequence size is so large. For *Mustela* species, the VNTRs size was reported to be approximately 350 bp, but the exact sequences were not successfully determined (Kurose *et al.*, 1999).

The Siberian weasel *Mustela sibirica*, which is the subject of the present study, has both Oriental and Palaearctic distributions (Wozencraft, 2005). In Japan, *M. sibirica* is indigenous to Tsushima Islands with an area of about 700 km², located between the Japanese islands and Korean Peninsula (Imaizumi, 1960; Abe *et al.*, 1994; Sasaki, 2009; Masuda *et al.*, 2012). The islands were reported to have been formed about 0.1 million years ago and currently surrounded by Korean and Tsushima Straits (Oshima, 1990, 1991). Masuda *et al.* (2012) examined the molecular phylogeny of the 5' sequence of about 600 bp (CCR) in the mtDNA control region of *M. sibirica*, and reported that three individuals from Tsushima Islands shared almost identical haplotypes when indel sites were excluded, having the closer phylogenetic relationships with haplotypes from Russian and Korean populations other than the haplotype of Taiwan (Masuda *et al.*, 2012).

In this chapter, I analyzed the complete mtDNA control region including the highly repetitious sequences of *Mustela sibirica* isolated in Tsushima Island, and discuss and the remarkably rapid evolution of the repeat units, and the implications of such unique sequences for the geographically isolated populations.

Materials and Methods

Samples and DNA extraction

Muscle tissue samples were obtained from 31 individuals of *M. sibirica* from Tsushima Islands, which were preserved in the Tsushima Wildlife Conservation Center. All individuals were road-killed on the islands. Sampling locations on the islands are shown in Fig. III-1. Total genomic DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN), following the manufacturer's protocol.

DNA amplification

Polymerase chain reaction (PCR) amplification of the mtDNA control region was performed using two primers, MUW1-6F (5'-GCCAACCAGTAGAACACCCA-3') and DTAN1-REV (5'-ATGGAGTCTTGTGACTCTTC-3'), which were newly designed in the present study. For confirmation of authenticity, two independent PCR amplification reactions were carried out separately. The PCR amplification was carried out in a reaction mixture of 20 µl containing 4 µl of 5X PrimeSTAR GXL DNA buffer (Takara), 1.6 µl of dNTP mixture (2.5 mM each dNTP, Takara), 0.4 µl of PrimeSTAR GXL DNA polymerase (1.25 units/µl, Takara), 0.2 µl from each of the two primers (25 pmol/µl), 2.0 µl of DNA extract and of distilled water. The PCR conditions were 35 cycles of 10 s at 98°C, 15 s at 60°C, and 1 min at 68°C; and then the PCR products were stored at 4°C. The molecular size of the PCR products was checked by electrophoresis on a 2% agarose gel followed by ethidium bromide staining, and observed under an ultraviolet illumination. Then, the PCR products were purified with the QIAquick Purification Kit (Qiagen), following the manufacturer's protocol.

Nucleotide sequencing

The DNA cycle sequencing was performed using the BigDye v3.1 Sequencing Kit (Applied Biosystems ABI) with sequencing primers newly designed in the present study.

Primers MUW1-6F and DL-REV (5'-GGCCATAGCTGAGTGATACC-3') were used as sequencing primers to sequence the 600-bp fragments of the control region. To determine the sequences of the VNTRs, the four sequencing primers were used; DTAN-2REV (5'-TAAGGGGGGTTTGACAAAGG-3'), MUW2-6F (5'-CCTCTCAAATGGGACATCTC-3'), MUWFTANC (5'-CCTTCATCATTTATCCGCAT-3'), and DTAN1-REV. The cycle PCR for sequencing was performed in 10 µl of the reaction mixture containing 1.75 µl of 5X BigDye v3.1 Sequencing Buffer (ABI), 1 µl of Ready Reaction Premix (ABI), 1.6 µl of any sequencing primer (1 pmol/µl), 1.0 µl of purified DNA template, and 4.65 µl of distilled water. As the PCR program, 30 cycles of 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C were performed. Amplified DNA fragments were purified with isopropanol, and then formamide was added. Nucleotide sequences were determined on an ABI 3730 DNA Analyzer. Each purified product from the two independent PCR amplifications were sequenced at least two times.

Sequence data analysis

Sequences obtained in the present study were aligned using MEGA version 6 (Tamura *et al.*, 2013). The 5'-end 600 bp against the repeats region were aligned and compared with the previously reported sequence of the Tsushima population of *M. sibirica* from Masuda *et al.* (2012) (Accession numbers AB007357, AB007359 and AB007360). About 250-bp fragments at the 3' end in the mtDNA control region, which were analyzed in the present study, were not polymorphic (see the conserved 250 bp in Fig. III-2), and aligned with the previously reported sequence of *M. sibirica* (NC_020673) from Yu *et al.* (2011). Repeat units in the VNTRs were classified according to Hoelzel *et al.* (1993a) and Hoelzel *et al.* (1994).

Results

Sequence variation within the mtDNA 5' end control region and polymorphism of C/T indels

Figure III-2 shows the basic structure within the mtDNA control region, based on our results in the present study. If C/T indel sites were excluded, the 31 individuals analyzed were clustered into one haplotype which is typically the same haplotype were reported among the three individuals *M. sibirica* previously reported by Masuda *et al.* (2012). When indel sites were included, the 31 individual were clustered into 17 haplotypes. The non-variable sequences of 250 bp following the VNTRs until the 3' end of the control region were identical among all individuals analyzed.

High polymorphism of VNTRs

Table III-1 shows the variation of repeat array structures in the VNTRs detected in the present study. In the VNTRs, all individuals analyzed included three of the main repeat units (#1, #2 and #3), four units (#h, #m, #k and #d) derived from #1, and one unit (#r) derived from #3, when the units were classified according to Hoelzel *et al.* (1993) and Hoelzel *et al.* (1994). The repeat unit “ACACG-” was newly found at the 3' end of the VNTRs of all individuals and named unit “#2*”, because this sequence had a one-nucleotide deletion at the 3' end of unit #2, which was originally reported by Hoelzel *et al.* (1994). Among all carnivores, unit #2* was reported for the first time.

The comparison of the VNTRs among the 31 individuals analyzed classified the VNTRs to seven patterns, based on the difference of the repeat numbers of unit “2h” (from 16 to 23 times), with the constant array “k3h” as the 5' starting units) and the 3' ending units “2*d1r” (Table III-1). On the seven different patterns, the smallest repeat number of units “2h” (pattern 1) was 16 times showing 195 bp as a total VNTRs length from one individual, whereas the largest number (pattern 7) was 23 times showing 265 bp from one individual. For other five patterns (patterns 2-6), 18 times of units “2h” in seven individuals, 19 times in six

individuals, 20 times in two individuals, 21 times in 11 individuals and 22 times in three individuals. Thus, the nucleotide length of VNTRs ranged from 195 to 265 bp.

Because Tsushima Islands are geographically separated into two main islands, northern and southern islands, the geographical distribution of the VNTR patterns on Tsushima was examined. Four patterns (18, 21, 19 and 22 times of “2h”) were shared by 17 individuals of the northern island and 11 individuals of the southern island. Two patterns (20 and 23 times of “2h”) were found in two individuals and in one individual only on the northern island, whereas one pattern (16 times of “2h”) was identified from one individual only on the southern island.

Two juvenile animals (MSI-TS13 and MSI-TS14), which were found at the same site on the same day, and thought to be siblings, because they shared an equal number of repeat units and also identical control region sequences, when including indels, the haplotypes are different from each other out of the VNTRs.

Four samples (TSW09-3, TSW09-4, TSW10-5 and TSW10-7) were collected from the same geographical locality during three months (December 2009, and January and February 2010), and they had different numbers of unit “2h” from each other. Two other individuals (TSW09-5 and TSW10-8) within the same year group from another geographical locality had two different patterns of VNTRs. Similarly, for the other year’s group, two individuals (MSI-TS07 and MSI-TS09), which have been sampled within years 1997-1998 from the same geographic locality, had two different patterns of the VNTRs.

High polymorphism of combined C/T indel sites and VNTRs

As shown in Table III-1, the combined sequences diversity of C/T indel of the 5’ 600 bp and VNTRs following that region showed a total of 27 variants from the studied 31 individuals. Even among individuals showing equal VNTRs, they showed remarkable

variation of C/T indels diversity. Only three pairs of individuals showed typically equal C/T indels/ VNTRs diversity; individuals Nos. 11 and 12 with a pattern 3 of VNTRs, individuals Nos. 24 and 25 with a pattern 5 of VNTRs and individuals Nos. 26 and 27 with a pattern 5 of VNTRs.

Discussion

Variation of mtDNA control region sequences

The present study showed that the 600-bp sequences at the 5' end of the mtDNA control region fell into from all the 31 individuals in the Tsushima population of *M. sibirica*. The results are congruent with the previously reported works showing the remarkably low mtDNA variations on island populations of carnivoran species: black bears (*Ursus americanus*) on Newfoundland Island, Canada (Paetkau & Strobeck, 1994); brown bears (*Ursus arctos*) on Kodiak Island, Alaska (Paetkau *et al.*, 1998); wolves (*Canis lupus*) on Banks Island, Canada (Carmichael *et al.*, 2001); American marten (*Martes americanus*) on Newfoundland Island, Canada (Kyle & Strobeck, 2003); pine martens in Scotland and Ireland (Kyle *et al.*, 2003); and Tsushima and Iromote wildcats (*Prionailurus bengalensis*) on Tsushima and Iromote Islands, Japan (Tamada *et al.*, 2008), and Japanese marten *Martes melampus* on Tsushima Islands (Sato *et al.*, 2009). All of the previous studies showed a kind of genetic distinctiveness of the insular populations, compared with their closest continental ones. As reported in Masuda *et al.* (2012), the Tsushima population of *M. sibirica* is phylogenetically closely related to the continental Russian population. With respect to the Tsushima populations of other carnivoran species, Sato *et al.* (2009) found only two haplotypes of mtDNA cytochrome *b* in the Tsushima population of *Martes melampus*, which is another member of Mustelidae distributed on Tsushima Islands. In addition, the Tsushima wildcat population shared identical sequences of the mtDNA control region (Tamada *et al.*, 2008).

On the other hand, the 3' end of the mtDNA control region (starting from CSB-3 and ending at the 3' end) was not polymorphic among all the 31 weasel individuals examined in the present study.

VNTRs structure and sequence variation

Although the VNTRs (named RS3 region) located between this non-variable region at the 3' end and the variable region at the 5' end in the control region have received less attention in the previous studies of mammalian mtDNA phylogeny, we revealed that the VNTRs evolve so rapidly and could show invaluable phylogenetic information even in a small insular population with a low level of genetic variations.

For different species of genus *Mustela*, Hoelzel *et al.* (1994) reported the occurrence of VNTRs in the American mink (*M. vison*), Siberian polecat (*M. erbarsmanni*) and black footed ferret (*M. nigripes*). The VNTRs of 10-bp units were reported from the least weasel (*M. nivalis*) (5'-TACGCATATG-3') and the ermine (*M. ermine*) (5'-TACGCACGCA-3'). But the exact number of repeats per individual could not be counted due to some technical difficulty in direct sequencing (Kurose *et al.*, 1999).

Compared with Hoelzel *et al.* (1994) reporting the VNTRs for 21 carnivore species, the present study shows the similar repeat patterns of VNTRs for the core region (2h)_n to two *Mustela* species, the Siberian polecat (*M. erbarsmanni*) and the black footed ferret (*M. nigripes*). By contrast, the American mink (*M. vison*) has some different pattern of the VNTRs (units 3321). The VNTRs of *M. sibirica* examined in the present study was different than the reported *M. erbarsmanni* and *M. nigripes* in the start and end repeat units; both had #k2h as a starting repeat unit. For the end unit, *M. erbarsmanni* had #24444 and #h444 and *M. nigripes* had #m223 and #m2334 end units.

The VNTRs of *M. sibirica* included units "3" and "d1", both were reported in Hoelzel *et*

al. (1994) but from relatively phylogenetically distal species; the former were reported in *M. vison*, while the later were reported in *Ursus arctos*, respectively. Although the relationships between the VNTR patterns and the species phylogeny may not be so strictly continuous, the overlapping turnover mechanisms could have affected the synthesis and evolution of VNTRs independently after carnivore divergence.

The newly described unit #2* (5'-ACACG-'3) could have been derived from unit #2 by one nucleotide deletion. This unit were always located at the 3' side of the units #d1r in all individuals examined in the present study. The occurrence of unit #2* and the number variation of core compound 10-bp unit "2h" suggests a step-wise continuous duplication of VNTRs due to slippage process during the DNA turnover.

On the other hand, Tamada et al. (2008) examined the VNTRs of the RS2 region for both eight individuals of the Tsushima wildcat and eight individuals of the Iromote wildcat, and reported the common pattern of repeat sequences among individuals in each population. Although our results were of the RS3 of *M. sibirica*, seven different patterns of the VNTRs were found in the 31 individuals within Tsushima Islands. This may indicate that the RS3 of *M. sibirica* have evolved rapidly more than the RS2 region of wildcats.

It is interesting to know that the individuals even from the same geographic locations within Tsushima Islands showed polymorphisms on different repeat numbers of compound unit "2h". As mtDNA is maternally inherited, the occurrence of the different repeat numbers among individuals from the same locality, which often share the identical 5'-end sequences in the control region, indicates that those animals may not belong to the same family (maternal lineage). This genetic feature shows that identification of the VNTR patterns in the mtDNA control region can be a useful tool to investigate maternal lineages for island populations, by counting how many families are present within a small population through the repeat number variations. The present study also showed that some patterns of the VNTRs in *M. sibirica*

were specific to the northern or southern island of Tsushima, although the other patterns were found in both islands. This result suggests some process of genetic differentiations between the northern and southern islands, in agreement with Kamada *et al.*'s (2012) report on the genetic separation of *Martes melampus* between the southern and northern Tsushima, based on a nuclear microsatellite analysis. Therefore, the polymorphism of mtDNA VNTRs found in the present study can be effectively applicable to the future pedigree studies using known captive or wild radio-collared individuals, in combination with use of variable nuclear DNA markers.

Combined genetic diversity of C/T indel sites and VNTRs

As regarded to be an insular island population, our studied individuals of the Siberian weasel showed high 17 variants of C/T indels at the 5' end of the mtDNA control region and seven patterns of VNTRs. Combining them together yielded 27 variants among 31 individuals. This could be an indication that combined C/T indels and VNTRs data could be used as an indicator to measure the genetic diversity of insular populations.

General Discussion

General Discussion

In the first chapter, complete mitochondrial genome sequences were first analyzed to investigate the phylogeography of the endemic Japanese weasel (*Mustela itatsi*) and the continental sister species, the Siberian weasel (*M. sibirica*). The results showed the genetic richness of *M. itatsi*, compared with *M. sibirica*. *Mustela itatsi* is distributed in a geographical area of about 1,500 km between the northern and southern locations. Within this geographical range, *M. itatsi* showed two monophyletic clades, each of which was separated to three subclades. By contrast, *M. sibirica* had only one subclade (Russian haplotypes) distributed within a geographical distance of about 3,500 km on the continent. This study also indicated that not only geographical barriers could have led to the present distribution of the Japanese endemic mammals on the Hondo islands, but also other factors (mainly ecological and behavioral factors) may play an important rule. For mammals with small size like these weasels, their home range could limit their divergence even in the case of having some land connections (i.e., this means that once diverged, they will have a little chance of migration and connections with other populations). This can explain why most endemic species of mammals on Hondo are of small home range and have an old divergence time from the continental populations despite having later land bridges and strait formations.

The second chapter showed that mitochondrial cytochrome *b*, which has been conventionally used for phylogenetic analysis of mammals, is not the highest parsimoniously informative protein-coding gene for both species. Among gene loci in mitochondrial genome, the control region had the highest variation in both species. Then, in the next chapter, molecular evolutionary characteristics of the control region were further examined.

In the third chapter, the genetic diversity of the control region was examined for an

insular population of *M. sibirica* on Tsushima Island, Japan. The results clearly showed that this region has two sites with remarkably high diversity: the C/T indel sites and the variable number tandem repeats (VNTRs). Combination of diversity patterns in both the sites yielded an overall high diversity among individuals. This indicates the genetic richness of the insular population and the possibility of the future use of the control region as a population genetic marker for isolated small populations

The present study revealed that the importance of complete mitochondrial genome sequences to further resolve the phylogenetic relationships between the Japanese and continental allopatric species. The results and indications from this study caused the emergence of other questions. First, where were the two species diverged? This question is still controversial not only for the studied pair, but also for other continental/Japanese endemic pairs. They all did not show any fossil records on any ancestral or intermediate forms both from the continent and the Japanese islands. The second question is why there is an incongruence between the estimated divergence times and the paleontological record? This problem is common also for other mustelid species. Fossils of *M. itatsi* and *M. sibirica* have been found not in the Early Pleistocene layers, but in the Middle Pleistocene ones. The third question goes to the Tsushima population of *M. sibirica*. The divergence time was calculated, which is much earlier than the last land connection between Tsushima Island and the Korean Peninsula, but it is still unclear how the Tsushima population was formed. To clarify such controversies, the following studies can provide some resolution:

(A) The comparison of the data of the present study with complete mitochondrial genome data on other continental/Japan-endemic species pairs. For example, *Martes* species/ Japanese marten (*Martes melampus*), the Asian badger (*Meles leucurus*)/Japanese badger (*Meles anakuma*). In addition, it is useful to conduct the

comparative study of non-mustelid pairs, which have more clear fossil records, i.e. the continental and Japanese populations of the Asiatic black bear (*Ursus thibetanus*).

(B) Ancient DNA studies for fossils of Japan and the continent. The estimated divergence times based on genetic data of the contemporary samples are sometimes not in concordance with the paleontological data. Recently, phylogeographic studies by ancient DNA analysis of fossils became practically possible and provide interesting results, using some individual genes, complete or near complete mitochondrial genome. Recently, the next generation sequencer technique known as the target capture sequencing was used to successfully sequence complete mitochondrial genome from Pleistocene human fossils (Posth et al., 2016). Analysis of such fossils is a powerful tool to understand the phylogeographic history of the species and populations with the past climate changes. In addition, the controversy of inconsistency between the divergence times and paleontological dating will be solved.

(C) Multi-approach genetic studies. To answer the questions like the founder population for clades of *M. itatsi*, a combination of maternal complete mitochondrial genome, autosomal nuclear loci, a large number of nuclear microsatellites and paternal Y-chromosome genes can be analyzed. The resulted summary statistics from all approaches separately can be compared to consider the best demographic scenario using the approximate Bayesian computation (ABC). Such kind of studies have the advantage for using different genes with different substitution rates and select the most acceptable demographic scenario among them to infer the phylogeographic history. Recently, the similar studies were already done for nonmodel species; the *Canis* species complex (Koepfli et al., 2015) and the Arabian camel (Almathen et al., 2016).

(D) For further study of the island populations of *M. sibirica* such as on Tsushima and Taiwan, showing their older Pleistocene divergence time, in relation to the continental

Eurasian population, genetic admixture studies should be done using large number of nuclear microsatellites and genome-wide SNPs. Analysis of mitochondrial DNA data using ABC modeling could suggest the mechanism driving to the current distribution of the species.

(E) Genome-wide analysis for the Japanese endemic mammals contributes to better understanding for the effect of endemism on the Japanese islands. In some cases, isolated populations are vulnerable to low genetic diversity and population decline. As land-bridge islands, the Japanese islands experienced multiple connections with the continent, allowing their fauna to have high genetic diversity. Genome-wide studies may show what kinds of genetic mechanisms lead to such peculiar diversity.

References

References

- Abe H, Ishii N. 1987.** Mammals of Tsushima Island. In Nagasaki Prefecture, Nagasaki, *Biogeography of Tsushima Islands*. 79-109 (in Japanese with an English summary)
- Abe H, Ishii N, Kaneko K, Maeda K, Miura S, Yoneda M. 1994.** A pictorial guide to the mammals of Japan. *Tokyo, Tokai Univ Press* (in Japanese).
- Almathen F, Charruau P, Mohandesan E, Mwacharo JM, Orozco-ter Wengel P, Pitt D, Abdussamad AM, Uerpmann M, Uerpmann HP, De Cupere B, Magee P. 2016.** Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. *Proceedings of the National Academy of Sciences*. **9**: 201519508
- Aranson U, Johnson E. 1992.** The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. *Journal of Molecular Evolution* **34**: 493-505.
- Avise JC, Ellis D. 1986.** Mitochondrial DNA and the evolutionary genetics of higher animals [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **312**: 325-342.
- Biju-Duval C, Ennafaa H, Denneboug N, Monnerot M, Mignotte F, Soriguer RC, E1 Gaaied A, E1 Hili A, Mounolou JC. 1990.** Mitochondrial DNA evolution in lagomorphs: origin of systematic heteroplasmy and organization of diversity in European rabbits. *Journal of Molecular Evolution* **33**: 92-102
- Carmichael LE, Nagy JA, Larter NC, Strobeck C. 2001.** Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. *Molecular Ecology* **10**: 2787-2798
- Chang DD, Wong TW, Hixson JE, Clyton DA. 1985.** Regulatory sequences for mammalian mitochondrial transcription and replication. In: Quagliariello E,

- Slater EC, Palmieri F, Saccone C, Kroon AM. *Achievements and Perspectives of Mitochondrial DNA Research* :135-144.
- Clayton DA 1991.** Nuclear gadgets in mitochondrial DNA replication and transcription. *Trends in Biochemical Sciences* **16**: 107-111
- Clayton DA 1992.** Transcription and replication of animal mitochondrial DNAs. *International review of cytology***141**: 217-232
- Conroy CJ, Cook JA. 2000.** Molecular systematics of the holarctic rodent (Microtus: Muridae). *Journal of Mammalogy* **81**: 344–359.
- Delisle I, Strobeck C. 2002.** Conserved primers for rapid sequencing of the complete mitochondrial genome from carnivores, applied to three species of bears. *Molecular Biology and Evolution* **19**: 357–361.
- Dobson M. 1994.** Patterns of distribution in Japanese land mammals. *Mammal Review* **24**: 91–111.
- Dobson M, Kawamura Y. 1998.** Origin of the Japanese land mammal fauna: allocation of extant species to historically-based categories. *Quaternary Research* **37**: 385–395.
- Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.
- Fu YX. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.
- Fumagalli L, Taberlet P, Favre L, Hausser J. 1996.** Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Molecular Biology and Evolution* **13**: 31-46.
- Ghivizzani SC, Mackay SLD, Madsen CS. 1993.** Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. *Journal of*

Molecular Evolution **37**: 36-47.

Hayasaka K, Ishida T, Horai S. 1991. Heteroplasmy and polymorphism in the major noncoding region of mitochondrial DNA in Japanese monkeys, associated with tandemly repeated sequences. *Molecular Biology and Evolution* **8**: 399-415

Heaney L.R. 1984. Mammalian species richness on islands on the Sunda Shelf, Southeast Asia. *Oecologia* **61**: 11-17.

Herman JS, Searle JB. 2011. Post-glacial partitioning of mitochondrial genetic variation in the field vole. *Proceedings of the Royal Society of London B: Biological Sciences* **278**: 601–3607.

Herman JS, McDevitt AD, Kawalko A, Jaarola M, Wójcik JM, Searle JB. 2014. Land-bridge calibration of molecular clocks and the post-glacial colonization of Scandinavia by the Eurasian field vole *Microtus agrestis*. *PloS one* **9**: p.e103949.

Hirata D, Mano T, Abramov AV, Baryshnikov GF, Kosintsev PA, Vorobiev AA, Raichev EG, Tsunoda H, Kaneko Y, Murata K, Fukui D, Masuda R. 2013. Molecular phylogeography of the brown bear (*Ursus arctos*) in northeastern Asia based on analyses of complete mitochondrial DNA sequences. *Molecular Biology and Evolution* **30**: 1644–1652.

Hirata D, Abramov AV, Baryshnikov GF, Masuda R. 2014. Mitochondrial DNA haplogrouping of the brown bear, *Ursus arctos* (Carnivora: Ursidae) in Asia, based on a newly developed APLP analysis. *Biological Journal of the Linnean Society* **111**: 627–635.

Ho SYM. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology* **38**: 409–414.

Ho SYW, Duchêne S. 2014. Molecular clock methods for estimating evolutionary

- rates and timescales. *Molecular Ecology* **23**: 5947-5965.
- Ho SYW, Phillips MJ, Cooper A, Drummond AJ. 2005.** Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* **7**: 1561–1568.
- Ho SYW, Saarma U, Barnett R, Haile J, Shapiro B. 2008.** The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS One* **3**: p.e1615.
- Holezel AR, Hancock JM, Dover JA. 1991.** Evolution of the cetacean mitochondrial D-loop region. *Molecular Biology and Evolution* **8**: 475-493.
- Holezel AR, Hancock JM, Dover JA. 1993.** Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. *Journal of Molecular Evolution* **37**: 190-197.
- Hoelzel AR, Lopez JV, Dover GA, O'Brien SJ. 1994.** Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. *Journal of Molecular Evolution* **39**: 191-199.
- Hosoda T, Suzuki H, Iwasa M, Hayashida M, Watanabe S, Tatara M, Tsuchiya K. 1999.** Genetic relationships within and between the Japanese marten *Martes melampus* and the Sable *M. zibellina*, based on variation of mitochondrial DNA and nuclear ribosomal DNA. *Mammal Study* **24**: 25-33.
- Hosoda T, Suzuki H, Harada M, Tsuchiya K, Han SH, Zhang Y, Kryukov AP, Lin LK. 2000.** Evolutionary trends of the mitochondrial lineage differentiation in species of genera *Martes* and *Mustela*. *Genes Genetics Systematics* **75**: 259–67.
- Hosoda T, Sato JJ, Lin LK, Chen YJ, Harada M, Suzuki H. 2011.** Phylogenetic history of mustelid fauna in Taiwan inferred from mitochondrial genetic loci.

- Canadian Journal of Zoology* **89**: 559–569.
- Imaizumi Y. 1960.** Coloured illustrations of the mammals of Japan. Hoikusha, Osaka (in Japanese).
- Imaizumi Y. 1970.** Land mammals of the Tsushima Islands, Japan. *Memoirs of the National Museum of Nature and Science* **3**: 159—176 (in Japanese).
- Jobb G, Von Haeseler A, Strimmer K. 2004.** TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology* **4**: 18.
- Jones MR, Good JM. 2016.** Targeted capture in evolutionary and ecological genomics. *Molecular ecology*. **25**: 185-202.
- Kamada S, Moteki S, Baba M, Ochiai K, Masuda R. 2012.** Genetic distinctness and variation in the Tsushima Islands population of the Japanese marten, *Martes melampus* (Carnivora: Mustelidae), revealed by microsatellite analysis. *Zoological Science* **29**: 827-833.
- Kawamoto Y, Shotake T, Nozawa K, Kawamoto S, Tomari K, Kawai S, Shirai K, Morimitsu Y, Takagi N, Akaza H, Fujii H, Hagihara K, Aizawa K, Akachi S, Oi T, Hayaishi S. 2007.** Postglacial population expansion of Japanese macaques *Macaca fuscata* inferred from mitochondrial DNA phylogeography. *Primates* **48**: 27–40.
- Kawamura Y, Kamei T, Taruno H. 1989.** Middle and late Pleistocene mammalian faunas in Japan. *Quaternary Research* **28**: 317–326.
- Kinoshita G, Sato JJ, Meschersky IG, Pishchulina SL, Simakin LV, Rozhnov VV, Malyarchuk BA, Derenko MV, Denisova GA, Frisman LV, Kryukov A.P. 2015.** Colonization history of the sable *Martes zibellina* (Mammalia, Carnivora) on the marginal peninsula and islands of northeastern Eurasia. *Journal of*

Mammalogy **96**: 172–184.

Knaus BJ, Cronn R, Liston A, Pilgrim K, Schwartz MK. 2011. Mitochondrial genome sequences illuminate maternal lineages of conservation concern in a rare carnivore. *BMC Ecology* **11**: 10.

Kopefli KP, Deere KA, Slater GJ, Begg C, Begg K, Grassman L, Lucherini M, Veron, G, Wayne RK. 2008. Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biology* **6**: 10.

Krause J, Unger T, Noçon A, Malaspinas AS, Kolokotronis SO, Stiller M, Soibelzon L, Spriggs H, Dear PH, Briggs AW, Bray SC. 2008. Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evolutionary Biology* **8**: 220.

Kurihara A, Hanna H, Takeaki H, Hiroshi K. 2016. Phylogeography of *Asparagopsis taxiformis* revisited: Combined mtDNA data provide novel insights into population structure in Japan. *Phycological Research* **64**: 95–101.

Kurose N, Masuda R, Yoshida MC. 1999. Phylogeographic variation in two mustelines, the least weasel *Mustela nivalis* and the ermine *M.erminea* of Japan, based on mitochondrial DNA control region sequences. *Zoological Science* **16**: 971-977

Kurose N, Masuda R, Siriaronrat B, Yoshida MC. 1999. Intraspecific variation of mitochondrial cytochrome *b* gene sequences of the Japanese marten *Martes melampus* and the sable *Martes zibellina* (Mustelidae , Carnivora , Mammalia) in Japan. *Zoological Science* **16**: 693–700.

Kurose N, Abramov AV, Masuda R. 2000. Intrageneric diversity of the cytochrome *b* gene and phylogeny of Eurasian species of the genus *Mustela* (Mustelidae ,

- Carnivora). *Zoological Science* **17**: 673–679.
- Kurose N, Kaneko Y, Abramov AV, Siriaronrat B, Masuda R. 2001.** Low genetic diversity in Japanese populations of the Eurasian badger *Meles meles* (Mustelidae , Carnivora) revealed by mitochondrial cytochrome *b* gene sequences. *Zoological Science* **18**: 1145–1151.
- Kurose N, Abramov AV, Masuda R. 2005.** Comparative phylogeography between the ermine *Mustela erminea* and the least weasel *M. nivalis* of Palaearctic and Nearctic regions, based on analysis of mitochondrial DNA control region sequences. *Zoological Science* **22**: 1069–1078.
- Kurose N, Masuda R, Tatara M. 2005.** Fecal DNA analysis for identifying species and sex of sympatric carnivores: A noninvasive method for conservation on the Tsushima Islands, Japan. *Journal of Heredity* **96**: 688-697.
- Kurose N, Abramov AV, Masuda R. 2008.** Molecular phylogeny and taxonomy of the genus *Mustela* (Mustelidae, Carnivora), inferred from mitochondrial DNA sequences: New perspectives on phylogenetic status of the back-striped weasel and American mink. *Mammal Study* **33**: 25–33.
- Kyle CJ, Strobeck C. 2003.** Genetic homogeneity of Canadian mainland marten populations underscores the distinctiveness of Newfoundland pine martens (*Martes americana atrata*). *Canadian Journal of Zoology* **81**: 57-66.
- Kyle CJ, Davison A, Strobeck C. 2003.** Genetic structure of European pine martens (*Martes martes*), and evidence for introgression with *M.americana* in England. *Conservation Genetics* **4**: 179-188.
- Koepfli KP, Pollinger J, Godinho R, Robinson J, Lea A, Hendricks S, Schweizer RM, Thalmann O, Silva P, Fan Z, Yurchenko AA. 2015.** Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species.

Current Biology. **25**: 2158-2165.

Lawlor T.E. 1986. Comparative biogeography of mammals on islands. *Biological Journal of the Linnean Society* **28**: 99–125.

Lebarbenchon C, Poitevin F, Arnal V, Montgelard C. 2010. Phylogeography of the weasel (*Mustela nivalis*) in the western-Palaeartic region: combined effects of glacial events and human movements. *Heredity* **105**: 449–462.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Lindqvist C, Schuster SC, Sun Y, Talbot SL, Qi J, Ratan A, Tomsho LP, Kasson L, Zeyl E, Aars J, Miller W. 2010. Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. *Proceedings of the National Academy of Sciences* **107**: 5053–5057.

Lüthi D, Le Floch M, Bereiter B, Blunier T, Barnola JM, Siegenthaler U, Raynaud D, Jouzel J, Fischer H, Kawamura K, Stocker TF. 2008. High-resolution carbon dioxide concentration record 650,000-800,000 years before present. *Nature* **453**: 379–382.

Lopez JV, Ceverio S, O'Brien SJ. 1996. Complete nucleotide sequence of the domestic cat (*Felis catus*) mitochondrial genome and the transposed mtDNA tandem repeats (Numt) in the nuclear genome. *Genomics* **33**: 229-246

Lunt DH, Whipple LE, Hyman BC. 1998. Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Molecular Ecology* **7**: 1441-1455.

Madsen CS, Ghivizzani SC, Hauswirth WW. 1993. In vivo and in vitro evidence for slipped mispairing in mammalian mitochondrial DNA. *Proceedings of Natural Academy of Science* **90**: 7671-7675

- Malyarchuk BA, Denisova GA, Derenko MV. 2015.** Molecular dating of intraspecific differentiation of stoats (*Mustela erminea*) based on the variability of the mitochondrial ND2 gene. *Russian Journal of Genetics Applied Research* **5**: 16–20.
- Marmi J, López-Giráldez JF, Domingo-Roura X. 2004.** Phylogeny, evolutionary history and taxonomy of the Mustelidae based on sequences of the cytochrome *b* gene and a complex repetitive flanking region. *Zoologica Scripta* **33**: 481–499.
- Masuda R. 2015a.** *Mustela nivalis* . In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 252–253.
- Masuda R. 2015b.** *Mustela erminea* . In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 254–255.
- Masuda R, Yoshida MC. 1994.** A molecular phylogeny of the family Mustelidae (Mammalia, Carnivora), based on comparison of mitochondrial *cytochrome b* nucleotide sequences. *Zoological Science* **11**: 605–612.
- Masuda R, Watanabe S. 2015.** *Mustela itatsi*. In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 250–251.
- Masuda R, Kurose N, Watanabe S, Abramov AV, Han S, Lin L, Oshida T. 2012.** Molecular phylogeography of the Japanese weasel, *Mustela itatsi* (Carnivora: Mustelidae), endemic to the Japanese islands, revealed by mitochondrial DNA analysis. *Biological Journal of the Linnean Society* **107**: 307–321.
- Matsubishi T, Masuda R, Mano T, Yoshida MC. 1999.** Microevolution of the mitochondrial DNA control region in the Japanese brown bear (*Ursus arctos*) population. *Molecular Biology and Evolution*, **16**: 676–684.
- Mcdevitt AD, Zub K, Kawalko A, Oliver MK, Herman JS, Wojcik JM. 2012.** Climate and refugial origin influence the mitochondrial lineage distribution of

- weasels (*Mustela nivalis*) in a phylogeographic suture zone. *Biological Journal of the Linnean Society* **106**: 57–69.
- Mignotte F, Gueride M, Champagne AM, Mounolou JC. 1990.** Direct repeats in the non-coding region of rabbit mitochondrial DNA involvement in the generation of intra- and interindividual heterogeneity. *European Journal of Biochemistry* **194**: 561-571.
- Millien-Parra V, Jaeger JJ. 1999.** Island biogeography of the Japanese terrestrial mammal assemblages: an example of a relict fauna. *Journal of Biogeography* **26**: 959–972.
- Mckay BD. 2012.** A new timeframe for the diversification of Japan’s mammals. *Journal of Biogeography* **39**: 1134–1143.
- Motokawa M. 2009.** Distribution patterns and zoogeography of Japanese mammals. In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 44-46.
- Murakami T. 2015.** *Martes zibellina* . In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 260–261.
- Nagata J, Masuda R, Tamate HB, Hamasaki SI, Ochiai K, Asada M, Tatsuzawa S, Suda K, Tado H, Yoshida MC. 1999.** Two genetically distinct lineages of the sika deer, *Cervus nippon*, in Japanese islands: comparison of mitochondrial *D-loop* region sequences. *Molecular Phylogenetics and Evolution* **13**: 511–519.
- Nagaoka S. 2001a.** The effect of glacial sea-level and climate change during the quaternary period in Kyushu island. In “*The Geomorphological History of Kyushu Island and Nansei Islands: The Topography of Japan. 7. Kyushu Island and Nansei Islands.*” Ed by Machida Y, Ohta Y, Kawana T, Moriwaki H, Nagaoka S, University of Tokyo Press, Tokyo, 298-301 (in Japanese).

- Nagaoka S. 2001b.** Tsushima island and Tsushima strait. In *Northern and Central Part of Kyushu Island: The Topography of Japan, 7. Kyushu Island and Nansei Islands* Ed by Machida Y, Ohta Y, Kawana T, Moriwaki H, Nagaoka S, Univ of Tokyo Press, Tokyo, pp 106-109 (in Japanese).
- Nei M. 1987.** *Molecular Evolutionary Genetics*. New York, NY: Colombia University Press.
- Nunome M, Torii H, Matsuki R, Kinoshita G, Suzuki H. 2010.** The influence of Pleistocene refugia on the evolutionary history of the Japanese hare, *Lepus brachyurus*. *Zoological Science* **27**: 746–754.
- Ogino S, Otsuka H, Harunari H, 2009.** The middle Pleistocene Matsugae fauna, northern Kyushu, West Japan. *Paleontological Research* **13**: 367–384.
- Ohdachi SD, Dokuchaev, NE, Hasegawa M, Masuda R. 2001.** Intraspecific phylogeny and geographical variation of six species of northeastern Asiatic *Sorex* shrews based on the mitochondrial cytochrome *b* sequences. *Molecular Ecology* **10**: 2199–2213.
- Ohdachi SD, Iwasa M, Nesterenko V, Abe H, Masuda R, Haberl W. 2004.** Molecular Phylogenetics of *Crocidura* shrews (Insectivora) in East and Central Asia. *Journal of Mammalogy* **85**: 396–403.
- Ohdachi SD, Hasegawa M, Iwasa MA, Vogel P, Oshida T, Lin L, Abe H. 2006.** Molecular phylogenetics of soricid shrews (Mammalia) based on mitochondrial cytochrome *b* gene sequences : with special reference to the Soricinae. *Journal of Zoology* **270**: 177–191.
- Oshida T, Masuda R. 2000.** Phylogeny and zoogeography of six squirrel species of the genus *Sciurus* (Mammalia , Rodentia), inferred from cytochrome *b* gene sequences. *Zoological Science* **3**: 405–409.

- Oshida T, Arslan A, Noda M. 2009.** Phylogenetic relationships among the Old World *Sciurus* squirrels. *Folia Zoologica* **58**: 14–25.
- Ohshima K. 1990.** The history of straits around the Japanese Islands in the late-Quaternary. *Quaternary Research* **29**: 193-208 (in Japanese with an English abstract).
- Ohshima K. 1991.** The late-Quaternary sea-level change of the Japanese islands. *Journal of Geography*. **100**: 967-975 (in Japanese).
- Ohshima K. 2000.** The history of strait formation around the Japanese Islands. *Kaiyou Monthly* **37**: 208-213 (in Japanese)
- Paetkau D, Strobeck C. 1994.** Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* **3**: 489-495.
- Paetkau D, Waits LP, Clarkson PL, Craighead L, Vyse E, Ward R, Strobeck C. 1998.** Variation in genetic diversity across the range of north American brown bears. *Conservation Biology* **12**: 418-429.
- Park SC, Yoo DG, Lee CW, Lee EI. 2000.** Last glacial sea-level changes and paleogeography of the Korea (Tsushima) Strait. *Geo-Marine Letters* **20**: 64–71.
- Posth, C., Renaud, G., Mittnik, A., Drucker, D.G., Rougier, H., Cupillard, C., Valentin, F., Thevenet, C., Furtwängler, A., Wissing, C., Francken, M., 2016.** Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a late glacial population turnover in Europe. *Current Biology*. **26**: 827-833.
- Rambaut A, Drummond AJ. 2007.** Tracer v1.4, available from <http://beast.bio.ed.ac.uk/Tracer>.
- Ronquist F, Huelsenbeck JP. 2003.** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

- Saccone C, Pesole G, Sbisa E. 1991.** The main regulatory region of mammalian mitochondrial DNA: structure- function model and evolutionary pattern. *Journal of Molecular Evolution* **33**: 83-91.
- Sasaki H. 2015.** *Mustela sibirica*. In: Ohdachi SD, Ishibashi Y, Iwasa MA, Saitoh T, eds. *The Wild Mammals of Japan*. Kyoto, Shoukadoh, 250–251.
- Sasaki H, Ohta K, Aoi T, Watanabe S, Hosoda T, Suzuki H, Abe M, Koyasu K, Kobayashi S, Oda SI. 2014.** Factors affecting the distribution of the Japanese weasel *Mustela itatsi* and the Siberian weasel *M. sibirica* in Japan. *Mammal Study* **39**: 133–139.
- Sato JJ. 2013.** Phylogeographic and feeding ecological effects on the mustelid faunal assemblages in Japan. *Animal Systematics, Evolution and Diversity* **29**: 99–114.
- Sato JJ, Hosoda T, Wolsan M, Tsuchiya K, Yamamoto Y, Suzuki H. 2003.** Phylogenetic relationships and divergence times among mustelids (Mammalia: Carnivora) based on nucleotide sequences of the nuclear interphotoreceptor retinoid binding protein and mitochondrial cytochrome *b* genes. *Zoological Science* **20**: 243–264.
- Sato JJ, Yasuda SP, Hosoda T. 2009.** Genetic diversity of the Japanese Marten *Martes melampus* and its implications for the conservation unit. *Zoological Science* **26**: 457–466.
- Sato JJ, Wolsan M, Prevosti FJ, Elía GD, Begg C, Begg K, Hosoda T, Campbell KL, Suzuki H. 2012.** Evolutionary and biogeographic history of weasel-like carnivorans (Musteloidea). *Molecular Phylogenetics and Evolution* **63**: 745–757.
- Savolainen P, Arvestad L, Lundeberg J. 2000.** MtDNA Tandem Repeats in Domestic Dogs and Wolves: Mutation Mechanism Studied by Analysis of the

- Sequence of Imperfect Repeats. *Molecular Biology and Evolution* **17**: 474-488.
- Shalabi MA, Abramov AV, Kosintsev PA, Lin L-K, Han S-G, Watanabe S, Yamazaki K, Kaneko Y, Masuda R. 2016.** Comparative Phylogeography of the endemic Japanese weasel (*Mustela itatsi*) and the continental Siberian weasel (*M. sibirica*), revealed by complete mitochondrial genome sequences. *Biological Journal of the Linnean Society* (in press).
- Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*: **btu033**.
- Stewart DT, Baker AJ. 1994.** Patterns of sequence variation in the mitochondrial D-loop region of shrews. *Molecular Biology and Evolution* **11**: 9-21.
- Strobeck C. 1998.** Variation in genetic diversity across the range of north American brown bears. *Conservation Biology* **12**: 418-429.
- Suzuki T, Yuasa H, Machida Y. 1996.** Phylogenetic position of the Japanese river otter *Lutra nippon* inferred from the nucleotide sequence of 224 bp of the mitochondrial *cytochrome b* gene. *Zoological Science* **13**: 621–626.
- Tajima F. 1989.** Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Tamada T, Siriaronrat B, Subramaniam V, Hamachi M, Lin LK, Oshida T, Rerkamnuaychoke W, Masuda R. 2008.** Molecular diversity and phylogeography of the Asian leopard cat, *Felis bengalensis*, inferred from mitochondrial and Y-chromosomal DNA sequences. *Zoological Science*, **25**: 154-163.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011.** MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.

- Tamura K, Stecher G, Peterson D, Filipski A. 2013.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.
- Tanabe AS. 2011.** Kakusan4 and Aminosan: Two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. *Molecular Ecology Resources* **11**: 914–921.
- Tashima S, Kaneko Y, Anezaki T, Baba M, Yachimori S, Abramov AV, Saveljev AP, Masuda R. 2011.** Phylogeographic sympatry and isolation of the Eurasian badgers (*Meles*, Mustelidae , Carnivora): Implications for an alternative analysis using maternally as well as paternally inherited genes. *Zoological Sciences* **4**: 293–303.
- Tatara M. 1994.** Ecology and conservation status of Tsushima marten, *Martes melampus tsuensis*. Ed by Buskirk SW, Harestad AS, Raphael MG, Powell RA. In *Martens, Sables, and Fishers: Biology and Conservation*. New York: Cornell Univ Press 272-279.
- Tatara M, Doi T. 1994.** Comparative analyses on food habits of Japanese marten, Siberian weasel and leopard cat in the Tsushima Islands, Japan. *Ecol Research* **9**: 99-107.
- Thompson JD, Higgins DG, Gibson TJ. 1994.** Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673–4680.
- Tsuchiya K, Suzuki H, Shinohara A, Harada M, Wakana S, Sakaizumi M, Han S, Lin L, Kryukov AP. 2000.** Molecular phylogeny of East Asian moles inferred

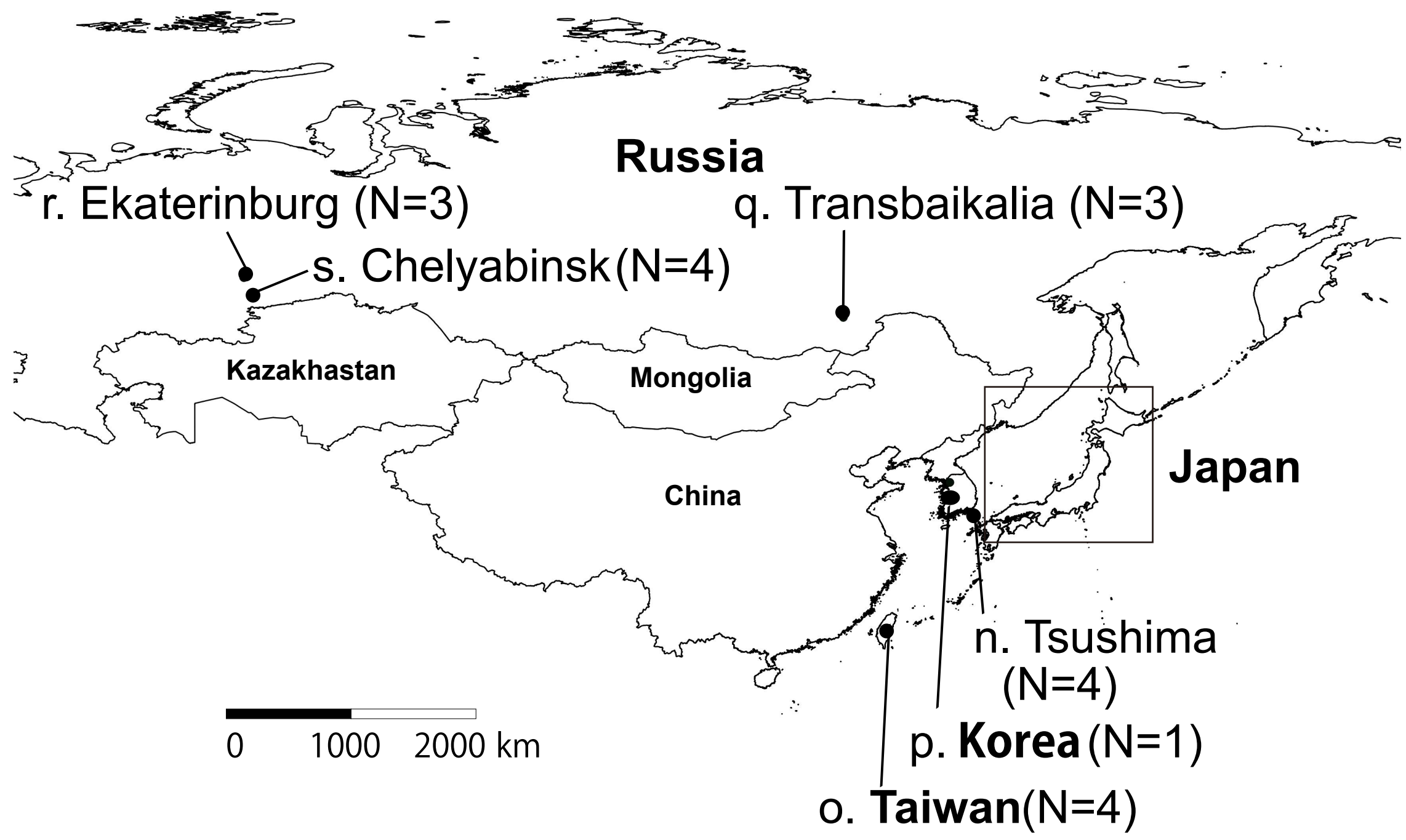
- from the sequence variation of the mitochondrial *cytochrome b* gene. *Genes Genetics Systematics* **75**: 17–24.
- Tsushima Development Bureau. 2011.** Encyclopedia of Tsushima. Showado Press, Isahaya (in Japanese)
- Uwai S, Kogame K, Yoshida G, Kawai, H, Ajisaka T. 2009.** Geographical genetic structure and phylogeography of the *Sargassum horneri/filicinum* complex in Japan, based on the mitochondrial *cox3* haplotype. *Marine biology* **156**: 901–911.
- Watanobe T, Ishiguro N, Nakano M. 2003.** Phylogeography and population structure of the Japanese wild boar *Sus scrofa leucomystax*: mitochondrial DNA variation. *Zoological Science* **20**: 1477–1489.
- Whittaker RJ, Fernández-Palacios JM. 2007.** *Island biogeography: ecology, evolution, and conservation*. Oxford University Press.
- Wozencraft WC. 2005.** Order Carnivora. Ed by Wilson DE, Reeder DM, In: *Mammal Species of the World. A Taxonomic and Geographic Reference*. The Johns Hopkins Univ Press, Baltimore, MD, 532-628.
- Wu J, Kohno N, Mano S, Fukumoto Y, Tanabe H, Hasegawa M, Yonezawa T. 2015.** Phylogeographic and demographic analysis of the Asian black bear *Ursus Thibetanus* based on mitochondrial DNA. *PLOSONe* **10**: 1–19.
- Yalden D. W. 1982.** When did the mammal fauna of the British Isles arrive? *Mammal Review* **12**: 1-57
- Yasukochi Y, Nishida S, Han SH, Kurosaki T, Yoneda M, Koike H. 2009.** Genetic structure of the Asiatic black bear in Japan using mitochondrial DNA analysis. *J. Heredity* **100**: 297–308.
- Yu L, Peng D, Liu J, Luan P, Liang L, Lee H, Lee M, Ryder OA, Zhang Y. 2011.**

On the phylogeny of Mustelidae subfamilies: analysis of seventeen nuclear non-coding loci and mitochondrial complete genomes. *BMC Evolutionary Biology* **11**:92.

List of Figures

Figure I-1. Sampling locations for the Japanese weasel (*Mustela itatsi*), on the Japanese islands (A) and the Siberian weasel (*M. sibirica*) on the continent (B); sample sizes were n=26 and n=21, respectively.

(A)



(B)

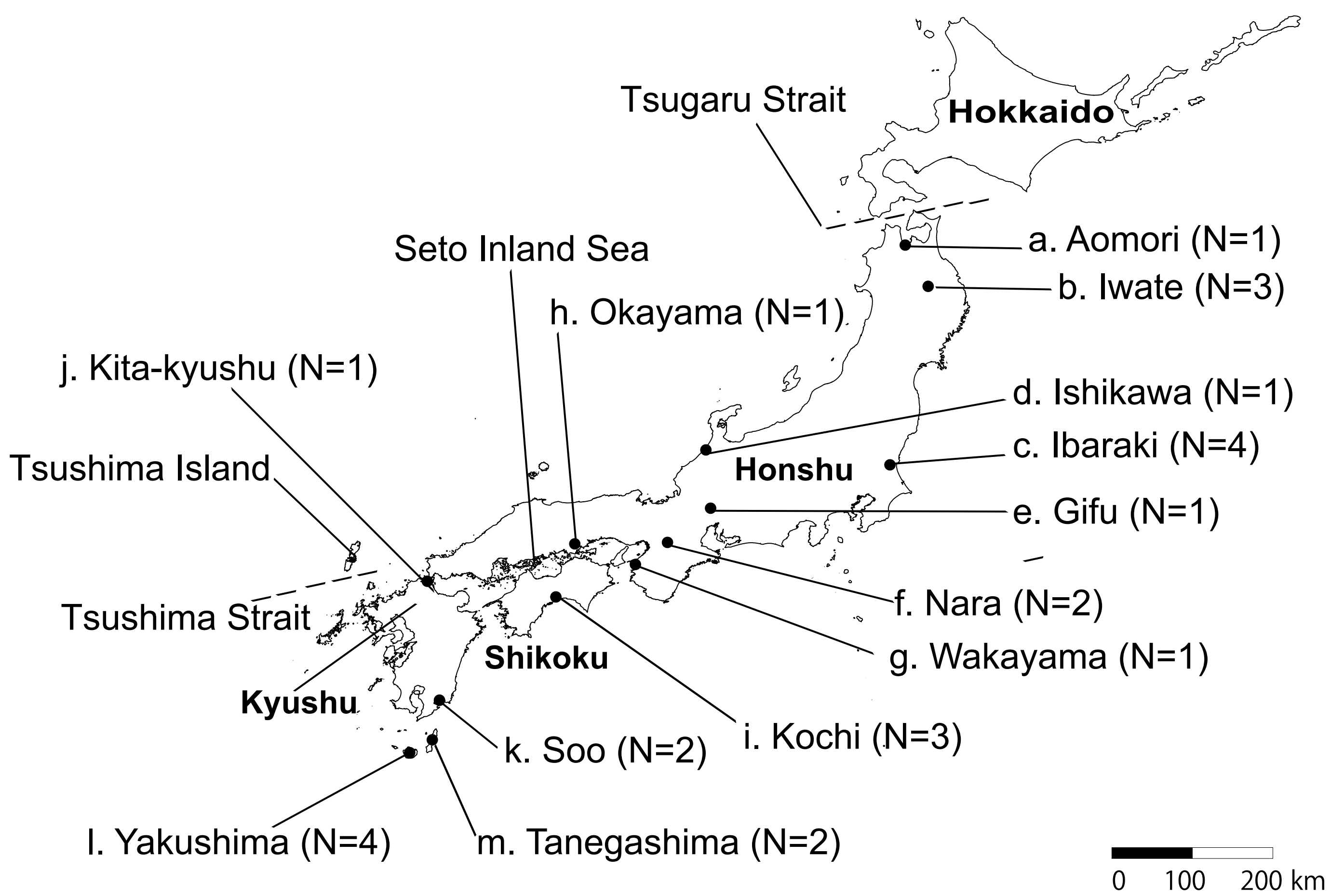


Figure I-1

Figure I-2. Bayesian inference (BI) tree for complete mtDNA sequences from *Mustela itatsi* and *M. sibirica*. Terminal taxa are indicated by haplotype name (this study) or previously reported accession number (NC_020637.1). Complete mtDNA sequences of the mountain weasel *M. altaica* (accession no. NC_021751.1), the least weasel *M. nivalis* (Accession no. NC_020639.1) and the ermine *M. erminea* (Accession no. NC_025516.1) were included as outgroup taxa. The topology of the maximum likelihood tree (ML) was identical to that of the BI tree. Nodal support is indicated by the bootstrap value (ML) followed by the posterior probability (BI). Major mtDNA clades and subclades and their corresponding geographical locations are indicated. Letters to the right of haplotype names correspond to sampling localities shown in Fig. 1 and Table 1.

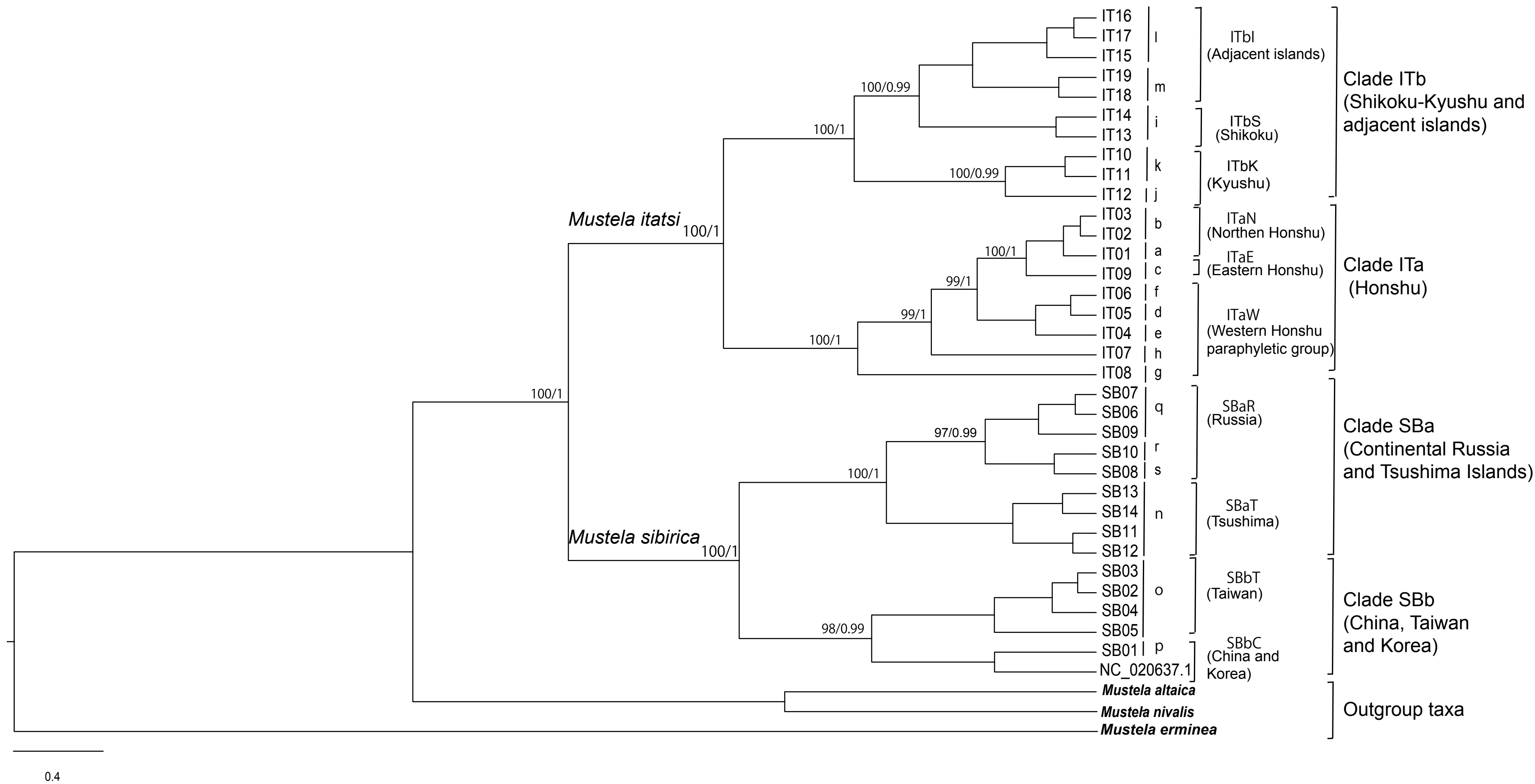


Figure I-2

Figure I-3. Maximum clade credibility tree from a BEAST Bayesian analysis of complete mtDNA sequences from *Mustela itatsi* and *M. sibirica*. Major mtDNA lineages and their geographic locations are labeled. Letters to the right of haplotype names correspond to sampling localities shown in Fig.I-1 and Table I-3. Numbers above nodes A to J are mean ages in Mya; numbers below the nodes are posterior probabilities. Node bars represent the 95% highest posterior density (HPD) of nodal age estimates. Terminal taxa correspond to those in Fig. I-2. Detailed information on nodal ages is given in Table I-5.

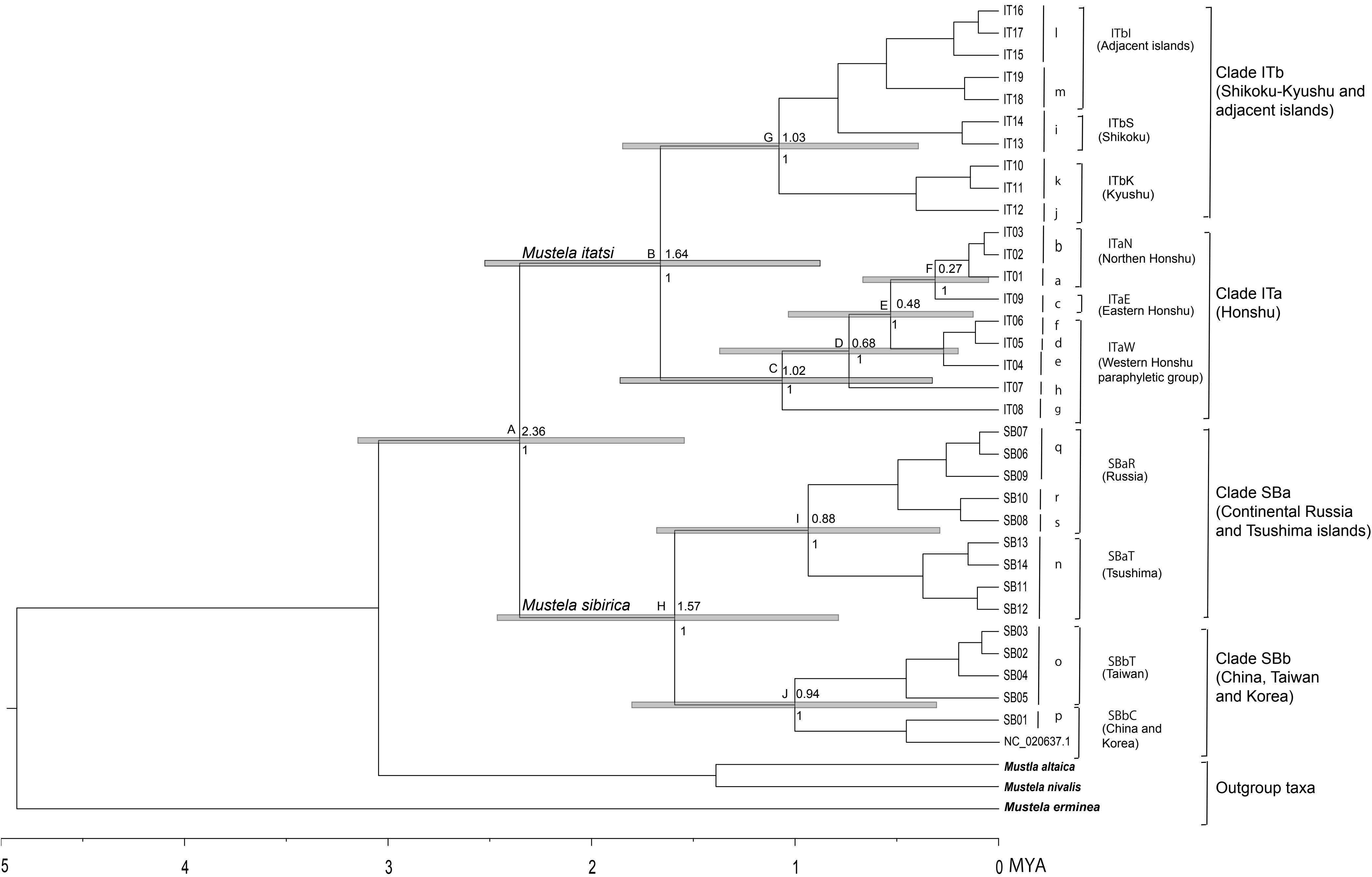


Figure I-3

Figure I-4 Maximum clade credibility tree from a BEAST Bayesian analysis of complete mtDNA sequences from *Mustela itatsi* and *M. Sibirica*, following method (i) of calibration (see, Material and Methods). All labels are the same as Fig. **I-3**, but only the calibration method and error bars are different, as indicated in the text and Table **I-5**.

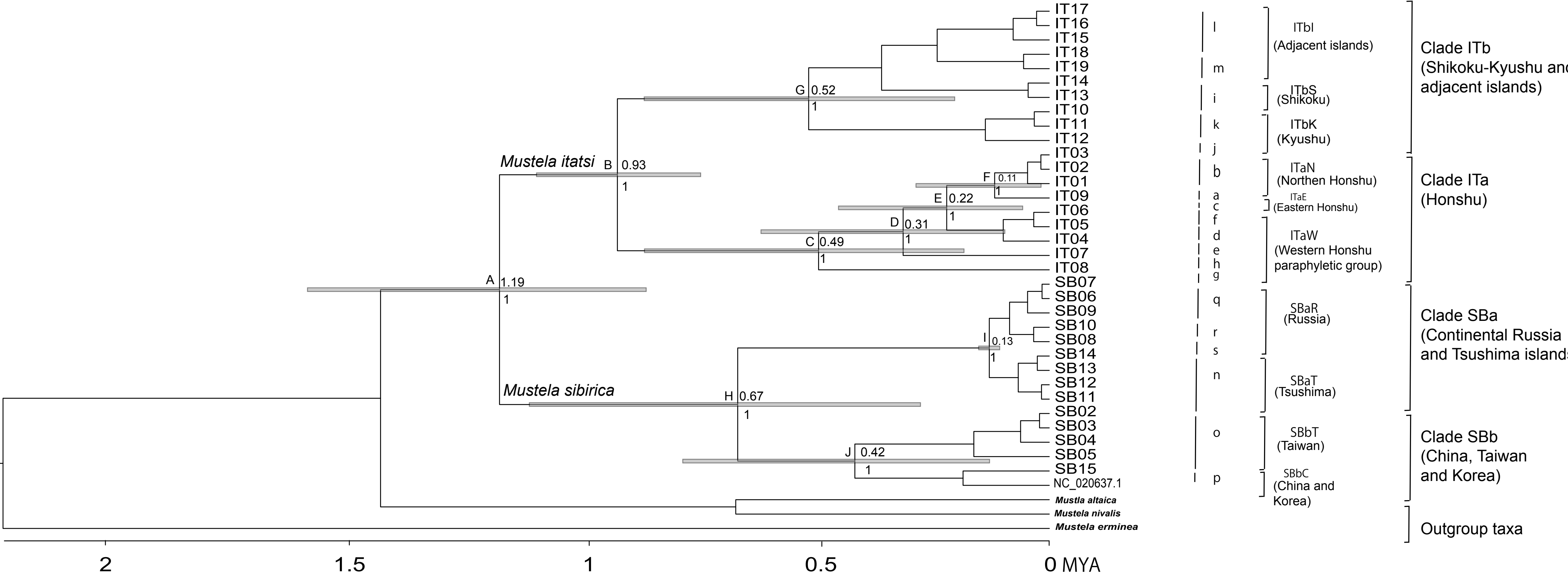
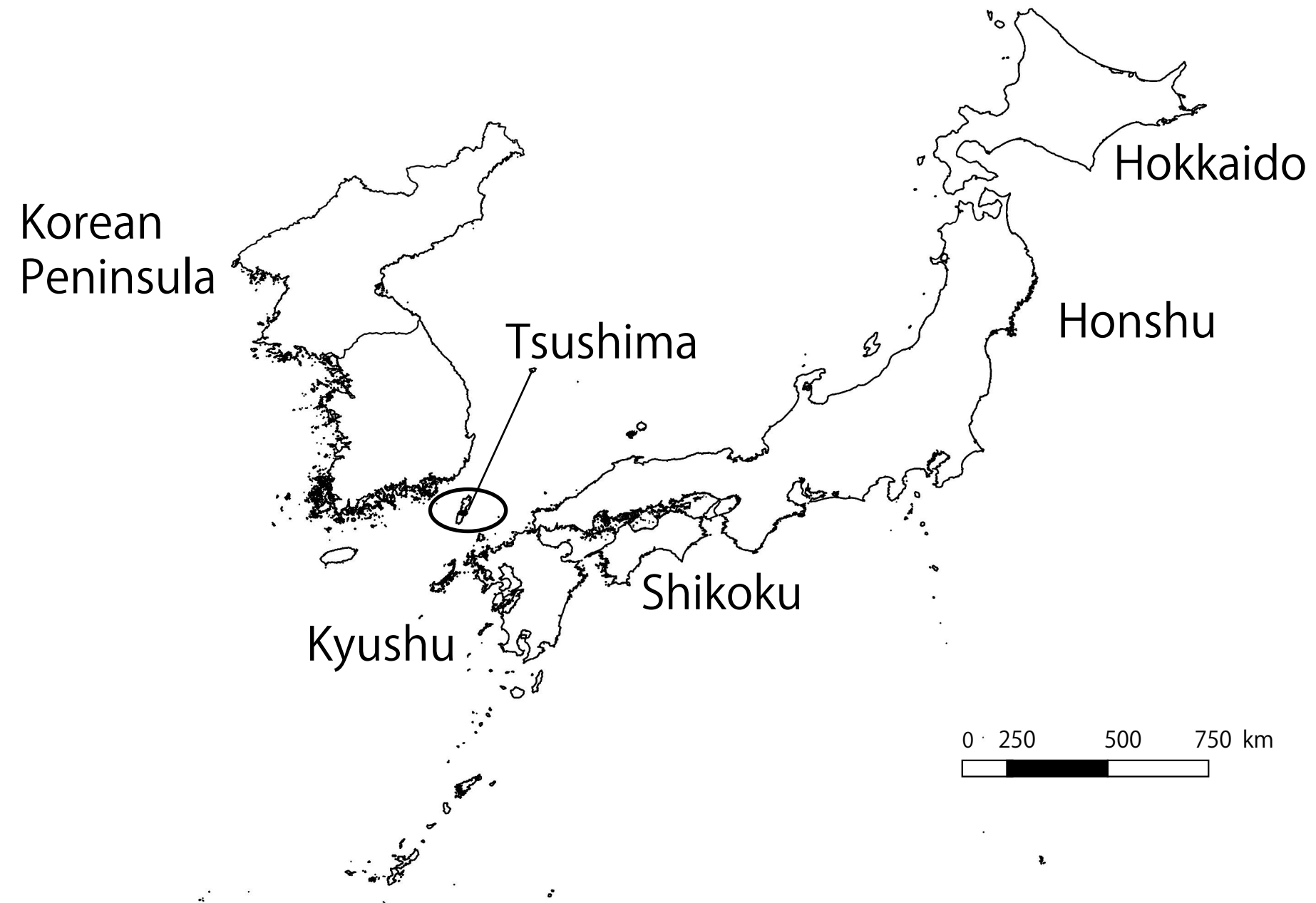


Figure I-4

Figure II I - 1 (A) Geographical location of Tsushima Islands between the Japanese islands and Korean Peninsula. (B) Sampling locations on Tsushima Islands.

(A)



(B)

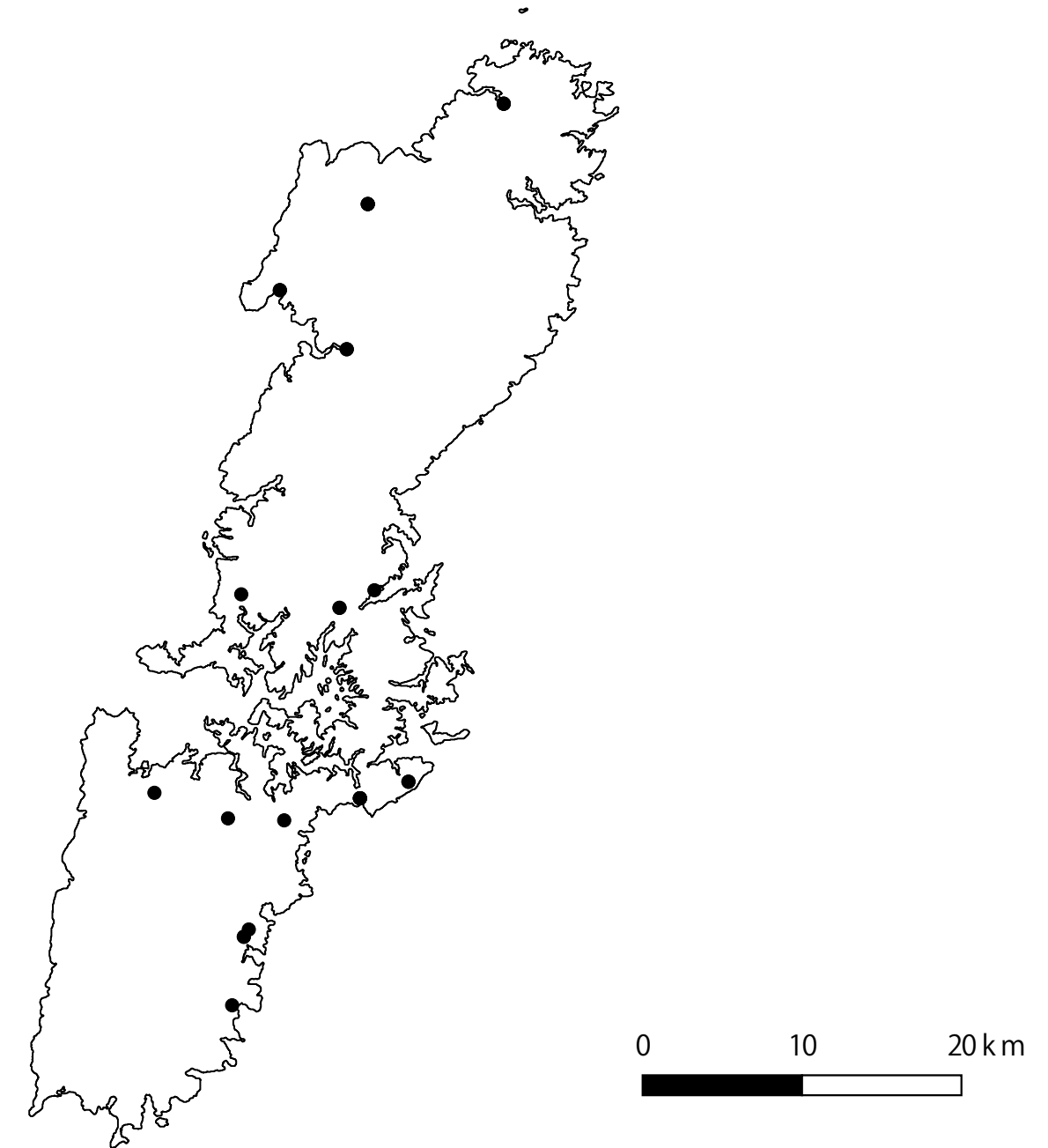


Figure III-1

Figure III-2. Schematic diagram showing the location of the weasel VNTRs in the control region of mtDNA, regions of the D-loop showing high polymorphism in this study were indicated in black color, other regions were indicated in dark-grey color.

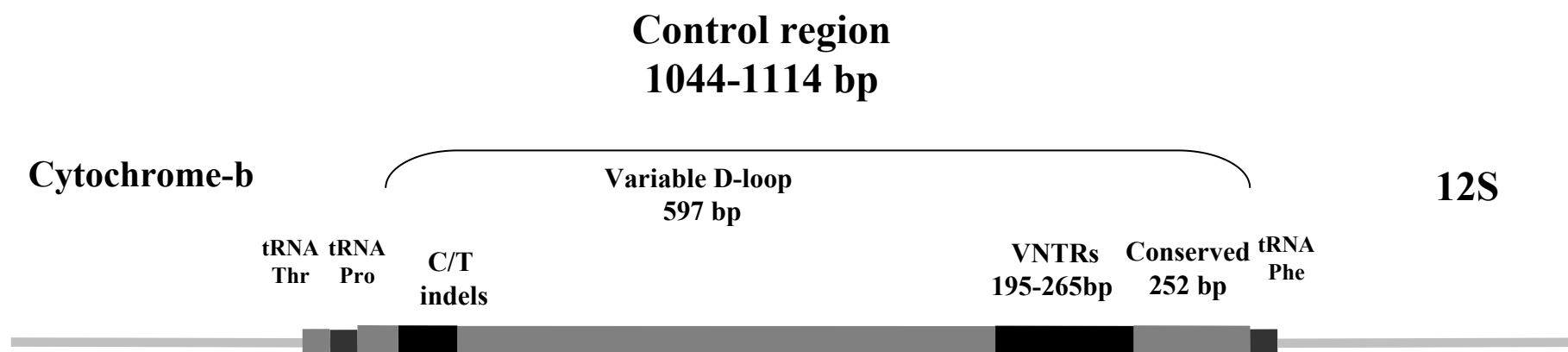


Figure III-2

List of Tables

Table I-1. List of haplotypes, numbers of individuals, sampling localities, and database accession numbers for the Japanese weasel, *Mustela itatsi* (IT), and the Siberian weasel, *M. sibirica* (SB).

Haplotype	No. of individuals	Sampling locality*	DDBJ/NCBI/EMBLAc cession no.
IT01	1	a. Aomori	AP017387
IT02	2	b. Iwate	AP017400
IT03	1	b. Iwate	AP017401
IT04	1	e. Gifu	AP017388
IT05	1	d. Ishikawa	AP017402
IT06	2	f. Nara	AP017403
IT07	1	h. Okayama	AP017404
IT08	1	g. Wakayama	AP017405
IT09	4	c. Ibaraki	AP017389
IT10	1	k. Soo	AP017390
IT11	1	k. Soo	AP017406
IT12	1	j. Kita-kyushu	AP017407
IT13	2	i. Kochi	AP017391
IT14	1	i. Kochi	AP017408
IT15	1	l. Yakushima	AP017409
IT16	2	l. Yakushima	AP017410
IT17	1	l. Yakushima	AP017411
IT18	1	m. Tanegashima	AP017392
IT19	1	m. Tanegashima	AP017412
SB01	1	p. Korea	AP017393
SB02	1	o. Taiwan	AP017413
SB03	1	o. Taiwan	AP017414
SB04	2	o. Taiwan	AP017415
SB05	1	o. Taiwan	AP017416
SB06	1	q. Transbaikalia, Russia	AP017394
SB07	1	q. Transbaikalia, Russia	AP017417
SB08	2	r. Ekaterinburg, Russia	AP017395
SB09	1	q. Transbaikalia, Russia	AP017418
SB10	5	s. Chelyabinsk, Russia	AP017396
SB11	1	n. Tsushima, Japan	AP017397
SB12	1	n. Tsushima, Japan	AP017419
SB13	1	n. Tsushima, Japan	AP017420
SB14	1	n. Tsushima, Japan	AP017421

*Letters for localities correspond to those in Fig. I-1.

Table I-2. List of oligonucleotide primers used to amplify and sequence complete mtDNA genome sequences for the Japanese weasel *Mustela itatsi* and the Siberian weasel *M. sibirica*.

Fragment no.	Forward primer name	Forward primer sequence (5'-3')	Reverse primer name	Reverse primer sequence (5'-3')	Amplicon size (bp)	Reference
1	MU1F	GCAAGGCACTGAAAATGCC	MU6R	TCTTCTGGGTGTAAGCCAGATGC	1,000	This study
2	Mt1F	GAATAGGGCCATGAAGCACG	MuW1c-1R	CCTTTCGTACTGGGAGAAGT	1,700	Knaus et al. (2011)
3	MUS1F	GACAAACCAGTCGAAGCGTC	MUS1R	ATGTTGGCGTATTCGGCTAG	1,140	This study
4	MU2F	ATGAATCCCCTACCAATG	MUS9R	GTATAGATGCAGTGGCTTGG	1,100	This study
5	MUW2-2F	CCCGTACTAATTAAACCCCC	MUW2c-2R	CTCTGTGGTGAAATATCACG	1,300	This study
6	MUS2F	GACCAAGGACCTTCAAAGCC	MUS2R	CCAAACCCTGGGAGAAT	1,100	This study
7	MUW1-2R	GCAGGAACTGGATGAACTGT	Mt2R	TACTTCTCGTTTGGATGCGA	1,000	This study
8	MUW1-3F	ACATTGTCCTTCACGATACG	MUS3R	TTGTAGGGGTAGTGAATGAGG	1,500	This study
9	MUW3-3F	ATGCTATCCCAGGACGCCT	MUW2-3R	CCGTCTGAGATTGTGAATGG	1,600	This study
10	MU4F	CTAGCCTCCGGAGTCTCTAT	MUS4R	AGTGATTTGCCAGGATGGT	1,000	This study
11	MUW1-4F	ATGAATGTGGTTTTGATCCC	MUW2c-4R	TCGCATCATTCCATATCCTC	1,300	This study
12	MUW3-4F	CTAGTGGGCTCCTTACCCCT	MUW2-4R	TTGTGTCTTGTCCAAGGTG	1,300	This study
13	MU5F	GAACTGCTAACTCATGCTCCC	MUS5R	ATGGAAGTGGCAACGAGGGT	1,300	This study
14	MUW2-5F	GCATTACCATTCACCACAAC	MUW1c-5R	CTGTGGCTGTGTCTGATGTA	1,500	This study
15	MUS7F	CACAACCAATAGCCCCACTA	MUS7R	CAGCTTTGGGTGCTGATG	1,400	This study
16	MUW1-6F	GCCAACCAGTAGAACACCCA	MUW1c-6R	CCCAGTTTGGGTCTCAGCTA	1,900	This study

Table I-3. Internal primers used to sequence the complete mtDNA genome in this study.

Primer name	Direction	Sequence (5' -3')	Reference
2ndW1F	Forward	CCACAGGGATACACTCATCT	This study
MUW3-1F	Forward	CACCAATTAAGAAAGCGTTC	This study
MU1R	Reverse	CTTCGTAGGAGATTGTTTGGGC	This study
MUW1-2F	Forward	CGATTCCGCTATGACCAGCT	This study
7thW2F	Forward	GAAGTAACTCAGGGAGTCCC	This study
MUW3-2F	Forward	CAATCGCACACATAGGATGA	This study
MUW1c-2R	Reverse	GGTCTACAGATGCTCCAGCA	This study
Mt2F	Forward	GACACACGAGCATATTTAC	Knaus et al. (2011)
3rdW1R	Reverse	CGTATCGTGAAGGACAATGT	This study
MUS3W1-F	Forward	CCTAAGACTGGTTTCAAGCCA	This study
MUW1c-3R	Reverse	TGGTGGCCTTGAAATGTTCC	This study
MUW2-4F	Forward	CCTGCGAAGCAGCATTAGG	This study
8thW2F	Forward	GCCTCATTCTCATGATCCAAC	This study
MUS5-W2F	Forward	GACCCCAACATCAATCGATTT	This study
MUW1-5F	Forward	GCCATCCTATATAACCGCA	This study
MUS6W2F	Forward	CCTTACAATGCAAAGCCTA	This study
MUW3-5F	Forward	CTACC AAAATCCATCTCCCA	This study
MU6F	Forward	GCAACTGCATTCATAGGTTACG	This study
MUW26F	Forward	CCTCTCAAATGGGACATCTC	This study
MUTANCF	Forward	CCTTCATCATTTATCCGCAT	This study
DTAN2R	Reverse	TAAGGGGGGTTTGACAAAGG	This study
DTAN1R	Reverse	ATGGAGTCTTGTGACTCTTC	This study

Table I-4. Genetic diversity indices for the Japanese weasel (*Mustela itatsi*) and Siberian weasel (*M. sibirica*). Asterisks indicate significant values.

Species	Clade	Sampl e size	No. of haplotypes	Haplotype diversity (SD)	No. of polymorphic sites	No. of parsimoniously informative sites	Nucleotide diversity	Tajima' s <i>D</i>	Tajima' s <i>P</i>	Fu' s <i>F_s</i>
<i>M. itatsi</i>	All	26	19	0.969 (0.021)	245	230	0.00619	2.12693	$P < 0.05^*$	10.11
	Honshu (ITa)	14	9	0.912 (0.059)	58	35	0.00115	1.78104	$0.10 > P > 0.05$	2.742
	Shikoku-Kyushu (ITb)	12	10	0.970 (0.044)	83	72	0.00205	1.04005	$P > 0.10$	8.354
<i>M. sibirica</i>	All	20	14	0.937 (0.043)	213	190	0.00519	2.17369	$P < 0.05^*$	6.676
	Russia, Tsushima (SBa)	14	9	0.879 (0.079)	42	39	0.00099	1.91616	$0.10 > P > 0.05$	3.263
	Korea, Taiwan clade (SBb)	6	5	0.933 (0.122)	61	9	0.00140	1.77661	$0.10 > P > 0.05$	0.612

Table I-5. Bayesian age estimates (millions of years ago, Mya) for time to the most recent common ancestor (TMRCA) for clades and subclades of the Japanese weasel (*Mustela itatsi*) and Siberian weasel (*M. sibirica*), based on the complete mtDNA genome. Nodes refer to those in Fig. 3.

Node MRCA	Nodal age (95% HPD)
A Between the Japanese weasel (IT) and the Siberian weasel (SB)	2.36 (1.54 – 3.14)
B Between the Honshu clade (ITa) and Shikoku-Kyushu clade (ITb)	1.64 (0.87 – 2.52)
C Within Honshu clade (ITa)	1.02 (0.32 – 1.85)
D Between haplotypes in the Honshu clade, excluding haplotype IT08 from Wakayama	0.68 (0.19 – 1.37)
E Between Haplotypes of Honshu clade excluding haplotypes IT07 of Okayama and IT08 of Wakayama	0.48 (0.12 – 1.03)
F Between the eastern (ITaE) and northern (ITaN) subclades in the Honshu clade	0.27 (0.049 – 0.66)
G Between the Kyushu (ITbK) and Shikoku – adjacent islands subclades (ITbS + ItBI)	1.03 (0.39 – 1.84)
H Between the Continental Russia-Tsushima clade (SBa) and TheChina-Korea-Taiwan clade (SBb)	1.57 (0.786 – 2.46)
I Between continental Russia (SBaR) and Tsushima (SBaT)	0.88 (0.28 – 1.67)
J Between the Taiwan (SBbT) and Korea-China (SBbC) subclades	0.94 (0.3 – 1.8)

Table I-6. Divergence times (millions of years ago, Mya) between Japanese and continental sister species or populations, calculated in previous studies.

Japanese/continental species or population pairs	Mitochondrial gene locus	Divergence time MYA (95% HPD)	Reference
Japanese squirrel (<i>Sciurus lis</i>)/red squirrel (<i>S.vulgaris</i>)	Cytochrome <i>b</i>	2.24 (0.16 – 8.45)	Oshida & Masuda (2000); Oshida <i>et al.</i> (2009); McKay (2012)
Japanese grass vole (<i>Microtus montebelli</i>)/tundra vole (<i>M. oeconomus</i>)	Cytochrome <i>b</i>	1.31 (0.08 – 5.61)	Conroy & Cook (2000); McKay (2012)
Japanese water shrew (<i>Chimarrogale platycephalaus</i>)/elegant water shrew (<i>C. himalyica</i>)	Cytochrome <i>b</i>	2.4 (0.31 – 4.15)	Ohdachi <i>et al.</i> (2006); McKay (2012)
Shinto shrew (<i>Sorex shinto</i>)/Laxmann’ s shrew (<i>S. caecutiens</i>)	Cytochrome <i>b</i>	2.28 (0.54 – 6.07)	Ohdachi <i>et al.</i> (2001, 2006); McKay (2012)
Japanese/continental populations of Asiatic black bear (<i>Ursus thibetanus</i>)	Control region	2.09 (0.19 – 4.08)	Yasukochi <i>et al.</i> (2009); McKay (2012)
Japanese/continental populations of <i>U. thibetanus</i>	Complete mtDNA	1.46	Wu <i>et al.</i> (2015)
Japanese otter (<i>Lutra nippon</i>)/common otter (<i>Lutra lutra</i>)	Cytochrome <i>b</i>	2.49 (0.00 – 4.87)	Suzuki <i>et al.</i> (1996); McKay (2012)
Japanese badger (<i>Meles anakuma</i>)/Asian badger (<i>M. leucurus</i>)	Cytochrome <i>b</i>	1.4 (0.16 – 2.86)	Kurose <i>et al.</i> (2001); McKay, (2012)
Japanese marten (<i>Martes melampus</i>)/sable (<i>M. zibellina</i>)	Cytochrome <i>b</i> , ND2, control region	1.6 – 1.8	Sato <i>et al.</i> (2009); Sato (2013)
Japanese lesser flying squirrel (<i>Pteromys momonga</i>)/ Russian flying squirrel (<i>P. volans</i>)	Cytochrome <i>b</i>	8.69 (0.67 – 16.78)	Oshida <i>et al.</i> (2000); McKay, (2012)
Dsinezumi shrew (<i>Crocidura dsinezumi</i>)/ Ussuri white-toothed shrew (<i>C. lasiura</i>)	Cytochrome <i>b</i>	0.5 (0.00 – 1.12)	Ohdachi <i>et al.</i> (2004); McKay, (2012)

Table II-1. The parsimony informative site Pi, length in base pairs and their ratio for each mtDNA genome protein coding genes, rRNA subunits and the control region for each of the Japanese weasel (*Mustela itatsi*) and the Siberian weasel (*M. sibirica*).

Gene	Length (bp)	<i>M. itatsi</i>	<i>M. sibirica</i>
		Pi/length	Pi/length
ND1	955	0.019	0.010
ND2	1044	0.022	0.020
CO1	1545	0.015	0.007
CO2	684	0.010	0.010
ATP8	204	0.010	0.024
ATP6	681	0.009	0.019
CO3	784	0.024	0.010
ND3	347	0.020	0.008
ND4L	271	0.011	0.022
ND4	1378	0.015	0.024
ND5	1830	0.013	0.010
Cytochrome <i>b</i>	1140	0.017	0.017
control region	591	0.045	0.024
12S rRNA	960	0.006	0.005
16S rRNA	1572	0.009	0.003

Table III-1. Patterns of both the C/T indel sites within the 5' end 600 bp and the VNTRs reported in the mtDNA control region of the Siberian weasel *Mustela sibirica*.

No Sample ID								VNTRs	Locality
								total	(Island)
		C/T indels		VNTRs			Pattern	length	
				Units				(bp)	
				k3h	2h	2*d1r			
1	MSI-TS4	-----TCCTCCCCCCCCCTCTTCCCCC	1	18	1	2		215	North
2	MSI-TS5	-----TCCTCCCCCCCCC---TTCCCCC	1	18	1	2		215	North
3	MSI-TS7	----TCCTCCCCCCCCCCCCCTTCCCCC	1	18	1	2		215	South
4	MSI-TS8	TCTTTTCCCCCCCCCTCTTCCCCCCC	1	18	1	2		215	South
5	MSI-TS9	-----TCCCCCCCCCCCCCTTCCCCC	1	21	1	5		245	South
6	MSI-TS10	-----CCCCCCCCCCCCCTTCCCCC	1	20	1	4		235	North
7	MSI-TS11	-----TCCTCCCCCCCCC--CTTCCCCC	1	21	1	5		245	North
8	MSI-TS12	T----CCCCCCCCCCCCCTCTTCCCCC	1	21	1	5		245	North
9	MSI-TS13	-----CCCCCCCCCCCCCTTCCCCC	1	18	1	2		215	North
10	MSI-TS14	-----TCCCCCCCCCCCC---TTCCCCC	1	18	1	2		215	North
11	MSI-TS15	-----TCCTCCCCCCCCC-CCTTCCCCC	1	21	1	5		245	North
12	MSI-TS16	-----TCCTCCCCCCCCCTCTTCCCCC	1	21	1	5		245	Unknown
13	MSI-TS17	-----TCCTCCCCCCCCCTCTTCCCCC	1	21	1	5		245	North
14	MSI-TS18	-----TCCTCCCCCCCCC--CTTCCCCC	1	19	1	3		225	North
15	MSI-TS19	-----TCCTCCCCCCCCC--CTTCCCCC	1	23	1	7		265	North
16	MSI-TS20	-----TCCTCCCCCCCCC---TTCCCCC	1	19	1	3		225	North
17	MSI-CTR1	-----TCCTCCCCCCCCC--CTTCCCCC	1	22	1	6		255	Unknown
18	MSI-CTR2	-----TCCTCCCCCCCCC-TCTTCCCCC	1	20	1	4		235	Unknown
19	TSW09-01	-----TCCTCCCCCCCCC---TTCCCCC	1	16	1	1		195	South
20	TSW09-02	-----TCCTCCCCCCCCC---TTCCCCC	1	19	1	3		225	North
21	TSW09-03	-----TCCTCCCCCCCCC--CTTCCCCC	1	21	1	5		245	South
22	TSW09-04	TC---CCCCCCCCCCCCCTCTTCCCCC	1	22	1	6		255	South
23	TSW09-05	-----TCCTCCCCCCCCC---TTCCCCC	1	20	1	4		235	North
24	TSW10-01	-----TCCTCCCCCCCCC--CTTCCCCC	1	21	1	5		245	South
25	TSW10-02	-----TCCCCCCCCCCCCCTTCCCCC	1	21	1	5		245	South
26	TSW10-03	-----TCTCCCCCCCCCCCCCTCCCCC	1	21	1	5		245	North
27	TSW10-04	-----TCTTCCCCCCCCC---TTCCCCC	1	19	1	3		225	South
28	TSW10-05	-----TCCTCCCCCCCCC-CCTTCCCCC	1	18	1	2		215	South
29	TSW10-06	-----TCCCCCCCCCCCCCTCTTCCCCC	1	22	1	6		255	North
30	TSW10-07	----TCCTCCCCCCCCC-CCTTCCCCC	1	19	1	3		225	South
31	TSW10-08	-----TCCTCCCCCCCCCTCTTCCCCC	1	21	1	5		245	North