Discoveries

Whole-genome sequencing revealed local genetic differences and multi-layered genetic structure in brown bear (*Ursus arctos*) populations

Yu Endo[1, *,§], Naoki Osada[2], Tsutomu Mano[3], Alexei V. Abramov[4], Ryuichi Masuda[5]

[1]Department of Natural History Sciences, Graduate School of Science, Hokkaido University, Sapporo, Japan

[2]Faculty of Information Science and Technology, Hokkaido University, Sapporo, Japan

[3]Research Institute of Energy, Environment and Geology, Hokkaido Research Organization, Sapporo, Japan

[4]Laboratory of Mammalogy, Zoological Institute, Russian Academy of Science, Saint Petersburg, Russia

[5]Department of Biological Sciences, Faculty of Science, Hokkaido University, Sapporo, Japan

*Corresponding author email: yuendo0822@gmail.com

§Present address: Ecological Genetics Laboratory, National Institute of Genetics, Mishima, Shizuoka, Japan

Abstract

A multi-layered genetic structure reflecting multiple demographic events across different time periods has been well documented particularly in modern humans. However, it is still unknown in wild species because of a lack of comprehensive genome-wide data on a global scale. The brown bear, *Ursus arctos*, exhibits a clear discrepancy between mitogenome- and nuclear genome-based phylogenetic trees. A previous study suggested that this discrepancy was caused by incomplete lineage sorting of the mitochondrial DNA lineages or recent expansion erasing the former genetic structure of the nuclear genome; however, neither scenario fully explained the discrepancy because genetic variation was observed across their habitat at different times. We performed whole-genome resequencing on nine brown bears from local populations around or on the Eurasian continent. The ancestral genetic characteristics have persisted in Western Asia and Central Asia, particularly in Kazakhstan and Tibet where minor mitochondrial haplotypes have been reported, whereas individuals from these areas shared few alleles with individuals speculated to possess many alleles from recent dispersal, which suggested that the recent ancestors of these populations had not experienced complete isolation from other populations. The Hokkaido and Etorofu (Iturup) Island populations obtained many alleles via gene flow from the polar bear *Ursus maritimus*; this finding was similar to that of a North American brown bear population previously reported to have hybridized with polar bears. This phenomenon may be attributed to less influence of recent expansion on island populations compared with that of continental populations. These results support our hypothesis and indicate that brown bears have a multi-layered genetic structure influenced by migration events at different times.

Introduction

Multi-layered genetic structure is an idea that modern genetic characteristic reflected multi demographic events in different times. This idea has been reported well in human genetic structure that the human genome is composed of multiple genetic traits from multiple migrations at different times (e.g., Tassi et al. 2015). For example, genome-wide genetic data revealed that the current Japanese population consists of two main genetic layers from Jomon hunter-gatherers and rice-farming Yayoi migrants, and more genetic layers from different migration events (Osada & Kawai 2021). Although the genetic structure of other species, such as the black rat, *Rattus rattus* (Puckett et al. 2020), the spruce, *Picea abies* and *P. obovata* (Zhou et al. 2024), and the Japanese wolf, *Canis lupus hodophilax* (Segawa et al. 2022), was also proposed to be multi-layered and reflect population demography at different times, the presence of a multi-layered genetic structure in wild species that distribute widely remains largely unknown because there is a lack of global genome-wide data and their limited distribution.

The brown bear, *Ursus arctos,* is a large mammal in Carnivora which distributes widely on the Northern Hemisphere and well known their global scale genetic structure and hybridization with related species by many previous studies. At the inter-species level, mitochondrial DNA (mtDNA)-based phylogenetic trees placed the polar bear, *Ursus maritimus*, within a brown bear lineage, clade 2, that is nowadays distributed in the Admiralty, Baranof, and Chichagof (ABC) Islands (Leonard et al. 2000). In contrast, phylogenetic trees estimated from the nuclear genome clearly showed their differentiation (Hailer et al. 2012; Miller et al. 2012). At the intra-species level, multiple mtDNA lineages with differing estimated divergence times are reported (Davison et al. 2011; Hirata et al. 2013; Lan et al. 2017), whereas the nuclear genome

did not exhibit a clear differentiation between those lineages (Bidon et al. 2014; Endo et al. 2021; de Jong et al. 2023). These discrepancy in the phylogenetic relationships has been attributed to many factors, including incomplete lineage sorting (Kutschera et al. 2014), male-biased dispersal (Bidon et al. 2014; Hirata et al. 2017), and hybridization with related species (Cahill et al. 2013, 2015, 2018; Lan et al. 2022; Liu et al. 2014; Miller et al. 2012; Wang et al. 2022), including the extinct cave bear (Barlow et al. 2018). Therefore, various factors and the complex relationships among them make it challenging to establish a clear and conclusive explanation.

At the intra-species levels, de Jong et al. (2023) revealed that admixture between populations led to erasure of the former genetic structure of the nuclear genome, and they proposed two possible hypotheses: *i*) the mtDNA haplotypes distribution was due to incomplete lineage sorting, or *ii*) the autosomal genome reflects recent population demography shaped by male-mediated gene flow. Although those two hypotheses can each partially explain the population demography, neither can fully explain the incongruences between brown bear mtDNA and nuclear genetic structure. Hypothesis *i* was suggested based on their multiple coalescent analysis result of the Hokkaido and Scandinavian populations, which represent only a portion of brown bear populations. The brown bear is widely distributed in the northern hemisphere and exhibits a genetic structure consisting of multiple mtDNA lineages (Davison et al. 2011; Hirata et al. 2013; Lan et al. 2017) with different divergence times (Figure 1). The appearance and divergence times of lineages have been consistently confirmed by many previous studies, including those that used ancient sample data from different areas (da Silva Coelho et al. 2023; Edwards et al. 2011; Kosintsev et al. 2022; Molodtseva et al. 2022; Segawa et al., 2021). Thus, it is highly possible that such a genetic structure was not only formed by incomplete lineage

sorting, and some local populations that have the mtDNA lineages with ancient divergence times might have retained ancestral genetic characteristics of the nuclear genome. Hypothesis *ii* cannot explain the genetic distinctiveness of brown bear subspecies in Himalaya, Mongolia, and Pakistan (de Jong et al. 2023; Tumendemberel et al. 2023), where an old divergent mtDNA haplotype was identified (Lan et al. 2017; Tumendemberel et al. 2019). Genetic distinctiveness of the nuclear genome might reflect both recent expansion, which mainly occurred by male-biased dispersal, and local differences before the expansion. Therefore, population genetic analysis is needed using samples from across this habitat, including ancient samples, to determine: *i*) the correlation between mtDNA lineage divergence time and the genetic distinctiveness of the nuclear genome, and *ii*) the differences in the influence of recent expansion, which was driven by male-biased dispersal, between populations.

In this study, we performed whole-genome resequencing on nine brown bears (Hirata et al. 2013, 2014) from the local Asian populations (Etorofu (Iturup) Island, Sakhalin, Northern Okhotsk, Kazakhstan, Tibet, Central Russia, and West Asia; Figure 1 and Supplementary table 1). These populations exhibit distinct genetic characteristics from those on the Eurasian and North American continents based on mtDNA analyses (Hirata et al. 2013, 2014); therefore, this difference might also exist in the nuclear genome. To infer local differences in phylogeny, demography, and gene flow, we conducted population genomic analysis using the data we produced and published modern genome data (Barlow et al. 2018; Benazzo et al. 2017; Cahill et al. 2013; Cahill et al. 2015; Endo et al. 2021; Kumar et al. 2017; Liu et al. 2014; Miller et al. 2012; Taylor et al. 2018; Wang et al. 2022) and ancient genome data (Barlow et al. 2018; Cahill et al. 2018; Lan et al. 2022; Segawa et al. 2021; Wang et al. 2022). Based on the results obtained,

this study verifies the correlation between genetic distance of mtDNA and that of nuclear genome, and recent dispersal produced significant discrepancy between the phylogenic trees at the intra-species level.

Materials and methods

Sampling and genome-sequencing

Nine skin or muscle samples were used that were collected by the Zoological Institute, Russian Academy of Sciences (ZIN, Saint Petersburg, Russia; Figure 1 and Supplementary table 1) from the local populations distributed in the Eurasian continent and islands: Etorofu (Iturup) Island, Sakhalin, Northern Okhotsk region, Kazakhstan, Tibet, Central Russia, and West Asia, across brown bear habitat (McLellan et al. 2017). From the samples, we extracted the total DNA using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and stored it at 4°C or −20°C until used.

Then, 150-bp pair-end whole-genome resequencing of these samples was conducted using an Illumina Novaseq 6000 sequencer by Macrogen Japan (Tokyo, Japan). The libraries were prepared using the TruSeq DNA Nano Sample Prep Kit (Illumina, San Diego, USA). The average of the total amount of sequence data per sample was ordered 90 Gb.

Mapping and genotyping

Whole-genome sequence data from the nine brown bears and 63 previously sequenced bears (Miller et al. 2012; Cahill et al. 2013, 2015, 2018; Benazzo et al. 2017; Kumar et al. 2017; Barlow et al. 2018; Taylor et al. 2018; Endo et al. 2021; Segawa et al. 2021; Lan et al. 2022;

Wang et al. 2022) were checked for sequence quality, and poor-quality sequences were removed using FastQC ver. 0.11.8 (Babraham Bioinformatics 2011) and Trimmomatic ver.0.39 (Bolger et al. 2014). We mapped raw sequence data to the reference sequence of the polar bear (GCF_017311325.1) using the BWA-MEM algorithm implemented in the BWA ver. 0.7.17 (Li & Durbin 2009) with the '-M' command option.

To determine sex for individuals (excluding ancient samples specified in Supplementary table 1), we calculated the depth ratio on the scaffolds of the X chromosome (n = 5; NW_024423319.1–NW_024423323.1) and the Y chromosome (n = 19; NW_024423324.1–NW_024423342.1) using the following formula. If the depth ratio was > 1, we set the sex as female.

$$Depth\ ratio = \frac{depth\ for\ X - link\ scaffolds}{depth\ for\ Y - link\ scaffolds}$$

After mapping, single nucleotide variants of modern samples (n = 61, Supplementary table 1) were called and then hard filtered under the following parameters: significant Fisher strand test (FS) > 60.0; variant confidence/quality by depth (QD) < 2.0; RMS mapping quality (MQ) < 40.0; strand odds ratio (SOR) > 9.0; MQRankSum < 12.5, and significant read position bias (ReadPosRankSum) < 8.0 in GATK ver. 4. 1. 2. 0 (McKenna et al. 2010). For single nucleotide variants on the X chromosome scaffolds, genotyping of males was performed with the option --ploidy 1, and that of females with the option --ploidy 2 in GATK. Genmap ver. 1.3.0 (Pockrandt et al. 2020) was used to calculate genome mappability and exclude SNPs with (30,

2)-mappability < 1.

A different method was used to genotype ancient samples because they had lower sequence quality compared with modern samples. Ancient samples (n = 11; Supplementary table 1) were genotyped by mpileup in samtools ver. 1.9 using the parameters -q 30 -Q 30 and pileupCaller (https://github.com/stschiff/sequenceTools) with the parameters --randomHaploid --skipTransitions after mapping to the reference sequence. Such samples are sometimes not suitable for population genetic analysis because of the small amount of sequence data, so we calculated the ratio of the number of covered bases with depth ≥ 1 to the total bases of the reference sequence using a samtools ver. 1.15 command, samtools coverage. We retained samples with a ratio of > 0.1.

EIGENSOFT ver. 7.2.1 (Patterson et al. 2006), which was used to perform PCA, does not accept data that contain more than 100 chromosomes. Therefore, for f_4 statistics and ancient DNA analysis, we used only scaffolds longer than 0.1 M bp and excluded the sex-linked scaffolds (n = 79) whose total length was approximately 99% of the total reference sequence length because most scaffolds were too short to be used for population genetic analysis.

Genetic diversity, genetic structure, phylogeny, and demography

We calculated heterozygosity and runs of homozygosity (ROH) per individual using PLINK ver. 1.9 (Purcell et al. 2007), using the parameters --het for heterozygosity, and --homozyg-kb 500 and --homozyg-kb 2000 for ROH. The genetic structure of the brown bears (n = 54) worldwide was estimated by PCA using EIGENSOFT ver. 7.2.1 (Patterson et al. 2006) and ADMIXTURE ver. 1.3.0 (Alexander et al. 2009). The relative scale of gene flow based on

genetic distance and sampling location was calculated by EEMS (Petkova et al. 2016). The number of modeled demes was set to be 400 and Markov chain Monte Carlo (MCMC) chains were each run for 20 million iterations;the first 10 million generations were discarded as burn-in and sampling occurred every 10,000th generation.

To construct phylogenetic trees based on autosomal and X-chromosomal SNPs, the identity-by-state of each pair was calculated using PLINK with the parameter (--distance square 1-ibs). We interpreted the non-sex-linked scaffolds as autosomal scaffolds (n = 3,875). After calculating the genetic distance matrix, phylogenetic trees were constructed with the neighbor-joining method (Saitou & Nei 1987) using the ape package in R (Paradis et al. 2004). We used the Asiatic black bear sample (Asiatic_black_bear; Kumar et al. 2017) as an outgroup. Pairwise sequentially Markovian coalescent analysis (PSMC; Li & Durbin 2011) was performed on all new samples and samples from the same or near sampling locations following by Endo et al. (2021).

Mitochondrial DNA and Y chromosomal DNA phylogeny

To obtain the whole mtDNA sequence, whole-genome sequence data from 37 individuals (Supplementary table 1) was assembled using getOrganelle with the parameters (-k 21,45,65,85,127). Then, 81 samples (Supplementary table 2) from the NCBI database were downloaded and added as sequence data. We extracted sequence data from two rRNA genes, 22 tRNA genes, and 12 coding genes, but we did not include NADH dehydrogenase subunit 6 sequences because of its nucleotide composition heterogeneity (Hirata et al. 2013), and gap sites were excluded. The final sequence length was 14,784 bp.

The phylogenetic tree was constructed by the maximum likelihood method using iqtree ver. 1.6.12 with the parameters (-m MFP -AICc -bb 1000 -alrt 1000) after alignment with the MUSCLE algorithm (Edgar 2004). The Asiatic black bear sample (Honshu; NCBI accession number: DRR250459) was used as an outgroup.

The divergence time for each mtDNA lineage was then estimated using BEAST ver. 2.4.3 (Bouckaert et al. 2014) under the GTR substitution model. The relaxed lognormal clock and the coalescent constant population model were used. To calibrate divergence times, radiocarbon dates for ancient sequences were set, which is a mean age of 120 thousands of years ago (kya) for the ancient polar bear (GU573488; Lindqvist et al. 2010), and 31.87 thousand years ago, 44 thousand years ago, and 409 thousand years ago for the cave bears (EU327344, NC011112, and KF437625; Bon et al. 2008, Krause et al. 2008, Dabney et al. 2013). Trees were sampled every 10,000 generations from a total of 1,000,000,000 generations; the first 100,000,000 generations were described as burn-in. The maximum credibility tree was generated using TreeAnnotator in BEAST. Other options were followed by the "Divergence Time Estimation" tutorial by Taming the BEAST (Barido-Sottani et al. 2018).

The Y-chromosome sequences were constructed from a .vcf file using bcftools ver. 1.9 (Danecek & McCarthy 2017). Phylogenetic trees based on haplotype data were constructed following the same method as that of mtDNA. We used the American black bear sample (American_black_bear; Kumar et al. 2017) as an outgroup for the Y chromosome tree.

Inter- and intra-species gene flow

TreeMix ver. 1.13 (Pickrell & Pritchard 2012) was used to show the phylogenetic relationships and gene flow between groups or individuals. After running TreeMix, we inferred the most likely number of migration events using OptM package in R (Fitak 2021).

Hybridization with related species and gene flow between brown bear populations were evaluated by f_4 statistics using Admixtools (Patterson et al. 2012). f_4(Outgroup, Polar bear; X, Kazakhstan[U13M]) and f_4(Outgroup, Cave bear; X, modern Polar bear), where X represents the ancient and modern brown bear individuals, were calculated to detect hybridization with polar bears and cave bears, respectively. An alternative evaluation to determine hybridization with related species, f_4 ratio (Patterson et al. 2012), was used to infer the mixing proportions of an admixture event. We calculated the ratio of f_4 values, α, for each related species as shown below.

$$\alpha_{Polar\ bear} = \frac{f_4(\text{Outgroup, Polar bear; X, Kazakhstan[U13M]})}{f_4(\text{Outgroup, Polar bear; West Asia, Kazakhstan[U13M]})}$$

$$\alpha_{Cave\ bear} = \frac{f_4(\text{Outgroup, Cave bear; X, modern Polar bear})}{f_4(\text{Outgroup, Cave bear; North America, modern Polar bear})}$$

The percentage of the admixture proportion was then calculated following the below formula.

$$f_4 ratio\ (\%) = (1 - \alpha) \times 100$$

We created mtDNA consensus sequences of 63 bears from the .vcf file using bcftools ver. 1.9 (Danecek et al. 2017) to detect modern brown bears whose mtDNA lineage was clade 3a1. Gene flow, driven by the recent expansion of mainly clade 3a1, was evaluated by the value of f_4(Outgroup, clade 3a1 in Europe; X, modern Brown bear excluding X) and of f_4(Outgroup, clade 3a1 in North America; X, modern Brown bear excluding X) using autosomal SNPs data. To evaluate the recent expansion, the dataset only including modern samples was used because it is important to use as many sites as possible is suitable for evaluation. We used the American black bear (American_black_bear; Kumar et al. 2017) and the Asiatic black bear (Asiatic_black_bear; Kumar et al. 2017) as outgroups.

Results

We genotyped 42,906,277 SNPs for modern samples and 10,537,259 SNPs for ancient samples. We used two datasets, one including only modern samples (n = 61, 42,906,277 SNPs) and the other including both modern and ancient samples (n = 72, 10,537,259 SNPs). The coverage, which is the proportion of covered bases with depth ≧ 1 to the total bases of the reference sequence (Supplementary table 1), was 10.14%–91.59% for both ancient and modern samples, but 81.26%–91.59% for only modern samples. The depth for modern samples was 6.85–55.11 (Supplementary table 1).

Genetic diversity, genetic structure, phylogeny, and demography

We computed heterozygosity and ROH for each individual to assess the genetic diversity (Supplementary figure 1). Bears from Central Asia exhibited high heterozygosity and low ROH, whereas those of Etorofu (Iturup) Island displayed the opposite pattern (Supplementary figure 1).

PCA showed genetic affinity between individuals from geographically close locations (Supplementary figures 2a and 3). The first principal component reflected the genetic difference between the North American individuals and other individuals, whereas the second principal comportment divided the East Asian individuals from the European and West and Central Asian individuals, this was supported by the result of $K = 2$ and $K = 3$ in ADMIXTURE (Figures 2a, b). In addition, the third and fourth principal components showed local distinction of the Kazakhstan and Tibetan individuals from other individuals (Supplementary figure 2b), and this was also shown in PCA based on the modern dataset (Supplementary figure 3) and $K = 4$ in ADMIXTURE (Figure 2c). The relative effective migration rate around Asia region especially central to east Asia was estimated as lower than other areas by EEMS (Figure 2d).

Phylogenetic trees based on mtDNA, X-chromosomal, autosomal, and Y-chromosomal DNA were constructed to investigate the identity of differentiation between the populations (Supplementary figure 4). Compared with trees based on the mitochondrial DNA showing the geographic discontinuities of mtDNA haplotypes (Davison et al. 2011; Figure 1B), the trees based on the nuclear genome reflected affinity between nearby populations (Supplementary figure 3); this was also, as previously reported by de Jong et al. (2023). Individuals from Kazakhstan and Western Asia, who had minor mtDNA haplotypes, tended to be distinct from other individuals, such as those from Europe, North America, and other Asian areas. Individuals from Tibet also exhibited this trend, but they were included within the same lineages as the

European, North American, and other Asian individuals. The mtDNA lineages found in West Asia (clade 1_West Asia and clade 3_West Asia) were different from other lineages reported in previous studies (e.g. Davison et al. 2011), and both diverged earlier than each related lineage (clade 1 and clade 3, respectively). Their divergence times were 253.2 thousand years ago (95%CI: 142.9–373.8 thousand years ago) and 138.6–145.8 thousand years ago (95%CI: 73.7–219.9 thousand years ago), respectively (Supplementary figure 5).

PSMC was performed on modern brown bear samples to estimate their population size (Supplementary figure 6). The effective population size of individuals from Tibet was notably different from that of other Central Asian individuals (Supplementary figure 6a). Nearly identical trends were observed between the Hokkaido and Etorofu (Iturup) Island individuals (Supplementary figure 6b), and between the West Asian and European individuals (Supplementary figure 6c). The population size changes in the Northern Okhotsk, Sakhalin, and Hokkaido were similar until approximately 100 thousand years ago; subsequently, but thereafter, the effective population sizes in the Northern Okhotsk and Sakhalin were slightly larger than that in Hokkaido (Supplementary figure 6d).

Inter- and intra-species gene flow

Population splits and mixtures were inferred using TreeMix (Supplementary figure 7). The phylogeny optimally estimated based on Δm (Supplementary figure 8) revealed a genetic affinity between nearby populations, which was consistent with the findings from the phylogenetic tree (Supplementary figure 4). The results of migration edges = 1, 2 indicated that migration occurred from the ABC island population to the polar bear population. Additionally,

the result of migration edges = 5 showed the migration from the polar bear population to the Hokkaido and Etorofu (Iturup) Island populations.

To compare the difference between populations regarding hybridization with related species, f_4(Outgroup, Polar bear; X, Kazakhstan[U13M]) and f_4(Outgroup, Cave bear; X, modern Polar bear), were calculated ,where X represents brown bear individuals,. The values of f_4(Outgroup, Polar bear; Hokkaido, Kazakhstan[U13M]) were more positive than f_4(Outgroup, Polar bear; North America, Kazakhstan[U13M]), but those were significantly more negative than f_4(Outgroup, Polar bear; East Asia excluded Etorofu (Iturup) Island, Kazakhstan[U13M]) (Supplementary figure 9 and Supplementary table 3). The values of f_4(Outgroup, Polar bear; ancient Brown bear, Kazakhstan[U13M]) were significantly more negative than those of f_4 values using modern samples from nearby areas as X. These trends were consistent regardless of the polar bear individual used and were also supported when using the only modern samples dataset (Supplementary table 4). The result of f_4(Outgroup, Cave bear; X, modern Polar bear) show that only f_4(American black bear, *kudarensis* Cave bear; Kazakhstan[U13M], modern Polar bear) and f_4(American black bear, *kudarensis* Cave bear; West Asia, modern Polar bear) were statistically significant (|Z score| > 3; Supplementary figure 10 and Supplementary table 5), although f_4(Asiatic black bear, *kudarensis* Cave bear; Kazakhstan[U13M], modern Polar bear) and f_4(Asiatic black bear, *kudarensis* Cave bear; West Asia, modern Polar bear) were not. In instances where X represented an individual with a statistically significant f_4 value, the proportions of gene flow estimated by f_4 ratio (Figure 3, Supplementary figure 11, Supplementary table 6, and 7) from polar bears or cave bears to brown bears was approximately 24.36% and 13.81%, respectively.

Among modern brown bears, 3 European (RF01, SJS01, and SLK1) and 7 North American (BB020, BB034, BB049, BB059, Den, GRZ, and WB039) individuals were identified belonging to clade 3a1 (Figure 1; Supplementary table 1). The value of f_4(Outgroup, clade 3a1 in Europe; X, the modern brown bear excluding X) and f_4(Outgroup, clade 3a1 in North America; X, the modern brown bear excluding X) were calculated to infer the scale of gene flow driven by the recent expansion of mainly a mtDNA lineage, clade 3a1 (Figure 4). The Kazakhstan brown bear population (only including U13M) has not been affected by the gene flow compared with other populations (Tukey's honestly significant difference test, $P < 0.05$). The West Asian brown bears also exhibited the same trend based on the values of f_4(Outgroup, clade 3a1 in North America; X, the modern brown bear excluding X). Both f_4 statistics supported the influence of recent gene flow on Tibetan individuals. The Hokkaido and Etorofu (Iturup) Island populations have not been affected by the recent gene flow than the Northern Okhotsk population, U34 (Tukey's honestly significant difference test, $P < 0.05$). On the other hand, no statistically significant differences were found between the Hokkaido and Etorofu (Iturup) Island populations and the Sakhalin population, AA1 which is closer to Hokkaido and Etorofu (Iturup) Island than Northern Okhotsk region.

Discussion

We performed population genomic analysis on brown bears and included genome sequence data from nine new individuals around the Eurasian Continent. This allowed us to verify the genetic differences between mtDNA and the nuclear genome and elucidate the contribution of genomic replacement driven by male-biased dispersal. Genetic differentiation

was observed between local populations, particularly around Kazakhstan, Tibet, and West Asia, where the ancient mtDNA lineages were confirmed. We detected geographic discontinuity in gene flow from polar bears, with more alleles from polar bears remaining in the Hokkaido and Etorofu (Iturup) Island populations compared with the findings in populations in Northern Okhotsk. These results could be attributed to differences in the influence of recent expansion, which was primarily driven by males, between the continental and island populations.

Genetic diversity of the brown bear

Genetic diversity in brown bears from Western and Central Asia was relatively higher compared with that in brown bears from other areas. Genetic structure analysis using EEMS showed that the low migration around Western and Central Asia restricted gene flow from others populations and contributed to the significant distinctiveness of these populations, especially in Kazakhstan, Tibet, and Western Asia; however, the population size around Western and Central Asia was inferred to have been constant, which retained genetic diversity. In contrast, Tumendemberel et al. (2023) reported low heterozygosity of brown bears from Mongolia and Pakistan located relatively near Kazakhstan and Tibet, and suggested that conservation is necessity. Even with high genetic diversity, comprehensive monitoring research is needed in Western and Central Asia to protect this species from extinction.

In contrast to the high genetic diversity of West and Central Asian brown bears, that of the Etorofu (Iturup) Island individual was significantly lower than that of the Hokkaido individuals and as low as the Apennine brown bear, which was recently listed as endangered; however genetic affinity was also found with the Hokkaido population and this result was

consistent with a prior mtDNA analysis (Hirata et al. 2013). These results indicate that the Etorofu (Iturup) Island population experienced a bottleneck after isolation from the Hokkaido population. Events such as the founder effect and human activity (Kostenko et al. 2004), could be causes of the low genetic diversity. There are approximately 60–260 brown bears on each island in the Kuril Archipelago (Kostenko et al. 2004). Therefore, conservation plans and additional genomic data are needed to monitor this population and evaluate its genetic diversity.

Demographic history of Western and Central Asian populations

PCA clearly distinguished the Tibet and Kazakhstan populations from the other populations, and this was also reflected in the phylogenetic tree. Previous studies (Hirata et al. 2013; Lan et al. 2017) reported minor mtDNA haplotypes: clade 5 in Tibet and clade 6 around Himalayas. Ancestral genetic traits from those mtDNA lineages have remained in this area because of geographical isolation from other populations, and this inference is partially supported by the result of our ADMIXTURE analysis and a previous study (de Jong et al. 2023). Comparison of these populations revealed that Tibetan individuals were slightly more distinct than the Kazakhstan individual, as indicated by the PCA and PSMC results that showed a distinct population size difference. By contrast, the result of f_4(Outgroup, clade 3a1 in Europe; X, the modern brown bear excluded X) showed the Tibetan population was more affected by the recent expansion more than the Kazakhstan population. Thus, the slightly higher genetic differentiation in Tibetan individuals may be attributed to a more ancient population demographic history predating the recent expansion. Further analyses are warranted to further elucidate these trends because the current sampling was limited.

PCA, ADMIXTURE, PSMC, and TreeMix indicated that the Western Asian population was relatively similar to the European populations. However, the phylogeny based on Y-chromosomal DNA sequences revealed that the Western Asian individuals diverged earliest among the brown bear lineages examined in this study (Supplementary figure 4b); this finding is consistent with that of de Jong et al. (2023). Additionally, the values of f_4(Outgroup, clade 3a1 in North America; X, the modern brown bear excluding X) did not support gene flow from the recent expansion. The minor mtDNA haplogroups, which were more closely related to the European lineages and not detected in other areas, were reported in previous studies (Ashrafzadeh et al. 2016; Hirata et al. 2014; Mizumachi et al. 2020). Consequently, we infer that the Western Asian population diverged from the European population before the recent expansion and has remained isolated from other populations.

The brown bear populations in Central Russia (Krasnojarsk Province, located in Siberia) were proposed to have resulted from admixture between other populations (Europe, North America, and Hokkaido populations). Both major mtDNA lineages, clade 3a1 and clade 3b, with clade 3b diverging earlier than clade 3a1, were found around this area (Hirata et al. 2014). In a nearby region, northern Mongolia, Tumendemberel et al. (2023) reported a mtDNA lineage that diverged earlier than clade 3, although there were low bootstrap values. Based on these results, the populations around Central Russia likely formed by admixture between ancestral and recently expanded populations.

f_4 and f_4 ratios detected hybridization between *kudarensis* cave bear, and the brown bear in Western Asia and Kazakhstan. The cave bear are widely distributed on the Eurasian continent several hundred thousand years ago (Knapp et al. 2009); therefore, hybridization could have

occurred during that period. Barlow et al. (2021) proposed that there was gene flow between cave bears excluding *kudarensis* cave bear and the common ancestor of the European and Western Asian populations. Our results may reflect additional gene flow from cave bears. However, f_4(Asiatic black bear, Cave bear; X, modern Polar bear) values were not statistically significant.

Gene flow between American black bears, brown bears, and polar bears was previously estimated by *D* statistics (Kumar et al. 2017). This gene flow could be a cause of the statistical significance of f_4(American black bear, Cave bear; X, modern Polar bear) because the value of f_4(W, X; Y, Z) statistics also become significantly negative if there is gene flow between W and Z. Although the statistical significance of f_4 values when other bear species were set as outgroups remains a question, alleles from hybridization with cave bears might have persisted, especially in these populations.

Demographic history of Etorofu (Iturup) Island and Hokkaido

PSMC analysis revealed a difference in population size between the islands (Hokkaido and Etorofu (Iturup) Islands) and the mainland (Sakhalin and the Northern Okhotsk) from approximately 100,000 years ago. This indicated a potential influence of the recent expansion. f_4 statistics estimated that the Northern Okhotsk individuals shared more alleles from the recently diverged lineage, 3a1, than other East Asia and Hokkaido individuals; however no statistical difference was observed between the Sakhalin and Hokkaido populations. These findings indicate that the ancestor of the Northern Okhotsk individuals was most affected by the recent

dispersal, whereas the ancestor of the Sakhalin individuals originated from admixture between recently dispersed individuals and individuals from Hokkaido.

TreeMix (migration edges = 5) and f_4 statistics confirmed more hybridization between polar bears and brown bears in Hokkaido and Etorofu (Iturup) Islands than in other Eastern Asian areas, which is partially consistent with the results of de Jong et al. (2023), who also showed gene flow between the Hokkaido population and polar bears. Previous studies (Cahill et al. 2013, 2015) identified that, among brown bear populations, the North American population, particularly individuals around the ABC Islands, harbored the most alleles from the polar bear. These finding indicate geographically discontinuous gene flow from the polar bears.

Two scenarios regarding this discontinuity are considered: *i*) gene flow by recent expansion, in which major alleles were replaced in the population on the Eurasian continent, as suggested by de Jong et al. (2023); and *ii*) gene flow from polar bears to only the Hokkaido and Etorofu (Iturup) Island populations. In this study, the PSMC and f_4(Outgroup, clade 3a1 in Europe; X, the modern brown bear excluding X) analyses demonstrated that scenario *i*) is more likely because these results showed that there has been very little effect of recent expansion on the Hokkaido and Etorofu (Iturup) Island populations. Additionally, ancient brown bears tended to share more alleles from the polar bear than modern brown bears distributed near where they were collected, and this was, also supported by Cahill et al. (2018). Thus, it is more likely that the brown bear ancestor possessed more alleles from polar bears than modern brown bears, and the recent expansion replaced those alleles with the alleles observed today.

Multi-layered genetic structure could explain the process of formation of the discrepancy

between phylogenic trees

de Jong et al. (2023) inferred incomplete lineage sorting of mtDNA based on the multiple coalescent analysis, which showed that individuals with different mtDNA lineages from the same populations coalesce more frequently than individuals with the same mtDNA lineage from different populations. Their finding was also supported by some of our results, which showed discrepancy between mtDNA and nuclear phylogenetic trees (Supplementary figure 4). However, we also emphasized local genetic differentiation, especially in Kazakhstan, Tibet, and West Asia, reflected in the contemporary genetic structure, population demography, and gene flow from related species. These populations do not include clade 3a1, which recently diverged, but mtDNA lineages that diverged before clade 3a1. Thus, the detected local genetic differentiation could be from genetic characteristics of the ancestral populations and relate to the difference of mtDNA lineages with the deep divergence times. Consequently, incomplete lineage sorting alone could not fully explain the observed differences between mtDNA and nuclear phylogenetic trees.

Using simulations, de Jong et al. (2023) proposed that gene flow driven by recent dispersal can rapidly erase genomic evidence of former population structure. They detected geographic discontinuity in the gene flow from polar bears and proposed that populations isolated from the continent (e.g., Hokkaido and ABC Islands) preserved polar bear genetic material. We tested for this trend using individuals from Etorofu (Iturup) Island, which is near Hokkaido. The results of f_4(Outgroup, clade 3a1; X, the modern brown bear excluding X) indicated that this pattern was formed because of recent male-biased dispersal. In addition, the f_4 statistics also showed that the individual from Kazakhstan had fewer alleles from the recently diverged lineage than other individuals and more alleles from cave bears. This finding indicates

that some populations even on the Eurasian continent, could have retained ancestral genetic traits. Thus, the impact of the recent expansion driven by male-biased dispersal differs by population.

Based on these findings, we propose multi-layered genetic structure of the brown bear genome (Figure 5); it includes both the retention of ancestral genetic traits (Figure 5a) and replacement of traits with new ones because of the recent expansion by male-biased dispersal (Figure 5b) is, reflected in the contemporary genetic structure (Figure 5c), is plausible to explain the incongruences between the phylogenic trees of mtDNA and the nuclear genome for the brown bear. The reason for the difference in influence of the recent expansion driven by male-biased dispersal could have primarily arisen because of climate changes. For example, climate change produces sea-level fluctuations that can lead to land bridges and glacial ice covering land. These environmental changes facilitated migration to new environments while occasionally imposing restrictions. The diverse environmental factors in each area underscore the significance of focusing on local demographic history to more fully elucidate brown bear evolution.

Although our study revealed the variation in brown bear populations, it is important to note that not all brown bear populations were included. For example, individuals from the northern part of Russia, Kunashiri Island, and around the border between East Russia and China were not included because of a lack of samples. Such sampling biases, referred to as Wallacean shortfall, have the potential to overlook genetic diversity and pose challenges in assessing future extinction risks (Lomolino et al. 2016). Additionally, this study and other studies (Cahill et al. 2018; Segawa et al. 2024) showed genetic differentiation between modern and ancient populations, and provided valuable insights into the evolution of the brown bear. Integrating

genomic data from individuals across various distribution areas and time periods would offer a more comprehensive understanding of brown bear evolutionary processes and contribute to the conservation of their genetic diversity.

Acknowledgments

We thank Mallory Eckstut, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript. This work was supported by the Japan Society of the Promotion of Science for Scientific Research (KAKENHI grant nos. 18H05508 and 22KJ0055), the Sasakawa Scientific Research Grant from The Japan Science Society (grant no. 2021-5022), the ZIN project no. 125012800908-0, and the establishment of university fellowships towards the creation of science technology innovation (grant no. JPMJFS2101).

Reference

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

Ashrafzadeh MR, Kaboli M, Naghavi MR. 2016. Mitochondrial DNA analysis of Iranian brown bears (*Ursus arctos*) reveals new phylogeographic lineage. *Mammalian Biology*. 81: 1–9.

Babraham Bioinformatics. 2011. FastQC: a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. last accessed 3 November, 2023.

Barido-Sottani J, et al. 2018. Taming the BEAST—A community teaching material resource for BEAST 2. *Systematic biology*. 67: 170–174.

Barlow A, et al. 2018. Partial genomic survival of cave bears in living brown bears. *Nature Ecology & Evolution*. 2: 1563–1570.

Barlow A, et al. 2021. Middle Pleistocene genome calibrates a revised evolutionary history of extinct cave bears. *Current Biology*. 31: 1771–1779.

Benazzo A, et al. 2017. Survival and divergence in a small group: the extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences*. 114: 9589–9597.

Bidon T, Janke A, Fain SR, Eiken HG, Hagen SB, Saarma U, Hallström BM, Lecomte N, Hailer F. 2014. Brown and polar bear Y chromosomes reveal extensive male-biased gene flow within brother lineages. *Molecular Biology and Evolution*. 31: 1353–1363.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30: 2114–2120.

Bon C, Berthonaud V, Maksud F, Labadie K, Poulain J, Artiguenave F, Wincker P, Aury JM, and Elalouf JM. 2012. Coprolites as a source of information on the genome and diet of the cave hyena. *Proceedings of the Royal Society B: Biological Sciences*. 279: 2825–2830.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology*. 10: e1003537.

Cahill JA, et al. 2013. Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. *PLoS Genetics*. 9: e1003345.

Cahill JA, Stirling I, Kistler L, Salamzade R, Ersmark E, Fulton TL, Stiller M, Green RE, Shapiro B. 2015. Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. *Molecular Ecology*. 24: 1205–1217.

Cahill JA, et al. 2018. Genomic evidence of widespread admixture from polar bears into brown bears during the last ice age. *Molecular Biology and Evolution*. 35: 1120–1129.

Dabney J, et al. 2013 Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*. 110: 15758–15763.

Danecek P, McCarthy SA. 2017. BCFtools/csq: Haplotype-aware variant consequences. *Bioinformatics*. 33: 2037–2039.

Davison J, et al. 2011. Late-Quaternary biogeographic scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quaternary Science Reviews*. 30: 418–430.

da Silva Coelho FA, et al. 2023. Ancient bears provide insights into Pleistocene ice age refugia in Southeast Alaska. *Molecular Ecology*. 32: 3641–3656.

de Jong MJ, et al. 2023. Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. *Communications Biology*. 6: Article 1.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32: 1792–1797.

Edwards CJ, et al. 2011. Ancient hybridization and an Irish origin for the modern polar bear matriline. *Current biology*, 21: 1251–1258.

Endo Y, Osada N, Mano T, Masuda R. 2021. Demographic history of the brown bear (*Ursus arctos*) on Hokkaido Island, Japan, based on whole-genomic sequence analysis. *Genome Biology and Evolution*. 13: evab195.

Feng S, et al. 2022. Incomplete lineage sorting and phenotypic evolution in marsupials. *Cell*. 185: 1646–1660.e18.

Fitak RR. 2021. OptM: estimating the optimal number of migration edges on population trees using TreeMix. *Biology Methods and Protocols*. 6: bpab017.

Hailer F, Kutschera VE, Hallström BM, Klassert D, Fain SR, Leonard JA, Ulfur A, Janke A. 2012. Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science*. 336: 344–347.

Hirata D, Mano T, Abramov AV, Baryshnikov GF, Kosintsev PA, Vorobiev AA, Raichev GV, Tsunoda H, Kaneko Y, Murata K, et al. 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.

Hirata D, Mano T, Abramov AV, Baryshnikov GF, Kosintsev PA, Murata K, Masuda R. 2017. Paternal phylogeographic structure of the brown bear (*Ursus arctos*) in northeastern Asia and the effect of male-mediated gene flow to insular populations. *Zoological Letters*. 3: 1–13.

Kambayashi, C, et al. 2022. Geography-dependent horizontal gene transfer from vertebrate predators to their prey. *Molecular Biology and Evolution*. 39: msac052.

Knapp M, et al. 2009. First DNA sequences from Asian cave bear fossils reveal deep divergences and complex phylogeographic patterns. *Molecular Ecology*. 18: 1225–1238.

Kosintsev PA, Bocherens H, Kirillova IV, Levchenko VA, Zazovskaya EP, Trofimova SS, Lan T, Lindqvist C. 2022. Palaeoecological and genetic analyses of Late Pleistocene bears in Asiatic Russia. Boreas. 51: 465–480.

Kostenko VA, Nesterenko VA, Trukhin AM. 2004. Mammals of the Kuril Archipelago. Vladivostok, Russia: Dalnauka. (in Russia)

Krause J et al. 2008 Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene–Pliocene boundary. *BMC Evolutionary Biology*. 8: 220.

Kumar V, Lammers F, Bidon T, Pfenninger M, Kolter L, Nilsson MA, Janke A. 2017. The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports*. 7: 1–10.

Kutschera VE, Bidon T, Hailer F, Rodi JL, Fain SR, Janke A. 2014. Bears in a forest of gene trees: Phylogenetic inference is complicated by incomplete lineage sorting and gene flow. *Molecular Biology and Evolution*. 31: 2004–2017.

Lan T, Gill S, Bellemain E, Bischof R, Nawaz MA, Lindqvist C. 2017. Evolutionary history of enigmatic bears in the Tibetan plateau–Himalaya region and the identity of the yeti. *Proceedings of the Royal Society B: Biological Sciences*. 284.

Lan T, et al. 2022. Insights into bear evolution from a Pleistocene polar bear genome. *Proceedings of the National Academy of Sciences*. 119.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25: 1754–1760.

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

Lindqvist C et al. 2010 Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. *Proceedings of the National Academy of Sciences*. 107: 5053–5057.

Liu S, et al. 2014. Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell*. 157: 785–794.

Leonard JA, Wayne RK, Cooper A. 2000. Population genetics of Ice Age brown bears. *Proceedings of the National Academy of Sciences*. 97: 1651–1654.

Lomolino MV, Riddle BR, Whittaker RJ. Biogeography. 5th ed. 2016. Sunderland: Sinauer Associates, Inc.

McKenna A, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. 20: 1297–1303.

McLellan BN, Proctor MF, Huber D, Michel S. 2017. *Ursus arctos* (amended version of 2017 assessment). The IUCN Red List of Threatened Species 2017. Version 2016.2.

Miller W, et al. 2012. Polar and brown bear genomes reveal ancient admixture and demographic

footprints of past climate change. *Proceedings of the National Academy of Sciences*. 109: 2382–2390.

Mizumachi K, Gorbunov SV, Vasilevski AA, Amano T, Ono H, Kosintsev PA, Hirata D, Nishita Y, Masuda R. 2021. Phylogenetic relationships of ancient brown bears (*Ursus arctos*) on Sakhalin Island, revealed by APLP and PCR-direct sequencing analyses of mitochondrial DNA. *Mammal Research*. 66: 95–102.

Molodtseva AS, et al. 2022. Phylogeography of ancient and modern brown bears from eastern Eurasia. *Biological Journal of the Linnean Society*. 135: 722–733.

Osada N, Kawai Y. 2021. Exploring models of human migration to the Japanese archipelago using genome-wide genetic data. *Anthropological Science*. 129: 45–58.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*. 20: 289–290.

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

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics*. 192: 1065–1093.

Petkova D, Novembre J, Stephens M. 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*. 48: Article 1.

Pickrell J, Pritchard J. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*. 8: e1002967.

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

Puckett EE, Orton D, Munshi-South J. 2020. Commensal rats and humans: integrating rodent phylogeography and zooarchaeology to highlight connections between human societies. *Bioessays*. 42: 1900160.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 81: 559–575.

Rogers J, et al. 2019. The comparative genomics and complex population history of *Papio* baboons. *Science Advances*. 5.

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 4: 406–425.

Segawa T, et al. 2021. Ancient DNA reveals multiple origins and migration waves of extinct Japanese brown bear lineages. *Royal Society Open Science*. 8: 210518.

Segawa T, Yonezawa T, Mori H, Kohno A, Kudo Y, Akiyoshi A, Wu J, Tokanai F, Sakamoto M, Kohno N, Nishihara H. 2022. Paleogenomics reveals independent and hybrid origins of two morphologically distinct wolf lineages endemic to Japan. *Current Biology*. 32: 2494–2504.

Segawa T, Rey-Iglesia A, Lorenzen ED, Westbury MV. 2024. The origins and diversification of Holarctic brown bear populations inferred from genomes of past and present populations. *Proceedings of the Royal Society B*. 291: 20232411.

Tassi F, Ghirotto S, Mezzavilla M, Vilaça ST, De Santi L, Barbujani G. 2015. Early modern human dispersal from Africa: genomic evidence for multiple waves of migration. *Investigative genetics*. 6: 1–16.

Taylor GA, et al. 2018. The genome of the North American brown bear or grizzly: *Ursus arctos*

ssp. *horribilis*. *Genes*. 9: 598.

Tumendemberel O, et al. 2019. Phylogeography, genetic diversity, and connectivity of brown bear populations in Central Asia. *PLoS One*. 14: e0220746.

Tumendemberel O, Hendricks SA, Hohenlohe PA, Sullivan J, Zedrosser A, Sæbø M, Proctor MF, Koprowski JL, Waits LP. 2023. Range-wide evolutionary relationships and historical demography of brown bears (*Ursus arctos*) revealed by whole-genome sequencing of isolated central Asian populations. *Molecular Ecology*. 32: 5156–5169.

Wang MS, et al. 2022. A polar bear paleogenome reveals extensive ancient gene flow from polar bears into brown bears. *Nature Ecology and Evolution*. 1–9.

Zhou Q, Karunarathne P, Andersson-Li L, Chen C, Opgenoorth L, Heer K, Milesi P. 2024. Recurrent hybridization and gene flow shaped Norway and Siberian spruce evolutionary history over multiple glacial cycles. *Molecular Ecology*. 33: e17495.

Data and resource availability

Raw sequence reads and mitochondrial DNA sequences are submitted in the DDBJ BioProject database (BioProject ID: PRJDB17730).

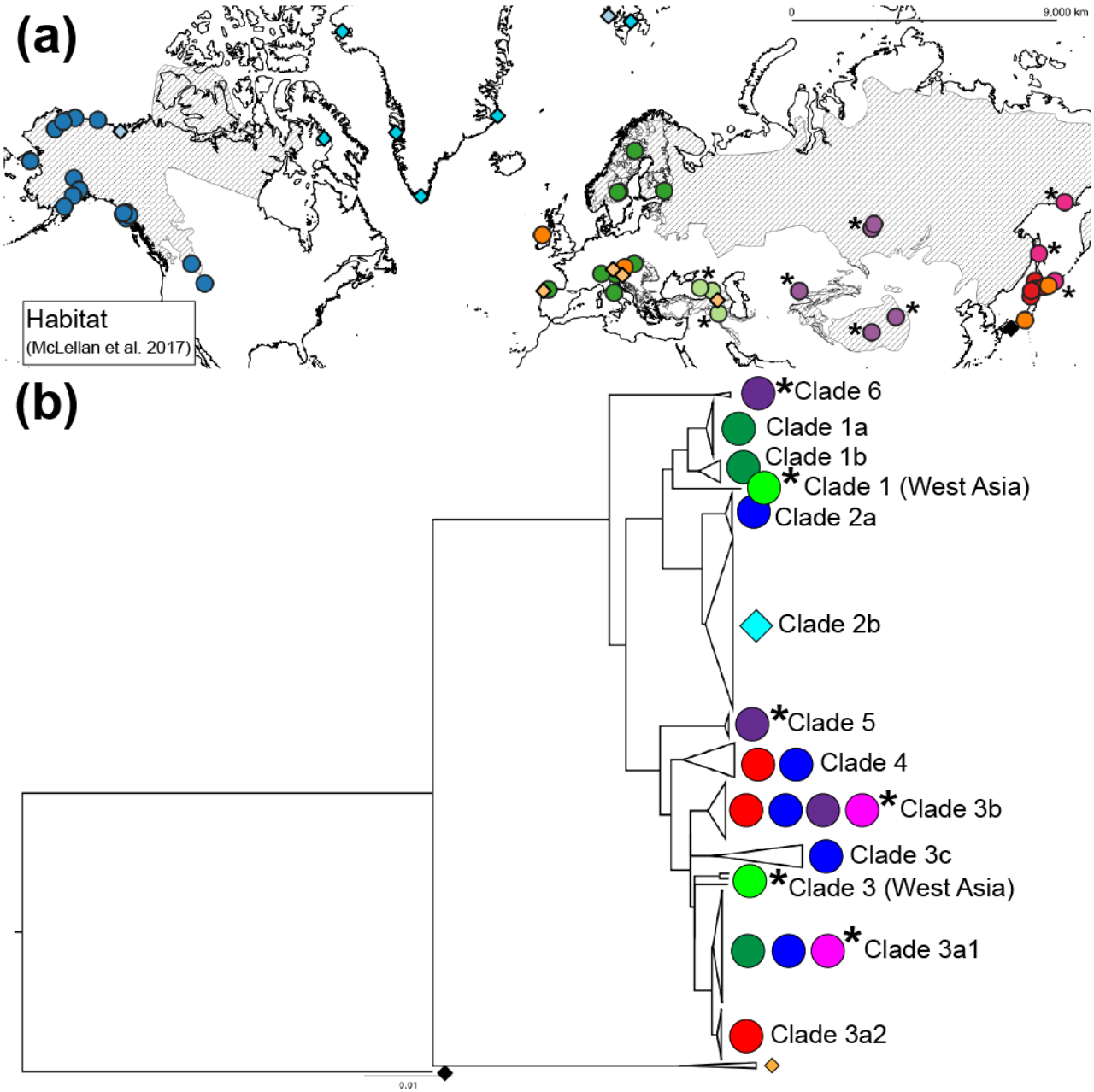

Figure 1. Sampling locations and mitochondrial DNA (mtDNA) lineages found in each location. Sampling locations for bears analyzed in this study excluding outgroups (a) and a phylogenetic tree of brown bear mtDNA lineages (b). The shaded area reflects the current habitat of the brown bear (McLellan et al. 2017). The dot colors represent correspond to the groups listed in Table 1: ancient brown bear, dark orange; ancient polar bear, light blue; cave bear, light orange; brown bears from West Asia, light green; brown bears from Central Asia, purple; brown bears from East Asia, pink; brown bears from Europe, green; brown bears from Hokkaido, red; brown bears from North America, blue. Shapes represent species: square, polar bear or cave bear; circle, brown bear. The asterisks shown near the dots indicate that new sequence data were reported in this article. These sampling locations were inferred based on the collected region (Hirata et al. 2014) and presence of brown bear habitat (McLellan et al. 2017). Two samples (sample name: GT_1 and Asiatic_black_bear) are not shown in this map because these individuals were from zoos.

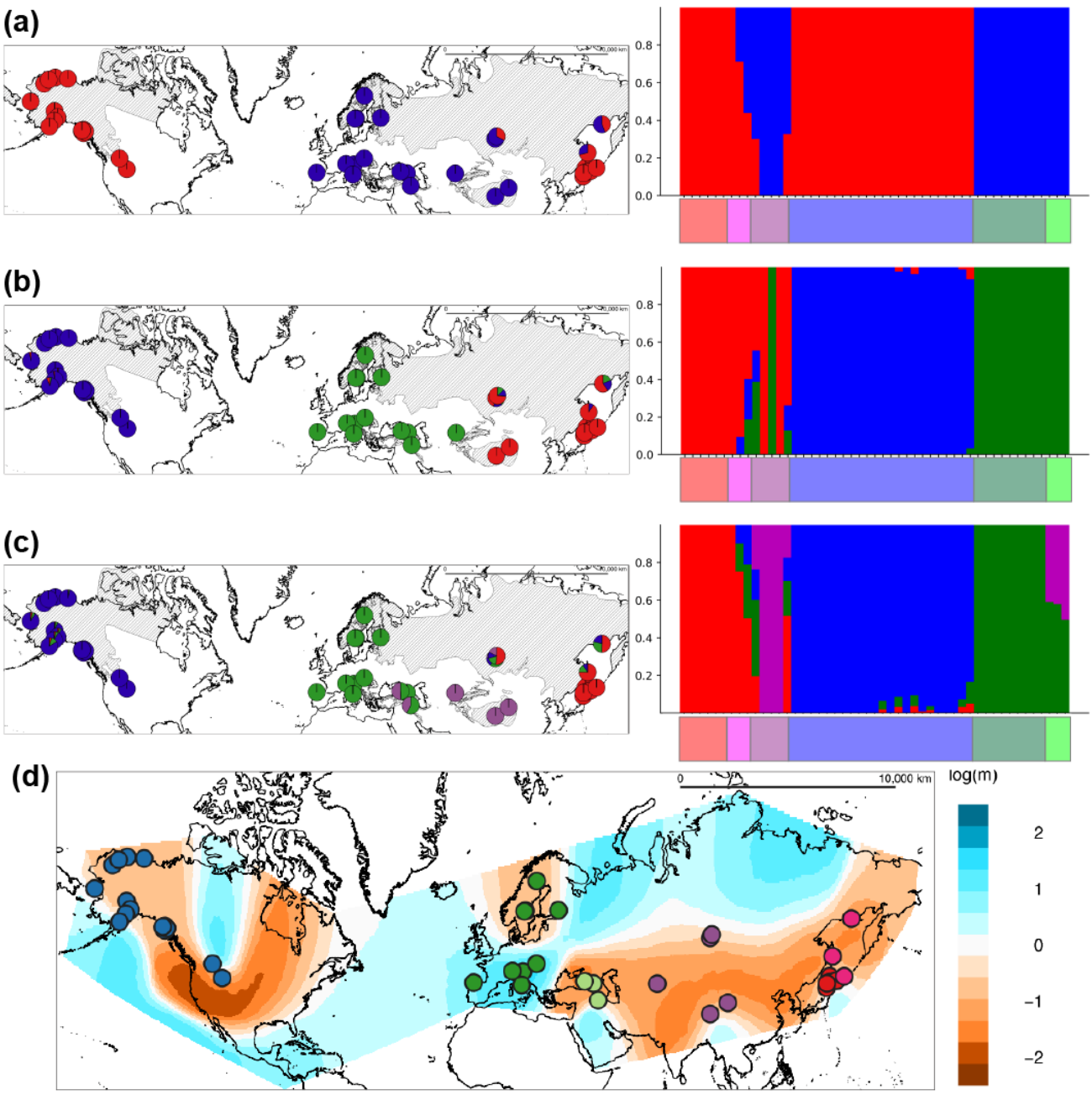

Figure 2. Genetic structure results for ADMIXTURE (*K*=2) (a), ADMIXTURE (*K*=3) (b), ADMIXTURE (*K*=4) (c), and EEMS (d) relative to sampling locations. The bar colors below the ADMIXTURE results correspond to the dot colors showing the groups listed in Table 1. One sample (GT_1) was excluded from EEMS because we could not find detailed information regarding where it was captured.

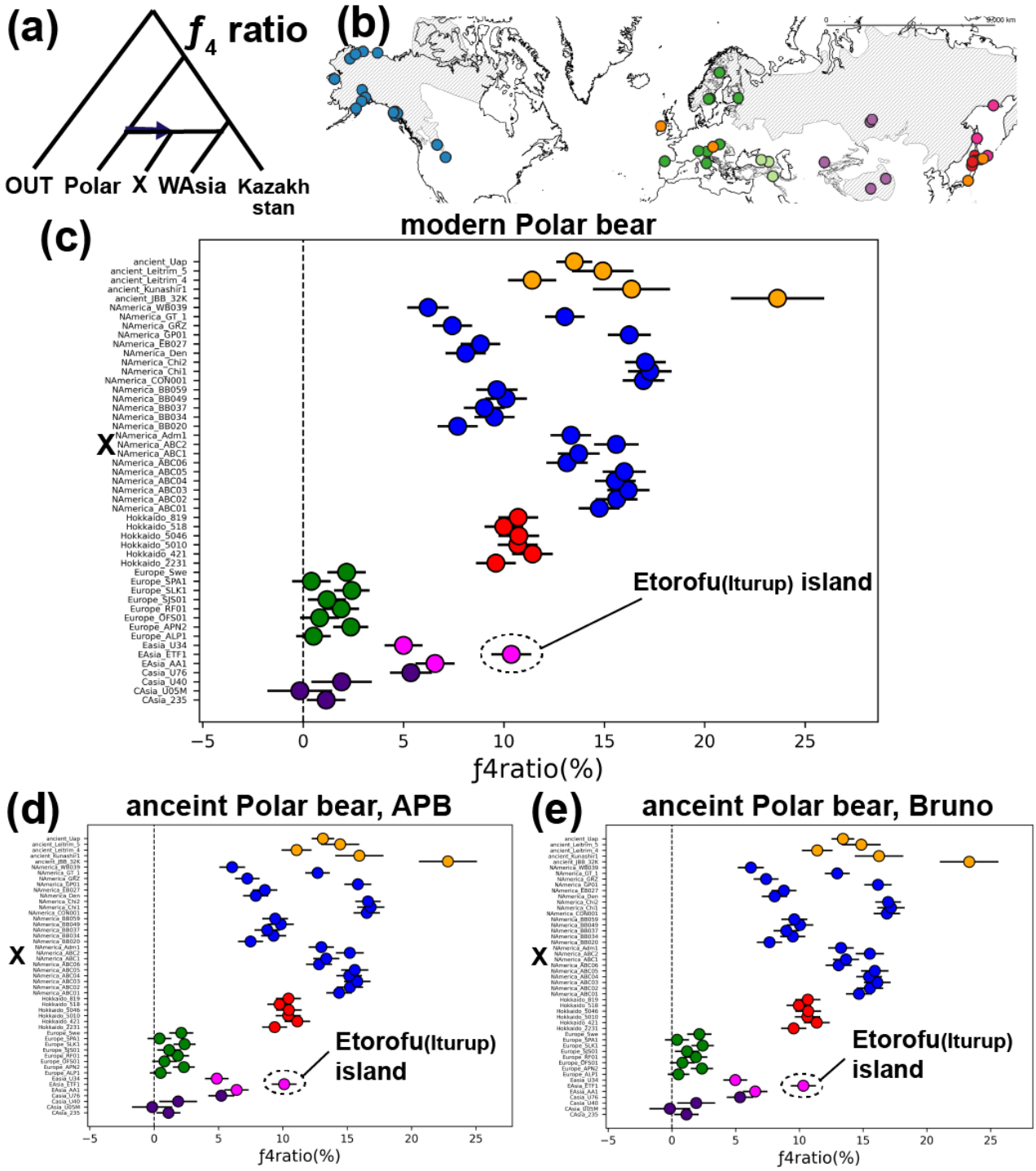

Figure 3. Percent of admixture proportions from polar bears based on f_4 ratio. f_4 ratio (%) reflects the proportion of gene flow from the polar bear to X, representing a brown bear individual (a). The dot colors correlate with those on the world map (b). Colors represent the groups listed in Table 1: ancient brown bear, dark orange; ancient polar bear, light blue; cave bear, light orange; brown bears from West Asia, light green; brown bears from Central Asia, purple; brown bears from East Asia, pink; brown bears from Europe, green; brown bears from Hokkaido, red; brown bears from North America, blue. Each figure shows the result when the polar bear is the modern polar bear (c), APB is an ancient polar bear (d), and Bruno is an ancient polar bear (e). The name of the ancient Polar bear corresponded to the sample name listed in Table 1.

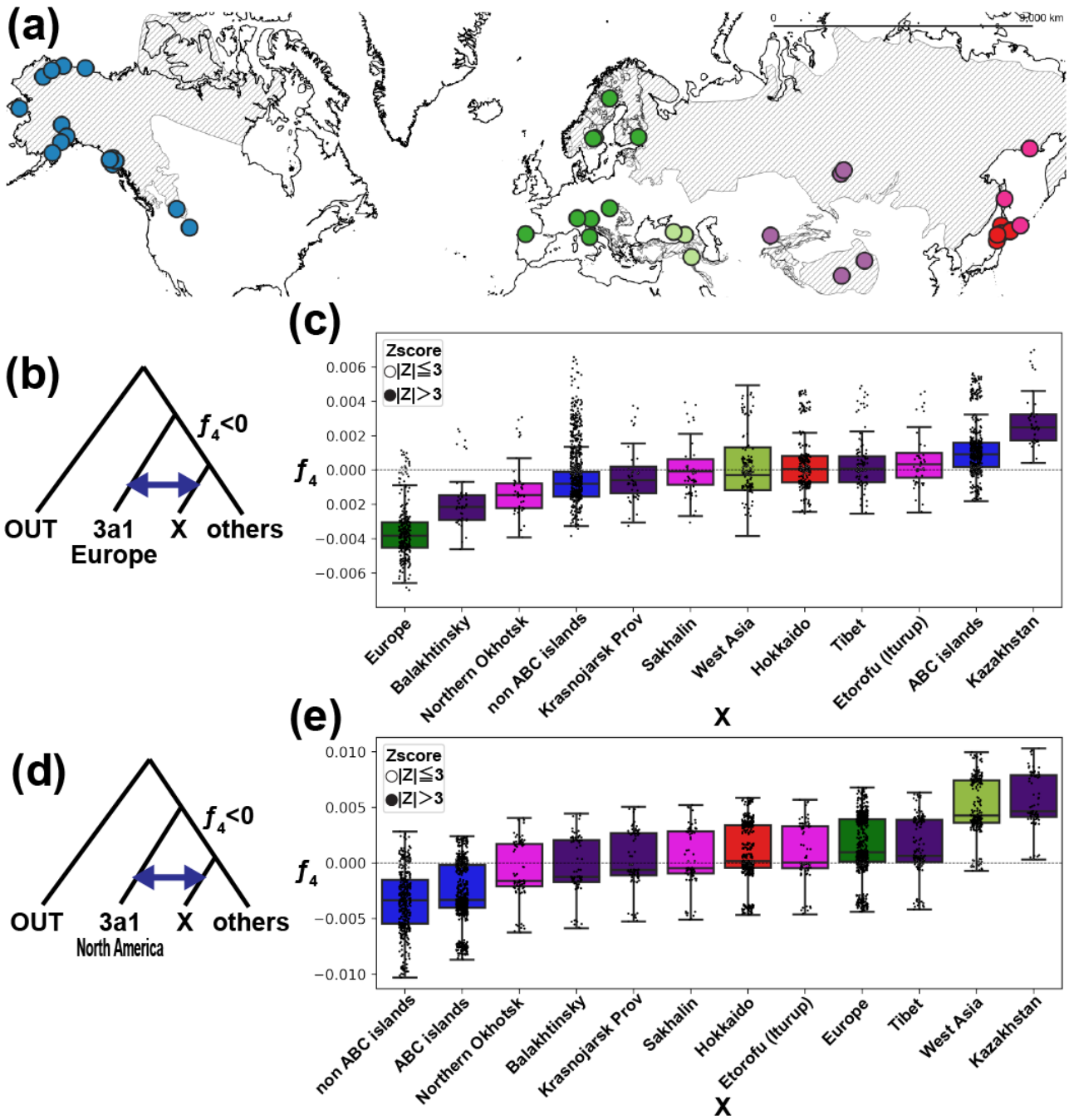

Figure 4. f_4(Outgroup, clade 3a1 in Europe or North America; X, the modern brown bear excluding X) values. f_4(Outgroup, clade 3a1 in Europe; X, the modern brown bear excluding X) topology and values shown in Figure 4b and 4c, respectively. f_4(Outgroup, clade 3a1 in North America; X, the modern brown bear excluding X) topology and values shown in Figure 4d and 4e, respectively. The boxplot colors correlate with those on the world map (a). Colors represent the groups listed in Table 1: ancient brown bear, dark orange; ancient polar bear, light blue; cave bear, light orange; brown bears from West Asia, light green; brown bears from Central Asia, purple; brown bears from East Asia, pink; brown bears from Europe, green; brown bears from Hokkaido, red; brown bears from North America, blue. If the f_4 value was less than 0, gene flow was supported between clade 3a1 in Europe or North America and X, which were brown bear individuals.

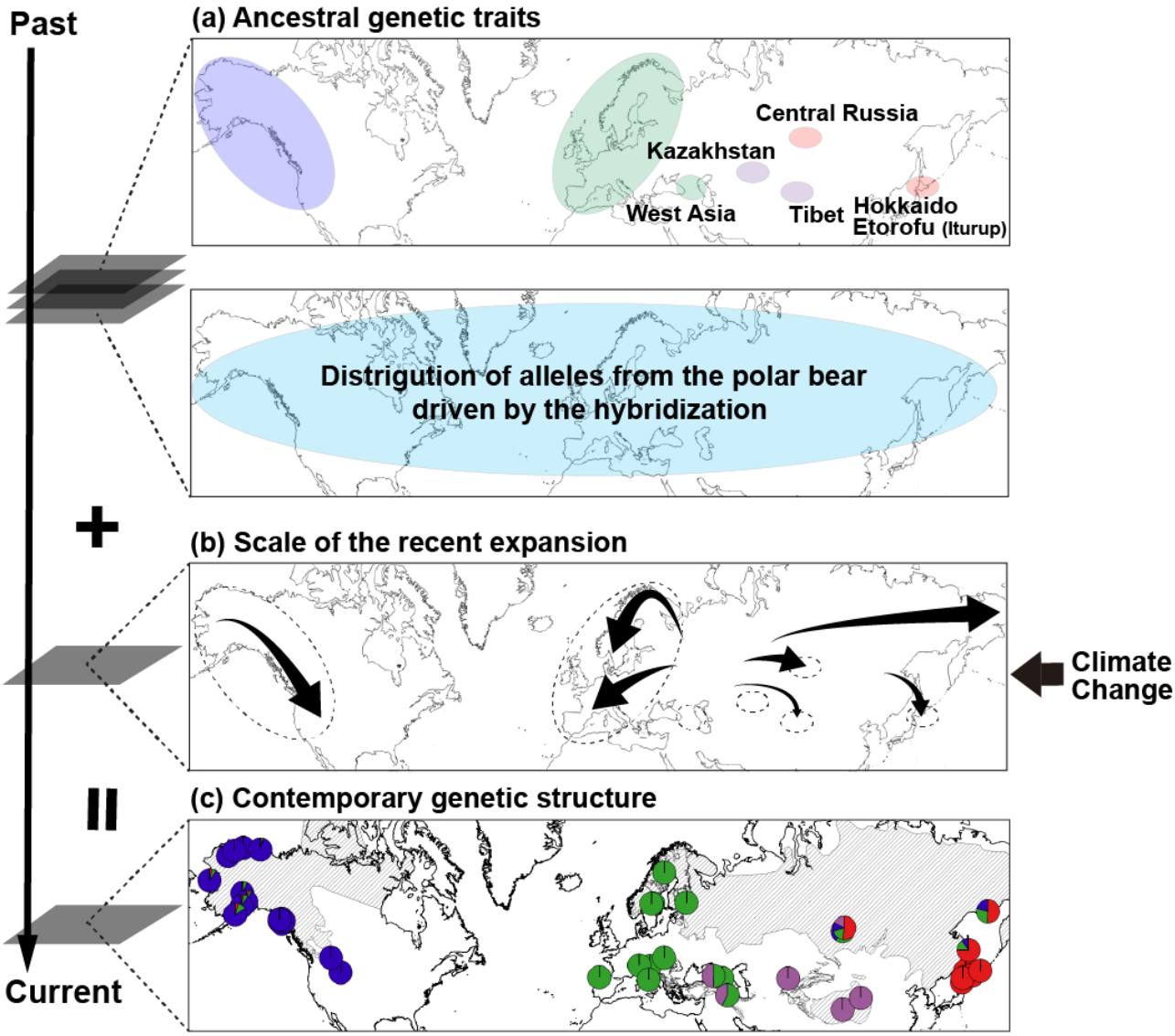

Figure 5. Multi-layered genetic structure of brown bear populations inferred by the results of this study. (a) Ancestral genetic traits before the recent expansion, (b) Scale of the recent expansion and (c) Contemporary genetic structure shown in Figure 5c. The contemporary genetic structure represent the ancestral genetic structure and allele frequency driven by hybridization with polar bears, respectively. Bolder arrows shown in (b) had a greater impact on the recent expansion estimated by f_4 statistics.