The origins and diversification of Holarctic brown bear populations inferred from genomes of past and present populations

Keywords: Brown bear, Evolution, Demography, Genome, Palaeogenome, Phylogenomics,

Takahiro Segawa[1], Alba Rey-Iglesia[2], Eline D Lorenzen[2], Michael V Westbury[2,*]

1. Center for Life Science Research, University of Yamanashi, Japan
2. Globe Institute, University of Copenhagen, Denmark

*Correspondence: m.westbury@sund.ku.dk

Abstract

The brown bear is one of the few surviving megafauna species that made it through the Late Quaternary megafaunal extinctions. It is widely distributed and has become a model organism for studying biogeographic patterns across Eurasia and North America. However, to date, most studies have focused on the mitochondrial genome with nuclear genomes only being used for evolutionary studies within the Ursusidae family or local population studies. Here we present novel genomic data from two Late Pleistocene brown bears from Japan and Siberia, and combine them with modern and ancient published genomes from across the Holarctic range, to investigate the species' genomic evolutionary history. We uncovered four differentiated populations corresponding to Europe, Siberia, Japan, and North America which all diverged ~100 kya. Palaeogenomic data reveals population continuity in all populations apart from in Europe where Late Pleistocene European individuals cluster with Siberia rather than contemporary/mid-late Holocene European individuals. We propose a replacement around the Pleistocene-Holocene transition in Europe, supported by results of a recent population expansion in modern European individuals. Population continuity in Japan and North America contrasts mitochondrial suggestions of multiple migration events. However, the rapid population diversification and increased likelihood of gene flow and incomplete lineage may have biased mitochondrial results. Demographic results show similar effective population declines in all populations ~50 kya which we suggest is due to the geographic dispersal of the newly diversified populations and a decrease of gene flow between populations.

Introduction

Megafaunal extinctions were a defining feature of the Late Pleistocene [1]. Despite this, many species still persist today, although they may have experienced local extirpations, range reductions, and losses of genetic diversity [2,3]. The brown bear (*Ursus arctos*) is one of the survivors of the Late Pleistocene megafauna extinctions. The species is one of the largest extant terrestrial carnivores, and is distributed across the Holarctic [4]. The ability of brown bears to survive may be due to their high level of ecological flexibility and extremely

broad dietary range, which has been observed in contemporary populations [5]. Brown bears have evolved a generalist omnivore strategy; although they possess morphological traits characteristic of carnivores, their diets in many ecosystems comprise primarily plant matter [6]. Despite their current dietary flexibility, brown bears experienced extensive range reductions during the Holocene, with extirpations in some geographic regions driven by climate and environmental change, and more recently, humans [7]. For example, based on fossil chronology, the relatively large steppe brown bear (*Ursus arctos priscus*) of Eurasia disappeared at the end of the Late Pleistocene. Although it was presumably well-adapted for survival in Europe during the Late Pleistocene, its size may have inhibited adaptive flexibility after the environment drastically changed [8].

The wide distribution, fossil chronology, and broad appeal of brown bears have led the species to become a model organism for understanding biogeographic patterns across the Holarctic [9]. The phylogeographic structure of contemporary brown bears in Europe has been considered consistent with the "expansion/contraction" model of postglacial recolonisation, where populations contracted into southern Europe during glacial periods and later recolonised the rest of Europe after the glacial period [9]. However, ancient DNA retrieved from Late Pleistocene brown bear fossil specimens have since questioned population contraction into southern refugia during the Last Glacial Maximum (LGM), and instead found relatively high levels of mitochondrial diversity in Late Pleistocene European brown bears [10]. In contemporary North American brown bears, three genetically divergent mitochondrial lineages have been described, leading to the hypothesis of three independent colonisation events from Eurasia [9]. In addition, independent colonisation events have also been hypothesised for Asia; timing of divergence estimated from mitochondrial genomes retrieved from Japanese brown bears have led to suggestions that brown bears migrated from mainland Eurasia to Japan at least twice [11]. However, the mitochondrial genome represents just a single loci, is exclusively maternally inherited, and may be heavily confounded by gene flow and incomplete lineage sorting [12,13], which can lead to erroneous inferences. Furthermore, the majority of genetic clades used to form biogeographic hypotheses for brown bears are paraphyletic and not unique to a single geographic region [9].

Nuclear data offer an opportunity to circumvent the challenges associated with mitochondrial genomes [12]. While nuclear genomic data are currently available for many brown bears across their range, they have thus far been used to understand the evolutionary relationships between brown bears and polar bears [14–16], or to elucidate local populations [17,18], but not to understand population relationships at a range-wide scale. To fill this knowledge gap, we investigate the evolutionary and demographic relationships among brown bear populations sampled across their Holarctic range (Fig 1A), using a combination of contemporary (n = 39) and ancient nuclear genomes (n = 12; including newly generated genomic datasets from two individuals). Specifically, we investigate (i) population structure among contemporary, Holocene, and Late Pleistocene individuals, (ii) phylogenomic relationships between contemporary populations, and (iii) comparative long-term (~500,000 years) demographic histories of contemporary individuals.

Methods

Data generation

We generated palaeogenomic data from a Late Pleistocene (32,719 - 32,209 cal. BP [11]) brown bear from the Japanese island of Honshu (JBB-32K [11]). Lab work was carried out in a clean room for ancient DNA at the National Institute of Polar Research, Japan. We extracted 100 mg of petrosal bone, and followed the DNA extraction procedure based on silica pellets in solution as described in Orlando et al. [19]. Double-stranded DNA was converted into an Illumina sequencing library using an NEBNext Quick DNA Library Prep Master Mix Set for 454 (New England BioLabs) with 20 µl of DNA extract without DNA fragmentation. Illumina sequencing adapters were added to the end-repair reaction. Indexing PCR was prepared in the ancient DNA laboratory, carried out in a physically separate modern laboratory. Each eluate (3 µl) was subjected to PCR amplification in 50-µl reactions containing 1 KAPA HiFi HotStart Uracil+ReadyMix (Kapa Biosystems) and 0.3 µM Dual Index primers of NEBNext Multiplex Oligos for Illumina (NEB). Three independent PCR reactions from the same library were combined after PCR and purified using a NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). Sequencing was carried out on an Illumina HiSeq XTen at Macrogen, Japan. Previously published data for this individual was downloaded from DDBJ DRA IDs: DRR308533 and DRR308534 [11].

We generated palaeogenomic data from a Late Pleistocene (aged beyond radiocarbon dating: >50 kya but molecularly dated to 61,826 years BP [HPD 95%: 27,866–96,783 years BP][20]) brown bear from eastern Siberia (CGG_1_020006 [20]). Lab work was carried out in a clean room for ancient DNA at Globe Institute, University of Copenhagen. We extracted DNA from ~50 mg of bone powder using a modified version of the protocol from [21]. Modifications included overnight incubation with the extraction buffer at 42 °C instead of at 37 °C, the bone powder was pelleted out of suspension, and the supernatant concentrated down to 150–200 µl for each sample using 30 kDa Amicon centrifugal filter unit (Millipore). The remaining supernatant was mixed with 13x its volume of binding buffer and DNA was purified with MinElute columns (Qiagen), following the manufacturer's instructions with the exception of a 15-minute incubation at 37 °C during the elution step. DNA extracts were transformed into sequencing libraries in 25 µl reactions following the protocol from Meyer and Kircher [22] with the modification that the initial DNA fragmentation was not performed as ancient DNA is already highly fragments and MinElute kit (Qiagen) was used for the purification steps. DNA libraries were indexed using KAPA HiFi uracil + premix (KAPA Biosystems). The number of cycles for index PCRs was determined from qPCR analysis. The resulting libraries were quantified on an Agilent 2100 Bioanalyser, pooled at equimolar concentration and sequenced on an Illumina HiSeq 2500 at the University of Copenhagen using single-end 80 base pairs (bp) reads.

We additionally downloaded the raw reads from 39 contemporary and 10 ancient brown bear individuals from across their current and past distributions [15–18,23–26] (Fig 1A and Supplementary Table 1). Samples (n = 7) from outside their current distribution are those from Ireland, Honshu, Japan, and Quebec in Canada.

Data processing and mapping

Contemporary brown bears

We processed the raw sequencing data from the contemporary brown bears (excluding the Hokkaido bears, detailed below) with the PALEOMIX v.1.2.12 [27] pipeline. Adapters, stretches of Ns, and low-quality bases in reads were trimmed and filtered with AdapterRemovalv2 [28] using default parameters. BWA-backtrack v0.7.15 [29] was used to map the cleaned reads to the polar bear pseudo-chromosome genome [14,30] (based on Genbank accession: GCA_000687225.1), which includes a mitochondrial genome (Genbank accession: NC003428), with default parameters. Reads with mapping quality < 30 were filtered using SAMtools [31]. Duplicates were removed with Picard v2.6.0 [32]. Possible paralogs were filtered using SAMtools v1.6. Finally, local realignment around indels was done using GATK (v 3.3) [33].

The six modern Hokkaido brown bears followed a similar protocol, but using Fastp v0.23.2 [34] for adapter trimming and duplicate removal using SAMtools. To be used as an outgroup/ancestral sequence, the raw data for the spectacled bear (*Tremarctos ornatus*) was downloaded (SRA accession code: SRR13602719) and mapped to the same polar bear reference genome using the BWA mem algorithm.

Ancient brown bears

We removed illumina adapter sequences, low-quality reads (mean q <25), short reads (<30bp), and merged overlapping read pairs with Fastp. We mapped the merged reads to the polar bear reference genome using BWA-backtrack v0.7.15, with the seed disabled (-l 999), and otherwise default parameters. Reads with mapping quality < 30 and duplicates were filtered using SAMtools [31]. Only individuals with >0.01x nuclear genome coverage were retained for downstream analyses.

Population structure

We evaluated the relationships between the contemporary and ancient brown bears included in this study by performing Principal Component Analyses (PCA). We performed these analyses using either pseudohaploid base calls or genotype likelihoods (GL) computed with ANGSDv0.921 [35]. We computed GL in ANGSD for all individuals specifying the parameters: minimum mapping and base qualities of 30 (-minmapQ 30 -minQ 30), calculate genotype likelihoods using the GATK algorithm (-GL 2), output a beagle GL file (-doGlf 2), calculate major and minor alleles based on GL (-doMajorMinor 1), remove transitions (-rmtrans 1), only include SNPs with a p-values <1e-6 (-SNP_pval 1e-6), only consider autosomal chromosomes (-rf), a minimum minor allele frequency of 0.05 (-minmaf 0.05), skip triallelic sites (-skiptriallelic 1), only consider reads mapping to one region uniquely (-uniqueonly 1), and only consider sites if at least 36 individuals have coverage (-minind 36). The pseudohaploid base call was computed using the same parameters with the addition of -doIBS 2, -doCov 1, and only included SNPs if at least two individuals had the minor allele (-minminor 2). To construct covariance matrices from the GL datasets for PCA we used

PCAngsd v0.98 [36]. We only included a single Apennines brown bear in these analyses due to their high levels of relatedness [18].

Phylogenetic relationships

Our PCA results suggested three or four main clusters, and to further investigate this we performed phylogenomic analyses. We generated consensus pseudo-haploid base calls using ANGSD (-dohaplocall 2) using the same filtering parameters as above but called all sites (-minminor 0) from three contemporary individuals per population, apart from Siberia which only comprised one sample. We used APP2, SPA1, and BGI-brownbear-20105434 to represent Mainland Europe; DRR276775, DRR276776, and DRR276777 to represent Japan; and ABC2_Lucky, BGI-brownbear-107596012, and BGI-brownbear-43157 to represent North America. The spectacled bear was included as an outgroup.

From the pseudo-haploid file, we computed sliding window phylogenies in PhyML-3.3.20211231 [37] using an available python script at https://github.com/simonhmartin/genomics_general (phyml_sliding_windows.py). We specified a window size of 20kb, a slide of 1Mb, a window must contain at least 100 sites, and PhyML to optimise the following parameters: tree topology (t), branch length (l), and substitution rate parameters (r). To infer the overall species tree and local posterior support values, the output trees were run through ASTRAL v5.7.8 [38] with the parameter -t 3 to output local posterior values. To further investigate the branching pattern of the brown bear populations, we ran ASTRAL again, but specified which individuals belonged to which predefined population based on our first run, and with the parameter -t 32 to output the topologies and corresponding normalised quartet scores. To test if the length of the internal branch is zero; thus, the species tree is a polytomy we used the -t 10 in ASTRAL [39].

Demographic analyses

We ran demographic analysis on diploid consensus genomes from single population representatives using PSMC [40]. These included Mainland Europe - BGI-brownbear-20105434, Japan - DRR276776, North America - ABC2_Lucky, and Siberia - S235. We first tested the impact of mapping approach and coverage on the PSMC results (Supplementary material). Based on these results, we remapped individual S235 using BWA mem to obtain higher coverage and downsampled all other individuals to 12.88x using SAMtools to ensure comparability. We called diploid genome sequences using SAMtools and BCFtools, specifying a minimum quality score of 20 and minimum coverage of 10 and specified only to use pseudochromosomes 1-18. We ran PSMC specifying specifying standard atomic intervals (4+25*2+4+6), a maximum number of iterations of 25 (-N), maximum 2N0 coalescent time of 15 (-t), and initial initial theta/rho ratio of 5 (-r), and performed 100 bootstrap replicates to investigate support for the resultant demography. PSMC results were checked for overfitting so that after 20 rounds of iterations, at least 10 recombinations are inferred to have occurred in the intervals each parameter spans. We plotted the output using a mutation rate of 1.82×10^{-8} per generation, assuming a generation time of 11 years [17,18].

We calculated Tajima's D for the three regional populations with data from multiple individuals available in ANGSD [41] using the SFS as input. We limited our dataset to

contemporary individuals, 14 individuals from Europe (we only included one individual from the Apennines (APP3) to avoid their recent local population history from influencing the population-wide results), six from Japan, and 12 from North America. We calculated the SFS from allele frequencies the autosomes 1-18 in ANGSD (-doSaf 1), and specifying the following parameters: only include sites covered in all individuals (-minind), the ancestral sequence as the spectacled bear (-anc), -dosaf 1 -uniqueOnly 1 -remove_bads 1 -minMapQ 30 -minQ 30 -doMajorMinor 1 -doMaf 2 -skipTriallelic 1 -GL 2 -minind (all individuals). To control for variation across the genome, we calculated Tajima's D in 10 Mb windows and extracted the mean values. To act as the ancestral allele, we used ANGSD to generate a consensus base call fasta file (-doFasta 2) of the spectacled bear with the filters: -mininddepth 5 -minq 25 -minmapq 25. Significance from 0 was calculated using a Wilcoxon signed rank test with continuity correction.

Heterozygosity

We estimated the genome-wide heterozygosity of each individual with >5x mean coverage based on previously published recommendations [42]. We used ANGSD to calculate site allele frequencies per individual (-doSaf 1) with the following filters: -minmapQ 30 -minQ 30 -skiptriallelic 1 -uniqueonly 1 -GL 2 -setMinDepthInd 5 -doMajorMinor 1, only use pseudochromosomes 1-18 (-rf). We specified we wanted to compute a folded SFS (-fold 1) by specifying the reference as the ancestral allele (-anc). Then we used realSFS, part of the ANGSD toolsuite to calculate the SFS in 20Mb windows. The count of heterozygous sites was divided by the total count of sites to obtain the individual estimate of genome-wide heterozygosity.

To estimate population levels of diversity, we pooled results from individuals belonging to the same population (Europe, Japan, North America - Siberia was represented by the single individual) and calculated the mean heterozygosity and standard deviations for each population using R v4.2.1 [43]. We also pooled all European individuals apart from those from the Apennine mountains due to their high levels of inbreeding and low diversity not representing their natural levels, as they were caused by human disturbances [18]. Significance of differences between the mean values in each population was calculated using a Welch unpaired two-sample t-test in R.

Runs of homozygosity

We calculated runs of homozygosity (ROH) in the modern individuals with >5x mean coverage using PLINK based on a previous approach [44]. We generated a PLINK file from the pseudochromosomes 1-18 from all individuals using ANGSD (-doPlink 2) and the following parameters: -minmapQ 30 -minQ 30 -GL 2 -doMajorMinor 1 -doMaf 2 -SNP_pval 1e-6 -minmaf 0.05 -skiptriallelic 1 -uniqueonly 1 -minind 34 -b ROH_bamlist -doPlink 2 -doGeno -4 -doPost 1 -postCutoff 0.95. We ran the resultant PLINK file in PLINK to calculate the RoH using the following parameters: --homozyg-snp 50 --homozyg-kb 500 --homozyg-density 50 --homozyg-gap 1000 --homozyg-window-snp 50 --homozyg-window-het 5 --homozyg-window-missing 5 --homozyg-window-threshold 0.05 --allow-extra-chr. We

filtered the output into four different ROH categories: 500kb - 1Mb, 1Mb - 2Mb, 2Mb - 5Mb, >5Mb.

We estimated the number of generations since inbreeding occurred using the calculation g = 100/(2rL)[45], where r = recombination rate, L = length of ROH in Mb, and g= number of generations. As genome-wide recombination rates for brown bears are unavailable, we present results based on the dog (1cM per Mb) [46]. Given this calculation, ROH>0.5Mb equates to inbreeding occuring within the last 100 generations, ROH>1Mb equates to inbreeding occuring within the last 50 generations, ROH>2Mb equates to inbreeding occuring within the last 25 generations, and ROH>5Mb equates to inbreeding occuring within the last 10 generations.

Results

Population structure

In the PCA analyses we observe clusters corresponding to three broad geographic regions consisting of Europe/Siberia, Japan and Kunashir Island, and North America, regardless of base call method (Fig 1B and Supplementary Fig S1). Assuming three clusters, all ancient individuals fall within the diversity of their modern counterparts, except one of the historical samples from Kunashir Island (Kunashir2). This individual showed an intermediate placement between Japanese and Siberian individuals in the PCA based on genotype likelihoods (Fig 1B), which is likely driven by the relatively high sequencing and DNA damage error rates observed in this individual, rather than reflecting a biological signal (Supplementary Fig S2). Therefore, this individual is not considered in interpretations of population structure.

The Europe/Siberia cluster consists of two subclusters along the PC2 axis: one cluster comprising contemporary individuals from continental Europe and Holocene individuals from Ireland; the other cluster containing Late Pleistocene individuals from mainland Europe, Ireland, and Siberia, and the modern Siberian individual. However, the placement of the Late Pleistocene Irish individuals slightly differs depending on base call methods. In the PCA based on genotype likelihoods, they cluster closer to the Siberian individuals (Fig 1B). In the PCA based on pseudohaploid data, they cluster intermediate between Siberian individuals and contemporary+Holocene European individuals (Supplementary Fig S1).

Phylogenomic analyses based on data from contemporary individuals showed four major groupings in our dataset, which group by geography; the three investigated European individuals clustered together, the three investigated Japanese individuals clustered together, and the three investigated North American individuals clustered together (Supplementary Fig S3). The Siberian individual S235 did not belong to either of the three monophyletic groups. At the root of the tree, we found low levels of local posterior support, suggesting high levels of discordance in relation to the branching pattern of the four groups. Further investigations into the branching pattern of the four groups, returned very similar counts for all topologies (Fig 1C). A further test found that all four groups represent a polytomy and therefore probably diverged at very similar times, and topologies are therefore highly influenced by incomplete lineage sorting and/or gene flow.

Demographic analyses

Based on the PSMC analysis, the four genomes show identical demographic trajectories until ~150 kya (Fig 2). Following the split of the European lineage from the others ~100 kya, it appears all four populations started to decline at ~60 kya.

Results from the genome-wide Tajima's D returned mean values of -0.25 for Europe, 0.45 for Japan, and 0.55 for North America; a positive Tajima's D indicates a recent demographic contraction, whereas a negative value indicates a demographic expansion [47]. We could not estimate Tajima's D for Siberia, as the analysis requires multiple individuals per population. All results were significantly different from 0 (p-value < 2.2e-16) as determined by a Wilcoxon signed rank test with continuity correction.

Genetic diversity

We found broadly overlapping levels of individual genome-wide heterozygosity across the species range (Fig 3A). The lowest heterozygosity levels were found in individuals from the Apennine mountains in Italy (Europe), which are already known to have very low diversity [18]. With their exclusion, the remaining European individuals had a relatively high mean heterozygosity (0.00204; SD=0.00059), second only to the Siberian individual (0.00232; SD=0.00037). The Japanese individuals had a mean value of 0.00188 (SD=0.00035), and the North American individuals had the lowest level of heterozygosity (0.00171; SD=0.00044). All populations had significantly different (p>>0.01) mean levels of heterozygosity (Supplementary table S2).

All individuals contained some levels of ROH, but some had more than others (Fig 3B). The highest levels of recent inbreeding events (ROH >5Mb) were found in European individuals from the Apennine mountains (Italy) and from Spain. The Siberian individual had the lowest levels of ROH.

Discussion

By combining contemporary and ancient (Late Pleistocene and Holocene) genomic data from brown bears sampled across their Holarctic range (Fig 1A), we uncover new patterns in the historical biogeography of Eurasia and North America. Utilising population structure (PCA), phylogenomic, and demographic analyses, we show clear phylogeographic subdivision in contemporary individuals, which comprises geographically structured populations: Europe, Siberia, Japan, and North America. Phylogenomic analyses suggest these four populations represent a polytomy, with demographic results indicating they diverged 150-100 kya. The age of diversification of the contemporary brown bear populations is relatively young, compared to the estimated divergence of brown bears from polar bears between ~1Ma [48] and ~400 kya [14], and brown bear fossils aged 500 kya recovered in Europe [49]. Therefore, the modern populations likely represent only a remnant of prior lineages that went extinct between the emergence of brown bears >400 kya and the divergence of contemporary populations ~150 kya. The timing of the contemporary population split is broadly consistent with the onset of the last glacial period ~115 kya [50], which witnessed significant environmental change, which may have led to a decline in the

global diversity of brown bears. The loss of brown bear lineages and the subsequent emergence of new areas of habitat may have led to the synchronous diversification of contemporary brown bears populations from a common ancestor. The origin of the modern North American population ~100 kya is in accordance with the fossil record; brown bears have a limited history in North America, appearing in eastern Beringia only ~70 kya [51,52]. Fossils in Eurasia are much older (~500 kya), suggesting a Eurasian origin of brown bears [53,54].

Late Pleistocene individuals from Japan and Siberia, and early Holocene North America, fall within or very close to their contemporary counterparts (Fig 1B). This suggests population continuity in these regions over at least the last 11,000 to 50,000 years. Population continuity in North America and Japan is in contrast with previous hypotheses of multiple migration events based on mitochondrial genomes [9,11]. However, as we uncovered large numbers of phylogenomic discordances across the genome (Fig 1C), likely linked to the synchronous divergence of contemporary lineages, the mitochondrial genome is very limited in resolution as it only represents a single loci and can be heavily biased by gene flow and incomplete lineage sorting.

A pattern of population continuity is not obvious in Europe, where Late Pleistocene individuals are more closely related to the modern Siberian individual than to their modern European counterparts (Fig 1B). This suggests the modern Siberian population may have been more widely distributed than today, and suggests a population replacement in Europe towards the end of the Late Pleistocene. This is in line with the previous hypothesis of Central Asia as the central point of origin and diversification of brown bears [55], and of bears from the Carpathian mountains recolonising Europe after the LGM. More genomes from western/central Asian brown bears, both contemporary and ancient, will help to resolve the origin of contemporary European bears.

Despite uncertainty in the geographic origin of contemporary European brown bears, the differences in the clustering of the Irish individuals based on their age (mid-late Holocene and Late Pleistocene) allows us to make inferences about the timing of their origins. Although our PCA results present the Late Pleistocene Irish individuals as either the same population as the Siberian/Late Pleistocene Austrian cluster (Fig 1B) or as intermediate between the contemporary European/Holocene Irish and the Siberian/Late Pleistocene Austrian clusters (Supplementary Fig S1) depending on base call method, making us uncertain of their exact position, it still shows them as different from the Holocene Irish individuals. Therefore, our findings indicate the population replacement in Europe occurred after the end of the Pleistocene, and before the mid-Holocene. The Late Pleistocene-Holocene transition was a period of major climatic upheaval, which influenced the diversity and distribution of extant taxa [56]. Higher δ^{15}N values, indicating more carnivorous diets, have been commonly documented in Late Pleistocene brown bears [20]. However, lower levels of δ^{15}N have also been uncovered towards the end of the Late Pleistocene and into the Holocene in Europe, which has been proposed to be due to a shift toward increasing herbivory [10]. Based on our findings, we suggest that instead of a shift in diet of Late Pleistocene

individuals, it may instead have been a shift in individuals themselves to another population, with a different diet that was more suited to the new environment. This theory is supported by our Tajima's D results showing a population expansion in Europe. This also supports previous results that brown bears recolonised Europe at the end of the Last Glacial Maximum (LGM) faster than any other carnivore of the Holocene fauna [57].

Our demographic results show a decrease in effective population size shortly after the divergence of the four modern populations ~50 kya. While this could be the result of climate change induced population declines, it could also be due to the reduction of gene flow between populations as they geographically separated from one another. In general, contemporary European individuals have the highest levels of genetic diversity, with relatively low levels of inbreeding, apart from a few outliers that were most impacted by humans [18] (Fig 3). High levels of diversity may have been key in adapting to the Holocene, allowing them to outcompete other bears. Unfortunately, we only had data from one contemporary Siberian individual, and can therefore only make population-level assumptions based on it. However, one of the highest levels of heterozygosity and the lowest levels of ROH are found in this individual, which could indicate a continually large effective population size. This is also supported by our PSMC, which shows higher recent effective population size than in the other populations.

Data accessibility

The raw Illumina sequence data for individual JBB-32K are available under DDBJ DRA ID: DRA015631. The raw Illumina sequence data for individual CGG_1_020006 are available under Bioproject PRJNA930648 on NCBI.

Funding

This study was supported by Grants-in-Aid for Scientific Research (no 22K18730) from the Japan Society for the Promotion of Science (JSPS), the Villum Fonden Young Investigator Programme, grant no. 13151.

References:

1. Stuart AJ. 2015 Late Quaternary megafaunal extinctions on the continents: a short review. *Geol. J.* **50**, 338–363.

2. Campos PF *et al.* 2010 Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (*Ovibos moschatus*) population dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5675–5680.

3. Campos PF *et al.* 2010 Ancient DNA sequences point to a large loss of mitochondrial genetic diversity in the saiga antelope (*Saiga tatarica*) since the Pleistocene. *Mol. Ecol.* **19**, 4863–4875.

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

5. Niedziałkowska M, Hayward MW, Borowik T, Jędrzejewski W, Jędrzejewska B. 2019 A meta-analysis of ungulate predation and prey selection by the brown bear *Ursus arctos* in Eurasia. *Mammal Research* **64**, 1–9.

6. Bojarska K, Selva N. 2012 Spatial patterns in brown bear *Ursus arctos* diet: the role of geographical and environmental factors. *Mamm. Rev.* **42**, 120–143.

7. Albrecht J, Bartoń KA, Selva N, Sommer RS, Swenson JE, Bischof R. 2017 Humans and climate change drove the Holocene decline of the brown bear. *Sci. Rep.* **7**, 10399.

8. Marciszak A, Schouwenburg C, Lipecki G, Talamo S, Shpansky A, Malikov D, Gornig W. 2019 Steppe brown bear *Ursus arctos 'priscus'* from the Late Pleistocene of Europe. *Quat. Int.* **534**, 158–170.

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

10. Ersmark E *et al.* 2019 Genetic turnovers and northern survival during the last glacial maximum in European brown bears. *Ecol. Evol.* **9**, 5891–5905.

11. Segawa T *et al.* 2021 Ancient DNA reveals multiple origins and migration waves of extinct Japanese brown bear lineages. *R Soc Open Sci* **8**, 210518.

12. Westbury MV *et al.* 2020 Hyena paleogenomes reveal a complex evolutionary history of cross-continental gene flow between spotted and cave hyena. *Science Advances* **6**, eaay0456.

13. Liu S *et al.* 2021 Ancient and modern genomes unravel the evolutionary history of the rhinoceros family. *Cell* **184**, 4874–4885.e16.

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

15. Miller W *et al.* 2012 Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2382–90.

16. Cahill JA *et al.* 2018 Genomic Evidence of Widespread Admixture from Polar Bears into Brown Bears during the Last Ice Age. *Mol. Biol. Evol.* **35**, 1120–1129.

17. Endo Y, Osada N, Mano T. 2021 Demographic History of the Brown Bear (Ursus arctos) on Hokkaido Island, Japan, Based on Whole-Genomic Sequence Analysis. *Genome Biol.*

18. Benazzo A *et al.* 2017 Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9589–E9597.

19. Orlando L *et al.* 2013 Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74–78.

20. Rey-Iglesia A, García-Vázquez A, Treadaway EC, van der Plicht J, Baryshnikov GF,

Szpak P, Bocherens H, Boeskorov GG, Lorenzen ED. 2019 Evolutionary history and palaeoecology of brown bear in North-East Siberia re-examined using ancient DNA and stable isotopes from skeletal remains. *Sci. Rep.* **9**, 4462.

21. Dabney J *et al.* 2013 Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15758–15763.

22. Meyer M, Kircher M. 2010 Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448.

23. Glocke I, Meyer M. 2017 Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome Res.* **27**, 1230–1237.

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

25. Barlow A *et al.* 2018 Partial genomic survival of cave bears in living brown bears. *Nat Ecol Evol* **2**, 1563–1570.

26. 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. *Mol. Ecol.* **24**, 1205–1217.

27. Schubert M *et al.* 2014 Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat. Protoc.* **9**, 1056–1082.

28. Schubert M, Lindgreen S, Orlando L. 2016 AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88.

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

30. Castruita JAS, Westbury MV, Lorenzen ED. 2020 Analyses of key genes involved in Arctic adaptation in polar bears suggest selection on both standing variation and de novo mutations played an important role. *BMC Genomics* **21**, 1–8.

31. Li H *et al.* 2009 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

32. Broad institute. 2019 Picard Toolkit. See http://broadinstitute.github.io/picard.

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

34. Chen S, Zhou Y, Chen Y, Gu J. 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890.

35. Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356.

36. Meisner J, Albrechtsen A. 2018 Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics* **210**, 719–731.

37. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321.

38. Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014 ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* **30**, i541–8.

39. Sayyari E, Mirarab S. 2018 Testing for Polytomies in Phylogenetic Species Trees Using Quartet Frequencies. *Genes* **9**. (doi:10.3390/genes9030132)

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

41. Korneliussen TS, Moltke I, Albrechtsen A, Nielsen R. 2013 Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics* **14**, 289.

42. Sánchez-Barreiro F *et al.* 2021 Historical population declines prompted significant genomic erosion in the northern and southern white rhinoceros (Ceratotherium simum). *Mol. Ecol.* **30**, 6355– 6369.

43. R Core Team. 2013 R: A language and environment for statistical computing.

44. Foote AD *et al.* 2021 Runs of homozygosity in killer whale genomes provide a global record of demographic histories. *Mol. Ecol.* **30**, 6162–6177.

45. Kardos M *et al.* 2018 Genomic consequences of intensive inbreeding in an isolated wolf population. *Nat Ecol Evol* **2**, 124–131.

46. Wong AK *et al.* 2010 A comprehensive linkage map of the dog genome. *Genetics* **184**, 595–605.

47. Tajima F. 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

48. Zou T, Kuang W, Yin T, Frantz L, Zhang C, Liu J, Wu H, Yu L. 2022 Uncovering the enigmatic evolution of bears in greater depth: The hybrid origin of the Asiatic black bear. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2120307119.

49. Sardella R, Palombo MR, Petronio C, Bedetti C, Pavia M. 2006 The early Middle Pleistocene large mammal faunas of Italy: An overview. *Quat. Int.* **149**, 104–109.

50. Rasmussen SO *et al.* 2014 A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy. *Quat. Sci. Rev.* **106**, 14–28.

51. Kurten B, Anderson E. In press. 1 980 Pleistocene Mammals of North America. *Columbia Uni-versity Press, New York*

52. Craighead JJ, Mitchell JA. 1982 Grizzly bear. *Mammals of North America. Johns Hopkins University Press, Baltimore, Maryland, USA* , 515–555.

53. Villalba de Alvarado M, Collado Giraldo H, Arsuaga JL, Bello Rodrigo JR, Heteren AH, Gómez-Olivencia A. 2022 Looking for the earliest evidence of*Ursus arctos*LINNAEUS, 1758 in the Iberian Peninsula: the Middle Pleistocene site of Postes cave. *Boreas* **51**, 159–184.

54. Wagner J. 2010 Pliocene to early Middle Pleistocene ursine bears in Europe: a taxonomic overview. *Journal of the National Museum (Prague), Natural History Series* **179**, 197–215.

55. Kurtén B. 1968 Pleistocene mammals of Europe: 1-317 (Weidenfeld & Nicolson, London)., 1971. *The age of mammals* , 1–250.

56. Lorenzen ED *et al.* 2011 Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359–364.

57. Sommer RS, Benecke N. 2005 The recolonization of Europe by brown bears Ursus arctos Linnaeus, 1758 after the Last Glacial Maximum. *Mamm. Rev.* **35**, 156–164.

Figures:

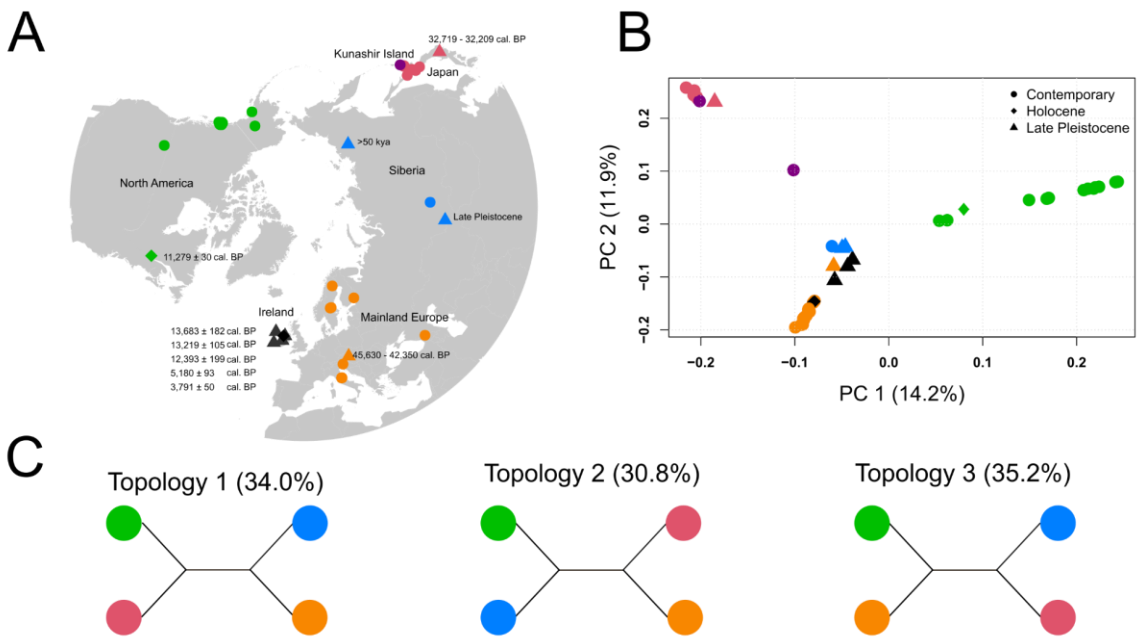


Figure 1: Population structure and phylogenomic results. A) Map showing the localities of the brown bear genomes included in this study. Ages of ancient samples are listed next to their locality. 'Late Pleistocene' indicates a direct age is unknown but it is listed as Late

Pleistocene in the original publication [23]. Sample >50 kya was dated to beyond the limit of radio carbon dating. B) PCA generated from genotype likelihoods. Percentage values in parentheses on the X and Y axes show the percentage of variation explained by the principal component. Colours and shapes represent geography and age respectively. C) Sliding window phylogenomic tree results using three individuals per population except for Siberia which was a single individual. Percentages show the proportion of windows supporting that topology. Colours in B and C align with the corresponding colours on the map.

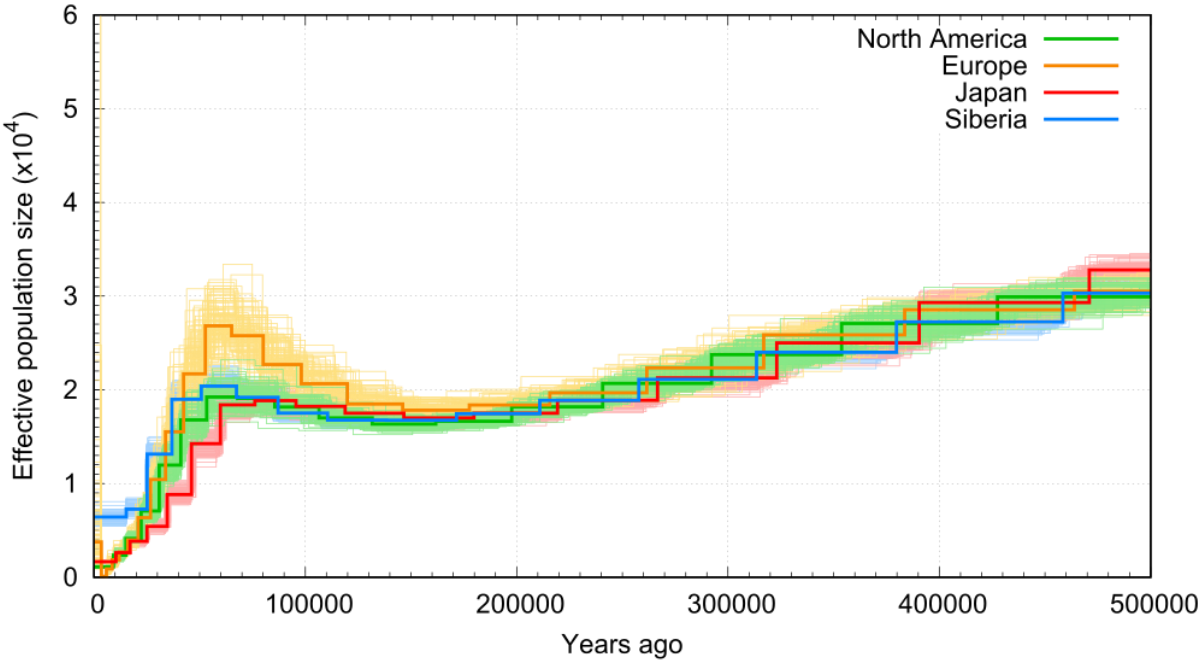


Figure 2: Effective population sizes over the past 500,000 years calculated using PSMC on single individual representatives for each population. Faded lines show the 100 bootstrap replicates produced for each species.

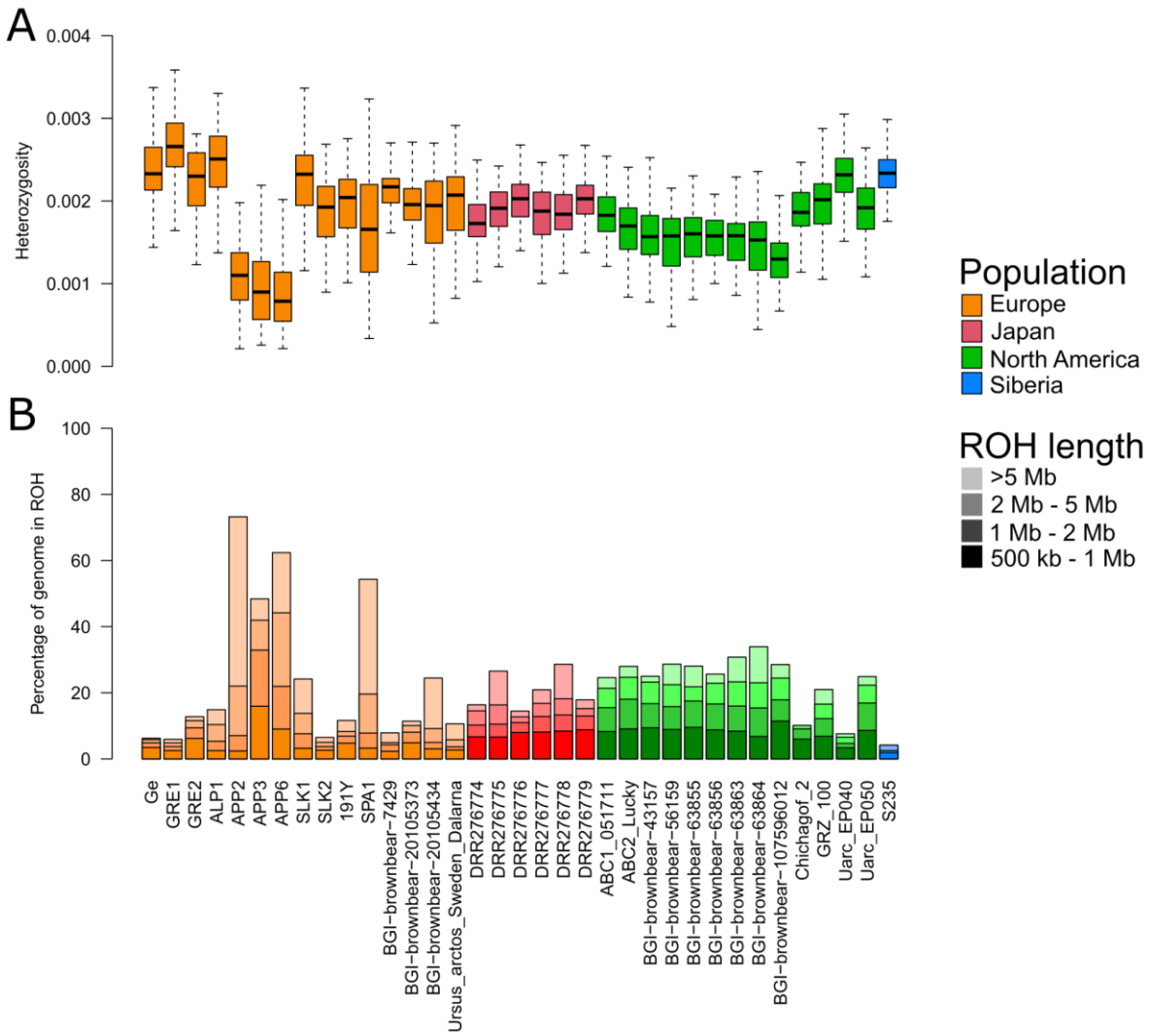
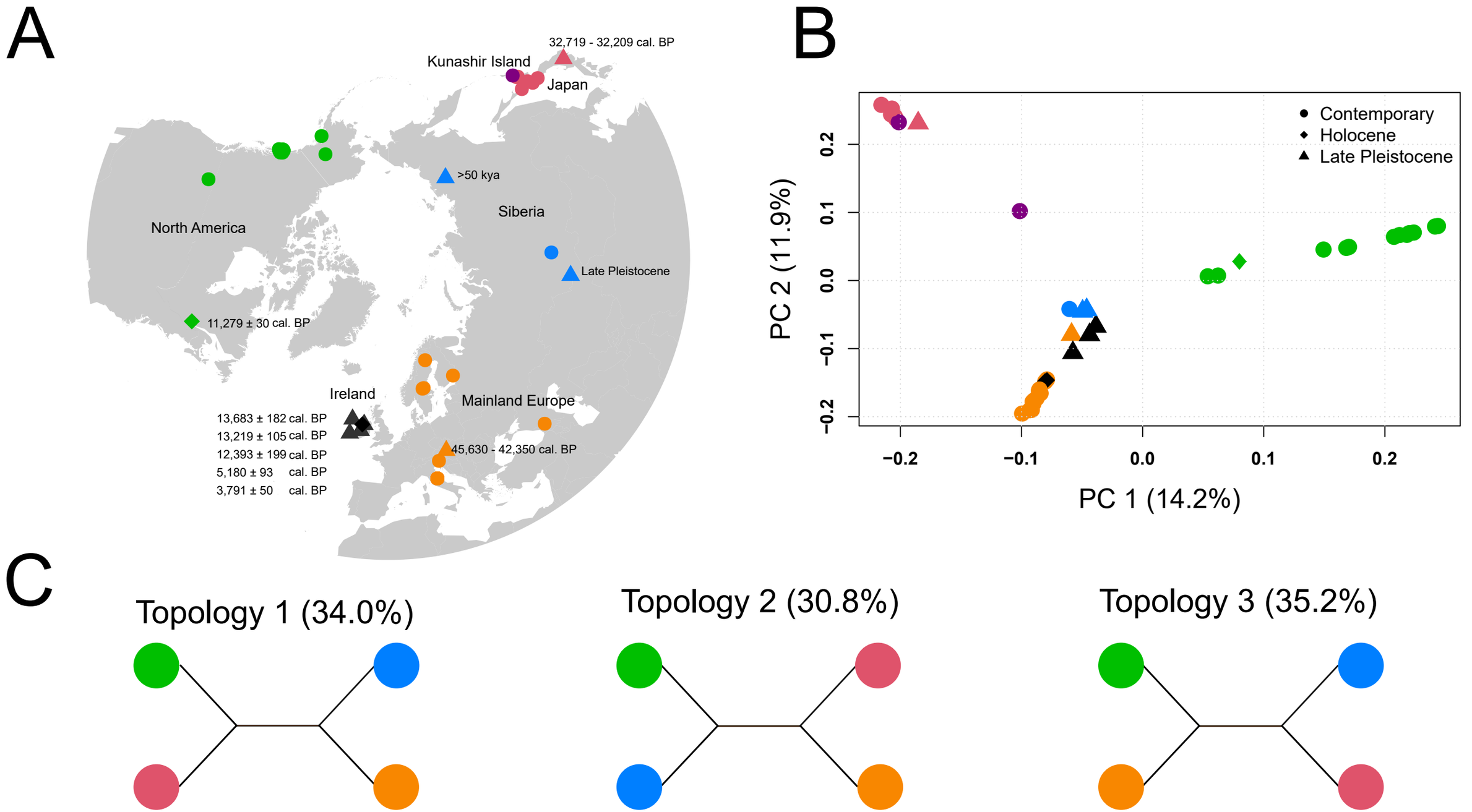
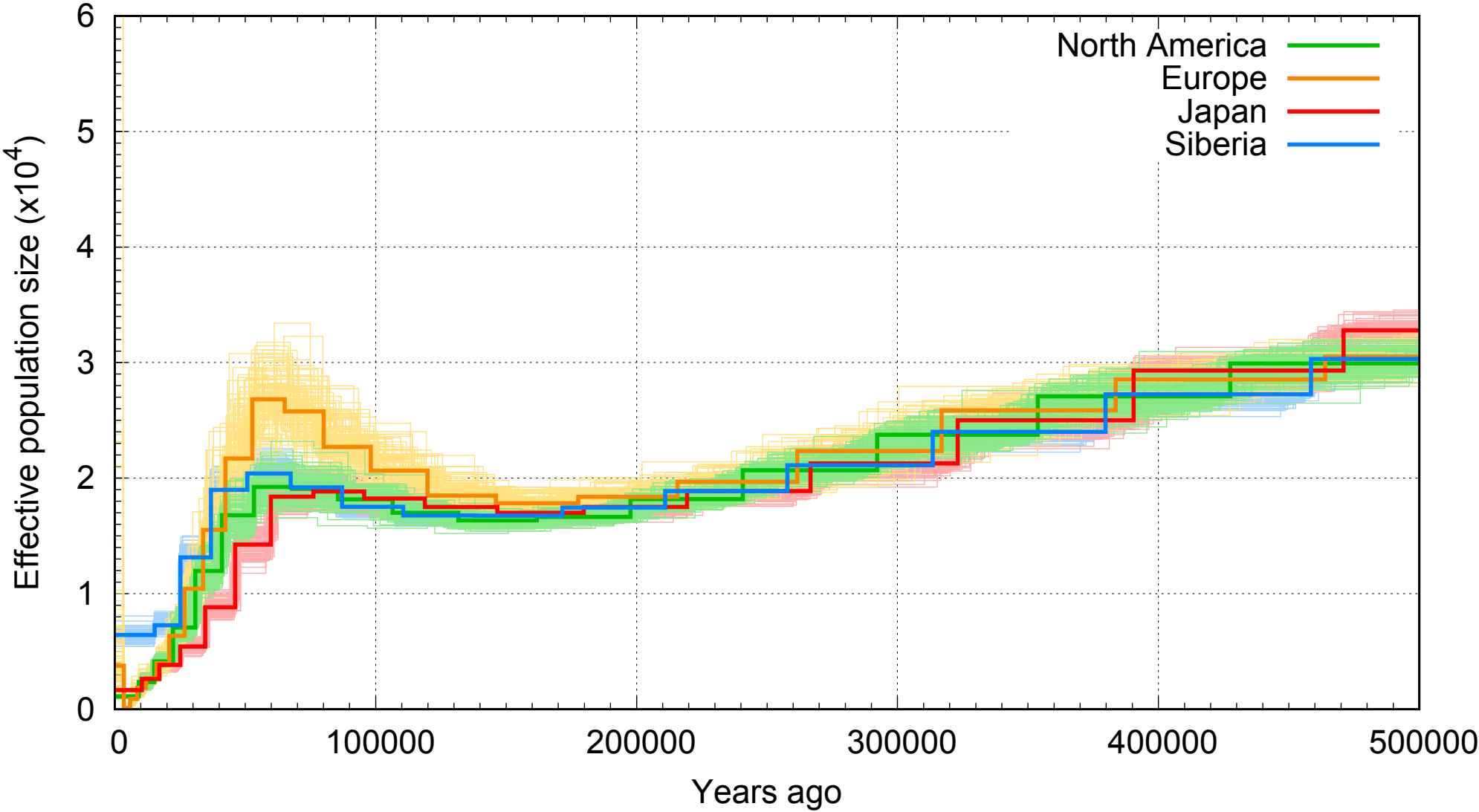
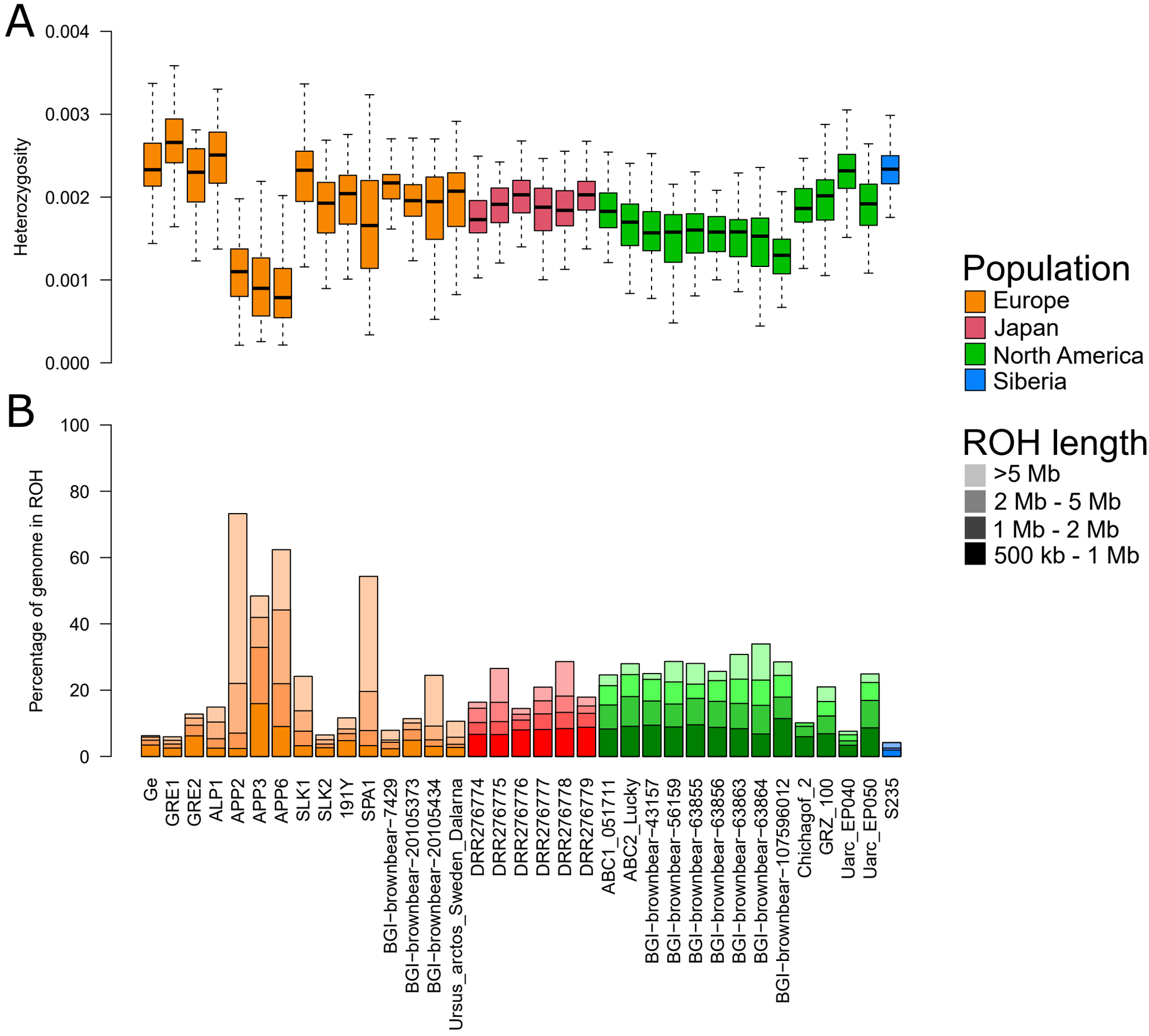


Fig 3: Genetic diversity and runs of homozygosity estimates. (A) Autosome-wide heterozygosity calculated using ANGSD for each individual with >5x mean coverage. Variation was calculated using 20Mb windows of covered bases. (B) Autosomal runs of homozygosity split into different ROH sizes for all individuals with >5x mean coverage calculated using PLINK.


A
32,719 - 32,209 cal. BP
Kunashir Island
Japan
>50 kya
Siberia
North America
Late Pleistocene
11,279 ± 30 cal. BP
Ireland
Mainland Europe
13,683 ± 182 cal. BP
13,219 ± 105 cal. BP
12,393 ± 199 cal. BP
5,180 ± 93 cal. BP
3,791 ± 50 cal. BP
45,630 - 42,350 cal. BP
B
Contemporary
Holocene
Late Pleistocene
PC 2 (11.9%)
PC 1 (14.2%)
−0.2
−0.1
0.0
0.1
0.2
C
Topology 1 (34.0%)
Topology 2 (30.8%)
Topology 3 (35.2%)


North America
Europe
Japan
Siberia
Effective population size (x10^4)
0
1
2
3
4
5
6
0
100000
200000
300000
400000
500000
Years ago


A
Heterozygosity
0.004
0.003
0.002
0.001
0.000
B
Percentage of genome in ROH
100
80
60
40
20
0
Ge
GRE1
GRE2
ALP1
APP2
APP3
APP6
SLK1
SLK2
191Y
SPA1
BGI−brownbear−7429
BGI−brownbear−20105373
BGI−brownbear−20105434
Ursus_arctos_Sweden_Dalarna
DRR276774
DRR276775
DRR276776
DRR276777
DRR276778
DRR276779
ABC1_051711
ABC2_Lucky
BGI−brownbear−43157
BGI−brownbear−56159
BGI−brownbear−63855
BGI−brownbear−63856
BGI−brownbear−63863
BGI−brownbear−63864
BGI−brownbear−107596012
Chichagof_2
GRZ_100
Uarc_EP040
Uarc_EP050
S235
Population
Europe
Japan
North America
Siberia
ROH length
>5 Mb
2 Mb - 5 Mb
1 Mb - 2 Mb
500 kb - 1 Mb