Classification: Biological Sciences, Evolution

Survival and divergence in a small group:
the extraordinary genomic history of the endangered Apennine brown bear stragglers

Short title: Extreme genomic effects in an isolated bear population

Andrea Benazzo[a,*], Emiliano Trucchi[a,b,*], James A. Cahill[c], Pierpaolo Maisano Delser[d,e,f], Stefano Mona[d,e], Matteo Fumagalli[g], Lynsey Bunnefeld[h,i], Luca Cornetti[j], Silvia Ghirotto[a], Matteo Girardi[k], Lino Ometto[l,m], Alex Panziera[a], Omar Rota-Stabelli[l], Enrico Zanetti[a], Alexandros Karamanlidis[n], Claudio Groff[o], Ladislav Paule[p], Leonardo Gentile[q], Carles Vilà[r], Saverio Vicario[s], Luigi Boitani[t], Ludovic Orlando[u], Silvia Fuselli[a], Cristiano Vernesi[k], Beth Shapiro[c], Paolo Ciucci[t], Giorgio Bertorelle[a,1].

[a]Department of Life Sciences and Biotechnology, University of Ferrara, 44121 Ferrara, Italy
[b]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo 1066, Norway
[c]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA
[d]Institute de Systematics, Evolution, Biodiversite, ISYEB-UMR 7205-CNRS, MNHN, UPMC, EPHE, Ecole Pratique de Hautes Etudes, CP39, 75005 Paris, France
[e]EPHE, PSL Research University, 75005 Paris, France
[f]Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland.
[g]Department of Life Sciences, Silwood Park Campus, Imperial College London, Ascot, SL5 7PY, United Kingdom
[h]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, EH9 3FL, United Kingdom
[i]Biological and Environmental Science, University of Stirling, Stirling, FK9 4LA, United Kingdom
[j]Zoological Institute, University of Basel, 4051 Basel, Switzerland
[k]Department of Biodiversity and Molecular Ecology, Fondazione Edmund Mach, 38010 San Michele all'Adige, Italy
[l]Department of Sustainable Agro-ecosystems and Bioresources, Fondazione Edmund Mach, 38010 San Michele all'Adige, Italy
[m]Independent researcher, 38016, Mezzocorona, Italy
[n]ARCTUROS, Protection and Management of Wildlife and the Natural Environment, 53075 Aetos, Florina, Greece
[o]Provincia Autonoma di Trento, 38100, Trento, Italy
[p]Department of Phytology, Faculty of Forestry, Technical University, 96053, Zvolen, Slovakia
[q]Veterinary Service, National Park of Abruzzo Lazio and Molise, 67032 Pescasseroli, Italy
[r]Doñana Biological Station (EBD-CSIC), 4102 Seville, Spain
[s]Institute of Atmospheric Pollution Research and Technologies, National Research Council, 70126 Bari, Italy
[t]Department of Biology and Biotechnologies 'Charles Darwin', University of Rome La Sapienza, 00185 Rome, Italy
[u]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 K Copenhagen, Denmark

[1]*Corresponding author:* Giorgio Bertorelle, Department of Life Sciences and Biotechnology, University of Ferrara, via Borsari 46, 44121 Ferrara, Italy. Phone: +390532455743. Email:ggb@unife.it

[*]*Equally contributing authors*

Keywords: balancing selection, genetic drift, genetic load, *Ursus arctos*, Apennine bear, Neolithic impact

Abstract

About 100 km east of Rome, in the Central Apennine mountains, a critically endangered population of approximately fifty brown bears live in complete isolation. Mating outside this population is prevented by several hundred kilometers of bear-free territories. We exploited this natural experiment to better understand the gene and genomic consequences of surviving at extremely small population size. First, we found that brown bear populations in Europe lost connectivity since Neolithic times, when farming communities expanded and forest burning was used for land clearance. In Central Italy, this resulted in a 40-fold population decline. The overall genomic impact of this decline included the complete loss of variation in the mitochondrial genome and along long stretches of the nuclear genome. Several private and deleterious amino acid changes were fixed by random drift; predicted effects include energy deficit, muscle weakness, anomalies in cranial and skeletal development, and reduced aggressiveness. Despite this extreme loss of diversity, Apennine bear genomes show non-random peaks of high variation, possibly maintained by balancing selection, at genomic regions significantly enriched for genes associated with immune and olfactory systems. Challenging the paradigm of increased extinction risk in small populations, we suggest that random fixation of deleterious alleles a) can be an important driver of divergence in isolation, b) can be tolerated when balancing selection prevents random loss of variation at important genes and c) is followed by or results directly in favorable behavioral changes.

Significance

A small and relict population of brown bears lives in complete isolation in the Italian Apennine mountains, providing a unique opportunity to study the impact of drift and selection on the genomes of a large endangered mammal and to reconstruct the phenotypic consequences and the conservation implications of such evolutionary processes. The Apennine bear is highly inbred and harbors very low genomic variation. Several deleterious mutations have been accumulated by drift. We found evidence that this is a consequence of habitat fragmentation in the Neolithic, when human expansion and land clearance shrank its habitat, and that retention of variation at immune system and olfactory receptor genes, as well as changes in diet and behavior, prevented the extinction of the Apennine bear.

\body
Introduction

Living in complete isolation at very small population size can drive a species to extinction because of several processes: (i) inbreeding, which unmasks recessive deleterious mutations, becomes unavoidable when only few potential mates are available (1); (ii) genetic drift tends to prevail over natural selection, limiting adaptation and allowing deleterious variants to increase in frequency, possibly until fixation (2, 3); and (iii) low levels of variation are expected, reducing the chances for an individual and for a population to have the genotypes best matching the environmental challenges (4).

This conservation paradigm, strictly related to the extinction vortex metaphor (5), is supported by empirical evidence (6–8), but it is challenged by studies showing that selection can be powerful also at small population sizes (9, 10) and that survival and even demographic expansion can occur with almost no genomic variation (11). Interestingly, if extinction does not occur, drift in small isolated groups can produce, or contribute to, genetic and phenotypic divergence, possibly leading to speciation (12, 13).

To explore the pattern of genomic variation and divergence in a large mammal living in isolation at small population size, but apparently not yet in the extinction vortex (14), we focused our attention on the last population of native Italian brown bear, the Apennine bear.

The Apennine bear consists of a critically endangered population of approximately fifty individuals living in the Apennine mountains in Central Italy (Fig. 1A; SI Appendix, Fig. S1), and is usually classified as the endemic subspecies *Ursus arctos marsicanus* (14–16). Several hundred kilometers preclude any opportunity for natural gene flow between the Apennine bear and the closest brown bear populations in the Alps, even considering the peripheral area of occasional occurrence (see SI Appendix, section S1.1). Human persecution has consistently reduced the geographic range and number of Apennine bears in the last few centuries. Even after the establishment of the Abruzzo Lazio and Molise National Park in 1923, and the introduction of legal protection of this species since 1939, the population failed to increase, with reported mortalities due mainly to deliberate or accidental killing (14). Mitochondrial DNA data support a strict genetic affinity of the Apennine bear with the geographically closest brown bear populations in the Alps and the Balkans (17–19). Nevertheless, some level of morphological divergence, due to drift or possibly associated with a diet shift, was inferred from the analyses of skull traits (16). How long Apennine bear has lived in isolation is not known, but historical records (20) and mtDNA and microsatellite diversity (17)

point to at least four to five centuries (40-50 generations).

In this paper we produced and analyzed whole genomes from brown bears sampled in the Apennine and in other European regions. We reconstruct the demographic history of fragmentation and population size reduction within Europe. We then take advantage of the Apennine bear as the result of a unique natural experiment designed to infer the effects of genetic drift on the genomes and phenotypes of a small and isolated population. As the Apennine bear is an iconic taxon of great interest for conservation, but the extinction risks are still unknown, our data will provide a genomic view into the current debate on the best strategy to protect this population.

Results and discussion

We performed whole-genome sequencing of six Apennine bears and six additional European brown bears from Greece (two individuals), Slovakia (two individuals), and one each from the Alps and Western Spain (Fig. 1A; SI Appendix, Table S1). The Alpine individual was born in South Slovenia but sampled about 400 km west in the Adamello National Park in Italy, where it was released in year 2000 during a reintroduction project (21). Paired-end Illumina sequences were aligned to the polar bear reference genome (22), and data analyses were performed on these and other available whole-genome data (SI Appendix, Table S1) from two Swedish brown bears, six polar bears, one black bear, and the panda. Four individuals (one from the Apennine, one from the Alps, one from Slovakia and one from Spain) were sequenced at an average depth of coverage of ~15X, versus ~5X for all other individuals. One Apennine individual was excluded from downstream analysis of nuclear sequences due to lower depth of coverage (2.4X). Genetic variation statistics and downstream statistical analyses, when based on ~5X genomes, were computed using genotype likelihoods.

Pattern of variation and inbreeding estimates

All the Apennine bears in our sample show the same mitochondrial genome sequence. Their nuclear genome has approximately 2/3 and 1/3 of the heterozygous sites observed in the Spanish bear and the other brown bears, respectively, and twice as much variation as the polar bear (Fig. 1B; SI Appendix, Tables S3 and S4). Direct estimates on >15X genomes and likelihood-based estimates on ~5X genomes produced very similar values. Apennine bears show strong patterns in the amount of variation along the nuclear genomes, with long stretches of several megabases of no or almost no variation, alternating with regions characterized by levels of variation similar to those observed in other brown bears (Fig. 1E). Long regions of homozygosity, identified as genomic regions >1Mb

4

where the number of heterozygous sites is never larger than 25 every 50kb, cover approximately three quarters of the Apennine genome (Fig. 1C), and have on the average 5 heterozygous sites every 200kb. This pattern is not seen in other brown bear or polar bear genomes.

The distribution of genetic variation along individual genomes is informative of recent inbreeding (23, 24). In particular, inbred individuals are expected to have long genomic fragments with both average and low heterozygosity values, producing a bimodal distribution of variation when estimated in windows across the genome (25). While all analyzed brown bear individuals produced a bimodal distribution of heterozygosity estimates, the estimated inbreeding coverage, F, which is based on estimate of the proportion of the genome with low heterozygosity, varied significantly across individuals (SI Appendix, section S2). Apennine bears are highly inbred, with F values between 0.69 and 0.77, with the Spanish bear as the next most inbred ($F = 0.57$). All other brown bears show F values lower than 0.29. Two points should be noted here: First, inbreeding coverage is known to be correlated with the inbreeding coefficient estimated with population genetics methods or from pedigrees (26), meaning that the F values estimated in the Apennine bears correspond to the value expected in a large population after six generations of full sib mating. Second, F values are not solely a consequence of the average variation levels. Polar bears have lower genetic variation than Apennine bears but they also have very different distribution of variation along the genome (Fig. 1C, E), with either zero or close to zero estimated inbreeding coefficients (SI Appendix, Fig. S4). This is probably a consequence of a longer history at low density and high connectivity in polar compared to Apennine bear (22, 27). Apennine bear genomes likely accumulated the effects of strong inbreeding occurring recently, but the fraction of their genomes at high variation still reveals a past history at much higher effective population size.

Demographic dynamic
The inferred demographic trajectory of the Apennine bear prior to 10 thousand years ago is similar to what is observed in other areas in Europe (Fig. 2A, SI Appendix, Fig. S6) and in Alaska (27). We observe a long term effective population size fluctuating around 20 to 40 thousand individuals, followed by a drastic decline by approximately one order of magnitude starting approximately 100 KY (thousand of years) ago. The simplest explanation for this decline is the climatic cooling associated with the last glacial period, which began after ~110,000 years ago (28). We find evidence for two different dynamics in brown bear populations at the onset of Holocene warming ~15-10 KY ago (Fig. 2A, SI Appendix, Fig. S6): 1) the "continental pattern", where the decline stops and the effective population sizes remain constant at around 2-3 thousand individuals in central and

5

northern Europe (SLK1, ALP1, and SWE1 in SI Appendix, Fig. S6); and 2) the "peninsular pattern" with a transitory expansion, which is evident in Spain but possible also in the Apennine bear based on the bootstrap analyses, followed by an additional and final decline to a very small population size (SPA1 and APN2 in SI Appendix, Fig. S6).

Demographic dynamics reconstructed from single genomes using PSMC proved very reliable under some simulation scenarios (29), but tend to smooth out steep changes (30), and should be considered with caution when the inference refers to recent events (31–33) or coverage is low (34). We therefore further explored the main demographic process that shaped the Apennine genomic variation by performing two additional analyses based on different statistical approaches and exploiting different information in the data. First, we used the approximate Bayesian coalescent approach (35), with site frequency spectra as summary statistics (SI Appendix, section S3.2). This analysis supported a rapid and approximately 40-fold decline of the Apennine bear effective size at approximately 2,500-8,800 years ago (peak probability 4,000 years ago), starting from a population size of several tens of thousands of individuals (Fig. 2B). Second, we simulated single genomes under different bottleneck scenarios, and compared the variation in 50 kb fragments with that observed in the Apennine high coverage genome. The best fit model assumed that population decline began few thousand of years ago, starting from several thousand of individuals and declining to a few hundreds (Fig. 2C). Considering these results and the shared demographic dynamics inferred among all brown bears prior to 10 KY ago, we conclude that the low level of genetic variation found in the Apennine bear is a consequence of its rapid and recent decline in population size. As discussed below, we believe that this result does not necessarily imply a large Apennine bear population in the past, but it is instead the signature of the process of fragmentation of a previously connected pan-European population.

Divergence and fragmentation history
We estimated a whole-genome Neighbor-Joining tree (Fig. 1D), which shows an almost instantaneous split separating brown bears originating in different geographic areas. This topology implies that Apennine bears are differentiated to the same degree from all the other European brown bears, regardless of their geographic distance from those bears. The tree also indicates that Apennine bears are very similar to each other, more than what is observed among individuals from other geographic areas (see also the STRUCTURE analysis in SI Appendix, Fig. S12 and Fig. S13).

We next inferred the history of connectivity among European bear populations using an

Approximate Bayesian Computation approach under a non-equilibrium island model (SI Appendix, section S5.2) (35). We assumed that a meta-population with a large number of demes originated from a single panmictic population. The connectivity parameter Nm (N=effective size, m=migration rate) controls the genetic divergence among demes, and we allowed it to vary in three time intervals. In our model, one of the demes represents the Apennine bear population, and this deme becomes completely isolated at some point in the past. Allele frequency spectra were computed from approximately 5,000 non-coding loci with a length of 4,000 base pairs each and used as summary statistics to compare observed and simulated data, to estimate the posterior densities of Nm, and estimate the timing of evolutionary events. The results of this analysis (Fig. 3) suggested: (i) an ancestral population split occurred a few hundred thousand years ago, compatible with the estimated arrival of the brown bear in Europe (36); (ii) a constant and high migration rate, not lower than several dozen animals per generation, existed between demes since the ancestral split and until a few thousands years ago when the migration rate dropped to a few individuals; and (iii) the Apennine bear has been completely isolated from other European bear populations since approximately 1,500 years ago (SI Appendix, Table S6), but the uncertainty interval for the estimated timing of this event and the timing of disruption of global connectivity overlap.

The switch from very high to very low (or 0) population connectivity can be seen as a fragmentation event with an age that corresponds approximately to the divergence time between groups. We therefore applied the F-statistic approach used to study the Human-Neanderthal split (37) to validate our results. This analysis, which is based on the fraction of derived alleles shared among individuals, suggests that the Apennine bear diverged between 2 and 3 thousand years ago, a date very close to the estimated fragmentation age (SI Appendix, Fig. S15). An additional ABC analysis based on the pairwise comparison among genomes support the view that the Apennine bear does not exchange genes with other populations since a few thousand years (SI Appendix, section 5.3).

In summary, our results (Figs. 1-3) jointly support the following scenario for the brown bear in Europe. After initial colonization, a large and still panmictic population was severely affected by the last glacial interval in terms of effective size but not connectivity. Population decline ceased at the onset of Holocene warming, when population sizes possibly increased in southern areas of Europe. In the last few thousand years, extreme contractions in habitat and population size occurred in Spain and in the Apennine bear, and fragmentation increased among bear populations across Europe. The estimated timing of this fragmentation is compatible with the trend of forest clearance

and land cover change related to the introduction and diffusion of agricultural technologies (38–40). In particular, the ice core levoglucosan flux (a biomarker of fire emissions) started to slowly increase in the Holocene, reached two peaks at around 5.5 and 2.5 KY ago, and then began to decline (40), consistent with the drop in the forest clearance rate in Europe (41). It seems, therefore, that this second demographic decline was caused mainly by the contraction of habitat due to farming, which profoundly modified the genetic structure among bear populations. Humans have further impacted bear populations in more recent times, driving this species to extinction in several regions (42).

mtDNA genomes

The unique and identical mtDNA sequence observed in all Apennine bears is differentiated from the mtDNA genomes found in the Alps (~20 mutations), Greece (~70), Spain (~100), and Slovakia (~300). Our data are compatible with the known phylogeographic structure in mtDNA found in Europe (Fig. 1D): three major clades can be identified, usually called 1a (Spain and Southern Sweden), 1b (Italy, Balkans, and Southern Carpathians), and 3a (North-Eastern Europe), and usually associated with different glacial refugia and postglacial recolonization processes (36). This pattern was not observed in the nuclear genomes, and it implies a strong genetic barrier in Sweden and a strict affinity between Apennine and Alpine bears. The discrepancy between nuclear and mitochondrial data is likely explained by male-biased dispersal (43–46), and the pattern of low genetic structure we observed at the Y-chromosome further supports this view (47, see SI Appendix, section S7) Until recent times, bear populations were geographically homogenized by males, but females philopatry resulted in some level of mtDNA structuring. Interestingly, habitat destruction and fragmentation has been suggested as a general factor that favors the increase of female philopatry (48, 49). Considering also the wide geographic distribution of mtDNA lineages in the past (50), it is possible that sex differences in the dispersal rate increased a few thousand years ago.

Adaptation and maladaptation

The pattern of genomic variation in coding regions is informative on the functional impact of the isolation history of the Apennine bear. We found evidence of two evolutionary processes with opposite outcomes: active maintenance of variation at specific families of genes and fixation by drift of several deleterious alleles.

Three lines of evidence support active maintenance of variation: the comparison between non-

synonymous and synonymous polymorphisms, the comparison between variation erosion along the genome, and the enrichment of specific pathways in the genomic regions that preserved higher variation than expected.

First, the Apennine bear retained more variation at non-synonymous than at synonymous sites (SI Appendix, section S8.1). In particular, when the heterozygous sites observed in a single non-Apennine genome are used as reference, approximately 40% of non-synonymous variants are still polymorphic in five Apennine bears, in contrast with 30% of synonymous variants (SI Appendix, Tables S12-13). This difference reflects a higher probability of retaining non-synonymous than synonymous polymorphisms, and it is not expected either under neutrality (similar fractions are predicted), nor under positive selection during divergence (more non-synonymous fixations are predicted), nor assuming reduced purifying selection due to drift (very few heterozygous sites under negative selection are expected in a single non-Apennine reference genome).

Second, in the genomic regions where variation is very high in non-Apennine bears, which may be due to functional or structural reasons, the loss of variation in the Apennine bears was less intense than in other regions (Fig. 4F). In fact, the fraction of 50kb genomic windows where the Apennine bears reached a level of variation similar to the value observed in the non-Apennine bears was ten times higher for windows with high compared to windows with average variation levels (SI Appendix, section S8.2).

Third, high variation windows fall in regions enriched for coding loci, and the significantly overrepresented pathways within these windows include six genes related to the adaptive immune system, 39 genes related to the olfactory signaling pathways, and two genes related to the digestion of dietary carbohydrate (Figs. 4A-D; SI Appendix, Table S14). These genes are present in several copies (51–53), and most of them have been suggested to be under balancing selection (54, 55). It seems therefore likely that evolutionary processes such as heterozygote advantage, assisted by a genetic mechanism such as non-allelic homologous recombination among multiple copies (56, 57), maintain high variation at relevant loci even under strong genetic drift, preserving their role in the interaction with the environment and pathogens. Direct sequencing of two MHC loci in additional individuals confirmed that the level of variation observed in the Apennine population is as high as in other European regions (SI Appendix, section S8.3).

Our results thus support the view that, even in small populations, the random loss of variation does

not affect all sites in the same way, and we further contribute to the general debate about the relative role of drift and selection when the effective population size is very small (9–11, 58).

We also find that Apennine and the non-Apennine bear genomes show fixed differences at approximately 1,000 nuclear genic positions. Even considering the possible under-estimation of rare variants due to low sample sizes and the ~5X coverage of most of our genomes, these are genomic positions showing extreme divergence in the Apennine population. Of these fixed differences, 411 produce non-synonymous changes in 360 genes, 40 of which are predicted to be deleterious in the Apennine bear (none in the non-Apennine group), and four others result in a premature stop codon (Figs. 4A-D; SI Appendix, section S9.1). Several of the fixed mutations were found in genes directly related to or associated with human monogenic disorders: severe anemia, craniofacial and ocular anomalies, small body size, proteinuria, cardiac and skeletal muscle related diseases, lactate dehydrogenase-B deficiency, and problems related to low levels of proteins in the blood. In the Apennine mtDNA genome, three non-synonymous substitutions occur in the NADH dehydrogenase subunit 5 (with one of these never observed in a large database of brown and polar bears) (SI Appendix, section S9.2). Several ND5 changes have been shown to reduce the activity of complex I, causing different health problems including muscle weakness. Finally, considering the popular belief and some observations regarding the rather docile temperament of the Apennine bear (no attacks on humans have ever been recorded in the last century (59)), we tested the pattern of divergence between Apennine and non-Apennine bears at 22 genes that were associated with tame/aggressive behavior (SI Appendix, section S9.3). Interestingly, we found a significant enrichment for fixed differences in these genes (Fig. 4E), suggesting that genetic drift or hunting of the more aggressive or bold individuals (60, 61) may have led to a genetically mediated shift in Apennine bear behavior.

In summary, random fixation of deleterious mutations probably increased the genetic load in the Apennine bear, with negative consequences such as a reduced ability to produce energy. These genetic changes also produced phenotypic divergence at traits usually used to identify this population, such as the small size, unique cranial morphology, and a less aggressive behavior. Additional behavioral and functional assays will be necessary to prove these hypothetical gene-phenotype links.

Conclusions and conservation perspectives

Why do a handful of Apennine bear stragglers survive, given their extremely low genomic variation, high inbreeding, and large number of fixed deleterious mutations? Despite the difficulty in assessing the role of genetic factors in past extinctions of small populations, the causative link between low genetic and genomic variation, fitness reduction, and high extinction risk is supported by theoretical arguments and empirical evidence (24, 62–65). Still, the long-term persistence of species at extremely low levels of genomic variation has also been documented (11, 66). We believe that our study, even if based on a limited number of whole genomes, can help explaining this apparent contradiction between expectation and observation. We identify an active process of maintenance of variation at crucial genes for pathogen defense and chemical perception. We hypothesize that the lack of competitors reduced the impact of many deleterious mutations (67), and that the highly diversified diet of the brown bear may have compensated for the energy production problems in the Apennine population by facilitating a switch from omnivory to an almost completely vegetarian diet (68, 69). In addition, we inferred a genetic component related to a behavioral change towards a less aggressive temperament, which potentially reduced the risk perceived by local human communities and thus limited persecution and attempts to eradicate the Apennine bear. Do these results imply that the extinction risk of the Apennine bear due to genetic factors is low, and therefore that invasive management options, such as genetic rescue via translocation of unrelated individuals (70), are unnecessary?

On the one hand, the benefits of such interventions to reduce inbreeding and favor demographic expansion, even when fitness decay is not reported and only a few individuals are introduced, are well documented and supported (71). Also, considering that the Apennine bear has been isolated from other bear populations for only several thousand years, adverse outbreeding effects related to hybridization among different populations should be minor, if they were to occur at all. Plausible candidates for the genetic rescue of the Apennine bear are the geographically and maternally (mtDNA) closer bear individuals from Slovenia or the Italian Alps. Alternatively, more adaptively similar individuals from Mediterranean areas could be introduced with even larger outbreeding benefits, an outcome supported by a recent study of experimental translocations in fish (72). If implemented, however, we warmly suggest additional genomic and non-genomic analyses for a careful choice of rescuers.

On the other hand, the recognition of the Apennine bear as an Italian iconic endangered taxon, the possible risk of introducing aggressiveness genes and deteriorating the relatively peaceful human-

bear coexistence in Central Italy, and the current levels of variation at relevant immune and olfactory genes, suggest avoiding genetic rescue. We call instead for an increase in conservation actions such as reducing incidental killing and poaching, securing food resources (e.g., protecting local wild fruit plants), and favoring natural dispersal in unoccupied but suitable areas. If direct evidence of inbreeding depression will be reported in the future, the genetic rescue option should be reconsidered. This approach would allow the maintenance and the monitoring of this ongoing natural experiment of evolution and divergence at small population size that produced, rephrasing (73), a cherished group, precisely because it represents a divergent evolutionary lineage.

Methods

Sequencing, mapping, SNPs and genotype calling
Genomic DNA was extracted from blood and muscle tissue of twelve individuals: six from Central Italy, one from Northern Italy, two from Slovakia, two from Greece and one from Spain. Samples were prepared for paired-end (2x100) Illumina sequencing following the protocol described in (74). A total of ~2.14 billion reads were uniquely mapped (SI Appendix, section S1.3) to 357 autosomal scaffolds (~2.16Gb: 95.3% of the whole assembly) of the polar bear reference genome (22) with high confidence (Q>=25). We included in our analyses two European brown bears from Sweden, six polar bears and one black bear data from previously published studies (22, 75). Eight individuals were sequenced at low coverage with an average sequencing depth of 2.4 to 6.3X, whereas the four remaining individuals were sequenced at higher coverage (from 14.4 to 16.5X). One sample from Central Italy (APN1) was excluded due to too low coverage. Single nucleotide variants and indels were jointly discovered in the 13 high coverage individuals (SI Appendix, Section S1.4) using the *UnifiedGenotyper* algorithm implemented in *GATK* and the software *ANGSD (76)* was used to compute several measure of genetic variation in low coverage samples integrating the genotype uncertainty (SI Appendix, section S1.5).

Long regions of homozygosity and inbreeding estimates
We calculated the Watterson estimator of θ in 50kb overlapping windows, with 10kb steps, over the 357 autosomal scaffolds, excluding windows with more than 30% missing sites. To quantify the proportion of the genome characterized by long regions of homozygosity in each individual, we explored the heterozygosity profile along the scaffolds (SI Appendix, section S1.6) in order to identify long regions (>=1Mb) composed by contiguous windows with less than 25 segregating sites (heterozygosity constantly lower than 0.0005). To quantify and compare the level of inbreeding across all individuals (SI Appendix, Section S2), we adopted the methodology presented in (25).

Demographic analyses
To study the past population size variation through time, we first applied the PSMC method (29) to the high coverage samples (SI Appendix, Section S3.1). We further studied the recent demographic dynamic of the Apennine brown bear population by Approximate Bayesian Computation (ABC-skyline;(34)) using the genomic data from the five Apennine individuals (SI Appendix, section S3.2). In order to identify the strength of the population contraction that could have produced the accumulation of homozygosity regions in the Apennine brown bear, we then performed coalescent simulations modeling the expected patterns of homozygosity regions in the genome under different demographic conditions (SI Appendix, Section S3.3).

Neighbor-Joining tree on nuclear and mitochondrial whole-genome distances
We used ANGSD to compute pairwise genomic distances between individuals (without calling genotypes) over 1,842,042,551bp, and a Neighbor-Joining tree was also computed (SI Appendix, section S4). In order to reconstruct the mtDNA phylogeny, one Illumina MiSeq lane (2x75bp kit) was used to sequence 11 additional samples: six from the Apennine, three from Slovakia and two from Greece, producing ~20 million paired-end reads. After aligning to the reference mitochondrial genome and quality filtering, 16,409bp were used to reconstruct a Neighbor-Joining tree based on HKY pairwise distances (SI Appendix, section S6).

Population structure analyses
We inferred the individual-based genetic structure using the Bayesian population model implemented in STRUCTURE 2.3.4 (77) including only SNPs at least 50 kb away from any known gene, not in repeated regions, separated by at least 150 kb to minimize linkage disequilibrium (7,971 unlinked SNPs; SI Appendix, section S5.1). To further explore the evolutionary history of the Apennine brown bear population and estimate the posterior density of relevant demographic parameters, we built a complex metapopulation model relating all European populations, followed by an ABC approach simulating a non-equilibrium 100 demes finite island model (SI Appendix, Section S5.2). A two-population isolation-with-migration model was also explored through an ABC approach to exclude recent or ongoing gene flow to or from the Apennine

population (SI Appendix, section S5.3). Using the same statistical approach employed to date the interruption of gene flow between modern humans and Neandertals (36), we estimated the time since complete isolation of the Apennine population from the Central European population counting the number of polymorphisms in one European individual for which an allele sampled at random from one Apennine individual was the derived one (*F*). To translate from *F* to time since population divergence, a calibration curve was built employing coalescent-based simulations under a demographic scenario derived from previous analyses (SI Appendix, section S5.4).

Genomic retention of non-synonymous SNPs
In coding regions, we selected all heterozygous sites in a single non-Apennine reference individual genotyped at high coverage, and we categorized them in non-synonymous (n-syn) and synonymous (syn) sites. We then estimated the fraction of sites that were also polymorphic (P_{n-syn} and P_{syn}) in the five Apennine samples, for syn and n-syn sites separately. Assuming that heterozygous sites in the reference genome are a proxy for the polymorphism in the Apennine ancestral population, P_{n-syn} and P_{syn} are estimates of the probability of retaining a polymorphism during the Apennine population divergence (SI Appendix, section S8.1). This is a McDonald-Kreitman-like test (78) on all genes during the Apennine population divergence. We predict that P_{n-syn} and P_{syn} should be the same if drift is mainly driving the fixation of n-syn substitutions in the Apennine brown bear genomes. However, if balancing selection played an important role during the divergence of the Apennine brown bear, we predict that P_{n-syn} should be higher than P_{syn}. On the contrary, if directional selection is a major force, we expect that P_{n-syn} should be lower than P_{syn}.

Genomic windows at high variation
We analyzed the distribution of θ_W along the genome, using 50 kb overlapping windows (with 10kb steps), in the alignment of five non-Apennine brown bears, identifying the windows with average θ values (AveWi) and the windows with top θ values (TopWi), which are the 50kb windows with a θ_W within 10% of the global mean and with a θ_W higher than the 99th percentile, respectively. We then analyzed θ_W in the same windows with respect to the alignment of the five Apennine individuals, thus identifying regions overlapping with the TopWi in the non-Apennine dataset (i.e., where high variation is retained more than expected by chance; SI Appendix, section S8.2). We tested for enrichment of specific pathways for genes in such regions using the software REACTOME (79).

Fixed differences between Apennine and non-Apennine brown bear individuals
The whole nuclear genome was screened for fixed differences, *i.e.,* sites where all five Apennine genomes and all eight non-Apennine genomes are monomorphic for a different allele (SI Appendix, Section S9.1). Fixed differences in the Apennine population in genes implying non-synonymous substitutions were tested for deleterious effect using two different bioinformatic tools (Panther (80) and Polyphen (81)), and relationship with human diseases was identified using the OMIM database. A similar analysis was performed on the mtDNA genome, comparing the unique Apennine population haplotype with 45 European sequences from public repositories and from our non-Apennine samples. In the case of the mtDNA, we looked for amino acids fixed in the Apennine sequence and present in the non-Apennine group at low frequency ((SI Appendix, section S9.2).

Enrichment of fixed differences in candidate genes for tame/aggressive behavior
Considering the behavioral differences suggested for the Apennine bear, we blindly tested if candidate genes previously associated with tameness/aggressiveness (TA) in other mammals showed a higher than expected signal of genetic divergence in comparison with other populations (SI Appendix, section S9.3). We analyzed 19 genes suggested as candidate determinants of a recent shift from an aggressive to a more docile temperament. For each gene, we counted the number of sites with fixed differences between the five Apennine bears genomes and five non-Apennine bears in a genomic region including the gene and two 10 kb windows upstream and downstream the gene (FD_{TA}). In order to statistically verify if FD_{TA} is larger

than expected, we randomly sampled 10,000 sets of 19 genes and computed the FD value for each random set.

ACKNOWLEDGMENTS

References

1. Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nat Rev Genet* 10:783–796.

2. Lande R (1994) Risk of population extinction from fixation of new deleterious mutations. *Evolution* 48:1460–1469.

3. Wright S (1931) Evolution in mendelian populations. *Genetics* 16:97–159.

4. Burger R, Lynch M (1995) Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution* 49:151–163.

5. Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of species extinction. *Conservation Biology: The Science of Scarcity and Diversity*, ed Soulé ME (Sinauer, Sunderland), pp 19–34.

6. Fagan WF, Holmes EE (2006) Quantifying the extinction vortex. *Ecol Lett* 9:51–60.

7. Blomqvist D, Pauliny A, Larsson M, Flodin LA (2010) Trapped in the extinction vortex? Strong genetic effects in a declining vertebrate population. *BMC Evol Biol* 10:33.

8. Norén K, Godoy E, Dalén L, Meijer T, Angerbjörn A (2016) Inbreeding depression in a critically endangered carnivore. *Mol Ecol* 25:3309–3318.

9. Aguilar A, et al. (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc Natl Acad Sci USA* 101:3490–3494.

10. Wood JLA, Yates MC, Fraser DJ (2016) Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evol Appl* 9:640–657.

11. Robinson JA, et al. (2016) Genomic flatlining in the endangered island fox. *Curr Biol* 26:1183–1189.

12. Butlin R, Debelle A, Kerth C, Snook RR (2012) What do we need to know about speciation? *Trends Ecol Evol* 27:27-39.

13. Wang L, Wan ZY, Lim HS, Yue GH (2016) Genetic variability, local selection and demographic history: genomic evidence of evolving towards allopatric speciation in Asian sea bass. *Mol Ecol* 25:3605–3621.

14. Ciucci P, Boitani L (2008) The Apennine brown bear: a critical review of its status and conservation problems. *Ursus* 19:130–145.

15. Ciucci P, et al. (2015) Estimating abundance of the remnant Apennine brown bear population using multiple noninvasive genetic data sources. *J Mammal* 96:206–220.

16. Colangelo P, et al. (2012) Cranial distinctiveness in the Apennine brown bear: Genetic drift effect or ecophenotypic adaptation? *Biol J Linn Soc Lond* 107:15–26.

17. Randi E, Gentile L, Boscagli G, Huber D, Roth HU (1994) Mitochondrial DNA sequence divergence among some west European brown bear (*Ursus arctos* L.) populations. Lessons for conservation. *Heredity* 73:480–489.

18. Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphisms, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proc R Soc Lond B Biol Sci* 255:195–200.

19. Swenson JE, Taberlet P, Bellemain E (2011) Genetics and conservation of European brown bears *Ursus arctos*. *Mamm Rev* 41:87–98.

20. Febbo D, Pellegrini M (1990) The historical presence of the brown bear in the Apennines. *Aquilo Ser Zool* 27:85–88.

21. Preatoni D, et al. (2005) Conservation of brown bear in the Alps: space use and settlement behavior of reintroduced bears. *Acta Oecol* 28:189–197.

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

23. Schubert M, et al. (2014) Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proc Natl Acad Sci USA*: 16991–16991.

24. Xue Y, et al. (2015) Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* 348:242–245.

25. Prüfer K, et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505:43–49.

26. Der Sarkissian C, et al. (2015) Evolutionary genomics and conservation of the endangered Przewalski's Horse. *Curr Biol* 25:2577–2583.

27. Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-reina OC (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc Natl Acad Sci U S A* 109:2382–2390.

28. Brauer A, et al. (2007) Evidence for last interglacial chronology and environmental change from Southern Europe. *Proc Natl Acad Sci USA* 104:450–455.

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

30. Bunnefeld L, Frantz LAF, Lohse K (2015) Inferring bottlenecks from genome-wide samples of short sequence blocks. *Genetics* 201:1157–1169.

31. Schiffels S, Durbin R (2014) Inferring human population size and separation history from multiple genome sequences. *Nat Genet* 46:919–925.

32. Liu X, Fu Y-X (2015) Exploring population size changes using SNP frequency spectra. *Nat Genet* 47:555–559.

33. Terhorst J, Kamm JA, Song YS (2017) Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet* 49:303–309.

34. Nadachowska-Brzyska K, Burri R, Smeds L, Ellegren H (2016) PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white Ficedula flycatchers. *Mol Ecol* 25:1058-1072.

35. Maisano Delser P, et al. (2016) Population genomics of *C. melanopterus* using target gene

capture data: demographic inferences and conservation perspectives. *Sci Rep* 6:33753.

36. 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.

37. Green RE, et al. (2010) A draft sequence of the Neandertal genome. *Science* 328:710–722.

38. Kaplan JO, et al. (2011) Holocene carbon emissions as a result of anthropogenic land cover change. *Holocene* 21:775–791.

39. Marlon JR, et al. (2013) Global biomass burning: a synthesis and review of Holocene paleofire records and their controls. *Quat Sci Rev* 65:5–25.

40. Zennaro P, et al. (2015) Europe on fire three thousand years ago: Arson or climate? *Geophys Res Lett* 42:5023-5033.

41. Fyfe RM, Woodbridge J, Roberts N (2015) From forest to farmland: pollen-inferred land cover change across Europe using the pseudobiomization approach. *Glob Chang Biol* 21:1197–1212.

42. Curry-Lindahl K (1972) The brown bear (*Ursus arctos*) in Europe: decline, present distribution, biology and ecology. *Bears: Their Biology and Management* 2:74–80.

43. Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Anim Behav* 28:1140–1162.

44. Støen O-G, Zedrosser A, Saebø S, Swenson JE (2006) Inversely density-dependent natal dispersal in brown bears *Ursus arctos*. *Oecologia* 148:356–364.

45. Costello CM, Creel SR, Kalinowski ST, Vu NV, Quigley HB (2008) Sex-biased natal dispersal and inbreeding avoidance in American black bears as revealed by spatial genetic analyses. *Mol Ecol* 17:4713–4723.

46. Jerina K, Adamič M (2008) Fifty years of brown bear population expansion: effects of sex-biased dispersal on rate of expansion and population structure. *J Mammal* 89:1491–1501.

47. Bidon T, Janke A, Fain SR, Eiken HG, Hagen SB, Saarma U, Hallstròm BM, Lecomte N, Hailer F (2014) Brown and polar bear Y chromosomes reveal extensive male-biased gene flow within brother lineages. *Mol Biol Evol* 31:1353-1363.

48. Henry RC, Coulon A, Travis JMJ (2016) The evolution of male-biased dispersal under the joint selective forces of inbreeding load and demographic and environmental stochasticity. *Am Nat* 188:423–433.49. Stow AJ, Sunnucks P, Briscoe DA, Gardner MG (2001) The impact of habitat fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence from allelic and genotypic analyses of microsatellites. *Mol Ecol* 10:867–878.

50. Valdiosera CE, et al. (2008) Surprising migration and population size dynamics in ancient Iberian brown bears (*Ursus arctos*). *Proc Natl Acad Sci USA* 105:5123–5128.

51. Perry GH, et al. (2007) Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 39:1256–1260.

52. Young JM, et al. (2008) Extensive copy-number variation of the human olfactory receptor gene family. *Am J Hum Genet* 83:228–242.

53. Flajnik MF, Kasahara M (2010) Origin and evolution of the adaptive immune system: genetic

events and selective pressures. *Nat Rev Genet* 11:47–59.

54. Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evol Biol* 16:363–377.

55. Alonso S, López S, Izagirre N, de la Rúa C (2008) Overdominance in the human genome and olfactory receptor activity. *Mol Biol Evol* 25:997–1001.

56. Ohta T (1982) Allelic and nonallelic homology of a supergene family. *Proc Natl Acad Sci USA* 79:3251–3254.

57. Ottaviani D, LeCain M, Sheer D (2014) The role of microhomology in genomic structural variation. *Trends Genet* 30:85–94.

58. Radwan J, Biedrzycka A, Babik W (2010) Does reduced MHC diversity decrease viability of vertebrate populations? *Biol Conserv* 143:537–544.

59. Naves J, et al. (2016) Brown bears attacks on humans in the world: a preliminary overview 2000-2015. 24th *International Conference on Bear Research and Management*, Anchorage, Alaska.

60. Ciuti S, et al. (2012) Human selection of elk behavioural traits in a landscape of fear. *Proc Biol Sci* 279:4407–4416.

61. Steyaert SMJG, et al. (2016) Ecological implications from spatial patterns in human-caused brown bear mortality. *Wildlife Biol* 22:144–152.

62. Frankham R (2005) Genetics and extinction. *Biol Conserv* 126:131-140.

63. Palomares F, et al. (2012) Possible extinction vortex for a population of Iberian lynx on the verge of extirpation. *Conserv Biol* 26:689–697.

64. Abascal F, et al. (2016) Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol* 17:251.

65. Rogers RL, Slatkin M (2017) Excess of genomic defects in a woolly mammoth on Wrangel island. *PLoS Genet* 13:e1006601.

66. Milot E, Weimerskirch H (2007) Surviving with low genetic diversity: the case of albatrosses. Proc Biol Sci 274:779–787.

67. Agrawal AF, Whitlock MC (2012) Mutation Load: The Fitness of Individuals in Populations Where Deleterious Alleles Are Abundant. *Annu Rev Ecol Evol Syst* 43:115–135.

68. Ciucci P, Tosoni E, Di Domenico G, Quattrociocchi F, Boitani L (2014) Seasonal and annual variation in the food habits of Apennine brown bears, central Italy. *J Mammal* 95:572–586.

69. Bojarska K, Selva N (2011) Spatial patterns in brown bear *Ursus arctos* diet: the role of geographical and environmental factors. *Mamm Rev* 42:120–143.

70. Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends Ecol Evol* 30:42–49.

71. Frankham R (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Mol Ecol*. 24:2610-2618.

72. Kronenberger JA, et al. (2017) Testing the demographic effects of divergent immigrants on small populations of Trinidadian guppies. *Anim Conserv* 20:3–11.

73. Tallmon DA (2017) Get a move on: the value of rescue. *Anim Conserv* 20:16–17.

74. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010:pdb.prot5448.

75. Miller W, et al. (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc Natl Acad Sci USA* 109:E2382–E2390.

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

77. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.

78. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in Drosophila. *Nature* 351:652–654.

79. Fabregat A, et al. (2016) The Reactome pathway Knowledgebase. *Nucleic Acids Res* 44:D481–7.

80. Thomas PD, et al. (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* 13:2129–2141.

81. Adzhubei IA, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249.

FIGURE LEGENDS

Figure 1. Phylogenetic relationships and patterns of homozygosity of the European brown bear. **(A)** Geographic distribution of the brown bear samples sequenced in this study: Slovakia (SLK, brown), Italian Alps (ALP, light brown), Spain (SPA, yellow), Italian Apennine (APN, red), and Greece (GRE, green). Samples from previous studies: polar bear from Greenland (POL, light blue), brown bear from Sweden (SWE, dark red), and black bear from Alaska (BLK, black).. The European brown bear range is shown as dark gray (wild) and light gray (reintroduced in recent times). [a] (22); [b] (74). **(B)** Average heterozygosity per site (θ_w) in brown bear individuals sequenced at different depths. Genome sequences from two Swedish brown bears and one polar bear individual (all polar bears have very similar levels of variation) from previously published studies are also included. Analyses are also replicated in high-coverage individuals after down-sampling the raw reads (striped bars). **(C)** Long regions of homozygosity (>1Mb with more than 99.95% of homozygous calls in contiguous 50kb windows), in ascending length order. Note that brown bear (top panel) and polar bear (bottom panel) tracks are on different x- and y-axis scale. The area under each plot is proportional to the fraction of the genome (shown in parentheses) that can be assigned to these fragments. **(D)** Neighbor-Joining tree using the whole genome sequence (left) or the whole mitochondrial genome sequence (right). The mitochondrial genome has also been sequenced in 11 additional samples, see SI Appendix, section S6. Note that Swedish samples have two different mitochondrial haplotypes clustering with the Slovakian samples and the Spanish sample, respectively. **(E)** Examples of the genomic variation along >40 Mb of scaffold1 in an Apennine brown bear individual (red), an Alpine brown bear (brown), and a polar bear (blue). Long windows of homozygosity in the Apennine sample are shaded in grey.
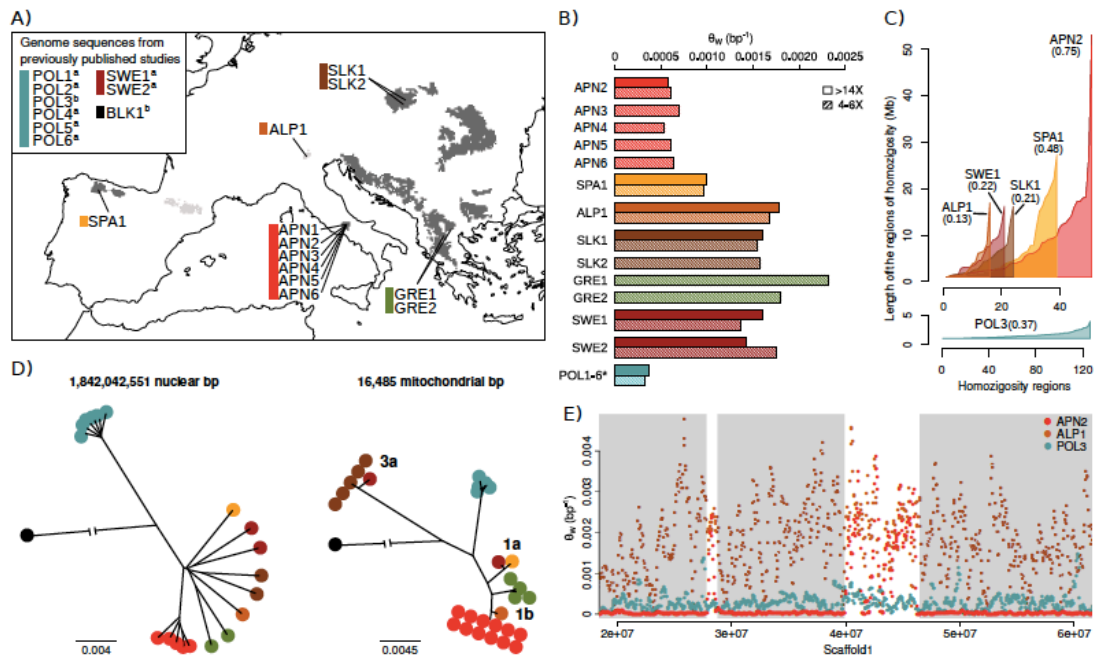
Figure 2. Inferred demographic history of the Apennine brown bear. **(A)** Demographic dynamics reconstructed applying the PSMC approach to the APN2 genome (solid line) with bootstrap-based uncertainty regions (shade). **(B)** Demographic dynamic of the Apennine brown bear estimated using the ABC-Skyline approach based on approximately 20Mb of non-coding regions from five individuals (median as a solid line and 95% confidence intervals are shaded). **(C)** Examples of patterns of the genomic distribution of θ_w per site in 1000 fragments of 50kb in 20 simulated individuals subjected to different bottlenecks (sketched, not to scale, in gray within each figure, width of the boxes is proportional to population size, and present time is at the bottom); a gray line represents each simulated individual, with θ_w values in ascending order whereas the red line represents the pattern in the observed APN2 genome in 50 Mb of the longest scaffold (scaffold1). N_0 and N_a correspond to the modern and the ancient population size, respectively, whereas T_b represents the time of the bottleneck (in years). The most supported scenario is reported in the top panel while three additional illustrative examples are reported below (see SI Appendix, section S3.3 and Fig S10 for details and more examples).

Figure 3. Inferred fragmentation scenario for the European brown bear. Posterior distributions on the right refer to the ABC meta-population model (see SI Appendix, section S5.2 for details); connectivity refers to Nm as the product of the effective population size of a deme, N, and the migration rate, m, per generation N_{anc}: ancestral population size before any fragmentation; T5: time of brown bear spread in Europe; N_1m_1: migration parameter among European brown bear populations before T3; T3: time of the decrease in connectivity among European brown bear populations. N_2m_2: migration parameter among European brown bear populations after T3. Dashed lines represent prior distributions.

Figure 4. Active maintenance and random loss of genetic diversity. Boxplot of the average genetic diversity (estimated with *ANGSD (76)*) within five Apennine brown bears **(A)**, five European brown bears **(B)** and between these two groups **(C)** grouped by gene categories and

including 10Kb upstream and downstream of each gene. The categories are as follows: "High" are genes in high-diversity genomic windows (see SI Appendix, section S8.2); "Del" are genes with non-synonymous deleterious mutations fixed in the Apennine population (see SI Appendix, section S9.1); "TA" are candidate genes associated with tame/aggressive behavior (see SI Appendix, section S9.3); and "All" are all the other genes. **(D)** Scatterplot of genetic diversity within and between groups in single genes of different categories; grey shades correspond to different number of All genes with specific combinations of variation levels. The most divergent gene for each category is indicated (High: RNPC3, Del: CMA1, and TA: GLTSCR2) as well as most of the genes with high diversity in APN population: immune response (HLA and IG-like), carbohydrates digestion (AMY1B/2B), aflatoxin detoxification (AKR7A1), control of DNA damage (HUS1), lipid transporter activity (APOL2/3), and olfactory receptors (marked with an asterisk). Note that some of the genes at high diversity in the APN population (e.g., APOL2/3) are not included in the High category as they were excluded due to marginally overlapping TopWi windows with unusual coverage (see SI Appendix, section S8.2). **(E)** Proportion of fixed differences in genes related to tame/aggressive behavior is significantly higher (yellow line) than expected at random (gray bars). **(F)** Genomic windows in the 99th percentile (dashed line), considering the genetic diversity in the European brown bear are also characterized by high diversity in the Apennine population and are enriched for immune system and olfactory receptor genes (brown shaded boxes). θ_w in the two groups along scaffold118 and scaffold80 is shown as an illustration. Positions of five MHC and six olfactory receptor genes are shown.
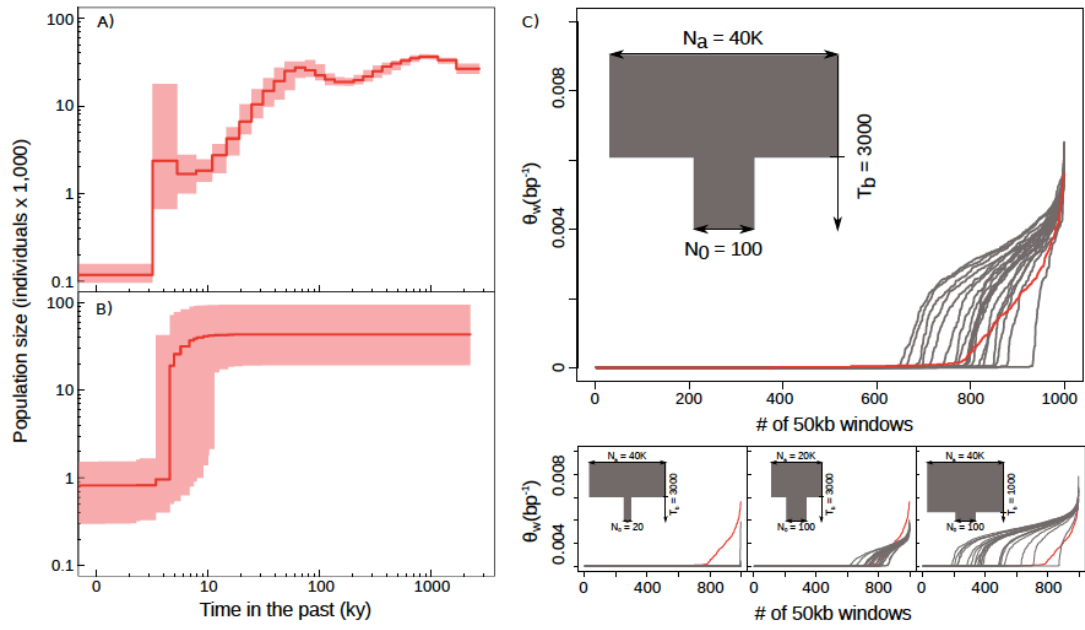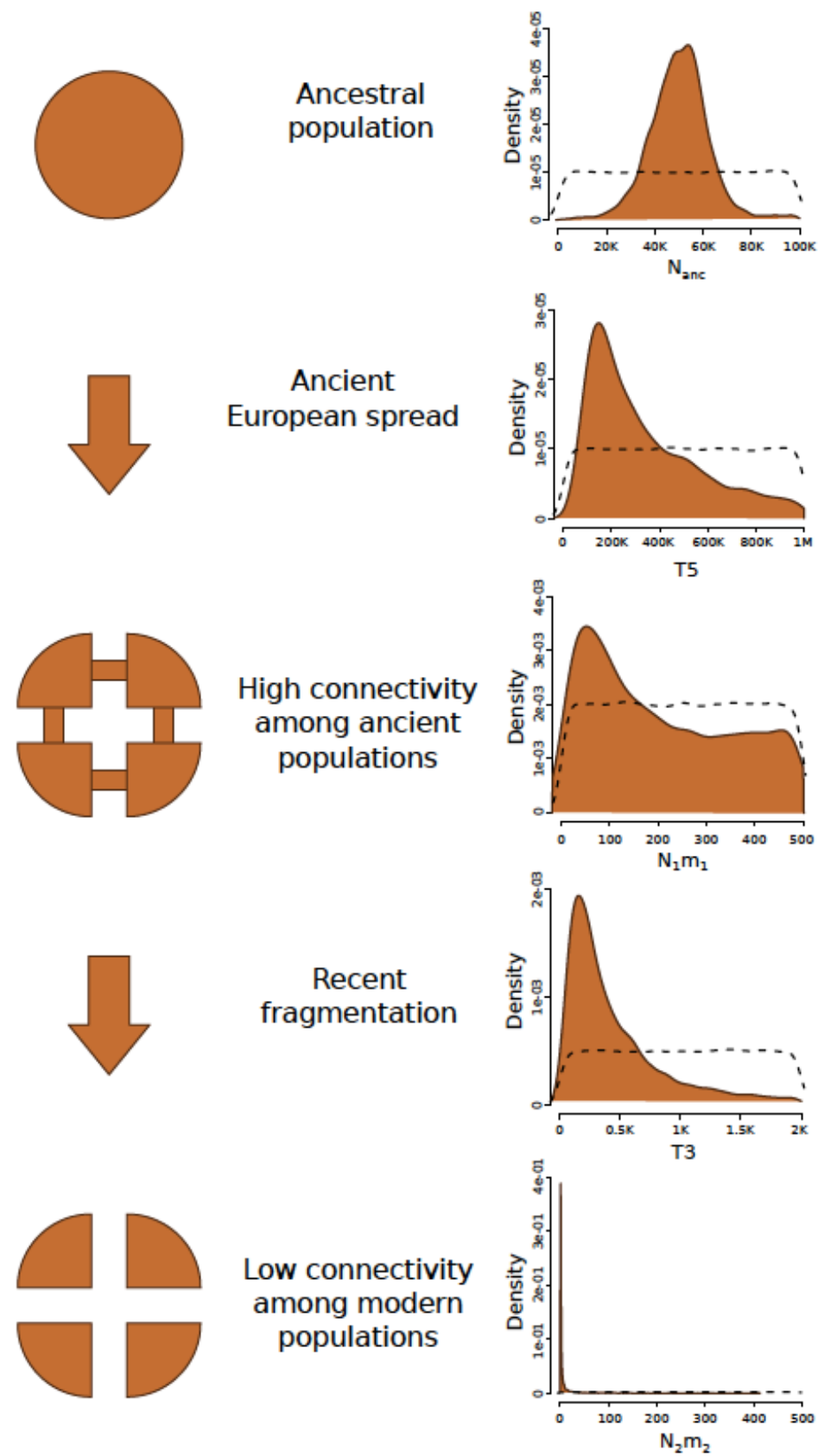
Figure 1

Figure 2

Figure 3

Figure 4