# Recent increase in species-wide diversity after interspecies introgression in the highly endangered Iberian lynx

Received: 30 June 2022

Accepted: 10 November 2023

Maria Lucena-Perez[1], Johanna L. A. Paijmans[2,3], Francisco Nocete[4], Jordi Nadal[5], Cleia Detry[6], Love Dalén[7,8], Michael Hofreiter[2], Axel Barlow[9] & José A. Godoy[1]

Genetic diversity is lost in small and isolated populations, affecting many globally declining species. Interspecific admixture events can increase genetic variation in the recipient species' gene pool, but empirical examples of species-wide restoration of genetic diversity by admixture are lacking. Here we present multi-fold coverage genomic data from three ancient Iberian lynx (*Lynx pardinus*) approximately 2,000–4,000 years old and show a continuous or recurrent process of interspecies admixture with the Eurasian lynx (*Lynx lynx*) that increased modern Iberian lynx genetic diversity above that occurring millennia ago despite its recent demographic decline. Our results add to the accumulating evidence for natural admixture and introgression among closely related species and show that this can result in an increase of species-wide genetic diversity in highly genetically eroded species. The strict avoidance of interspecific sources in current genetic restoration measures needs to be carefully reconsidered, particularly in cases where no conspecific source population exists.

Genetic diversity is an intrinsic and critical component of biodiversity as it determines adaptive potential and thus influences, together with other genetic and non-genetic factors, the extinction probability of a species under environmental change. However, genetic diversity is being rapidly lost as populations become small and isolated as a consequence of human actions, a process that is often accompanied by increased genetic load, reduced fitness and increased extinction probabilities[1,2]. It is thus not surprising that the analyses of ancient and historical DNA, which enable past and current diversity to be compared directly, typically find a net loss of genetic diversity through time in declining species and populations[3–6].

Lost genetic diversity can be eventually restored by mutation, but this is a slow process that is also dependent on population size. De novo DNA mutation is, however, not the only mechanism by which novel genetic variants can be introduced into a population's gene pool. Gene flow from other conspecific populations can reintroduce lost variation and slow down or even reverse diversity loss, and reinforcement of gene flow has proved an effective management strategy to restore diversity and, eventually, fitness and adaptive potential[7]. As for interspecific gene flow, the situation is less straightforward because the species is generally regarded as a closed system, so that the potential for restoration by hybridization is often negated. However, it is now widely recognized that the occurrence of admixture between species

[1]Department of Ecology and Evolution, Estación Biológica de Doñana, CSIC, Seville, Spain. [2]Evolutionary Adaptive Genomics, University of Potsdam, Potsdam, Germany. [3]Department of Zoology, University of Cambridge, Cambridge, UK. [4]Grupo de Investigación MIDAS, Departamento Historia I (Prehistoria), Universidad de Huelva, Huelva, Spain. [5]SERP, Departament de Prehistoria, Historia Antiga i Arqueologia, Universitat de Barcelona, Barcelona, Spain. [6]UNIARQ – Centro de Arqueologia da Faculdade de Letras da Universidade de Lisboa, Alameda da Universidade, Lisbon, Portugal. [7]Centre for Palaeogenetics, Stockholm, Sweden. [8]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden. [9]School of Environmental and Natural Sciences, Bangor University, Bangor, Gwynedd, UK. e-mail: godoy@ebd.csic.es

is much more common than previously thought. Particularly for mammals, a spate of recent genomic studies has uncovered widespread admixture within several major clades, involving both extant and extinct species[8–12]. These interspecific admixture events will increase genetic variation in the recipient species' gene pool, which may persist and become widespread across the species' populations. Thus, it is possible, at least in theory, for interspecies admixture to reverse the genetic loss caused by recent population size declines, although its use as a restoration measure is generally discouraged due to the risk of outbreeding depression and the loss of species 'identity'[13,14].

The Iberian lynx (*Lynx pardinus*) is one of the four species in the genus *Lynx*, which also includes the bobcat (*Lynx rufus*), the Canada lynx (*Lynx canadensis*) and the Eurasian lynx (*Lynx lynx*). Eurasian and Iberian lynx were once separated only at the subspecies level[15–17], although their coexistence in Southern Europe during most of the Pleistocene and the absence of intermediate forms support their recognition as distinct species[18–21]. The most common phylogeny across autosomal windows, and thus the consensus autosomal phylogeny, place the Iberian and the Eurasian lynx as sister species diverging around 1.0 million years ago (Ma); however, low recombination regions in autosomes, and particularly in the X chromosome, support an alternative tree where Eurasian and Canada lynx are sister species, with Iberian lynx diverging earlier (around 1.5 Ma (refs. 22,23)). These contrasting patterns indicate extensive introgression between the Eurasian and Iberian lynx distorting the divergence history in high recombination regions of the genome[23]. The Iberian lynx experienced a severe population bottleneck during the twentieth century, which resulted in two small and isolated populations by the end of the century[24]. Genome-wide studies have shown that current Iberian lynx genetic diversity is among the lowest recorded for any mammal[25], and the comparison of microsatellite data of historical Iberian lynx and mitogenomic data of historical and ancient samples with that of modern samples revealed a decline in genetic diversity associated with the recent bottleneck[24]. However, ancient nuclear genetic data that would allow this recent decline to be placed within the broader context of Iberian lynx evolution is currently lacking. In this Article, we generate palaeogenomic data from ancient Iberian lynx dated ~2–4 thousand years ago (ka). We find ancient lynx to exhibit even lower genetic diversity than their modern conspecifics, which can be explained by a recent pulse of admixture with the closely related Eurasian lynx (*L. lynx*).

## Results

### Genome sequencing of ancient Iberian lynx

We generated palaeogenomic data from three Iberian lynx samples previously investigated for mitochondrial DNA[24] from Eras del Alcázar, Andújar, Spain (4,270 ± 30 years before present (BP)); La Moleta del Remei, Alcanar, Catalonia, Spain (2,520 ± 30 BP); and Monte Molião, Algarve, Portugal (dated by archaeological context approximately 2,070 BP) (Fig. 1a and Extended Data Table 1). Sample bone powder was pre-treated with sodium hypochlorite to reduce contamination[26]. Approximately 400–900 million sequences were then generated from each sample using Illumina sequencing and mapped to the reference genome assemblies of the Iberian lynx and the domestic cat. This provided a depth of genome coverage of 2.3× (Algarve), 4.3× (Catalonia) and 2.5× (Andújar) and 1.8× (Algarve), 3.5× (Catalonia) and 2.0× (Andújar), when mapped to the Iberian lynx and the domestic cat reference genomes, respectively. All samples show typical ancient DNA damage patterns of C → T substitutions at the fragment ends (Extended Data Fig. 1). Two of the ancient lynx were males (Catalonia and Andújar), and one was a female (Algarve).

Genomic data from modern Eurasian lynx[27,28] and from an ancient (2,570 ± 30 BP) Eurasian lynx from the north of the Iberian Peninsula (Sima de Pagolusieta in Vizcaya, Spain[29]) were included in our study for comparative purposes. Our final dataset comprised 30 modern Iberian lynx from the two remaining populations, Andújar ($n = 18$) and Doñana ($n = 12$) (refs. 30,31), and 12 Eurasian lynx from six different populations distributed along an east–west gradient (Balkans, Carpathians, Kirov, Caucasus, Yakutia and Primorsky Krai; $n = 2$ for each population). Sample and sequencing details, including depth of coverage, are presented in Supplementary Tables 2 and 3, respectively.

### Population structure of Iberian lynx

Previous studies have shown that the two surviving Iberian lynx populations show a high degree of genetic differentiation[24]. Microsatellite data from historical populations across their former distribution have shown a lower overall degree of genetic structure in historical times, with modern Andújar being less differentiated from other historical populations than the more isolated modern Doñana population, consistent with Andújar maintaining larger population sizes and higher connectivity until very recently[24]. Our novel genome-based results show a relative homogeneity of the ancient Iberian lynx, when analysed together with Andújar and Doñana contemporary populations, on both principal component analysis (PCA) and individual-based clustering analyses (Fig. 1b,c). The ancient Iberian lynx are also closer to the Andújar population than to the Doñana population, both in the PCA and the individual-based clustering analysis when two a priori subpopulations are specified ($K = 2$) (Fig. 1b,c). When specifying $K = 3$, the ancient populations emerge as a separate cluster, with varying degrees of shared ancestry with the Andújar population depending on the run considered (Fig. 1c). However, we observe that PC3 of the PCA separates the ancient Catalonian individual from the other Iberian lynx (Fig. 1b), suggesting the occurrence of some geographical or temporal structure. This ancient structure is, in any case, shallower than that observed between the two surviving populations and shallower than potential substructuring within the Andújar population, as shown by the dispersion of contemporary Andújar samples in PC2 and PC3, and subsequent partitions on the individual clustering analysis, which split the contemporary Andújar population at $K = 4$ and the Doñana population at $K = 5$, but maintain the ancient populations as a single cluster (Extended Data Fig. 2).

### Genetic diversity of Iberian lynx

We compared the genetic diversity of modern and ancient Iberian lynx using both individual variables (that is, observed autosomal heterozygosity) and population variables (that is, nucleotide diversity, $\pi$, and Watterson estimator, $\theta_W$). Unexpectedly, the genetic diversity of ancient lynxes is lower than that of contemporary lynxes, as reflected in observed individual heterozygosity (Fig. 1d and Extended Data Table 2) and population diversity measures, both $\pi$ and $\theta_W$ (Extended Data Fig. 3 and Table 3). For instance, $\theta_W$ diversity in the ancient population (ancient $\theta_W = 9.50 \times 10^{-6}$) represents ~62% and ~38% of $\theta_W$ in Doñana and Andújar, respectively (Extended Data Fig. 3 and Table 3).

The observed pattern is opposite to that expected from artefacts associated with the use of ancient degraded materials, where sequencing errors and post-mortem damage are expected to artefactually increase genetic diversity, and is robust to changes in reference genome, mapping algorithms and the inclusion/exclusion of transitions (Extended Data Fig. 3 and Table 3). Lower diversity in ancient samples also does not seem to be caused by differences in depth of sequencing coverage, as contemporary samples were subsampled to a depth of coverage similar to the one shown by ancient samples (~2.5×), for both individual and population diversity measures (Methods). We nevertheless investigated the potential relationship between observed heterozygosity and depth of coverage further by using linear regression (Fig. 1d). Although a positive correlation between observed heterozygosity and depth of coverage is observed, the heterozygosity of subsampled modern genomes are consistently higher than those of ancient samples (Fig. 1d), and their ranges do not overlap after the ancient sample of Catalonia (highest ancient depth of 4.3×) is subsampled to a comparable depth of 2.58×, resulting in an observed heterozygosity

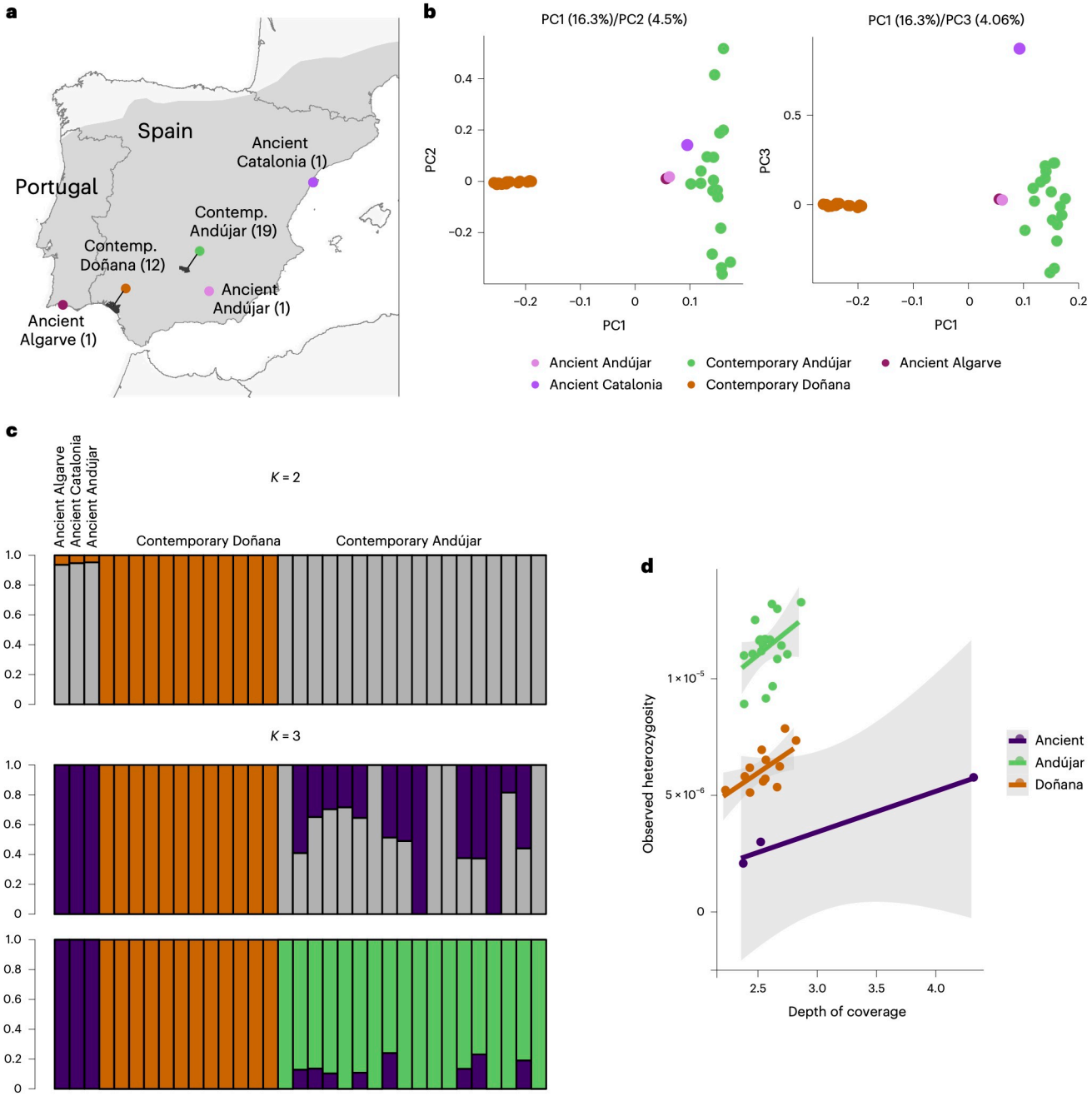


**Fig. 1 | Geographic distribution, genetic structure and genetic diversity of ancient and contemporary Iberian lynx genomes. a**, Contemporary and ancient sampling, with the number of individuals sampled in parenthesis. Ancient Iberian lynx distribution, based on ref. 87, is shown in grey, and contemporary (Contemp.; around year 2002, following the decline of the twentieth century and previous to any translocation and reintroduction) distribution is shown in black[24]. **b**, PCA plots of PC1–PC2 and PC2–PC3, including the percentage of the variance explained by each axis (in parenthesis). **c**, Relationships among individuals based on individual-based clustering. Results are shown for $K = 2$ and $K = 3$. For $K = 2$, all runs converge to the same result, while for $K = 3$, runs result in two different solutions represented here. **d**, Observed heterozygosity per sample versus depth of coverage of the sample. The shading represents the confidence interval around the regression line (thick line). Differences in diversity among the ancient samples might be attributed to differences in coverage, as the one with higher coverage shows similar heterozygosity to other ancient samples when subsampled to a similar depth of coverage (2.58×, $3.15 \times 10^{-6}$; Extended Data Table 2), but differences between ancient and modern samples cannot be explained only by differences in coverage.

of $3.15 \times 10^{-6}$. This rejects sequencing depth as the sole factor driving differences in observed heterozygosity.

Differences in population genomic diversity between ancient and modern Iberian lynx are consistent across different genomic categories such as promoters, untranslated regions (UTRs), coding sequences (CDS), introns or small RNAs (Fig. 2a). Genomic regions showing the largest diversity differences between ancient and modern Iberian lynx show no evidence of enrichment for coding regions (Fig. 2b), suggesting that the difference in diversity is not related to the relaxation of purifying selection in contemporary lynx.

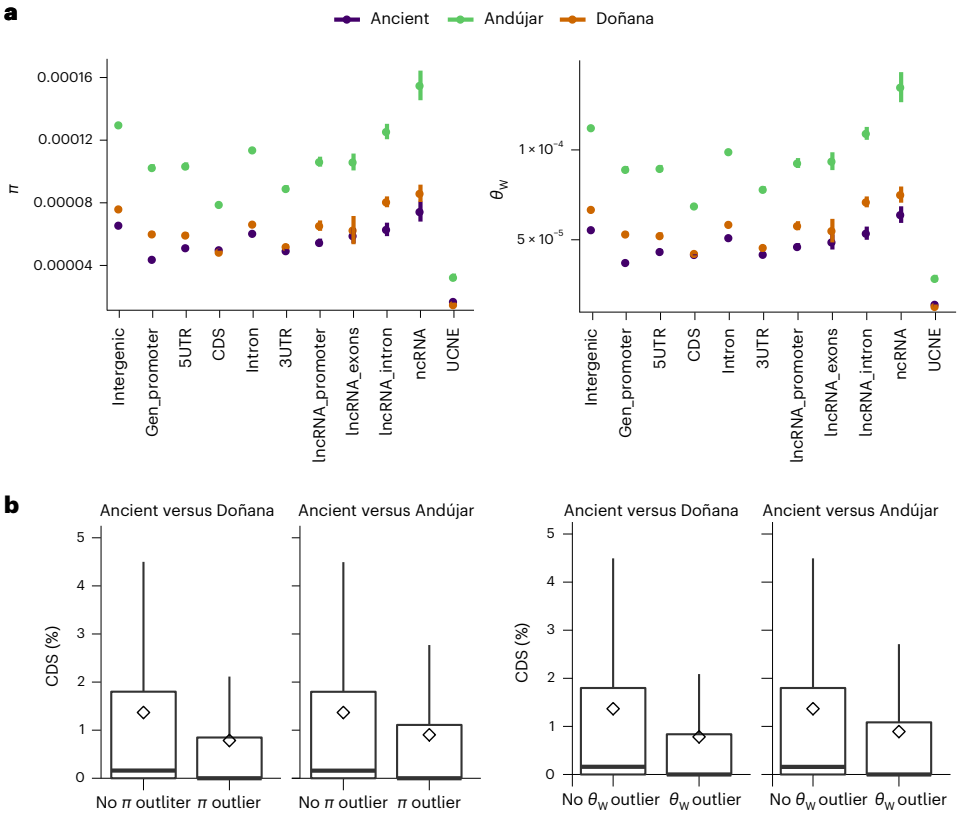


**Fig. 2 | Genetic diversity across genomic features and check for purifying selection. a**, The genetic diversity (nucleotide diversity, $\pi$, left; Watterson's theta, $\theta_w$, right) in different genomic features for contemporary (Doñana and Andújar) and ancient populations. The points represent the mean, and bars represent the s.d. calculated over 10 kb windows using 100 iterations. **b**, The percentage of CDS in genomic windows with lowest diversity difference (outliers) compared with the rest of the windows for $\pi$ (left) and $\theta_w$ (right). Outlier windows were defined as those showing extreme diversity difference between ancient and contemporary samples (that is, ancient diversity minus contemporary diversity) lower than the average minus 2 × s.d. Comparisons include ancient versus current Doñana and ancient versus current Andújar and both $\pi$ and $\theta_w$. The diamonds represent the mean value, the box delimits the first and third quartiles, the thick horizontal lines represent the median and the whiskers extend to 1.5 times the distance between the first and third quartiles. Estimates shown in **a** and **b** were obtained from the sample of three ancient individuals and one random subsample of three individuals from each of the two contemporary populations.

## Admixture with Eurasian lynx

The increase in genetic diversity observed in modern relatives of ancient Iberian lynx is incompatible with a demographic history of major population declines and extirpation over most of their historical range in the last few centuries. Nor is it easily explained by an inadvertent sampling of three isolated and low-diversity ancient populations, since population analyses suggest low levels of ancient genetic structure and a close affinity of the three ancient Iberian lynx to the Andújar population. We therefore tested a hypothesis of interspecies admixture in the intervening time period between the ancient and modern Iberian lynx as a potential source of novel genetic variation, leading to an increase in diversity.

*D* statistic tests showed that all modern Iberian lynxes share a significant excess of derived alleles with their sister species, the Eurasian lynx (*L. lynx*), relative to all ancient Iberian lynxes, consistent with gene flow between Eurasian lynx and Iberian lynx in the last 2 kyr (Fig. 3a). We also find varying levels of admixture with Eurasian lynx occur among the ancient Iberian lynx individuals in a gradient, in which more recent individuals show more derived alleles with Eurasian lynx (Fig. 3b–d). Precisely, the more recent ancient individuals from Catalonia (2.5 ka) and Algarve (2 ka) share similar derived alleles with Eurasian lynx (Fig. 3b), while the oldest individuals from Andújar (4.2 ka) share fewer derived alleles with Eurasian lynx than those two (Fig. 3c,d). Gene flow from Eurasian lynx has permeated the entire modern Iberian lynx distribution, as the proportion of admixed alleles in modern individuals is similar in the two surviving populations (Extended Data Fig. 4a).

As for the donor population, we find that western Eurasian individuals share an excess of derived alleles with contemporary Iberian lynx relative to Eurasian lynx from Asia (Fig. 3e), whereas there is a relative homogeneity in interspecies admixture levels among different western or different eastern populations (Extended Data Fig. 4b,c). To further investigate the timing and geographic source of admixture, we used an ancient Eurasian lynx that inhabited the Iberian Peninsula 2 ka (ref. 29). Contemporary Iberian lynx share more alleles with contemporary Eurasian lynx from both eastern and western populations than with the ancient Eurasian lynx inhabiting the Iberian Peninsula (Extended Data Fig. 4d,e). Also, the ancient Iberian lynx do show evidence of admixture with contemporary Eurasian lynx, but less than with the ancient Eurasian lynx from Iberia: they share more alleles with western contemporary individuals than with the ancient Eurasian lynx, the amounts shared with the latter being similar to those shared with the eastern contemporary populations (Extended Data Fig. 4f,g).

## Directional gene flow from Eurasian lynx into Iberian lynx

*D* statistics do not allow the direction of gene flow to be identified. To investigate this, we used an approach developed for low-coverage palaeogenomes that examines the occurrence of phylogenies compatible

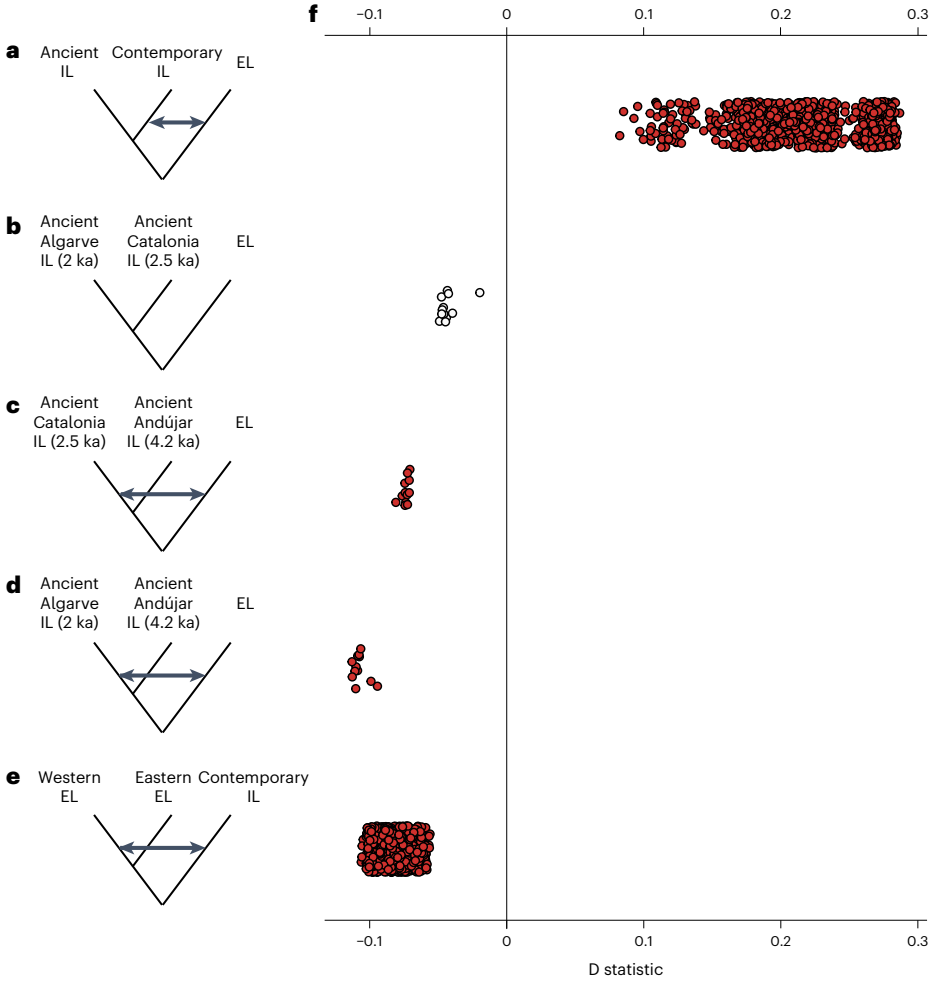


**Fig. 3 | *D* statistics for different combinations of Iberian and Eurasian lynx genomes. a**–**e**, Five different tree topologies tested are shown along with their corresponding *D* values, with double-headed arrows indicating the admixing lineages supported by significant tests. Western EL included genomes sampled in Kirov, Caucasus, Balkans and Carpathians, while Eastern EL genomes were sampled in Primorsky–Krai and Yakutia. Red and white points show significant (absolute *Z* score >3) and non-significant (absolute *Z* score) *D* values, respectively. EL, Eurasian lynx; IL, Iberian lynx.

with admixture along non-overlapping genomic blocks[8]. The test involves two individuals of each admixing species that vary in their amount of admixture. We test the hypothesis that a species is the donor or the recipient of gene flow by comparing the frequency of genome blocks where the more-admixed individual groups with the opposing species' clade, relative to the frequency where the less-admixed individual groups with the opposing species' clade. An increase in the former provides evidence that the species is the recipient of gene flow. We selected the highest-coverage ancient Iberian lynx (Catalonia, 2,520 ± 30 BP, 3.5× when mapped to the domestic cat reference) for these tests, which were repeated using all modern Iberian lynx individuals. For Eurasian lynx, we used an individual from the Balkans and one from Yakutia to represent the two major Eurasian lynx clades. In particular, we examined for an increase in the number of genomic blocks where the more-admixed modern Iberian lynx groups with Eurasian lynx (with the less-admixed ancient Iberian lynx in a basal position), relative to the opposite pattern, with the less-admixed ancient Iberian lynx grouping with Eurasian lynx (and the more-admixed modern Iberian lynx in a basal position). For all modern Iberian lynx, we found that genomic blocks where the more-admixed modern Iberian lynx groups with Eurasian lynx outnumber the opposite pattern by a factor of 1.4–2.6 (Extended Data Fig. 5). Although we do not rule out bidirectional gene flow, these results indicate that the identified admixture event(s) between the two species transferred alleles from Eurasian lynx into Iberian lynx, consistent with the observed increase in genetic diversity.

### Quantifying admixture fractions in modern Iberian lynx

Having verified the occurrence of directional gene flow from Eurasian into Iberian lynx, we estimated the admixed fraction of the genomes of modern Iberian lynx relative to ancient Iberian lynx using the $\hat{f}$ statistic. $\hat{f}$ assumes unidirectional admixture between donor and recipient species, and will tend to underestimate the admixed fraction if gene flow was bidirectional, and is thus a conservative measure. The estimated admixture fractions are remarkably consistent among modern Iberian lynx irrespective of the western Eurasian lynx individuals used in the comparisons (Fig. 4). Consistent variation is also observed when different ancient Iberian lynx are used, leading to an estimated excess admixture in modern Iberian lynx of around 2% when compared with the ancient individuals from Catalonia (2.5 ka) and Andújar (4.2 ka) and around 1.2% when compared with the more recent individual from Algarve (2 ka).

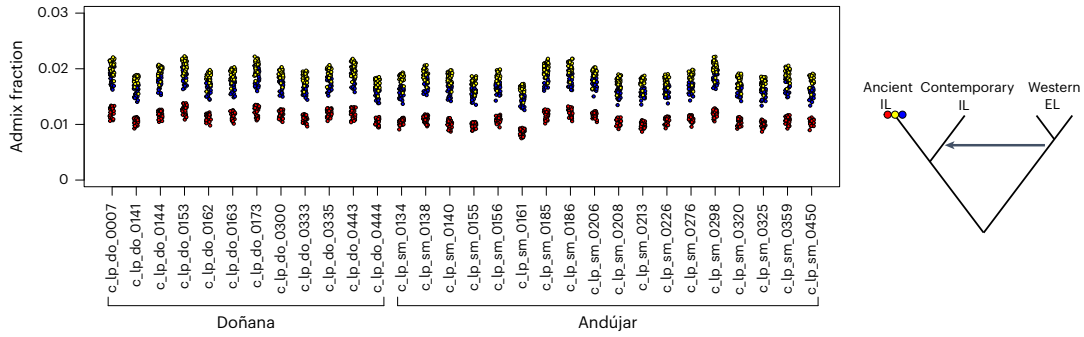


**Fig. 4 | $\hat{f}$ tests (left) on a five-taxon topology ((ancient Iberian lynx, contemporary Iberian lynx), (western Eurasian lynx, western Eurasian lynx), outgroup) to estimate the genomic fraction of Eurasian lynx introgression in contemporary Iberian lynx above that occurring in ancient Iberian lynx, assuming unidirectional gene flow (right).** Each sampled contemporary Iberian lynx is indicated on the *x* axis, with clouds of points showing results generated using different combinations of the other three ingroup individuals. Points corresponding to the three sampled ancient lynx are indicated by different colours: red (Algarve, 2 ka), blue (Catalonia, 2.5 ka) and yellow (Andújar, 4.2 ka). Western Eurasian lynx (Kirov, Caucasus, Balkans and Carpathians) were used in these tests since, based on geographic proximity to the Iberian lynx, they are likely to best represent the introgressing population. All comparisons provided a statistically significant (absolute *Z* score >3) signal of higher admixture in modern than in ancient lynx.

## Discussion

A general trend of reduced genetic diversity in current compared with ancient populations is well documented among species that have undergone postglacial or historical population declines where ancient or historical diversity is directly compared with current diversity[32–38]. The Iberian lynx is one good example of species that lost substantial amounts of both mitochondrial and nuclear genetic diversity following a process of steep decline, fragmentation and local extirpation in the twentieth century[24]. Recent genetic erosion, evidence of other serial bottlenecks and low effective sizes throughout most of the species' history have been invoked to explain the extremely low genome-wide and species-wide genetic diversity of Iberian lynx[25]. Here, we report the unexpected observation of two to three times lower genomic diversity in ancient (2,000–4,000 years ago) than in contemporary Iberian lynx individuals and postulate admixture with its sister species, the Eurasian lynx, as the most likely cause of the observed increase. Although similar processes have been reported following intraspecific gene flow in small isolated populations[39], here we demonstrate an increase in genetic diversity across an entire species following an event of gene flow from a closely related species.

Genomic evidence for extensive admixture between Eurasian and Iberian lynxes since their divergence is abundant and comes from phylogenomic, *D* statistics and model-based approaches[11,23,25]. Although the two species show non-overlapping disjunct distributions now, this was not the case in the past. The Iberian lynx extended its range during the late Pleistocene and early Holocene into southern France and northern and, possibly, southern Italy, where it probably coexisted with the Eurasian lynx, and the Eurasian lynx inhabited the north of the Iberian Peninsula until the early twentieth century[40–44]. Opportunities for encounter and thus hybridization did indeed occur in historical times in Iberia, although direct genetic evidence for this has been absent until now.

Our results provide this lacking evidence and some indication on the timing and geographic patterns of previous admixture events. First, ancient Iberian lynx from Catalonia (2.5 ka) and Algarve (2 ka) were already admixed with Eurasian lynx, although to a lesser extent than contemporary lynxes, and our oldest ancient Iberian lynx (Andújar, 4.2 ka) showed less introgression than younger ancient lynxes (Catalonia, 2.5 ka; Algarve, 2 ka), suggesting a continuous or recurrent process rather than a sporadic event. Second, consistent with geography, western Eurasian lynx have contributed more alleles to the Iberian lynx genome than the eastern lineage; however, some signal of introgression is present from eastern lynx. This may not be surprising given eastern and western Eurasian lineages started to diverge around 100 ka and gene flow was maintained until 22 to 15 ka (refs. 27,28). Signals of introgression from eastern Eurasian lynx may thus be attributed to admixture events occurring before this divergence and to a high proportion of derived alleles being shared across eastern and western populations from their ancestral population. Third, we found no major differences in introgression levels across contemporary populations of western Eurasian or Iberian lynx, indicating that introgression predated the differentiation of current populations, which has been estimated to have occurred 200 years ago for Iberian lynx, and during the Holocene, with a drastic intensification in historical times due to drift, for Eurasian lynx populations in Western Europe[24,28]. Surprisingly, the single ancient genome of Eurasian lynx from Iberia showed less evidence of admixture with Iberian lynx than other contemporary western lynx populations, indicating that the main source of introgression during the last few millennia may have been a differentiated ancestral population more closely related to contemporary western populations. A more extensive sampling of ancient Eurasian and Iberian lynx genomes may further elucidate the geographic and temporal patterns of admixture between these two species.

The concurrence of patterns of higher autosomal diversity and higher introgression in contemporary compared with ancient Iberian lynx individuals supports the hypothesis that the former is the consequence of the latter. We nevertheless tested the possibility that increased genetic diversity is due to the accumulation of mildly deleterious variation caused by the relaxation of purifying selection during the species' decline. However, increased diversity in modern Iberian lynx is observed in both constrained (for example, coding) and supposedly neutral (for example, intergenic) sequences, and the lack of enrichment of coding regions among windows with largest diversity increases discards the relaxation of purifying selection as the sole cause of the observed differences. Finally, the estimated amount of differential introgression seems enough to produce such a large proportional increase in diversity, given that with an observed rate of divergence between modern Eurasian and Iberian lynx of 1.19 fixed differences per kilobase, a 2% introgression would introduce ca. 60,000 new variants, almost ten times the number of heterozygous positions observed in ancient samples. Future studies will further assess the relationship between local introgression and heterozygosity along the genome.

Beyond providing insights on the evolutionary history of the Iberian lynx, one of the world's rarest cat species, our results are also of wider interest for the conservation of species, particularly in the context of genetic rescue. This widely debated and sometimes

controversial conservation strategy involves the introduction of distantly related individuals into imperilled and genetically depauperate populations to restore their fitness and evolutionary potential[45–48]. The use of closely related species for genetic rescue is generally discouraged owing to presumed risks of outbreeding depression caused by either intrinsic or extrinsic incompatibilities. Nonetheless, our results add to the accumulating evidence for natural admixture and introgression in the genomes of many species, and they show that this can result in a substantial increase in the standing genetic diversity in highly eroded populations. The strict avoidance of interspecific sources in genetic restoration measures may need to be carefully reconsidered, particularly in cases such as the Iberian lynx where no additional conspecific source population exists, so a closely related species might be the only source of novel genetic diversity.

## Methods

### Samples

Ancient DNA data was obtained from Iberian lynx archaeological and palaeontological remains. We tested 20 ancient samples selected based on their sequencing performance among the 58 samples analysed in ref. 24. Three of them yielded enough genomic data for further analyses and are the main subjects of this study (Extended Data Table 1). Data from modern samples were obtained from refs. 28,31, and data for the ancient Eurasian lynx sample were obtained from ref. 29 (Extended Data Table 4).

### Laboratory methods

All laboratory work was carried out in dedicated ancient DNA facilities at the University of Potsdam, following established procedures to prevent contamination with modern or synthetic DNA[49]. We performed 70 DNA extractions from 20 archaeological samples, selected based on mitogenome sequence yield[24]. Negative controls were included in all experiments. We drilled the remains to get a variable amount of bone powder up to a maximum of 50 mg. For Algarve and Catalonia sample extraction, we pre-treated the samples with 1 ml of 0.5% bleach for 15 min at room temperature to minimize DNA contamination[26]. For all three ancient samples, bone powder was then digested for ~18 h at 37 °C in 1 ml extraction buffer (0.45 M ethylenediaminetetraacetic acid, 0.25 mg $ml^{-1}$ Proteinase K) with constant rotation. DNA in the resulting supernatant was purified following the protocol in ref. 50, which maximizes degraded DNA recovery by increasing the ratio of binding buffer to sample (13:1).

DNA extracts were converted into sequencing libraries using the single-stranded protocol described in ref. 51. Before library preparation, DNA extracts were treated with uracil–DNA glycosylase and endonuclease VIII to remove deoxyuracils resulting from post-mortem DNA damage. Then, double-stranded DNA was denatured and a biotinylated adaptor oligo was ligated to the 3′ end of each molecule. The resulting products were immobilized on streptavidin-coated beads to repair blunt ends, and the P5 adaptor was ligated to the template molecule. We performed quantitative PCR (qPCR) to evaluate library concentration and used this information to estimate the minimal number of amplification cycles to use in the indexing PCR to avoid clonality (ancient Andújar: six cycles; ancient Algarve: eight cycles; and ancient Catalonia: nine cycles). Indexing of single-stranded libraries incorporated an index sequence next to the P7 and P5 adaptors using PCR with AccuPrime Pfx DNA polymerase (ThermoFisher, cat. no. 12344024) for its ability to read over uracil[52]. Indexed libraries were quantified using Qubit 2.0 fluorometer (ThermoFisher Scientific) and a 2200 TapeStation Instrument (Agilent Technologies) and pooled in an equimolar ratio.

Test sequencing to evaluate endogenous content was done using the Illumina MiSeq sequencing platform, producing 2× 70 bp paired-end reads. Out of the 70 libraries prepared, four libraries, coming from three samples (Fig. 1a), showed an endogenous content above 10%. These libraries were checked for complexity using preseq software[53], before major sequencing using the Illumina HiSeq X sequencing platform (2× 150 bp paired-end reads) at the Swedish Museum of Natural History.

### Data processing

Contemporary and ancient data were processed following the same pipeline. Whole-genome resequencing data were quality checked using FastQC[54]. Overlapping reads were merged, and adaptors and low-quality reads removed using SeqPrep[55], with default parameters with reads shorter than 30 bp removed. Merged reads coming from ancient samples and merged and unmerged contemporary data were mapped to the 2.8 Gb *Lynx pardinus* genome[25] and the 2.5 Gb *Felis catus* genome v9.0 (GCF_000181335.3) (ref. 56) using bwa aln 7.17 (ref. 57). After removing reads with mapping quality below 30 and sorting using samtools[58], we added read groups[59], merged sequencing runs on an individual basis[58], marked duplicates using picard tools[59] and realigned using GATK[60]. Coverage, mapping statistics and read-length distributions were calculated using samtools[58]. The authenticity of the ancient data was checked by the small fragment length distribution and the existence of ancient DNA damage—excess of C to T substitutions at the ends of the reads—as revealed by mapDamage[61] (Extended Data Fig. 1). Mapping to the Iberian lynx reference genome was used for structure and diversity analysis, while mapping to the domestic cat reference genome was used for *D* statistics tests to avoid previously reported reference bias[62]. For all our analyses, we excluded the previously identified low-complexity and low-mappability regions in lynx or in cat, as well as X and Y chromosome when specified[25,56]. We also sexed the samples by calculating the ratio of depth of coverage on the X chromosome to that of the A1 autosome. This method was validated using contemporary samples of known sex.

To avoid depth of coverage bias, we subsampled our contemporary Iberian lynx mapped samples to a depth of coverage similar to the ancient samples (approximately 2.5×) using samtools[58] and used these subsampled datasets for population structure analysis (PCA and clustering) and genetic diversity analysis. For admixture analysis (*D* statistics, gene flow direction test and $\hat{f}$ analysis), we used pseudohaplodized data (details in each section). The percentage of duplicated reads, original and subsampled depth of coverage, when it applies, are presented in Extended Data Table 5.

### Population structure

PCA was done for contemporary Iberian lynx data using ANGSD. We calculated the genotype posterior probabilities using ANGSD[63,64] and NGSTools/ngsPopGen/ngsCovar[65,66] with the following filters (-uniqueOnly 1 -remove_bads 1 -only_proper_pairs 1 -baq 1 -C 50 -minMapQ 30 -minQ 20 -doCounts 1 -minInd (number of individuals in the population/2) -setMaxDepth (3*sqrt(average (AVR) depth for the population)) -skipTriallelic 1 -SNP_pval 1e-3. To infer the ancestral state, we used a fasta sequence of *L. rufus* covering 97% of our bases obtained after mapping reads to the Iberian lynx genome, calling variants using SAMtools mpileup (-q 30) (ref. 58) and pseudohaploidizing using pu2fa (-C45)[67]. As heterogeneity along the genome in depth of coverage in the ancient samples can artificially distort PCA axes, we performed four different PCAs: one including only contemporary data and three PCAs including contemporary samples plus one different ancient sample each. Then, we projected each individual ancient PCA, that is, the coordinates of contemporary and ancient samples obtained in each of the PCA including one ancient individual, onto the axes obtained in the PCA using only contemporary data. We did that using Procrustes analysis, following ref. 68, and using the package MCMCpack[69] in R[70]. This produced new transformed coordinates for the contemporary individuals. After repeating the procedure for all three PCAs including an ancient sample, we finally plotted the mean of all the transformed and the original contemporary coordinates, plus the ancient coordinates, together.

Genotype likelihoods calculated with ANGSD[63,64] were used for a genomic clustering analysis using NGSadmix[64,71] with the same filters used for PCA analysis. NGSadmix was run ten times from $K = 1$ to $K = 5$, and the results were plotted using R. Optimal $K$ was evaluated following[72], using CLUMPAK[73].

### Nuclear genomic diversity

We calculated individual heterozygosity using ANGSD[64,74,75] and realSFS[75] on an individual-by-individual basis. Individual heterozygosity is the second value of the SFS when calculated for a diploid individual. We explored the relationship between diversity and depth of coverage by plotting the correlation between heterozygosity and depth.

At a population level, two measures of genetic diversity per site (nucleotide diversity $\pi$ and Watterson's estimator $\theta_W$) were calculated using ANGSD[64,74,75] and realSFS[75]. The analysis was run for five different groups of three random individuals in the contemporary populations. Using thetaStat[75], we performed a sliding-window approach with a window size set to 50,000 bp and a step size of 10,000. All filters described for PCA were also used for diversity calculations, except –SNP_pval (not applicable). We calculated the mean diversity weighted by the number of informative sites in each window, and s.d. by bootstrapping over windows as implemented in the 'boot' package for R[76], to account for correlation among nearby sites due to linkage disequilibrium. Finally, we computed the average of all iterations for the contemporary populations and calculated the s.d.

To check whether diversity differences were conspicuous along the genome, we calculated average diversity ($\pi$ and $\theta_W$) using ANGSD[64,74,75] and realSFS[75] across different genomic features representing neutral versus evolutionary constrained regions, namely, intergenic regions, coding gene promoters, 5′ UTRs, CDS, introns, 3′ UTRs, long non-coding (lnc) RNA promoters, lnc RNA exons, lnc RNA introns, non-coding (nc) RNA (mostly miRNAs, snRNAs and snoRNAs) and ultra-conserved non-coding elements, for the ancient and the contemporary populations. To do so, we concatenated positions of the genomic feature considered and did a window-based analysis with a window size equal to 10 kb and a window step of 5 kb. To avoid biases due to differences in sample size, we randomly selected one of the five contemporary iterations of size $n = 3$ for this analysis. S.d. was calculated using a bootstrapping over windows procedure implemented in the R package 'boot'[76] with 100 iterations. We tested whether topological windows with largest differences in diversity ($\pi$ and $\theta_W$) in contemporary versus ancient comparisons were enriched or depleted in CDS sites. To do so, we calculated differences in diversity between the ancient and the contemporary population by windows and calculated mean and s.d of the difference distribution. Diversity outlier windows were defined as windows with a difference in diversity lower than (typically negative) value of the mean −2 × s.d. Then, we calculated the percentage of CDS sites for each window and compared the distribution and mean of CDS percentage in outliers versus non-outlier windows.

### Admixture between species

We tested for admixture between Iberian lynx and Eurasian lynx populations with the $D$ statistic[77,78], using the data mapped to the domestic cat genome to avoid reference bias[8,79]. This method uses haploidized sequences of four individuals representing: two sister populations (P1 and P2), a potential introgressing population (P3) and an outgroup (P4) to define ancestral alleles (A). Then, it identifies how many derived alleles (B) are shared between P1 and P3, and between P2 and P3, leading to two different patterns: ABBA, derived allele shared by P2 and P3, and BABA, derived allele shared between P1 and P3. The $D$ statistic then is calculated as (sum(ABBA) − sum(BABA))/(sum(ABBA) + sum(BABA)). Under incomplete lineage sorting, we expect that ABBA occurred at the same frequency as BABA and, hence, the $D$ value would be 0. Excess of ABBA or BABA results in positive or negative non-zero $D$ values and is interpreted as admixture between P2 and P3 and between P1 and P3, respectively.

The standard practice for generating pseudohaploid sequences for $D$ statistic analysis of palaeogenome data has been to identify alleles by randomly selecting a single read from the read stack, to overcome any differences in sequencing coverage. However, this can be problematic when the ancient dataset is in P1 or P2, because excessive errors cause derived alleles that define the ingroup (BBBA) to be converted into $D$ statistic informative sites[80]. We therefore used the consensify method, which calls a majority consensus from a random sample of three reads, to generate haploidized sequences[80]. This method greatly reduces the error rate of the inferred alleles, while maintaining the reduction of coverage bias provided by single read sampling, and hence reduces the rate of false positives in $D$ statistic analysis.

We carried out allele counts for each dataset in ANGSD, with map and base quality filters applied (-minMapQ 30, -minQ 30). Consensify[81] was run on the resulting output files specifying a maximum depth filter of the integer read number below the 95th percentile of coverage, calculated in advance using ANGSD. We then tested for admixture on topologies compatible with the sister group relationship of Iberian and Eurasian lynx, as indicated in Fig. 3, using the program D_stat.cpp and the Python script D-stat_parser.py, available at ref. 82. The domestic cat (*Felis catus*) was used as the outgroup. Significance was assessed using a weighted block jack-knife based on 5 megabase genome windows, with absolute $Z$-scores >3 being considered statistically supported[77].

### Phylogenetic test of gene flow direction

The $D$ statistic does not provide an explicit test of the direction(s) of gene flow. To test the donor and recipient relationship in the identified admixture between Eurasian and Iberian lynx, we applied a method based on the observed frequencies of opposing tree topologies along a set of 100 kb non-overlapping windows. This method is described in ref. 8 and is conceptually identical to the test of gene flow direction provided by DFOIL statistics[83], except that it uses counts of phylogenetic tree topologies, as opposed to single-nucleotide polymorphisms (SNPs), which have been shown to be more robust to distortion by high error rates typical of palaeogenomic datasets[80]. This method analyses four individuals with a symmetrical species tree: ((Iberian+, Iberian−), (Eurasian+, Eurasian−)), with + or − denoting individuals that are more or less admixed, as identified by prior $D$ statistics analysis. In this experimental design, if the variability in admixture among the Iberian individuals is explained by gene flow from Eurasian into Iberian lynx, this will cause an excess of non-symmetrical tree topologies with Iberian− in the basal position, relative to the number with Iberian+ in the basal position. To maximize sensitivity, the analysis was carried out for each contemporary Iberian lynx, using representative individuals of ancient Iberian lynx and contemporary Eurasian lynx selected to maximize the observed $D$ values and the level of sequencing coverage. Pseudohaploid sequences were generated using random single-read sampling in ANGSD, with base (-minQ 30) and map (-minMapQ 30) quality filters applied. A custom bash script making use of the SNP-sites program[84] was then used to divide the aligned sequences into 100 kb non-overlapping windows and to calculate the maximum-likelihood phylogeny of each using the BINGAMMA model in RaxML[85] and the domestic cat as outgroup to root the trees (available at ref. 86).

### $\hat{f}$ analysis

After determining the predominant direction of gene flow was from Eurasian into Iberian lynx, we estimated the introgressed genome fraction in contemporary Iberian lynx above that occurring in ancient Iberian lynx, assuming unidirectional gene flow using the $\hat{f}$ method[77]. The $\hat{f}$ statistics measures the excess of shared derived alleles between the admixed individual and candidate introgressor standardized by the maximum excess of shared derived alleles expected in an entirely (100%) introgressed individual. These analyses used the same consensify sequences as used in $D$ statistics analysis and the F_hat.cpp program and the Python script F_hat_parser.py (available at ref. 82). We ran these

tests using all combinations of contemporary and ancient Iberian lynx and western Eurasian lynx. The latter were selected as best representing the introgressing Eurasian lynx population, which probably occurred at the contact zone between the two species in Western Europe. Outgroup and tests of significance were as described for the *D* statistics analyses.

## References


1. Allendorf, F. W., Funk, W. C., Aitken, S. N., Byrne, M. & Luikart, G. *Conservation and the Genomics of Populations* 3rd edn (Oxford Univ. Press, 2012).
2. Frankham, R., Ballou, J. D. & Briscoe, D. A. *Introduction to Conservation Genetics* 2nd edn (Cambridge Univ. Press, 2010).
3. de Bruyn, M., Hoelzel, A. R., Carvalho, G. R. & Hofreiter, M. Faunal histories from Holocene ancient DNA. *Trends Ecol. Evol.* **26**, 405–413 (2011).
4. Diez-del-Molino, D., Sanchez-Barreiro, F., Barnes, I., Gilbert, M. T. P. & Dalen, L. Quantifying temporal genomic erosion in endangered species. *Trends Ecol. Evol.* **33**, 176–185 (2018).
5. Hofreiter, M. & Barnes, I. Diversity lost: are all Holarctic large mammal species just relict populations? *BMC Biol.* **8**, 46 (2010).
6. Leigh, D. M., Hendry, A. P., Vázquez-Domínguez, E. & Friesen, V. L. Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. *Evol. Appl.* **12**, 1505–1512 (2019).
7. Ralls, K. et al. Call for a paradigm shift in the genetic management of fragmented populations. *Conserv. Lett.* **11**, e12412 (2018).
8. Barlow, A. et al. Partial genomic survival of cave bears in living brown bears. *Nat. Ecol. Evol.* **2**, 1563–1570 (2018).
9. Iacolina, L., Corlatti, L., Buzan, E., Safner, T. & Sprem, N. Hybridisation in European ungulates: an overview of the current status, causes, and consequences. *Mamm. Rev.* **49**, 45–59 (2019).
10. Kumar, V. et al. The evolutionary history of bears is characterized by gene flow across species. *Sci. Rep.* **7**, 46487 (2017).
11. Li, G., Davis, B. W., Eizirik, E. & Murphy, W. J. Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res.* **26**, 1–11 (2016).
12. Palkopoulou, E. et al. A comprehensive genomic history of extinct and living elephants. *Proc. Natl Acad. Sci. USA* **115**, E2566–E2574 (2018).
13. Chan, W. Y., Hoffmann, A. A. & van Oppen, M. J. H. Hybridization as a conservation management tool. *Conserv. Lett.* **12**, e12652 (2019).
14. Quilodrán, C. S., Montoya-Burgos, J. I. & Currat, M. Harmonizing hybridization dissonance in conservation. *Commun. Biol.* **3**, 391 (2020).
15. Corbet, G. B. & Hill, J. E. *A World List of Mammalian Species* 2nd edn (British Museum of Natural History, 1986).
16. Tumlison, R. Felis lynx. *Mamm. Species* **269**, 1–8 (1987).
17. Weigel, I. *Das Fellmuster der wildlebenden Katzenarten und der Hauskatze in vergleichender und stammesgeschichtlicher Hinsicht Säugetierkundliche Mitteilungen* (München, 1961).
18. Van den Brink, F.-H. Distribution and speciation of some carnivores. *Mamm. Rev.* **1**, 67–79 (1970).
19. Kurten, B. & Granqvist, E. Fossil pardel lynx (*Lynx pardina spelaea Boule*) from a cave in southern France. *Ann. Zool. Fenn.* **24**, 39–43 (1987).
20. Matjuschkin E. N. *Der Luchs* (A. Ziemsen Verlag, 1978).
21. Werdelin, L. The evolution of lynxes. *Ann. Zool. Fenn.* **18**, 37–71 (1981).
22. Harris, A., Foley, N., Williams, T. & Murphy, W. Tree house explorer: a novel genome browser for phylogenomics. *Mol. Biol. Evol.* **39**, msac130 (2022).
23. Li, G., Figueiró, H. V., Eizirik, E. & Murphy, W. J. Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. *Mol. Biol. Evol.* **36**, 2111–2126 (2019).
24. Casas-Marcé, M. et al. Spatio-temporal dynamics of genetic variation in the Iberian lynx along its path to extinction reconstructed with ancient DNA. *Mol. Biol. Evol.* **34**, 2893–2907 (2017).
25. Abascal, F. et al. Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol.* **17**, 251 (2016).
26. Korlević, P. et al. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques* **59**, 87–93 (2015).
27. Bazzicalupo, E. et al. History, demography and genetic status of Balkan and Caucasian *Lynx lynx* (Linnaeus, 1758) populations revealed by genome-wide variation. *Divers. Distrib.* **28**, 65–82 (2022).
28. Lucena-Perez, M. et al. Genomic patterns in the widespread Eurasian lynx shaped by late quaternary climatic fluctuations and anthropogenic impacts. *Mol. Ecol.* **29**, 812–828 (2020).
29. Lucena-Perez, M. et al. Ancient genome provides insights into the history of Eurasian lynx in Iberia and Western Europe. *Quat. Sci. Rev.* **285**, 107518 (2022).
30. Kleinman-Ruiz, D. et al. Purging of deleterious burden in the endangered Iberian lynx. *Proc. Natl Acad. Sci. USA* **119**, e2110614119 (2022).
31. Lucena-Perez, M. et al. Bottleneck-associated changes in the genomic landscape of genetic diversity in wild lynx populations. *Evol. Appl.* **14**, 2664–2679 (2021).
32. D'Elia, J., Haig, S. M., Mullins, T. D. & Miller, M. P. Ancient DNA reveals substantial genetic diversity in the California Condor (*Gymnogyps californianus*) prior to a population bottleneck. *Condor* **118**, 703–714 (2016).
33. Dufresnes, C. et al. Howling from the past: historical phylogeography and diversity losses in European grey wolves. *Proc. Royal Soc. B Biol. Sci.* **285**, 20181148 (2018).
34. Dussex, N., von Seth, J., Robertson, B. C. & Dalen, L. Full mitogenomes in the critically endangered kakapo reveal major post-glacial and anthropogenic effects on neutral genetic diversity. *Genes* **9**, 220 (2018).
35. Feng, S. et al. The genomic footprints of the fall and recovery of the crested ibis. *Curr. Biol.* **29**, 340–349.e347 (2019).
36. Sánchez-Barreiro, F. et al. Historical population declines prompted significant genomic erosion in the northern and southern white rhinoceros (*Ceratotherium simum*). *Mol. Ecol.* **30**, 6355–6369 (2021).
37. Sheng, G. L. et al. Ancient DNA from giant panda (*Ailuropoda melanoleuca*) of south-western China reveals genetic diversity loss during the Holocene. *Genes* **9**, 198 (2018).
38. van der Valk, T., Diez-del-Molino, D., Marques-Bonet, T., Guschanski, K. & Dalen, L. Historical genomes reveal the genomic consequences of recent population decline in eastern gorillas. *Curr. Biol.* **29**, 165–170 e6 (2019).

39. Vila, C. et al. Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proc. Royal Soc. B Biol. Sci.* **270**, 91–97 (2003).
40. Clavero, M. & Delibes, M. Using historical accounts to set conservation baselines: the case of lynx species in Spain. *Biodivers. Conserv.* **22**, 1691–1702 (2013).
41. Jiménez, J., Clavero, M. & Reig-Ferrer, A. New old news on the 'lobo cerval' (*Lynx lynx*?) in NE Spain. *Galemys* **30**, 1–6 (2018).
42. Mecozzi, B. et al. The tale of a short-tailed cat: new outstanding late Pleistocene fossils of *Lynx pardinus* from southern Italy. *Quat. Sci. Rev.* **262**, 107028 (2021).
43. Rodríguez-Varela, R. et al. Ancient DNA reveals past existence of Eurasian lynx in Spain. *J. Zool.* **298**, 94–102 (2016).
44. Rodríguez-Varela, R. et al. Ancient DNA evidence of Iberian lynx palaeoendemism. *Quat. Sci. Rev.* **112**, 172–180 (2015).
45. Bell, D. A. et al. The exciting potential and remaining uncertainties of genetic rescue. *Trends Ecol. Evol.* **34**, 1070–1079 (2019).
46. Tallmon, D. A., Luikart, G. & Waples, R. S. The alluring simplicity and complex reality of genetic rescue. *Trends Ecol. Evol.* **19**, 489–496 (2004).
47. Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C. & Tallmon, D. A. Genetic rescue to the rescue. *Trends Ecol. Evol.* **30**, 42–49 (2015).
48. Detry, C. & Arruda, A. M. A fauna da Idade do Ferro e da Época Romana de Monte Molião (Lagos, Algarve): continuidades e rupturas na dieta alimentar. *Revista Portuguesa de Arqueologia* **16**, 213–226 (2013).
49. Fulton T. L. in *Ancient DNA: Methods and Protocols* (eds Shapiro B. & Hofreiter M.) (Humana, 2012).
50. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
51. Gansauge, M. T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).
52. Dabney, J. & Meyer, M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques* **52**, 87–94 (2012).
53. Preseq software. *Smith Lab Research* https://github.com/smithlabcode/preseq (2014).
54. FastQC software. *Babraham Bioinformatics* https://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
55. SeqPrep software. *John St. John* https://github.com/jstjohn/SeqPrep (2016).
56. Buckley, R. et al. A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *PLoS Genet.* **16**, e1008926 (2020).
57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
58. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
59. Picard software. *Broad Institute* https://broadinstitute.github.io/picard (2014).
60. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
61. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
62. Sheng, G.-L. et al. Paleogenome reveals genetic contribution of extinct giant panda to extant populations. *Curr. Biol.* **29**, 1695–1700 e6 (2019).
63. Kim, S. Y. et al. Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC Bioinf.* **12**, 231 (2011).
64. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
65. Fumagalli, M. Assessing the effect of sequencing depth and sample size in population genetics inferences. *PLoS One* **8**, e79667 (2013).
66. Fumagalli, M., Vieira, F. G., Linderoth, T. & Nielsen, R. ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics* **30**, 1486–1487 (2014).
67. Chrom-Compare software. *Paleogenomics* https://github.com/Paleogenomics/Chrom-Compare (2014).
68. Borg I. & Groenen P. J. F. *Modern Multidimensional Scaling: Theory and Applications* (Springer, 1997).
69. Martin, A. D., Quinn, K. M. & Park, J. H. MCMCpack: Markov chain Monte Carlo in R. *J. Stat. Softw.* **42**, 1–21 (2011).
70. R_Core_Team R: a language and environment for statistical computing (R Foundation for Statistical Computing, 2019).
71. Skotte, L., Korneliussen, T. S. & Albrechtsen, A. Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693–702 (2013).
72. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
73. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
74. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinf.* **15**, 356 (2014).
75. Korneliussen, T. S., Moltke, I., Albrechtsen, A. & Nielsen, R. Calculation of Tajima's *D* and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinf.* **14**, 289 (2013).
76. Canty A. & Ripley B. boot: bootstrap R (S-Plus) functions. R package version 1.3–11 (2014).
77. Durand, E. Y., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
78. Green, R. E. et al. A draft sequence of the neandertal genome. *Science* **328**, 710–722 (2010).
79. Gunther, T. & Nettelblad, C. The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genet.* **15**, e1008302 (2019).
80. Barlow, A., Hartmann, S., González, J., Hofreiter, M. & Paijmans, J. Consensify: a method for generating pseudohaploid genome sequences from palaeogenomic datasets with reduced error rates. *Genes* **11**, 50 (2020).
81. Consensify software. *Barlow, A. and Paijmans, J.L.A.* https://github.com/jlapaijmans/Consensify (2018).
82. Admixture workflow. *Cahill, J.A.* https://github.com/jacahill/Admixture (2018).
83. Pease, J. B. & Hahn, M. W. Detection and polarization of introgression in a five-taxon phylogeny. *Syst. Biol.* **64**, 651–662 (2015).
84. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016).
85. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
86. TreeHacker software. *Paijmans, J.L.A and Barlow, A.* https://github.com/jlapaijmans/treehacker (2023).
87. Rodríguez, A. & Delibes, M. Internal structure and patterns of contraction in the geographic range of the Iberian lynx. *Ecography* **25**, 314–328 (2002).

## Acknowledgements

This research was funded by the Spanish Dirección General de Investigación Científica y Técnica through projects CGL2013-47755-P and CGL2017-84641-P to J.A.G and is an extension of a project on ancient lynx genetics granted to Miguel Delibes de Castro by the Fundación BBVA. M.L.-P. was supported by a PhD contract from Programa Internacional de Becas 'La Caixa-Severo Ochoa'. J.N. received financial support through projects HAR2014- 55131 from the Ministerio de Ciencia e Innovación and SGR2014-108 from the Generalitat de Catalunya. We acknowledge support from Science for Life Laboratory, the Knut and Alice Wallenberg Foundation, the National Genomics Infrastructure funded by the Swedish Research Council and Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. We also acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund via the Welsh Government. Logistical support was provided by the Laboratorio de Ecología Molecular certified to ISO9001:2015 and ISO14001:2015 quality and environmental management systems. Data processing and most calculations and analyses were carried out in the Genomics servers of Doñana's Singular Scientific-Technical Infrastructure, with additional computing and storage resources provided by Fundación Pública Galega, Centro Tecnolóxico de Supercomputación de Galicia. Logistical laboratory support was provided by the Laboratorio de Ecología Molecular certified to ISO9001:2015 and ISO14001:2015 quality and environmental management systems. Data processing and most calculations and analyses were carried out in the Genomics servers of Doñana's Singular Scientific-Technical Infrastructure, with additional computing and storage resources provided by Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia.

## Author contributions

J.A.G. conceived the project. A.B., J.A.G. and M.L.-P. designed the study. C.D., F.N. and J.N. provided ancient samples and critical input on archaeological context. M.L.-P. performed the laboratory work under the supervision of J.L.A.P. and M.H. A.B., J.L.A.P. and M.L.-P. analysed the data. A.B., J.A.G., J.L.A.P. and M.L.-P. interpreted the results, with critical input from M.H. and L.D. M.L.-P. drafted the manuscript with support from A.B. and J.A.G. and input by J.L.A.P., C.D., J.N., M.H. and L.D. All authors approved the final version of the manuscript.