

# Bottleneck-associated changes in the genomic landscape of genetic diversity in wild lynx populations

Maria Lucena-Perez<sup>1</sup> | Daniel Kleinman-Ruiz<sup>1,2</sup> | Elena Marmesat<sup>1</sup> |  
Alexander P. Saveljev<sup>3</sup> | Krzysztof Schmidt<sup>4</sup> | José A. Godoy<sup>1</sup> 

<sup>1</sup>Departamento de Ecología Integrativa, Estación Biológica de Doñana (CSIC), Sevilla, Spain

<sup>2</sup>Departamento de Genética, Facultad de Biología, Universidad Complutense, Madrid, Spain

<sup>3</sup>Department of Animal Ecology, Russian Research Institute of Game Management and Fur Farming, Kirov, Russia

<sup>4</sup>Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland

## Correspondence

Maria Lucena-Perez and José A. Godoy, Departamento de Ecología Integrativa, Estación Biológica de Doñana (CSIC), Sevilla, Spain.

Emails: lucenaperezmaria@gmail.com (M.L.-P.) and godoy@ebd.csic.es (J.A.G.)

## Funding information

Dirección General de Investigación Científica y Técnica, Grant/Award Number: CGL2013-47755-P, CGL2017-84641-P, PGC2018-095810-B-I00 and SEV-2012-0262; Russian Science Foundation, Grant/Award Number: 18-14-00093; "la Caixa" Foundation, Grant/Award Number: LCF/BQ/SO14/52250035 and LCF/BQ/SO15/52260006; Narodowe Centrum Nauki, Grant/Award Number: 2014/15/B/ NZ8/00212; FP7 People: Marie-Curie Actions, Grant/Award Number: PIRSEGA-2009-247652

## Abstract

Demographic bottlenecks generally reduce genetic diversity through more intense genetic drift, but their net effect may vary along the genome due to the random nature of genetic drift and to local effects of recombination, mutation, and selection. Here, we analyzed the changes in genetic diversity following a bottleneck by comparing whole-genome diversity patterns in populations with and without severe recent documented declines of Iberian (*Lynx pardinus*,  $n = 31$ ) and Eurasian lynx (*Lynx lynx*,  $n = 29$ ). As expected, overall genomic diversity correlated negatively with bottleneck intensity and/or duration. Correlations of genetic diversity with divergence, chromosome size, gene or functional site content, GC content, or recombination were observed in nonbottlenecked populations, but were weaker in bottlenecked populations. Also, functional features under intense purifying selection and the X chromosome showed an increase in the observed density of variants, even resulting in higher  $\theta_W$  diversity than in nonbottlenecked populations. Increased diversity seems to be related to both a higher mutational input in those regions creating a large collection of low-frequency variants, a few of which increase in frequency during the bottleneck to the point they become detectable with our limited sample, and the reduced efficacy of purifying selection, which affects not only protein structure and function but also the regulation of gene expression. The results of this study alert to the possible reduction of fitness and adaptive potential associated with the genomic erosion in regulatory elements. Further, the detection of a gain of diversity in ultra-conserved elements can be used as a sensitive and easy-to-apply signature of genetic erosion in wild populations.

## KEYWORDS

bottleneck, endangered species, genetic diversity, genetic drift, genetic erosion, genetic load, genomic landscape, purifying selection

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Genetic diversity is one of the three main components of biodiversity, together with species and ecosystem diversity, which we need to preserve in the current biodiversity crisis. Genetic diversity is intimately related to adaptive potential and the loss of genetic diversity is generally associated with reduced fitness and increased extinction risks, so the importance of genetic variation transcends to species diversity and is also connected to ecosystem resilience (Allendorf et al., 2013). Despite this, genetic diversity is currently largely neglected in conservation legislation and practice and this has given rise to urgent calls to revert this situation and to incorporate the routine monitoring of genetic diversity, including wild species, in conservation programs (Des Roches et al., 2021; Hoban, Bruford, et al., 2021; Hoban, Campbell, et al., 2021; Laikre et al., 2020).

Genetic diversity in populations is the result of the interplay between mutation, genetic drift, recombination, selection, and gene flow (Ballenghien et al., 2017; Ellegren & Galtier, 2016). In demographically stable populations, the site frequency spectrum (SFS) and genetic diversity varies along the genome and across genomic features, following variations in the strength of recombination, selection and mutation, and the random action of genetic drift, creating a genomic landscape of genetic diversity. Changes in effective population size alter these equilibrium patterns of genetic diversity across the genome. In particular, a demographic bottleneck (i.e., a sudden reduction in  $N_e$ ) will increase the action of genetic drift and cause a general loss of genetic diversity, in a magnitude determined by its severity and duration (Garza & Williamson, 2001), and will distort the SFS through the preferential loss of low-frequency alleles (Nei & Maruyama, 1975). However, we might expect that the effect of demographic bottlenecks on genetic diversity will locally depart from predictions based on equilibrium or simple drift models because evolutionary forces interact in complex ways between them and with genetic drift (Duret & Arndt, 2008; Duret & Galtier, 2009; Halldorsson et al., 2019; Pratto et al., 2014; Smith et al., 2018; Terekhanova et al., 2017; Williams et al., 2015).

Genetic diversity is ultimately created by mutation, and mutation rates, often estimated by divergence rates between species, vary extensively along and among chromosomes (Gonzalez-Perez et al., 2019). For example, it is known that GC content is positively correlated with genetic diversity, partly due to the hyper-mutability of CpG dinucleotides (Duret & Arndt, 2008; Duret & Galtier, 2009; Smith et al., 2018). Recombination is also positively correlated with genetic diversity due to its associated mutagenic effect (Duret & Arndt, 2008; Duret & Galtier, 2009; Halldorsson et al., 2019; Pratto et al., 2014; Smith et al., 2018; Terekhanova et al., 2017) and to its reduction of genetic hitchhiking and background selection, as described below (Stephan, 2010).

Natural selection impacts genetic diversity at selected sites by either reducing (purifying or negative selection, and positive selection) or maintaining it (balancing selection). The effects of selection on diversity may extend to neighboring linked neutral sites especially in regions of low recombination, a process often referred to

as hitchhiking or linked selection. Selective sweeps and background selection are the consequence of linked positive and negative selection, respectively (Cutter & Payseur, 2013). Furthermore, the effectiveness of selection may be hampered if multiple linked loci experience selective pressures simultaneously via Hill-Robertson interference (Felsenstein, 1974; Hill & Robertson, 1966).

A reduction in  $N_e$  would impact regions under selection leading to two predictions. Firstly, variants with selection coefficients lower than the inverse of  $N_e$  will behave as effectively neutral (Kimura, 1962), resulting in the accumulation of moderately deleterious variation in small populations due to the reduced efficacy of purifying selection (Charlesworth et al., 1993). Secondly, the higher inbreeding rates in small populations may expose in homozygosity the deleterious effects of (partially) recessive mutations, thus facilitating their elimination by selection (i.e., purging) (García-Dorado, 2007; Garcia-Dorado, 2012; Hedrick & Garcia-Dorado, 2016).

Levels of genetic diversity vary extensively along and among chromosomes following variations in the relative intensity of these evolutionary processes; these correlations are often captured through proxies. For example, diversity is known to vary regionally with gene density, related to the action of natural selection (Martin et al., 2016; Payseur & Nachman, 2002), or with GC content, which correlates with mutation. Meanwhile, the low and high genetic diversity in centromeric and telomeric regions is largely attributed to low and high recombination rates, respectively. Sex chromosomes usually show highly contrasting patterns of diversity with respect to autosomes, as they are impacted quite differently by these processes. The X chromosome, for example, is expected to show  $\frac{3}{4}$  of the diversity of the autosomes on the basis of its lower effective size due to male hemizyosity, but deviations from this ratio can arise due to several causes, including demographic changes (with bottlenecks reducing and population growth increasing the ratio; Pool & Nielsen, 2007), sex-biased demography (e.g., reproductive success or migration rates), recombination only in females (excluding the pseudo-autosomal region), sex-biased mutation, and differences in selection regime due, among other things, to the exposure of variants in hemizygous males (reviewed in Schaffner, 2004; Wilson Sayres, 2018).

Genetic diversity has been traditionally assessed in wild species through the use of anonymous and purportedly neutral genetic markers, but genome-wide assessments are becoming increasingly feasible and common (Hohenlohe et al., 2021). The transition from markers to genomes offers the opportunity to incorporate functional variation into the genetic assessment of endangered species, and this has prompted calls to abandon neutral variation in favor of functional variation and even to question the relevance of neutral diversity in conservation (Teixeira & Huber, 2021). Whereas there are many arguments and practical considerations for maintaining a focus on neutral variation in conservation genomic studies, it is clear that the field greatly needs a better understanding of the interaction of functional genetic diversity and demographic history (DeWoody et al., 2021; García-Dorado & Caballero, 2021). Furthermore, the incorporation of functional variation might provide novel and much needed parameters to effectively detect populations where genetic diversity might be compromising fitness and population viability.

Despite the increasing interest and decreasing costs of genomic analyses in conservation, the lack of the required resources in most endangered species, such as a reference genome, and/or appropriate samples, has limited progress. We find, on one hand, few empirical studies assessing differences in diversity along the genome in stable populations of nonmodel organisms (e.g. Corcoran et al., 2017; Dutoit et al., 2017; Wang et al., 2016), and on the other, simulations exploring the dynamics of diversity in nonequilibrium scenarios, including demographic bottlenecks (Torres et al., 2020). To our knowledge, the few empirical studies analyzing patterns of diversity following a population bottleneck on wild conservation-relevant populations mostly deal with genome-wide averages, and at most contrast putatively neutral and selected coding sites (e.g., synonymous vs. nonsynonymous variation) (Grossen et al., 2020; Robinson et al., 2016, 2018; Saremi et al., 2019; Xue et al., 2015).

Here, we characterize the changes in the genomic landscape of diversity brought about by a demographic bottleneck in the two lynx species present in Eurasia, the highly endangered Iberian lynx (*Lynx pardinus*) and the broadly distributed Eurasian lynx (*Lynx lynx*). By taking advantage of the availability of a *de novo* assembled and annotated Iberian lynx reference draft genome (Abascal et al., 2016), we measure genetic diversity across chromosomes, chromosomal regions, and genomic features, in populations with contrasting recent demographic histories. To do so, we assess genomic diversity patterns in one population per species that best represent an ancestral unaffected (i.e., prebottleneck) scenario (from now on referred to as nonbottlenecked populations, NB), and we compare them with those of genetically eroded populations of the same lineage (bottlenecked populations, B), two populations of Eurasian lynx and one population of Iberian lynx. We also assess the role of recombination, mutation, and selection in shaping postbottleneck diversity patterns by comparing local values between regions or features with contrasting genetic diversity dynamics. By focusing on the variance of bottleneck effects across the genome, we aim to: (a) assess how the magnitude of loss of genetic diversity caused by a bottleneck varies among chromosomes, chromosomal regions, and genomic features, and (b) evaluate whether this magnitude is related to differences in local levels of selection, mutation, and recombination. By providing snapshots of the changes in genetic diversity in different functional categories, this study contributes to a better understanding of the fitness consequences of population bottlenecks and species declines. In particular, it alerts on the reduction of fitness and adaptive potential associated to the genomic erosion in regulatory elements, and proposes the gain of diversity in ultra-conserved elements as a sensitive and easy-to-apply signature of genetic erosion.

## 2 | MATERIALS AND METHODS

### 2.1 | Study populations

The Iberian lynx was recognized as the most endangered felid in the world (Nowell & Jackson, 1996). In 2002, when the species was classified as critically endangered, only two populations persisted with <100 individuals in total, the only remnants of a likely

panmictic population thousands of years ago, which became progressively contracted and fragmented in the last two or three centuries (Casas-Marce et al., 2017; Ferreras et al., 2010; Palomares et al., 2011; Rodríguez & Delibes, 2003). The highly eroded and peripheral Doñana population remained isolated during two centuries at an estimated  $N_e$  of 20 individuals, reaching 10 in the last few decades (Casas-Marce et al., 2013, 2017). In contrast, Andújar is the remnant of the Eastern Sierra Morena subpopulation, which remained large and genetically connected with other populations until ca. 1950, and has since decreased in size to a minimum of  $N_e \approx 21$  estimated for the year 2002 (Casas-Marce et al., 2013, 2017) (see also Figure S4 in [35] for a demographic reconstructions of Iberian lynx populations from 1950 to 2015). This contrasting recent history is reflected in the overall genomic diversity in Doñana being about half of Andújar's, with the latter representing about 84% of that of the historical Iberian lynx metapopulation (Casas-Marce et al., 2017).

Unlike the Iberian lynx, the Eurasian lynx is one of the most broadly distributed felids in the world (Breitenmoser et al., 2015). In Europe, there are large differences between populations in terms of neutral genetic differentiation and diversity (Lucena-Perez et al., 2020). Particularly, NE-Poland and the Scandinavian populations, represented here by Norway, went through pronounced declines and remained isolated from other lynx populations during the last century. Specifically, NE-Poland and Norway went through sharp population declines 200–300 ya and maintained effective sizes below 100 for the last 100–150 years (Lucena-Perez et al., 2020) (Bazzicalupo et al., 2021). On the contrary, eastern European populations remained well connected and relatively large, and Kirov represents the population of the European lineage with the highest genetic diversity (Lucena-Perez et al., 2020; Ratkiewicz et al., 2014). This population has apparently maintained effective sizes well over 1,000 at least until 500 ya and over 500 until present. These contrasting histories are reflected in average genomic diversity, as reported previously, Poland and Norway showing around 80% and 74% of that found in Kirov, respectively (Lucena-Perez et al., 2020).

For the sake of simplicity, and in light of their contrasting recent history, we refer to Doñana, Norway, and NE-Poland as bottlenecked (B) populations, and Andújar and Kirov as nonbottlenecked (NB) populations, although the latter have also declined in size to a much lower extent.

### 2.2 | Sampling, DNA extraction, library preparation, and sequencing

Whole-genome resequencing data were obtained from 60 Iberian and Eurasian lynx (Figure 1, Table S1). We sampled and processed 20 Iberian samples, which together with available sequence data from 11 additional individuals (Abascal et al., 2016), resulted in a total of 31 Iberian lynx from the two remnant populations: Andújar ( $n = 19$ ) and Doñana ( $n = 12$ ). We used 29 whole-genome sequences from Eurasian lynx (Lucena-Perez et al., 2020) from three different populations: Kirov

region, Russia ( $n = 13$ ), NE-Poland (Białowieża and Knyszyn Primeval Forests) ( $n = 8$ ), and Norway ( $n = 8$ ). Samples were digested overnight using proteinase K and DNA was extracted using NucleoMag B-beads (NucleoMag DNA from tissue kit) in LEM-EBD facilities (Seville, Spain). gDNA samples were used for preparing Illumina sequencing compatible paired-end libraries. The libraries were prepared, quantified, and sequenced in Illumina HiSeq2000 flowcell v3 (Illumina Inc.), in  $2 \times 101$  bp paired-end mode, following standard Illumina procedures at Centro Nacional de Análisis Genómico (CNAG-CRG).

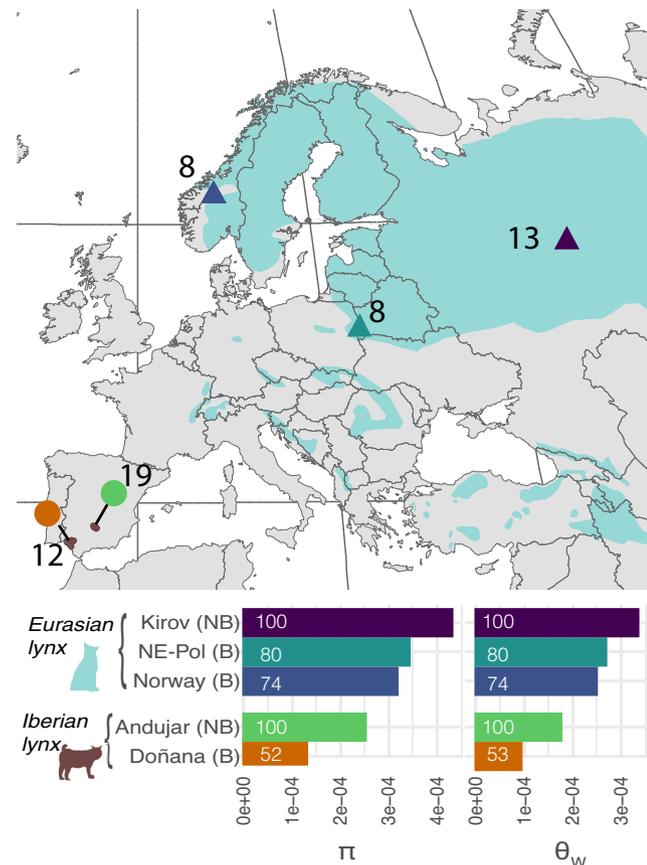
## 2.3 | Quality control, trimming, and mapping

Iberian and Eurasian lynx data were quality-controlled using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>), and adaptors were removed when necessary using SeqPrep (<https://github.com/jstjohn/SeqPrep>). All short-read sequence data were mapped to a 2.8 Gb Iberian lynx LYPA1.0 genome assembly (Abascal et al., 2016) using BWA-MEM (Li, 2013) with default parameters. We

added read groups to each sample using picard-tools (<https://broadinstitute.github.io/picard/>) and merged the bam files from the same individual with SAMtools merge (Li et al., 2009). We marked PCR duplicates using picard-tools (<https://broadinstitute.github.io/picard/>) and performed a local realignment and a base quality recalibration of the data using GATK 3.4 (McKenna et al., 2010). We calculated overall mapping stats using SAMtools flagstat (Li et al., 2009) and average depth using SAMtools depth (Li et al., 2009). We standardized sample depth to avoid any bias, so medium–high depth samples were subsampled to a depth within the range of the newly sequenced samples, using SAMtools view -s (Li et al., 2009). Average depth of coverage for Eurasian lynx samples was  $5.9 \times$  (range 5.0–7.6  $\times$ ).

## 2.4 | Genome annotation

The Iberian lynx reference genome has been annotated for several genomic features, namely, 3' UTR, CDS, introns, and 5' UTR (defined on the principal isoform of each gene), lncRNAs, and ncRNAs (Abascal et al., 2016). Here, we also included promoters of protein-coding genes and lncRNA, defined as sequences 1000-bp upstream of the gene or lncRNA, respectively, and Ultra Conserved Non-Coding Elements (UCNEs). UCNEs (Dimitrieva & Bucher, 2013) were annotated by translating human coordinates (<https://ccg.vital-it.ch/UCNEbase/>) into domestic cat *Felis\_catus\_5.0* coordinates using LiftOver (<http://rohsdb.cmb.usc.edu/GBshape/cgi-bin/hgLiftOver>), and these to lynx LYPA1.0 coordinates using lynx to cat synteny (Abascal et al., 2016). To define intergenic regions we added a security buffer of  $\pm 1000$  bp to any annotated region to avoid the influence of adjacent areas, using bedtools subtract and bedtools intersect (Quinlan & Hall, 2010). Each single 3' UTR, CDS, intron, 5' UTR, lncRNA, ncRNA, promoter (for genes and lncRNA), UCNE, and intergenic region—from now on, each single unit—was also assigned to chromosomes when possible using lynx to cat synteny, and excluding from the analysis units that overlapped more than one chromosome. We annotated as subtelomeric and pericentromeric those contigs containing more than 1000 bp syntenic to these regions in the domestic cat genome, that is, 2 Mb away from the telomeres and 10 Mb around the centromere, respectively (Abascal et al., 2016). Then, we calculated the percentage of sites in each unit overlapping subtelomeric, pericentromeric, or interstitial regions (defined as regions not pericentromeric or subtelomeric), and assigned units to regions where the overlap was  $\geq 75\%$ . The number of units, along with their average size, concatenated length, and number of total and informative sites, for each population, feature, and chromosome type is given in Table S12.



**FIGURE 1** Sampling and global diversity ( $\pi$  and  $\theta_w$ ) of the five studied populations. Sampled populations of the Eurasian lynx (triangle) and the Iberian lynx (circle) are marked within the distribution of the species (light blue for the European part of the Eurasian lynx distribution; dark red for the Iberian lynx). Diversity estimates for each population are represented below. For B populations, the diversity percentage relative to the reference NB population is indicated in the bar plot

## 2.5 | Genetic diversity per unit

For each population, genetic diversity per site was calculated using ANGSD (Korneliussen et al., 2013) with the following filters: -unique-Only 1 -remove\_bads 1 -only\_proper\_pairs 1 -baq 1 -C 50 -min-MapQ 30 -minQ 20 -doCounts 1 -minInd (number of individuals in

the population/2) –setMaxDepth (average [AVR] depth for the population + [0.95\*SD depth for the population]) –setMinDepth (AVR depth for the population – [0.95\*SD depth for the population]) –skip-Triallelic 1). First, we calculated the site allele frequency (SAF) and the site frequency spectrum (SFS) for each population using ANGSD and NGSTOOLS/realSFS (Korneliussen et al., 2013). This SFS plus its corresponding SAF were used to calculate nucleotide diversity ( $\pi$ ), and Watterson theta estimator ( $\theta_W$ ) per site (Korneliussen et al., 2013). It must be noted that ANGSD reports values of diversity per site whenever a site passes the filters, without distinguishing between variable and invariant sites, and that site diversity is not zero even for invariant sites because it considers the likelihood of all genotypes in the calculation. Then, we calculated diversity per unit by averaging diversity across sites, considering only informative sites (defined as all sites remaining after filtering). To avoid biases due to very few informative sites, we only considered units with more than 50 informative sites and with information for at least 20% of the positions. Results are presented for autosomes (A), unless otherwise noted. For the X chromosome,  $\theta_W$  was corrected for the sample size of this chromosome based on the number of males and females sampled in each population (Lucena-Perez et al., 2020). Additionally, we excluded the X chromosome pseudoautosomal region (PAR) delimited here as 10 Mb from the extreme of the chromosome, based on an estimate of 9 Mb for the domestic cat X chromosome PAR (Li et al., 2016).

We averaged  $\pi$ , and  $\theta_W$  for the different populations, chromosomes, regions (subtelomeric regions vs. pericentromeric regions vs. interstitial), and features (Intergenic, Gen promoter, 5' UTR, CDS, Intron, 3' UTR, lncRNA promoter, lncRNA exons, lncRNA intron, ncRNA, UCNE), weighting by the number of informative sites. Then, we used  $\pi$ , and  $\theta_W$  weighted means to calculate a measure of the skewness of the SFS toward rare variants ( $S$ ) as  $1 - (\pi/\theta_W)$ , where a positive  $S$  value indicates excess, and a negative value deficit, of rare variants over the mutation-drift equilibrium for neutral alleles under the infinite sites model (Becher et al., 2020). Dispersion in diversity estimates ( $SD$ ) was calculated by bootstrapping units using the R package boot (Canty & Ripley, 2017) with 100 iterations. We then computed ratios of diversity in B vs. NB populations in different chromosomes, regions, and features.

We also computed the ratios of  $\pi$ , and  $\theta_W$  diversity between the X chromosome and the autosomes (X/A diversity ratios) for each population and feature. Dispersion in X/A ratios was calculated from dispersion in X and dispersion in A, using a propagation of the uncertainties formula, where  $SD$  refers to the standard deviation

$$\sqrt{\frac{A SD^2}{A^2} + \frac{X SD^2}{X^2}} * \left(\frac{X}{A}\right)^2$$

## 2.6 | Relative diversity differences between B and NB populations across features

We quantified relative diversity differences between pairs of one B population and its corresponding NB population in different genomic features. By doing so, we take NB as the closest representation of

the ancestral population and assume that observed differences in diversity are mostly due to population contraction in the B population. Comparisons included Norway vs. Kirov and NE-Poland vs. Kirov for Eurasian lynx, and Doñana vs. Andújar for Iberian lynx. We used fix-size adjacent windows on a concatenation of units of each feature to avoid distortions caused by differences in unit length across features. Additionally, we subsampled populations within species to the same number of individuals (i.e., eight for Eurasian, and 12 for Iberian lynx) to avoid biases due to differences in sample size.

We filtered concatenated sites for each population and feature within each species so that it contained the same positions in the two populations being compared, and computed diversity for each population and feature in nonoverlapping 1 kb windows. We calculated differences in diversity over these nonoverlapping 1 kb windows to be able to compute statistical tests without the bias introduced by different features having different average length. Relative differences in diversity ( $\delta$ ) were computed as:

$$\frac{B \text{ diversity} - NB \text{ diversity}}{B \text{ diversity} + NB \text{ diversity}}$$

In comparison with a direct diversity ratio (B diversity/NB diversity), this expression reduces the occurrence of the denominator being zero, shows a reduced variance, is bound to values between –1 and 1, and shows an approximately symmetrical distribution around the mean within the bounds. A positive  $\delta$  means that the B population shows more diversity than the NB population and vice versa. Window diversity, computed as the arithmetic mean of site diversities estimated by ANGSD taking into account genotype uncertainty, is never zero. As a consequence, the distribution of the genetic diversity in windows shows a clearly bimodal distribution, with the lowest mode (usually  $10^{-5}$  to  $10^{-7}$ ) corresponding to windows with probably no diversity (Figure S1). As the diversity estimated for these windows with no diversity might still differ in the two populations, rendering irrelevant and probably biased  $\delta$  values, we transformed  $\delta_{\theta_W}$  and  $\delta_{\pi}$  to zero for windows where the diversity in both populations was below the empirically determined diversity threshold separating the two modes (Figure S1). We plotted  $\delta$ , and its dispersion calculated by bootstrapping over windows as implemented in Hmisc package in R (Harrell Jr, 2019) with 1,000 iterations. Also, in order to directly compare the magnitude of the changes in the density of variable sites with changes in their allele frequencies, we plotted  $\delta_{\theta_W}$  against  $\delta_{\pi}$ . Due to non-normality of the  $\delta$  statistic revealed by Shapiro–Wilk normality test, we compared the  $\delta$  distribution among different features using a Wilcoxon signed-rank test for paired samples and applied a strict Bonferroni correction for multiple tests to assess significance. All tests were run in R as implemented in stats package.

## 2.7 | Genomic variables

We first annotated units for the following genomic variables: average recombination rate, divergence, as a proxy of mutation rate,

GC content, and two different scores related to selection strength (Residual Variation Intolerance Score (RVIS), and functional site percentage). Secondly, nonoverlapping 1 kb windows used for calculating diversity differences were also annotated for these genomic variables. We did this by averaging the values annotated for the units included in each 1 kb window weighting by the number of overlapping sites.

To calculate recombination rate, we used the latest domestic cat linkage map (Li et al., 2016). Original coordinates, which are referred to *Felis\_catus\_8.0* assembly, were first translated to *Felis\_catus\_5.0* and then to LYPA1.0 coordinates using lynx to cat synteny (Abascal et al., 2016). We assigned to each unit the average recombination rate of the 2 MB window where it was included. Divergence was computed as the number of observed substitutions between the Iberian or the Eurasian lynx and the bobcat (*L. rufus*), divided by the number of informative sites in each unit, as reported by Lucena-Perez et al. (2020). To estimate GC content per unit, we used the Iberian lynx LYPA1.0 genome assembly, or a consensus Eurasian lynx genome (Abascal et al., 2016). Using BEDTools nuc (Quinlan & Hall, 2010), we parsed a bed file with units coordinates and estimated the GC percentage of each unit. The level of tolerance to variation of each gene was estimated by RVIS, a gene-based score computed for human sequence data that reaches more negative values for intolerant genes (Petrovski et al., 2013). Lynx genes were annotated using a lynx-to-human orthologs database (Abascal et al., 2016). For functional sites percentage, we annotated the genome based on 2 Mb windows. For each 2 Mb window, we summed up the number of sites belonging to 3' UTR, CDS, 5' UTR, lncRNAs, ncRNAs, promoters, or UCNEs, divided it by 2 Mb and multiplied it by 100. Windows were then intersected with units using bedtools intersect (Quinlan & Hall, 2010).

## 2.8 | Testing relationships between genetic diversity and genomic variables and differences in genomic variables between windows with contrasting behavior

To test how the relationships between different variables—namely chromosome size, gene content, recombination, divergence, GC content, RVIS, or functional sites percentage—and genomic diversity variables in the populations considered, we did linear regressions between those variables and genetic diversity at the chromosome level for each population. Chromosome length, gene content, and recombination rates were obtained from domestic cat reference genome, downloaded from [https://www.ncbi.nlm.nih.gov/genome/?term=txid9685\[orgn\]](https://www.ncbi.nlm.nih.gov/genome/?term=txid9685[orgn]).

Then, for each of the three pairwise NB-B comparisons we analyzed the linear regression of genetic diversity in the NB population with genetic diversity in the corresponding B population, separately for centromeric, interstitial, and telomeric regions. We report both the predictor (i.e., the slope) and the proportion of variance explained ( $R^2$ ) as heat maps.

We then focused on genic—CDS and introns—and intergenic regions to assess whether there are significant differences between windows that showed higher diversity in B than NB population ( $\delta_{\theta_{W}} > 0.1$ ) and the rest of the windows in the average values of the following genomic variables: NB diversity ( $\pi$ ,  $\theta_{W}$ , and  $S$ ) as a proxy of the ancestral diversity, recombination rate, RVIS and functional sites percentage as a proxy of selection, divergence as proxy of mutation rate, and GC content. After rejecting normality of these variables of interest, we used a Wilcoxon signed-rank test for unpaired samples. The significance threshold was adjusted by applying a strict Bonferroni correction. To get an estimation of the effect of these variables, we calculated effect size ( $r$ ), as implemented in (Mangiafico, 2020). Next, for each NB-B pairwise comparison we focused on the subset of windows with no  $\theta_{W}$  diversity in the NB population, and compared windows with no  $\theta_{W}$  diversity in the B population either, that is, no diversity in NB and no diversity in B windows ( $ND_{NB}$ - $ND_B$  windows), against those with some  $\theta_{W}$  diversity in the B population ( $ND_{NB}$ - $D_B$  windows) (Figure S1). Again, we used a Wilcoxon signed-rank test for unpaired samples and Bonferroni correction to assess significant differences. All statistics were calculated using R (R Core Team, 2019), and results were graphed using ggplot2 R package (Wickham, 2009).

## 3 | RESULTS

### 3.1 | Bottlenecks reduce average genomic diversity

As a first approach to genomic diversity patterns, we estimated genome-wide averages across populations and species. Their comparison between NB and B populations provides an overall indication of the relative intensity and duration of bottlenecks in the different populations. At the species level, the Eurasian lynx shows about twice the genome-wide diversity of the Iberian lynx (Eurasian lynx:  $\pi$  weighted mean (wm) =  $3.3 \times 10^{-4}$ ;  $\theta_{Wwm}$  =  $2.8 \times 10^{-4}$ ); Iberian lynx  $\pi_{wm}$  =  $1.9 \times 10^{-4}$ ;  $\theta_{Wwm}$  =  $1.4 \times 10^{-4}$ ) (Figure 1). Both species are among the mammals with the lowest overall genomic diversity reported so far (Abascal et al., 2016; Lucena-Perez et al., 2020). Within Eurasian lynx, whole-genome diversity statistics are  $\pi_{wm}$  =  $3.4 \times 10^{-4}$  and  $\theta_{Wwm}$  =  $3.3 \times 10^{-4}$  in the NB Kirov population, and these are reduced to 80% and 74% in the B NE-Poland and Norway populations, respectively, for both  $\pi$  and  $\theta_{W}$  (Figure 1, Table S2). In Iberian lynx, both  $\pi$  and  $\theta_{W}$  diversity in the Doñana B population are ~52% of that found in the NB Andújar population ( $\pi_{wm}$  =  $2.5 \times 10^{-4}$ ;  $\theta_{Wwm}$  =  $1.8 \times 10^{-4}$ ) (Figure 1, Table S2). Based on the relative diversity of each B population with respect to their NB population counterpart, bottlenecks rank from less to more extreme (longer duration and/or smaller  $N_e$ ) in the order: NE-Poland, Norway, and Doñana. All populations show a negative  $S$  (an indication of a generalized scarcity of rare alleles) (Becher et al., 2020), particularly the Iberian lynx ones (average across populations of  $S$  =  $-0.41$  vs.  $S$  =  $-0.28$  in Eurasian lynx) (Table S2).

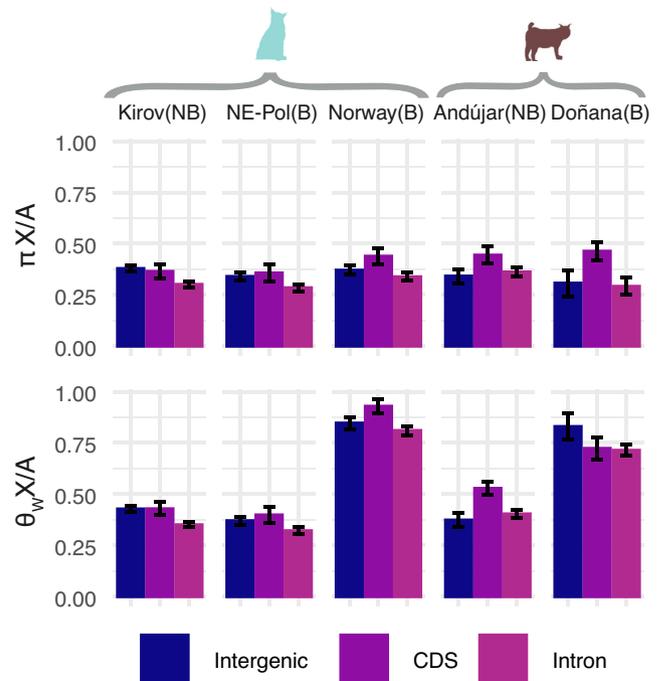
### 3.2 | Bottlenecks lead to a larger accumulation of low-frequency variants in the X chromosome than in the autosomes

We then compared genetic diversity across chromosomes in NB and B populations, with a particular focus in the comparison of autosomes to the X chromosome. Chromosomes differ widely in diversity within populations (Figure S2; Table S4). Largest diversity differences occur between any of the autosomes (A) and the X chromosome, with the latter showing lower diversity than A consistently across populations (Figure S2; Tables S3 and S4). Global X/A ratios range across populations from 0.31 to 0.37 for  $\pi_X/\pi_A$ , and from 0.36 to 0.84 for  $\theta_{WX}/\theta_{WA}$  (Table S3). While  $\pi_X/\pi_A$  is similar to  $\theta_{WX}/\theta_{WA}$  in NB populations and NE-Poland,  $\theta_{WX}/\theta_{WA}$  is notably larger than  $\pi_X/\pi_A$  in the B Doñana and Norway populations (Figure 2). This pattern is consistent across different genomic features, namely, CDS, intergenic, and introns (Figure 2). Also, the diversity reduction in B relative to NB populations is globally similar for A and X ( $\pi_{B,X}/\pi_{NB,X}$  is similar to  $\pi_{B,A}/\pi_{NB,A}$ ); however, the larger (above 1)  $\theta_{WB,X}/\theta_{WNB,X}$  than  $\theta_{WB,A}/\theta_{WNB,A}$  for Norway and Doñana, when compared to Kirov and Andújar, respectively, indicates a larger density of variants in the X chromosome in B relative to NB populations. Accordingly,  $S$  values are negative in the NB Andújar and Kirov, and also in the B NE-Poland populations, but positive in the B Doñana and Norway populations (Table S2).

### 3.3 | Bottlenecks blur expected differences in diversity among and along autosomes

In order to better understand the consequences of bottlenecks in the patterns of genetic variation, we estimated the level of correlation of the genetic diversity in different autosomes to several known genomic correlates, and compared them in NB and B populations. Among the autosomes, average chromosomal genetic diversity is positively related to divergence, GC content, gene content, functional sites percentage, RVIS (a gene-based measure of the deficit of standing functional variation, used here as a proxy for the intensity of ongoing purifying selection), and recombination, and negatively related to chromosome size in all populations of both species (Figure S3, Table S4). However, these correlations show lower predictors and smaller fractions of explained variance ( $R^2$ ) in the B populations, especially so for  $\theta_W$  (Figure S3).

To assess the contribution of recombination to changes in genetic diversity following a bottleneck, we compared genetic diversity in NB and B populations in different chromosomal regions that differ widely in recombination rates. Along autosomes, subtelomeric regions show the highest diversity, followed by the interstitial regions, and lastly the pericentromeric regions (Figure S4) in all populations. Subtelomeric regions show 1.52–1.66, and 1.29–1.62 times the interstitial  $\pi$  and  $\theta_W$  diversity, respectively, with lower ratios of  $\theta_W$  occurring in the B Norway and Doñana



**FIGURE 2** X/A ratio for  $\pi$  and  $\theta_W$  diversity for the five populations analyzed, and three different genomic features (CDS, introns, and intergenic)

populations, indicating that  $\theta_W$  diversity in B pops is still higher in subtelomeric regions than in other chromosomal regions, but not as high as in NB populations (Table S5). Pericentromeric regions show 0.83–0.94 and 0.87–0.94 of interstitial  $\pi$  and  $\theta_W$  diversity, respectively, both ratios tending to be higher in B populations than in their NB counterparts. Skewness of SFS is moderate and similar across chromosomal regions in NB populations, but the more extremely bottlenecked populations show more negative  $S$  values in subtelomeric regions, indicating a relatively larger overall scarcity of low-frequency alleles.

Diversity across chromosomal regions ranks according to recombination rates (subtelomeric = 3.7 cM/Mb, interstitial = 3.5 M/Mb, and pericentromeric = 2.3 cM/Mb), and divergence (subtelomeric = 7.9e-03, interstitial = 5.5e-03, and pericentromeric = 5.2e-03), and inversely to functional sites percentage (subtelomeric = 3.3, interstitial = 6.0, and pericentromeric = 6.3). Also, subtelomeric regions show higher GC content (subtelomeric = 0.58, interstitial = 0.47, and pericentromeric = 0.48), and higher RVIS (subtelomeric = 0.2, interstitial = -0.31, and pericentromeric = -0.28) (Table S5).

Finally, we assessed the level of correlation between autosomal genetic diversity in NB with that of their respective B populations. Autosomal diversities in NB and B populations were generally correlated, but the more intense the bottleneck, the weaker these correlations became both in terms of predictor and  $R^2$  (Figure S5), reflecting the random effects of genetic drift. In addition, these correlations were weaker in subtelomeric regions compared to pericentromeric or interstitial regions (Figure S5).

### 3.4 | Smaller relative diversity reductions in B populations in selectively constrained features

To explore the influence of natural selection on the changes in genetic diversity occurring in B populations, we compared genetic diversity in NB and B populations across genomic features that differ in biological function and their expected level of purifying selection. For all the features considered, namely, intergenic regions, coding gene promoters, 5' UTR, CDS, introns, 3' UTR, long-non-coding RNA promoters, long-non-coding RNA exons, long-non-coding RNA introns, non-coding RNA (mostly miRNAs, snRNAs, and snoRNAs), and ultra-conserved-non-coding-elements (UCNE), patterns of genetic diversity are negatively correlated with their anticipated selective pressure (Figure 3, Table S6). Taking intergenic as a reference, the most selectively constrained features present the lowest diversity, in particular UCNE (0.19–0.20 and 0.21–0.56 of intergenic  $\pi$  and  $\theta_W$ , respectively), followed by CDS ( $\pi = 0.58$ –0.6;  $\theta_W = 0.59$ –0.79; both relative to intergenic). With regard to the NB-B comparisons, regulatory regions such as gene promoters, noncoding RNAs, but also 5' UTR and 3' UTR, and even introns, show smaller reductions of diversity in B populations with respect to NB than intergenic regions.  $S$  is negative for all populations and features, except for UCNE in Norway and Doñana, which shows a high positive value ( $S = 0.53$  and  $S = 0.50$ , respectively), and for CDS in Norway ( $S = 0.04$ ), suggesting an accumulation of low-frequency variants in these constrained features in the most bottlenecked populations (Table S6). Notably, despite the lower diversity in B than in NB populations in most features, diversity in UCNE in the B Doñana and Norway populations is actually higher than in their reference NB population (Andújar and Kirov, respectively) when measured as Watterson's theta ( $\theta_W$ ), but not as nucleotide diversity ( $\pi$ ) (Figure 3).

Sliding window analysis of relative diversity differences between NB and B populations ( $\delta$ ) shows the lower overall diversity in the B population as a negative global mean of  $\delta$  for most features, both for  $\pi$  ( $\delta_\pi$ ) and  $\theta_W$  ( $\delta_{\theta_W}$ ) (Figure 4). For intergenic regions, Doñana-Andújar shows the more extreme diversity difference, followed by Norway-Kirov, and finally NE-Poland-Kirov. Regarding other features, and consistently among comparisons,  $\delta_{\theta_W}$  and  $\delta_\pi$  for UCNE are less negative, and  $\delta_{\theta_W}$  becomes even positive for the Doñana-Andújar and Norway-Kirov comparisons (Figure 4). A similar trend of less negative  $\delta$  values is also noticeable for CDS, and more subtly for 3' UTR, and 5' UTR, indicating again a smaller diversity loss for selectively constrained genomic features (relative to putatively neutral ones) in bottlenecked populations (Figure 4).

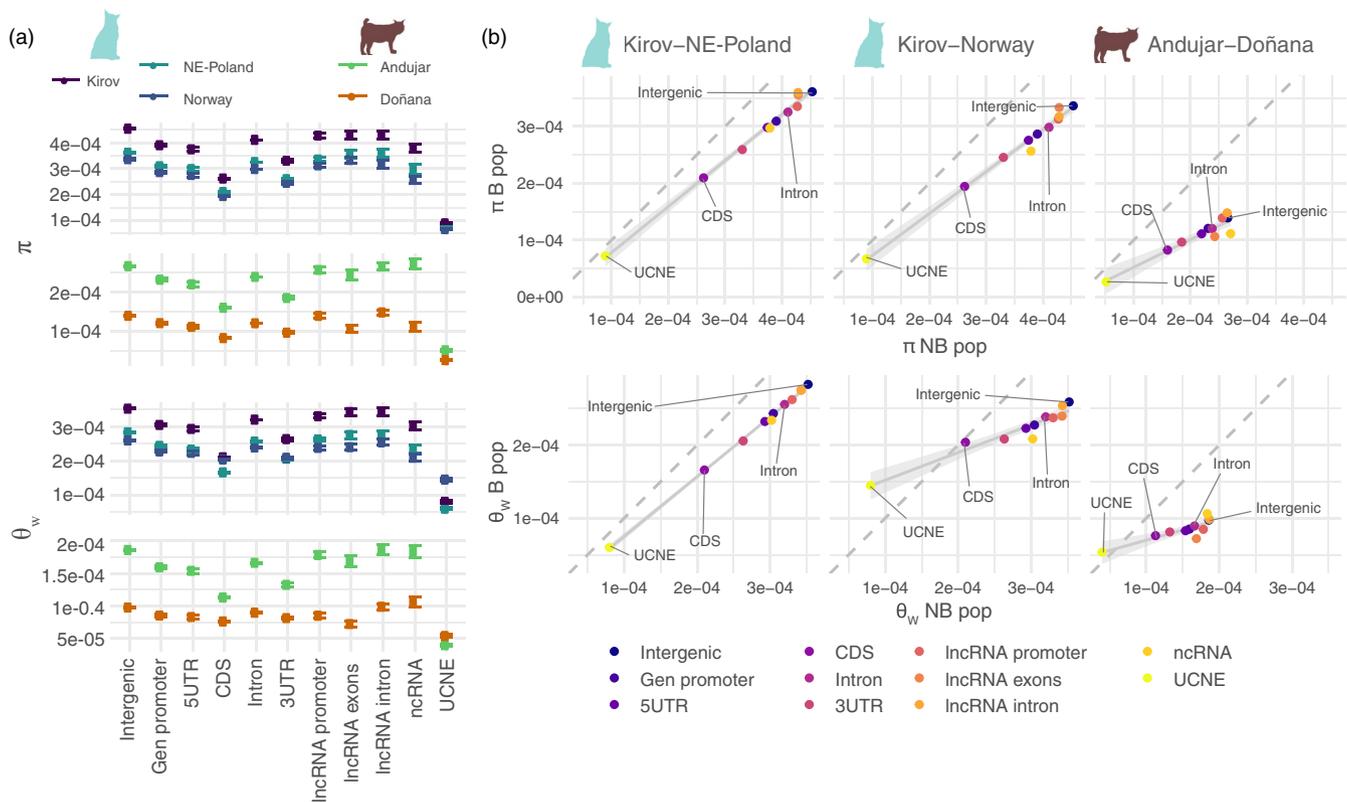
$\delta_{\theta_W}$  and  $\delta_\pi$  are generally correlated, but  $\delta_{\theta_W}$  tends to be less negative than  $\delta_\pi$  for most features and in the most bottlenecked populations (Doñana and Norway). This pattern is particularly notable for UCNE followed by CDS, 3' UTR and 5' UTR. Although with lower  $\delta$  than intergenic, ncRNA  $\delta_{\theta_W}$  tends to be larger than  $\delta_\pi$  in the Norway-Kirov comparison, and to a lesser extent in the Doñana-Andújar comparison (Figure S6).

### 3.5 | Gains in diversity in B populations occur in regions of higher mutational input and in genes more tolerant to variants

In order to gain further insights into the processes driving diversity gains in B populations, we compared average values of different genomic variables between windows of the same feature that show net  $\theta_W$  diversity gains with windows that do not. Windows with  $\delta_{\theta_W} > 0.1$ , that is, with larger  $\theta_W$  in B than in NB populations, show significantly lower diversity in NB populations ( $\pi$ ,  $\theta_W$ ) and larger  $S$ , that is, SFS more skewed toward rare alleles, than the rest of windows ( $p$ -value  $< 3e-04$ ; strict Bonferroni correction; Figure S7, Table S8), the effect sizes of these variables measured as a correlation coefficient ( $r$ ) being larger for Norway-Kirov (range: |0.15|–|0.22|), followed by Doñana-Andújar (range: |0.09|–|0.14|), and finally NE-Poland-Kirov comparison (|0.01|–|0.04|) (Table S9). They also show significant differences in other genomic variables like divergence, recombination, or selective constraints, which often are opposite to that expected from their effect in diversity (Figure S7, Table S8). For instance, windows with  $\delta_{\theta_W} > 0.1$  (i. e. which gain diversity after the bottleneck) tend to show lower diversity in the corresponding NB population, but are in regions with higher divergence, GC content, and recombination, which are generally associated to high diversity. Consequently, the magnitude and significance of these contrasts become larger when we limit the set of windows being considered to those showing no diversity in the NB population, that is, when we compare windows that lack  $\theta_W$  diversity in both NB and B populations ( $ND_{NB}$ - $ND_B$  windows) to those with no diversity in NB but with some diversity in the B population ( $ND_{NB}$ - $D_B$  windows) (Figure 5, Tables S10 and S11). Therefore,  $ND_{NB}$ - $D_B$  windows tend to be in regions with higher divergence and GC content than  $ND_{NB}$ - $ND_B$  windows. Also, recombination rate is significantly higher in  $ND_{NB}$ - $D_B$  windows, for nonselectively constrained features in the Norway-Kirov (introns), and Doñana-Andújar comparisons (intergenic and introns). Regarding selection, coding, but also intronic,  $ND_{NB}$ - $D_B$  windows tend to be in genes with higher RVIS, that is, more tolerant to change, in moderate NB-B pairwise comparisons, but not so in the most extreme Doñana-Andújar comparison. More generally, differences in genomic variables between  $ND_{NB}$ - $ND_B$  and  $ND_{NB}$ - $D_B$  windows are smaller overall when the bottleneck is more intense (Norway and Doñana) and the features are selectively constrained (CDS).

## 4 | DISCUSSION

Here, we present one of the most comprehensive empirical studies of the genomic consequences of bottlenecks in natural populations. By directly comparing genomic diversity in Eurasian and Iberian lynx populations with contrasting recent demography, we assessed how the impact of recent bottlenecks on population genetic diversity varies across chromosomes, chromosomal regions, and genomic



**FIGURE 3** (a) Average  $\theta_w$  and  $\pi$  diversity for different features in nonbottlenecked (NB: Kirov and Andújar) and bottlenecked (B: Norway, NE-Poland, Doñana) populations of Eurasian lynx (upper subpanel) and Iberian lynx (lower subpanel). Error bars represent the standard deviation (SD) obtained from bootstrapping. (b) Relationship between diversity ( $\theta_w$  or  $\pi$ ) of the NB (X axis) vs. B populations (Y axis) for different features. The dashed line represents the diagonal, where the diversity of NB would be equal to B. Diversity is generally lower in B populations than NB populations across features, but the difference is smaller for selected features (e. g. CDS) and is even reverted in UCNE in two out of the three comparisons

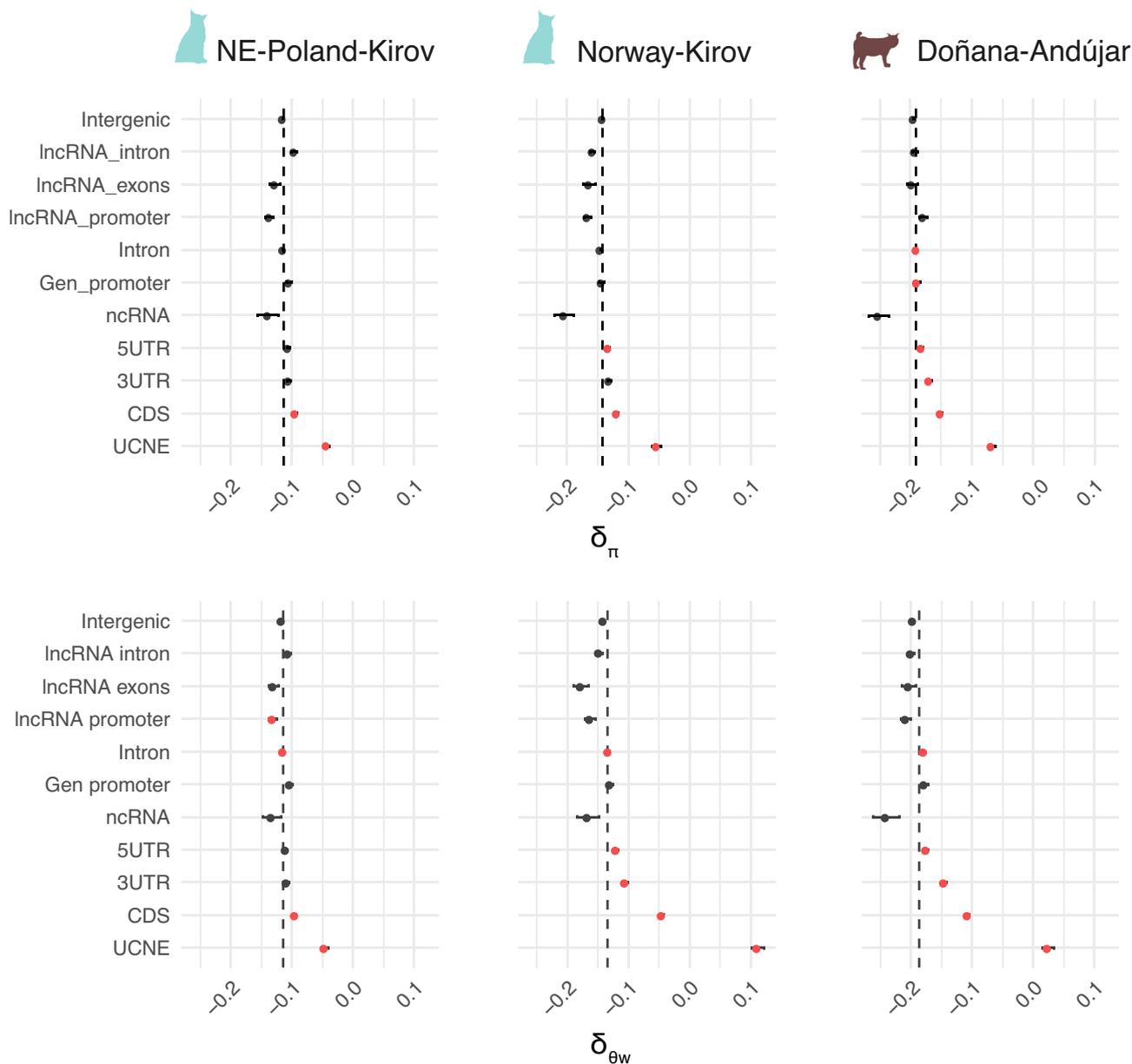
features. Also, we explored how different genomic variables, such as mutation and selection, as well as the initial SFS, might be affecting the different outcomes of genetic drift and, in consequence, the changes in diversity following a demographic decline. Our study provides empirical evidences of the conspicuous action of drift across the genome in replicated natural bottleneck scenarios. Comparing three bottlenecks of different intensity allowed us to assess the consistency of the results across replicates, and to account for differences regarding drift intensity.

Recent bottlenecks are expected to globally reduce the genomic diversity in a magnitude determined by the intensity and duration of the bottleneck, and our estimates grossly rank according to records and demographic reconstructions from genetic and genomic data (Abascal et al., 2016; Casas-Marce et al., 2017; Lucena-Perez et al., 2020). Skewness of the SFS is negative in all populations, indicating a general scarcity of low-frequency variants, which is however more pronounced in the Iberian lynx populations, consistent with the expectation for its smaller effective size at mutation-drift equilibrium.

One major finding of our study is that the magnitude of diversity reductions occurring in bottlenecked populations varies extensively across chromosomes, chromosomal regions, and genomic features. One first consequence of this is that correlations between the genetic diversities in NB and B populations across chromosomes and

chromosomal regions weaken as bottleneck intensity increases, especially in subtelomeric regions. Also, the correlations of genetic diversity with genomic variables like recombination, divergence, or gene content, which are expected at migration-mutation-selection-drift equilibrium and still observed in the NB populations, also weaken or disappear in the B populations. Both observations indicate that, in B populations, the reduction in mean diversity is accompanied by an increased variance across chromosomes and chromosomal regions, making genetic diversity patterns to progressively depart from those expected at equilibrium. Whereas the stochastic nature of genetic drift is a major contributor to the increased variance in genetic diversity in bottlenecked populations, our results also suggest an important contribution of its interaction with mutation and selection, as discussed below.

Regarding the X chromosome and autosomes comparison, results suggest a larger accumulation of low-frequency alleles in X relative to A (X/A ratios) in B populations. The X chromosome differs from the A in their response to a bottleneck because: (a) its effective size is  $\frac{3}{4}$  of A, which predicts a ratio of diversity X/A of 0.75 at mutation-drift equilibrium; (b) it has a lower recombination rate; (c) it has a lower mutation rate; and (d) recessive deleterious variation is continuously exposed to selection in hemizygous individuals (i.e., males in mammals) (Arbiza et al., 2014). Besides, under population

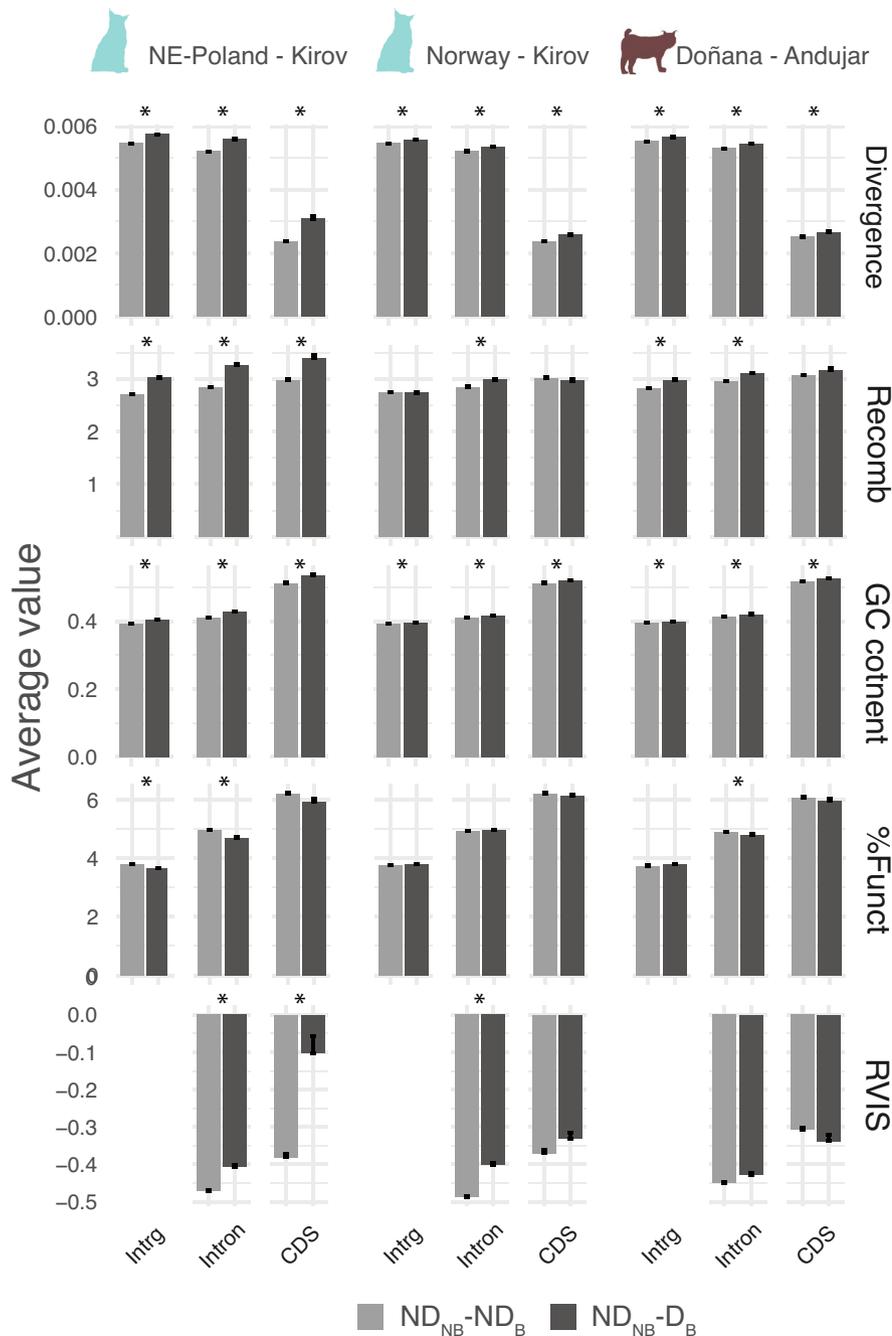


**FIGURE 4**  $\delta_{\theta_w}$  and  $\delta_{\pi}$  across features for the three comparisons of B vs. NB populations. Positive values indicate higher diversity in the B population. The dashed lines represent the average  $\delta$  value considering all features. Errors bars are the SD obtained from bootstrapping the data. The colored circle, triangles, and squares represent significant  $p$ -values when compared to intergenic after Bonferroni correction (Table S7)

size change scenarios, the dynamics of the two types of chromosomes might differ, with X diversity undergoing faster changes than autosomal diversity (Pool & Nielsen, 2007). It is unclear which of these distinctive characteristics is responsible for the observed marked accumulation of low-frequency variants in X in B populations. Although its lower effective size will translate to a larger number of selected sites behaving as neutral, the fact that the pattern is observed also at neutral regions suggests processes other than relaxed purifying selection are involved.

An accumulation of low-frequency alleles is also observed in some selectively constrained features, such as CDS and UCNE in the autosomes. Features that are subject to unequal selective pressures behave very differently in populations that underwent different

degrees of population size decline. Whereas some features, like intergenic regions, ncRNA and promoters show a similar relative reduction of  $\theta_w$  and  $\pi$  diversity in B populations, others such as UCNE and CDS and, to a lesser extent, 5' and 3' UTR, show a relatively smaller reduction in  $\theta_w$  than in  $\pi$ , which in the extreme case of UCNE results in higher  $\theta_w$  (suggestive of an excess of low-frequency variants) in the B than in the corresponding NB populations. Such increase in  $\theta_w$  in selectively constrained features is, in principle, unexpected from purely random processes, since random genetic drift in bottlenecked populations would result in the preferential loss of alleles in the low-frequency side of the spectrum, where most selectively constrained alleles will be (Nei, 1975). The pattern observed for the X chromosome and some selectively constrained features



**FIGURE 5** Comparison of average values of different genomic variables in windows with no diversity in both NB and B populations ( $ND_{NB}-ND_B$  windows) vs. windows with no diversity in NB but some diversity in B populations ( $ND_{NB}-D_B$  windows). The asterisk represents significant differences between both sets of windows.  $ND_{NB}-D_B$  windows tend to be in regions with higher recombination rate, divergence, and content. They also tend to be associated with genes more tolerant to changes in the moderate, but not in the most extreme bottlenecks

could thus be the consequence of (a) methodological limitations related to limited sampling, (b) various evolutionary processes other than genetic drift, like mutation and selection, acting by themselves or through their interaction with drift.

Regarding methodological limitations, it must be noted that the estimates of genetic diversity, and especially  $\theta_W$  diversity, are conditioned by both the sample sizes and the SFS. Given our sample sizes of 16 and 24 genomes, alleles segregating at frequencies below the detection threshold will be underrepresented in our samples. One consequence is that their likely loss in bottlenecked populations will remain largely unnoticed, whereas a few haplotypes increasing in frequency by drift will result in an increase in the observed  $\theta_W$  diversity. This is especially true for those regions and features with the highest skew in SFS, that

is, the most selectively constrained features like UCNE and CDS. On the other hand, the possibility of observing an increase in  $\theta_W$  diversity following a bottleneck will also depend on the SFS and thus on starting  $\pi$  diversity. Indeed, we found that windows with higher diversity tend to lose diversity whereas those with low diversity are more likely to gain diversity after the bottleneck. Interestingly, windows that gain diversity also tend to show higher divergence and GC content, which could be directly related to a higher mutational input generating a larger number of alleles segregating at low frequencies in the prebottleneck population and favoring the accumulation of de novo mutations during the bottleneck. Given the relatively short separation of the populations (i.e., likely less than 60 generations in the case of the Eurasian lynx (Lucena-Perez et al., 2020), and around 40 generations in Iberian

lynx (Casas-Marce et al., 2017), assuming a generation time of 5 years (Lucena-Perez et al., 2018), and their small effective size (in the range of tens), it is unlikely that a large fraction of the observed accumulation of variants at low frequencies in the B populations is due to *de novo* mutations since the start of the bottleneck. Thus, most of these alleles increasing in frequency during and after the bottleneck would be pre-existing variants that remain undetected before the bottleneck and which increase in frequency and are thus recorded and result in higher observed  $\theta_W$  after the bottleneck. Recombination could also contribute to this pattern through its direct mutagenic effect (Duret & Arndt, 2008; Duret & Galtier, 2009; Halldorsson et al., 2019; Pratto et al., 2014; Smith et al., 2018; Terekhanova et al., 2017; Williams et al., 2015).

The accumulation of additive deleterious variants due to the relaxation of purifying selection could also be contributing to the observed increase in  $\theta_W$  in selectively constrained regions (Balick et al., 2015; Kirkpatrick & Jarne, 2000). Empirical evidence for this has been found in domesticated species (Makino et al., 2018; Marsden et al., 2016), while evidence in humans remains controversial (Do et al., 2015; Henn et al., 2016; Lohmueller, 2014; Simons & Sella, 2016; Simons et al., 2014). The reduced efficacy of purifying selection in bottlenecked populations would affect those variants with  $1/2N_{e,NB} < s < 1/2N_{e,B}$ , that is, alleles with  $s$  between these values would be efficiently selected in the NB but will drift in frequency as neutral variants in the B population; for example, in the case of Andújar ( $N_e = 20$ ) and Doñana ( $N_e = 10$ ) the corresponding range would be  $0.025 < s < 0.05$ , which are rather high values. Moreover, it must be noted that while Doñana has remained at such low size for many decades, Andújar was much larger a few generations ago, so the range of mutations that behave as neutral in Doñana and not in Andújar may be substantially wider (with a lower inferior  $s$  limit), so that the actual number of variants involved may be relatively large (Casas-Marce et al., 2017). The contribution of the reduced efficacy of purifying selection is further supported by the observation that in shallower demographic declines the smaller relative loss of diversity occurs mainly in genes with higher RVIS (i.e., more tolerant to changes). However, under more extreme bottlenecks the relationship with RVIS weakens and becomes nonsignificant, suggesting that the relaxation of purifying selection results in the accumulation of potentially deleterious variants even in highly intolerant genes, and alerting of possible fitness reductions. Finally, empirical and theoretical studies show that selectively constrained diversity approaches a new equilibrium faster than neutral diversity (Brandvain & Wright, 2016; Gordo & Dionisio, 2005; Pennings et al., 2014; Song & Steinrücken, 2012). We can thus expect that the diversity in CDS and UCNE that is due to additive mutations reaches its new increased equilibrium level faster than intergenic or intronic diversity reaches their new reduced equilibrium level.

#### 4.1 | Possible fitness consequences for the B populations

Our results suggest that overall diversity reductions in bottlenecked populations affect regulatory elements, such as promoters, or

ncRNA, in a similar fashion as intergenic regions. This loss of diversity is likely removing pre-existing variation in the regulation of gene expression. The loss of variation in regulatory elements of transcription likely compromises these populations' potential for acclimation to the new environments through phenotypic plasticity, reducing their viability under environmental fluctuations. Furthermore, there is increasing evidence that regulatory variation can be an important source of adaptations, possibly a more important source than changes in protein sequence for rapid evolution, so its loss can have a major impact on adaptive potential (Harrisson et al., 2014). These observations call for greater attention to regulatory variation and patterns of gene expression in conservation genomics.

For other selectively constrained features, our results suggest that overall genetic diversity reductions in bottlenecked populations are accompanied by the accumulation of possible deleterious alleles in functional regions such as CDS, including genes highly intolerant to changes (in the most extreme bottlenecks), and extending beyond coding regions to include UCNE, and to a lesser extent 5' UTR and 3' UTR. The latter two features are essential for efficient transcription and for the post-transcriptional regulation of gene expression, and variants in both regions have been linked to several diseases (Hindorff et al., 2009). Regarding UCNE, an accumulation of deleterious variation in these regions is likely to affect their prominent function as regulatory *cis* elements, especially during development (Marcovitz et al., 2016; Polychronopoulos et al., 2017), leading to a reduction in fitness.

Our results call for extending the analyses of genetic variation in endangered species to include functional variation beyond coding sequences. Including regulatory elements will broaden our understanding of the effects of drift on the genome and allow for a more comprehensive assessment of the possible short-term fitness reductions and the long-term loss of adaptive potential in endangered species. As the genome-wide assessment of genetic variation in non-model species is becoming increasingly feasible, and more readily facilitated by the increasing availability of reference genomes, the widespread implementation of genetic monitoring in wild species demands simple and cost-effective methods, and informative and easy to implement indicator variables. In this regard, the accumulation of variants in highly constrained UCNEs may provide an easily detectable and reliable signal of recent demographic declines.

#### ACKNOWLEDGEMENTS

This project was funded by the National Science Center, Poland, through project no 2014/15/B/ NZ8/00212, project "Biodiversity of East-European and Siberian large mammals on the level of genetic variation of populations" (BIOGEAST), contract No. PIRSEGA-2009-247652, and by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovación (MICINN) through projects CGL2013-47755-P and CGL2017-84641-P, and by Russian Science Foundation (project No. 18-14-00093). MLP and DKR were supported by PhD contracts from the Programa Internacional de Becas "La Caixa-Severo Ochoa" from "La Caixa" Foundation (ID 100010434), under agreement LCF/BQ/

SO14/52250035 and LCF/BQ/SO15/52260006. Also, DKR was supported by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovación by the PGC2018-095810-B-I00 project to Aurora García-Dorado. Logistical support was provided by Laboratorio de Ecología Molecular (LEM-EBD) (certified to ISO9001:2015 and ISO14001:2015 quality and environmental management systems), and Laboratorio de SIG y Teledetección (LAST-EBD) in Estación Biológica de Doñana. Calculations were carried out in the Genomics servers of the Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD). Authors also acknowledge the computing and storage resources provided by Fundación Pública Galega Centro Tecnológico de Supercomputación de Galicia (CESGA). EBD-CSIC received support from the Spanish Ministry of Economy and Competitiveness under the "Centro de Excelencia Severo Ochoa 2013-2017" program, grants SEV-2012-0262. We also acknowledged Jerez Zoo (Spain) for providing samples. We are grateful to Ingrid Reinkind for kindly providing the Eurasian lynx samples from Norway. We also want to thank Enrico Bazzicalupo, Pedro Jordano, Rubén Bernardo-Madrid, Isabel García-Barón, and Elena del Águila-Pérez for insightful discussion about the analysis.

#### CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

#### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in European Nucleotide Archive (ENA) at <https://www.ebi.ac.uk/ena>, accession PRJEB44874.

#### ORCID

José A. Godoy  <https://orcid.org/0000-0001-7502-9471>

#### REFERENCES

- Abascal, F., Corvelo, A., Cruz, F., Villanueva-Cañas, J. L., Vlasova, A., Marcet-Houben, M., Martínez-Cruz, B., Cheng, J. Y., Prieto, P., Quesada, V., Quilez, J., Li, G., García, F., Rubio-Camarillo, M., Frias, L., Ribeca, P., Capella-Gutiérrez, S., Rodríguez, J. M., Câmara, F., ... Godoy, J. A. (2016). Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biology*, 17(1), 251. <https://doi.org/10.1186/s13059-016-1090-1>
- Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2013). *Conservation and the genetics of populations*, 2nd ed. Wiley-Blackwell.
- Arbiza, L., Gottipati, S., Siepel, A., & Keinan, A. (2014). Contrasting X-linked and autosomal diversity across 14 human populations. *The American Journal of Human Genetics*, 94(6), 827–844. <https://doi.org/10.1016/j.ajhg.2014.04.011>
- Balick, D. J., Do, R., Cassa, C. A., Reich, D., & Sunyaev, S. R. (2015). Dominance of deleterious alleles controls the response to a population bottleneck. *PLoS Genetics*, 11(8), 1–23. <https://doi.org/10.1371/journal.pgen.1005436>
- Ballenghien, M., Faivre, N., & Galtier, N. (2017). Patterns of cross-contamination in a multispecies population genomic project: Detection, quantification, impact, and solutions. *BMC Biology*, 15(1), 25. <https://doi.org/10.1186/s12915-017-0366-6>
- Bazzicalupo, E., Lucena-Perez, M., Kleinman-Ruiz, D., Pavlov, A., Trajçe, A., Hoxha, B., Sanaja, B., Gurielidze, Z., Kerdikoshvili, N., Mamuchadze, J., Yarovenko, Y. A., Akkiev, M. I., Ratkiewicz, M., Saveljev, A. P., Melovski, D., Gavashelishvili, A., Schmidt, K., & Godoy, J. A. (2021). History, demography and genetic status of Balkan and Caucasian *Lynx lynx* (Linnaeus, 1758) populations revealed by genome-wide variation. *Diversity and Distributions*, in press.
- Becher, H., Jackson, B. C., & Charlesworth, B. (2020). Patterns of genetic variability in genomic regions with low rates of recombination. *Current Biology*, 30(1), 94–100.e3. <https://doi.org/10.1016/j.cub.2019.10.047>
- Brandvain, Y., & Wright, S. I. (2016). The limits of natural selection in a nonequilibrium world. *Trends in Genetics*, 32(4), 201–210. <https://doi.org/10.1016/j.tig.2016.01.004>
- Breitenmoser, U., Breitenmoser-Würsten, C., Lanz, T., von Arx, M., Antonevich, A., Bao, W., & Avgan, B. (2015). *Lynx lynx* (errata version published in 2017). *The IUCN Red List of Threatened Species*.
- Canty, A., & Ripley, B. D. (2017). boot: Bootstrap R (S-Plus) functions. [R statistical package].
- Casas-Marce, M., Marmesat, E., Soriano, L., Martínez-Cruz, B., Lucena-Perez, M., Nocete, F., Rodríguez-Hidalgo, A., Canals, A., Nadal, J., Detry, C., Bernáldez-Sánchez, E., Fernández-Rodríguez, C., Pérez-Ripoll, M., Stiller, M., Hofreiter, M., Rodríguez, A., Revilla, E., Delibes, M., & Godoy, J. A. (2017). Spatiotemporal dynamics of genetic variation in the Iberian Lynx along its path to extinction reconstructed with ancient DNA. *Molecular Biology and Evolution*, 34(11), 2893–2907. <https://doi.org/10.1093/molbev/msx222>
- Casas-Marce, M., Soriano, L., López-Bao, J. V., & Godoy, J. A. (2013). Genetics at the verge of extinction: Insights from the Iberian lynx. *Molecular Ecology*, 22(22), 5503–5515. <https://doi.org/10.1111/mec.12498>
- Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134(4), 1289–1303. <http://dx.doi.org/10.1093/genetics/134.4.1289>
- Corcoran, P., Gossmann, T. I., Barton, H. J., Slate, J., & Zeng, K. (2017). Determinants of the efficacy of natural selection on coding and noncoding variability in two passerine species. *Genome Biology and Evolution*, 9(11), 2987–3007. <https://doi.org/10.1093/gbe/evx213>
- Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, 14(4), 262–274. <https://doi.org/10.1007/s10450-013-9579-3>
- Des Roches, S., Pendleton, L. H., Shapiro, B., & Palkovacs, E. P. (2021). Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution*, 5(5), 574–582. <https://doi.org/10.1038/s41559-021-01403-5>
- DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, 30(17), 4147–4154.
- Dimitrieva, S., & Bucher, P. (2013). UCNEbase—a database of ultraconserved non-coding elements and genomic regulatory blocks. *Nucleic Acids Research*, 41(D1), D101–D109. <https://doi.org/10.1093/nar/gks1092>
- Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., & Reich, D. (2015). No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nature Genetics*, 47(2), 126–131. <https://doi.org/10.1038/ng.3186>
- Duret, L., & Arndt, P. F. (2008). The impact of recombination on nucleotide substitutions in the human genome. *PLoS Genetics*, 4(5), e1000071. <https://doi.org/10.1371/journal.pgen.1000071>
- Duret, L., & Galtier, N. (2009). Biased gene conversion and the evolution of mammalian genomic landscapes. *Annual Review of Genomics and Human Genetics*, 10(1), 285–311. <https://doi.org/10.1146/annurev-genom-082908-150001>
- Dutoit, L., Burri, R., Nater, A., Mugal, C. F., & Ellegren, H. (2017). Genomic distribution and estimation of nucleotide diversity in natural

- populations: Perspectives from the collared flycatcher (*Ficedula albicollis*) genome. *Molecular Ecology Resources*, 17(4), 586–597. <https://doi.org/10.1111/1755-0998.12602>
- Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, 17(7), 422–433. <https://doi.org/10.1038/nrg.2016.58>
- Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, 78(2), 737–756. <https://doi.org/10.1093/genetics/78.2.737>
- Ferreras, P., Rodríguez, A., Palomares, F., & Delibes, M. (2010). Iberian lynx: the uncertain future of a critically endangered cat. In D. Macdonald & A. J. Loveridge (Eds.), *Biology and conservation of wild felids* (pp. 507–520). Oxford University Press.
- García-Dorado, A. (2007). Shortcut predictions for fitness properties at the mutation–selection–drift balance and for its buildup after size reduction under different management strategies. *Genetics*, 176(2), 983–997. <https://doi.org/10.1534/genetics.106.065730>
- García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: Inbreeding, purging, mutation, and standard selection. *Genetics*, 190(4), 1461–1476. <https://doi.org/10.1534/genetics.111.135541>
- García-Dorado, A., & Caballero, A. (2021). Neutral genetic diversity as a useful tool for conservation biology. *Conservation Genetics*, 22(4), 541–545. <https://doi.org/10.1007/s10592-021-01384-9>
- Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, 10(2), 305–318. <https://doi.org/10.1046/j.1365-294x.2001.01190.x>
- Gonzalez-Perez, A., Sabarinathan, R., & Lopez-Bigas, N. (2019). Local determinants of the mutational landscape of the human genome. *Cell*, 177(1), 101–114. <https://doi.org/10.1016/j.cell.2019.02.051>
- Gordo, I., & Dionisio, F. (2005). Nonequilibrium model for estimating parameters of deleterious mutations. *Physical Review E*, 71(3), <http://dx.doi.org/10.1103/physreve.71.031907>
- Grossen, C., Guillaume, F., Keller, L. F., & Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, 11(1), 1001. <https://doi.org/10.1038/s41467-020-14803-1>
- Halldorsson, B. V., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., Gunnarsson, B., Oddsson, A., Halldorsson, G. H., Zink, F., Gudjonsson, S. A., Frigge, M. L., Thorleifsson, G., Sigurdsson, A., Stacey, S. N., Sulem, P., Masson, G., Helgason, A., Gudbjartsson, D. F., ... Stefansson, K. (2019). Characterizing mutagenic effects of recombination through a sequence-level genetic map. *Science*, 363(6425), eaau1043. <https://doi.org/10.1126/science.aau1043>
- Harrell, F. E. Jr, Dupont, with contributions from C., & many others. (2019). Hmisc: Harrell Miscellaneous.
- Harrison, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, 7(9), 1008–1025. <https://doi.org/10.1111/eva.12149>
- Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution*, 31(12), 940–952. <https://doi.org/10.1016/j.tree.2016.09.005>
- Henn, B. M., Botigué, L. R., Peischl, S., Dupanloup, I., Lipatov, M., Maples, B. K., & Bustamante, C. D. (2016). Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 113(4), E440–E449. <https://doi.org/10.1073/pnas.1510805112>
- Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, 8(3), 269–294. <https://doi.org/10.1017/S001667230800949X>
- Hindorf, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*, 106(23), 9362–9367. <https://doi.org/10.1073/pnas.0903103106>
- Hoban, S., Bruford, M. W., Funk, W. C., Galbusera, P., Griffith, M. P., Grueber, C. E., & Vernesi, C. (2021). Global commitments to conserving and monitoring genetic diversity are now necessary and feasible. *BioScience*, 71(9), 964–976. <https://doi.org/10.1093/biosci/i/biab054>
- Hoban, S., Campbell, C. D., da Silva, J. M., Ekblom, R., Funk, W. C., Garner, B. A., Godoy, J. A., Kershaw, F., MacDonald, A. J., Mergeay, J., Minter, M., O'Brien, D., Vinas, I. P., Pearson, S. K., Pérez-Espona, S., Potter, K. M., Russo, I.-R., Segelbacher, G., Vernesi, C., & Hunter, M. E. (2021). Genetic diversity is considered important but interpreted narrowly in country reports to the Convention on Biological Diversity: Current actions and indicators are insufficient. *Biological Conservation*, 261, 109233. <https://doi.org/10.1016/j.biocon.2021.109233>
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62–82. <https://doi.org/10.1111/mec.15720>
- Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*, 47(6), 713–719. <https://doi.org/10.1093/genetics/47.6.713>
- Kirkpatrick, M., & Jarne, P. (2000). The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist*, 155(2), 154–167. <https://doi.org/10.1086/303312>
- Korneliusson, T. S., Moltke, I., Albrechtsen, A., & Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics*, 14(1), 289. <https://doi.org/10.1186/1471-2105-14-289>
- Laikre, L., Hoban, S., Bruford, M. W., Segelbacher, G., Allendorf, F. W., Gajardo, G., Rodríguez, A. G., Hedrick, P. W., Heuertz, M., Hohenlohe, P. A., Jaffé, R., Johannesson, K., Liggins, L., MacDonald, A. J., Orozco-Wengel, P., Reusch, T. B. H., Rodríguez-Correa, H., Russo, I.-R., Ryman, N., & Vernesi, C. (2020). Post-2020 goals overlook genetic diversity. *Science (New York, N.Y.)*, 367(6482), 1083–1085. <https://doi.org/10.1126/science.abb2748>
- Li, G., Hillier, L. D. W., Grahn, R. A., Zimin, A. V., David, V. A., Menotti-Raymond, M., Middleton, R., Hannah, S., Hendrickson, S., Makunin, A., O'Brien, S. J., Minx, P., Wilson, R. K., Lyons, L. A., Warren, W. C., & Murphy, W. J. (2016). A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3 Genes|Genomes|Genetics*, 6(6), 1607–1616. <https://doi.org/10.1534/g3.116.028746>
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv:1303.3997v1 [q-Bio.GN]*.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Lohmueller, K. E. (2014). The distribution of deleterious genetic variation in human populations. *Current Opinion in Genetics & Development*, 29, 139–146. <https://doi.org/10.1016/j.gde.2014.09.005>
- Lucena-Perez, M., Marmesat, E., Kleinman-Ruiz, D., Martínez-Cruz, B., Węcek, K., Saveljev, A. P., Seryodkin, I. V., Okhlopkov, I., Dvornikov, M. G., Ozolins, J., Galsandorj, N., Paunovic, M., Ratkiewicz, M., Schmidt, K., & Godoy, J. A. (2020). Genomic patterns in the widespread Eurasian lynx shaped by Late Quaternary climatic fluctuations and anthropogenic impacts. *Molecular Ecology*, 29(4), 812–828. <https://doi.org/10.1111/mec.15366>
- Lucena-Perez, M., Soriano, L., López-Bao, J. V., Marmesat, E., Fernández, L., Palomares, F., & Godoy, J. A. (2018). Reproductive biology and genealogy in the endangered Iberian lynx: Implications

- for conservation. *Mammalian Biology*, 89, 7–13. <https://doi.org/10.1016/j.mambio.2017.11.006>
- Makino, T., Rubin, C.-J., Carneiro, M., Axelsson, E., Andersson, L., & Webster, M. T. (2018). Elevated proportions of deleterious genetic variation in domestic animals and plants. *Genome Biology and Evolution*, 10(1), 276–290. <https://doi.org/10.1093/gbe/evy004>
- Mangiafico, S. (2020). rcompanion: Functions to support extension education program evaluation [R statistical package].
- Marcovitz, A., Jia, R., & Bejerano, G. (2016). 'reverse Genomics' predicts function of human conserved noncoding elements. *Molecular Biology and Evolution*, 33(5), 1358–1369. <https://doi.org/10.1093/molbev/msw001>
- Marsden, C. D., Del Vecchio, D. O., O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., & Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences of the United States of America*, 113(1), 152–157. <https://doi.org/10.1073/pnas.1512501113>
- Martin, S. H., Möst, M., Palmer, W. J., Salazar, C., McMillan, W. O., Jiggins, F. M., & Jiggins, C. D. (2016). Natural selection and genetic diversity in the butterfly heliconius melpomene. *Genetics*, 203(1), 525–541. <https://doi.org/10.1534/genetics.115.183285>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29(1), 1–10.
- Nowell, K., & Jackson, P. (1996). *Wild cats: A status survey and conservation action plan*. International Union for the Conservation of Nature and Natural Resources (IUCN).
- Palomares, F., Rodríguez, A., Revilla, E., López-Bao, J. V., & Calzada, J. (2011). Assessment of the conservation efforts to prevent extinction of the Iberian Lynx. *Conservation Biology*, 25(1), 4–8. <https://doi.org/10.1111/j.1523-1739.2010.01607.x>
- Payseur, B. A., & Nachman, M. W. (2002). Gene density and human nucleotide polymorphism. *Molecular Biology and Evolution*, 19(3), 336–340. <https://doi.org/10.1093/oxfordjournals.molbev.a004086>
- Pennings, P. S., Kryazhimskiy, S., & Wakeley, J. (2014). Loss and recovery of genetic diversity in adapting populations of HIV. *PLoS Genetics*, 10(1), e1004000. <https://doi.org/10.1371/journal.pgen.1004000>
- Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S., & Goldstein, D. B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genetics*, 9(8), e1003709. <https://doi.org/10.1371/journal.pgen.1003709>
- Polychronopoulos, D., King, J. W. D., Nash, A. J., Tan, G., & Lenhard, B. (2017). Conserved non-coding elements: developmental gene regulation meets genome organization. *Nucleic Acids Research*, 45(22), 12611–12624. <https://doi.org/10.1093/nar/gkx1074>
- Pool, J. E., & Nielsen, R. (2007). Population size changes reshape genomic patterns of diversity. *Evolution*, 61(12), 3001–3006. <https://doi.org/10.1111/j.1558-5646.2007.00238.x>
- Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G. V., & Camerini-Otero, R. D. (2014). Recombination initiation maps of individual human genomes. *Science*, 346(6211), 1256442. <https://doi.org/10.1126/science.1256442>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Ratkiewicz, M., Matosiuk, M., Saveljev, A. P., Sidorovich, V., Ozolins, J., Männil, P., Balciuskas, L., Kojola, I., Okarma, H., Kowalczyk, R., & Schmidt, K. (2014). Long-range gene flow and the effects of climatic and ecological factors on genetic structuring in a large, solitary carnivore: The Eurasian lynx. *PLoS One*, 9(12), 1–29. <https://doi.org/10.1371/journal.pone.0115160>
- Robinson, J. A., Brown, C., Kim, B. Y., Lohmueller, K. E., & Wayne, R. K. (2018). Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in island foxes. *Current Biology*, 28(21), 3487–3494.e4. <https://doi.org/10.1016/j.cub.2018.08.066>
- Robinson, J. A., Ortega-Del Vecchio, D., Fan, Z., Kim, B. Y., vonHoldt, B. M., Marsden, C. D., Lohmueller, K. E., & Wayne, R. K. (2016). Genomic flatlining in the endangered island fox. *Current Biology*, 26(9), 1183–1189. <https://doi.org/10.1016/j.cub.2016.02.062>
- Rodríguez, A., & Delibes, M. (2003). Population fragmentation and extinction in the Iberian lynx. *Biological Conservation*, 109(3), 321–331. [https://doi.org/10.1016/S0006-3207\(02\)00158-1](https://doi.org/10.1016/S0006-3207(02)00158-1)
- Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., Figueiró, H. V., Johnson, W. E., Milne, H. J., O'Brien, S. J., O'Connell, B., Onorato, D. P., Riley, S. P. D., Sikich, J. A., Stahler, D. R., Villela, P. M. S., Vollmers, C., Wayne, R. K., Eizirik, E., ... Shapiro, B. (2019). Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, 10(1), 4769. <https://doi.org/10.1038/s41467-019-12741-1>
- Schaffner, S. F. (2004). The X chromosome in population genetics. *Nature Reviews Genetics*, 5(1), 43–51. <https://doi.org/10.1038/nrg1247>
- Simons, Y. B., & Sella, G. (2016). The impact of recent population history on the deleterious mutation load in humans and close evolutionary relatives. *Current Opinion in Genetics & Development*, 41, 150–158. <https://doi.org/10.1016/j.gde.2016.09.006>
- Simons, Y. B., Turchin, M. C., Pritchard, J. K., & Sella, G. (2014). The deleterious mutation load is insensitive to recent population history. *Nature Genetics*, 46(3), 220–224. <https://doi.org/10.1038/ng.2896>
- Smith, T. C. A., Arndt, P. F., & Eyre-Walker, A. (2018). Large scale variation in the rate of germ-line de novo mutation, base composition, divergence and diversity in humans. *PLoS Genetics*, 14(3), e1007254. <https://doi.org/10.1371/journal.pgen.1007254>
- Song, Y. S., & Steinrücken, M. (2012). A simple method for finding explicit analytic transition densities of diffusion processes with general diploid selection. *Genetics*, 190(3), 1117–1129. <https://doi.org/10.1534/genetics.111.136929>
- Stephan, W. (2010). Genetic hitchhiking versus background selection: the controversy and its implications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1544), 1245–1253. <https://doi.org/10.1098/rstb.2009.0278>
- Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences of the United States of America*, 118(10), e2015096118–<https://doi.org/10.1073/pnas.2015096118>
- Terekhanova, N. V., Seplyarskiy, V. B., Soldatov, R. A., & Bazykin, G. A. (2017). Evolution of local mutation rate and its determinants. *Molecular Biology and Evolution*, 34(5), 1100–1109. <https://doi.org/10.1093/molbev/msx060>
- Torres, R., Stetter, M. G., Hernandez, R. D., & Ross-Ibarra, J. (2020). The temporal dynamics of background selection in nonequilibrium populations. *Genetics*, 214(4), 1019–1030. <https://doi.org/10.1534/genetics.119.302892>
- Wang, J., Street, N. R., Scofield, D. G., & Ingvarsson, P. K. (2016). Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related populus species. *Genetics*, 202(3), 1185–1200. <https://doi.org/10.1534/genetics.115.183152>
- Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer-Verlag.
- Williams, A. L., Genovese, G., Dyer, T., Altemose, N., Truax, K., Jun, G., Patterson, N., Myers, S. R., Curran, J. E., Duggirala, R., Blangero,

- J., Reich, D., & Przeworski, M. (2015). Non-crossover gene conversions show strong GC bias and unexpected clustering in humans. *Elife*, 4, e04637. <https://doi.org/10.7554/eLife.04637>
- Wilson Sayres, M. A. (2018). Genetic diversity on the sex chromosomes. *Genome Biology and Evolution*, 10(4), 1064–1078. <https://doi.org/10.1093/gbe/evy039>
- Xue, Y., Prado-Martinez, J., Sudmant, P. H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D. N., de Manuel, M., Hernandez-Rodriguez, J., Lobon, I., Siegmund, H. R., Pagni, L., Quail, M. A., Hvilson, C., Mudakikwa, A., Eichler, E. E., ... Scally, A. (2015). Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, 348(6231), 242–245. <https://doi.org/10.1126/science.aaa3952>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Lucena-Perez, M., Kleinman-Ruiz, D., Marmesat, E., Saveljev, A. P., Schmidt, K., & Godoy, J. A. (2021). Bottleneck-associated changes in the genomic landscape of genetic diversity in wild lynx populations. *Evolutionary Applications*, 14, 2664–2679. <https://doi.org/10.1111/eva.13302>