Beyond population size: whole-genome data reveal bottleneck legacies in the peninsular Italian wolf

Daniele Battilani (1,3,4), Roberta Gargiulo (2), Romolo Caniglia (3), Elena Fabbri (3), Jazmín Ramos-Madrigal (4), Claudia Fontsere (4), Marta Maria Ciucani (4), Shyam Gopalakrishnan (4), Matteo Girardi (5), Ilaria Fracasso (5), Matteo Mastroiaco (1), Paolo Ciucci (1), Cristiano Vernesi (5)

1: Università di Roma La Sapienza, Dept. Biology and Biotechnologies "Charles Darwin", Roma, Italy; 2: Royal Botanical Gardens, Kew, United Kingdom; 3: Area per la Genetica della Conservazione, ISPRA, Ozzano dell'Emilia Bologna, Italy; 4: Center for Evolutionary Hologenomics, The Globe Institute, University of Copenhagen, Copenhagen, Denmark; 5: Research and Innovation Centre-Fondazione Edmund Mach, S. Michele all'Adige, Italy.

Contacts

Corresponding authors: daniele.battilani@uniroma1.it, romolo.caniglia@isprambiente.it

Co-authors:

r.gargiulo@kew.org, romolo.caniglia@isprambiente.it, elena.fabbri@isprambiente.it, jazmin.madrigal@sund.ku.dk, claudia.fontsere@sund.ku.dk, ciucani@palaeome.org, shyam.gopalakrishnan@sund.ku.dk, matteo.girardi@fmach.it, ilariaf93@gmail.com, mastroiaco.1808065@studenti.uniroma1.it, paolo.ciucci@uniroma1.it, cristiano.vernesi@fmach.it

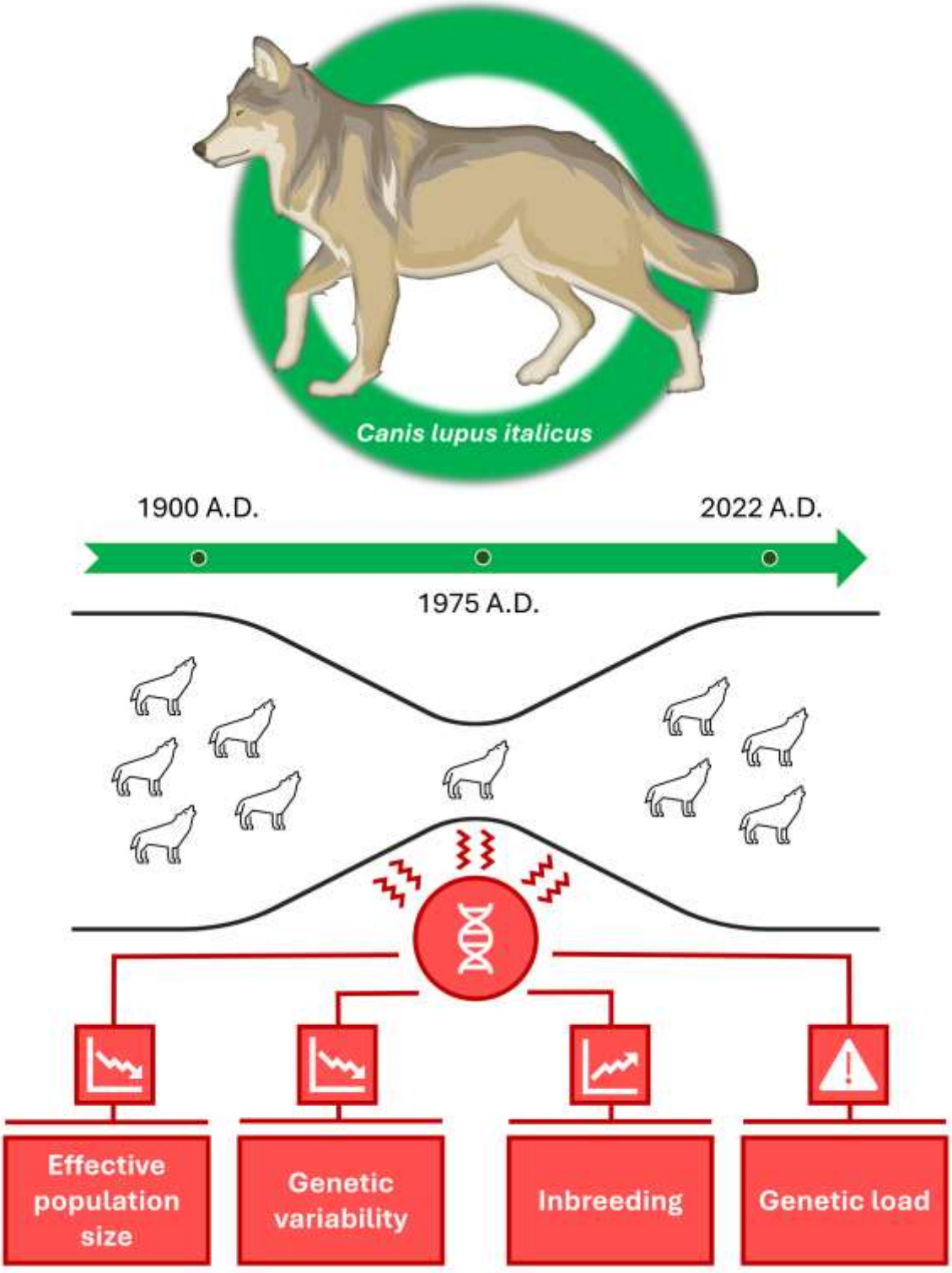
Abstract

Preserving genetic diversity and adaptive potential while avoiding inbreeding depression is crucial for the long-term conservation of natural populations. Despite demographic increases, traces of past bottleneck events at the genomic level should be carefully considered for population management. From this perspective, the peninsular Italian wolf is a paradigmatic case. After being on the brink of extinction in the late 1960s, peninsular Italian wolves rebounded and recolonized most of the peninsula aided by conservation measures, including habitat and legal protection. Notwithstanding their demographic recovery, a comprehensive understanding of the genomic consequences of the historical bottleneck in Italian wolves is still lacking. To fill this gap, we sequenced whole genomes of thirteen individuals sampled in the core historical range of the species in Central Italy to conduct population genomic analyses, including a comparison with wolves from two highly-inbred wolf populations (i.e., Scandinavia and Isle Royale). We found that peninsular Italian wolves, despite their recent recovery, still exhibit relatively low genetic diversity, a small effective population size, signatures of inbreeding, and a non-negligible genetic load. Our findings indicate that the peninsular Italian wolf population is still susceptible to bottleneck legacies, which could lead to local inbreeding depression in case of population reduction or fragmentations. This study emphasizes the importance of considering key genetic parameters to design appropriate long-term conservation management plans.

Keywords

Canis lupus; conservation genomics; genetic diversity; N_e estimation; inbreeding; genetic load

Graphical abstract



1. Introduction

Genetic diversity is one of the three biodiversity pillars [1], having recently received full recognition at the policy level in the Kunming-Montreal Global Biodiversity Framework [2]. Preserving genetic diversity and avoiding inbreeding depression is needed to maintain long-term thriving natural populations [3]. Therefore, key genetic parameters should be carefully taken into account to evaluate the conservation status of natural populations, especially if they have experienced a severe bottleneck, regardless of any subsequent demographic recovery [1,3]. One key parameter is the effective population size (N_e), which represents an idealized population of randomly mating individuals [4] and is linked to genetic diversity loss over time [5]. Factors deviating Wright-Fisher model assumptions can cause estimated N_e significantly different from population size (N) [6,7,5,8]. Severe bottlenecks may increase genetic drift and inbreeding, potentially impacting population standing genetic variation and lead to inbreeding depression. Without inter-population connectivity, new genetic variation can only arise through de-novo mutations, which are more likely to quickly drift in isolated small populations, leading to persistently low genetic diversity [9,10]. Additionally, an increase in homozygosity for deleterious recessive alleles [11] may result in genetic load transitioning from a 'masked' to 'realized' status, thereby reducing individual fitness with potentially negative effects at the population level [12], and thus representing an additional serious conservation challenge in the absence of gene flow [13].

Genetic diversity of wild species has strongly declined in the last century, mainly due to anthropogenic activities that triggered demographic declines and habitat fragmentation [14]. Due to their life history traits and propensity to generate conflicts with humans, large carnivores are particularly exposed to these negative effects [15]. Nevertheless, due to recent legal protection, conservation efforts, and improved habitat suitability, large carnivores are currently recovering throughout Europe [16]. The wolf (*Canis lupus*), is a paradigmatic case of such a natural recovery [16], due to its ability to disperse rapidly over long distances and anthropogenic landscapes [15,17,18]. However, while Eastern European wolves have maintained relatively large populations, functionally connected with counterparts in Russia and Asia [19], populations in Southern Europe, which started to diverge during the Pleistocene, have remained isolated for centuries [20].

In particular, prolonged isolation south of the Alps and recurrent bottlenecks from the last glacial maximum until the last century have made the Italian wolf population the most morphologically and genetically differentiated among the European wolf populations [21–23], to be recognized as a distinct subspecies (*Canis lupus italicus*) [24–26]. As expected, genetic drift and inbreeding acting during such prolonged isolation, as well as continuous demographic declines , caused a drastic loss of genetic variability in the population [23]. Nevertheless, the last historical bottleneck, which occurred after the Second World War, mainly caused by human persecutions, was dramatic from a conservation perspective. The species was brought to the brink of extinction in the late 1960s-1970s with only a few

individuals surviving in the Central-Southern Apennines [27]. However, due to conservation measures and overall positive attitudes from most of the society, wolves have naturally and rapidly recolonized much of their original range throughout the Italian peninsula over the following five decades [28], finally reaching the western Alps in the 1990s [29] and the eastern Alps in the 2010s [30]. Currently, both Alpine and peninsular wolf populations are numerically increasing and are considered as two distinct management units because of their strongly different ecological and socio-economic contexts [31–33]: the first is transboundary, and directly connected with the Dinaric-Balkan-Pindus wolf populations, while the second still remains isolated.

However, the peninsular population is still threatened by critical issues such as poaching, illegal killings, and anthropogenic hybridization with the domestic dog, which might affect the gene pool of contemporary wolves [34–41]. The genetic consequences of these population dynamics and threats for the peninsular Italian wolves can be considered well studied through mitochondrial DNA, microsatellite, and high-resolution SNP panel analyses [21–24,28,42–44]. Although a few studies have used a single peninsular Italian wolf genome [20,45], no study has yet provided information at the population level with whole-genome data.

Therefore, to verify if even a recovered wolf population might still reveal bottleneck legacies, that should not be ignored in long-term conservation and management planning, we sequenced good-coverage whole genomes of thirteen individuals sampled in the historical stronghold of the peninsular population in Central Italy. We used these newly sequenced genomes to perform, for the first time, comprehensive population genomic analyses, with the aims to: (i) evaluate the current genomic variability and estimate Ne in the peninsular Italian wolf population; (ii) assess trends in the historical Ne; and (iii) investigate if inbreeding has significantly affected the peninsular Italian wolf genome-wide variation and its genetic load after the historical bottleneck. Moreover, for comparative purposes, we also included two other highly-inbred wolf populations in Scandinavia and Isle Royale (USA). In fact, similarly to Italian wolves, these two wolf populations underwent strong bottlenecks and founder effects [46,47], and currently show whole-genomic signals of deep inbreeding and increased genetic load [48–50].

2. Materials & Methods

2.1 Sample collection & DNA extraction

Tissue samples were collected from 13 peninsular Italian wolves between 2007 and 2012 from found-death individuals in the Central Apennines, where historical strongholds of wolves in Italy are located [27]. The tissue samples were stored in ethanol at -20°C and subsequently processed in the Conservation Genomics Research Unit at the Fondazione

Edmund Mach (FEM). Small fragments of tissue of around 25 mg were extracted using the DNeasy Blood and Tissue Kit (Qiagen) with overnight digestion at 56°C. The elution was performed at the GLOBE Institute (University of Copenhagen) using two washes of 50 µL of AE buffer, with 10 minutes of incubation at 37°C. Until the elution, samples were stored at -20°C inside the DNeasy Mini spin columns.

2.2 Library preparation, amplification & whole genome sequencing

Extracts were fragmented in the Covaris LE220 plus Focused-ultrasonicator with the parameters set for 350-bp fragment length. The extracts were diluted to obtain 100 ng concentration and BGI libraries for Italian wolves were constructed following previously optimized protocols [51,52] and using 10 µM adaptors. Libraries were purified using MinElute columns using PE buffer (Qiagen) and eluted in 60 µL of EB buffer. The PCR mixture for the peninsular Italian wolf libraries consisted of: 20 µL of purified library, 0.2 µM of forward and reverse BGI primers, 2.5 U/µL PfuTurbo Cx HotStart DNA Polymerase, 10 µL of Buffer 10X, 0.08 mg/mL BSA, 0.5 mM of dNTPs (25 µM) and 61.2 µL AccuGene molecular biology water (Lonza, Basel, CH). The amplification of peninsular Italian wolf samples was performed as follows: initial denaturation at 95°C for 2 minutes followed by 10 to 12 cycles of 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 110 seconds, and a final elongation step at 72°C for 10 minutes. Peninsular Italian wolf samples were sequenced on ⅛ of a lane each on MGIseq2000 PE150 and DNBSEQ PE150, respectively. Four out of 13 samples have been previously published by [53] (Supplementary Table 1) at a low coverage (ca 3.6x). We resequenced these samples aiming to reach a higher coverage for the purpose of this study.

2.3 Dataset

In addition to the 13 peninsular Italian wolves (WIT) genomes sequenced here, we also retrieved genomic data from public databases (NCBI GenBank) (Supplementary Table 1). Therefore, the final dataset compiled for this study comprised 101 modern samples, including 13 WIT (*Canis lupus italicus*), 10 Scandinavian wolves (*Canis lupus*) (WSC) [48], 11 North American wolves from Isle Royale (*Canis lupus*) (WUS) [49], and 67 modern domestic dogs (*Canis lupus familiaris*) belonging to 67 different breeds of medium or large size (DOG) (Supplementary Table 1). The 10 Scandinavian wolves were a random subsample of those available (n=96). One African wolf (*Canis lupaster*) and one Golden Jackal (*Canis aureus*) were used as outgroups (OUT) for the genetic load analyses.

2.4 Quality control & alignment

We applied a quality control procedure on the sequencing reads, using FastQC [54] to check for possible issues such as low quality scores and anomalous GC content, and we used multiQC [55] to visualize them. The reads were mapped both onto the wolf reference genome [56] and onto the dog reference genome (CanFam3.1[57]), as the "admixture analyses" and the "genetic load analyses" required genomic regions from dogs and outgroups to be also mapped to the dog reference genome. To perform mapping, we set and ran the automated PALEOMIX BAM pipeline [58]: first, it indexed each read prefix using SAMtools 'faidx' [59]; then it removed the specified BGI adapters using AdapterRemoval [60]; the mapping was done using BWA 'mem' algorithm that is suggested for modern samples [61], setting the minimum mapping quality to 0 to retain all the reads in this step; to conclude PCR duplicates were removed using Picard MarkDuplicates (https://broadinstitute.github.io/picard/). After that, we used SAMtools [59] to remove non primary alignment reads (samtools view -F 256).

2.5 Genotype processing

We used GATK v 4.3.0.0 and referred to GATK Best Practice Workflow to call high quality genotypes [62]. Then we applied two additional GATK tools for 'hard filtering' our genotypes: VariantFiltration (QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5 and ReadPosRankSum < −8.0; settings taken from alternative protocol 2 in GATK Best Practices) to mark the filters, and SelectVariants to apply them. Finally, we applied other filters using VCFtools [63] to keep only biallelic SNPs (flags --remove-indels --max-alleles 2 --min-alleles 2) and to filter for minor allele frequency (MAF), missingness, minimum quality and minimum and minimum average depth (flags --maf 0.05 --max-missing 0.9 --minQ 30 --minDP 5 --min-meanDP 5). We used those filters on different datasets according to the assumptions of the downstream analyses (Supplementary Table 2). As "admixture analyses" rely on the assumption that SNPs are not in physical linkage, we performed LD pruning using PLINK v 1.90b6.21 [64], setting a window size of 10 kb, a step size of 5 bp and an r2 threshold of 0.5 (flag --indep-pairwise 10 5 0.5). Moreover, to avoid violating the assumptions of random mating when carrying out most population genomic analyses , we used NgsRelate2 to identify and eventually remove closely related individuals [65] applying thresholds of KING-robust kinship ≥ 0.20, R0 ≤ 0.1, and R1 ≥ 0.5 [66].

2.6 Admixture analyses

We explored patterns of genetic differentiation among samples through a preliminary non-model Principal Components Analysis (PCA), calculating principal components with PLINK v 1.90b6.21 (flag --pca). The percentage of variance explained was calculated from the '.eigenval' output, and the first two principal components (PCs) were used for plotting in R v

4.2.1 using 'ggplot2' [67]. Then, we used ADMIXTURE v1.3.0 that uses a maximum likelihood approach [68] to estimate the proportions of a given number of ancestries (K) for each individual. We assumed K values from 2 to 6 and set the --cv flag to calculate cross-validation errors (CV) for each K. For each K, we ran 15 independent iterations with different starting seeds and chose the iterations with the highest likelihood. The best K was then chosen based on the lowest CV value among the best iterations. If individuals potentially admixed with dogs were identified, we re-ran the analysis excluding them. Admixed individuals were further investigated within each specific wolf population through a supervised ADMIXTURE analysis (flag --supervised) to confirm their status after reapplying the filters for 'admixture analyses'. In case some of these individuals were confirmed as admixed (i.e., hybrids or introgressed with dog), we conducted the downstream analyses by both keeping and removing them to highlight potential differences. As Stefanović et al. [69] demonstrated pervasive jackal-dog hybridization across the *Canis aureus* range, we also checked dog ancestry in the chosen outgroups (OUT) applying an additional supervised ADMIXTURE analysis.

2.7. Genomic variability analyses

To compare the patterns of genomic variation among the three wolf populations, we estimated the observed heterozygosity (Ho) and nucleotide diversity (π). We used ANGSD [70] to estimate the heterozygosity of each sample, by calculating the folded site frequency spectrum (SFS) on autosomes. We estimated genotype likelihoods using ANGSD's GATK (-GL 1) model (doSaf 1), removing bases with quality score lower than 20 (-minQ 20) and reads with mapping quality lower than 30 (-minmapq 30). The dog reference genome was used both as reference and as ancestral (-ref and -anc options). The genome-wide SFS was estimated using the realSFS utility tool provided in ANGSD and subsequently the final heterozygosity was calculated as the ratio of heterozygous sites/total sites. Since most of the individual genomes did not exhibit a high depth of coverage (> 20x), and a recent study has pointed out that low sequencing depth can bias the specific downstream analyses we are interested in (e.g. inbreeding and N_e estimates; [71]), we tested the possible correlation between individual heterozygosity and read depth using R v 4.2.1 (Pearson's correlation test). Subsequently, we tested for possible discrepancies between heterozygosity estimated from genotype likelihood and heterozygosity estimated from variant calling procedures. Nucleotide diversity (π) was estimated on variant sites only, using VCFtools [63]. We applied a sliding window approach with a window size of 100 kb (--window-pi 100000), including all the samples for each population.

2.8 Demographic analyses

To understand the demographic trend in the three wolf populations over time, we estimated recent historical N_e in the three wolf populations using the linkage disequilibrium (LD) method as implemented in GONE [72] GONE exploits the genetic distances among SNPs to estimate N_e in more recent generations (relying on the information associated with loosely linked SNPs) and historical N_e (relying on the information associated with tightly linked SNPs), providing reliable estimates of N_e up to 200 generations ago. In GONE, input parameters were set to their default values, except the average recombination rate that was set to 1.3459 CentiMorgans per Megabase, as obtained from [73] (CanFam 3.1). We considered generation length in wolves to be 3-4 years [74,75] and ran the analysis on a maximum of 50,000 SNPs per chromosome (maximum value accepted by GONE) with no additional MAF filtering. Analyses in GONE were repeated 20 times to obtain empirical confidence intervals. Although this method may not fully capture the uncertainty introduced by the sampling process (as, for instance, methods based on jackknifing), especially given our reduced sample size, it is able to provide an indication of the uncertainty around our N_e estimates in GONE [72]. Point estimates were summarized using the geometric mean across the values obtained in each replicate. We also used the software currentNe, which is more accurate than GONE for contemporary N_e (i.e. last 2-3 generations) even with small sample sizes [76]. CurrentNe produces confidence intervals for N_e based on artificial neural networks without the need for iterating the analysis. We subsampled the datasets for WIT, WUS and WSC to 50,000 random SNPs prior to the analyses in current N_e, and used the N_e estimate obtained only based on LD between chromosomes. To understand how the pedigree of the sampled individuals would affect the N_e estimation, we carried out the analyses by either including or excluding highly-related individuals, under the expectation that in a random sample, the frequency of related individuals is a determinant of the genetic drift signal in the population and therefore, the exclusion of putative relatives from the analyses can upwardly bias the N_e estimates [77,78].

2.9 Inbreeding analyses

We identified runs of homozygosity (ROH), indicative of putative identity-by-descent chromosome segments, in WIT, WSC, and WUS whole-genome data using the window-based approach implemented in PLINK v 1.90b6.21 (--homozyg). We employed a sliding window of 100 kb (--homozyg-kb 100), a threshold that has found favor in population genetics studies of non-model species [49,79,80]. For the remaining parameters, we kept PLINK default values. A minimum of 100 SNPs (--homozyg-snp 100) at a minimum density of 1 SNP per 50 kb was required to call a ROH (--homozyg-density 50). To account for genotyping errors, we allowed up to 1 heterozygous site per 1000 kb window within called ROHs (--homozyg-window-het 1), as per [81], and 5 missing calls per 1000 kb window within called ROHs (--homozyg-window-missing 5). A length of 1000 kb between two SNPs was

required in order for them to be considered in two different ROHs (--homozyg-gap 1000). We calculated the total length of all ROHs for each individual (SROH) and plotted it on the 'x' axis against the number of ROHs for each individual (NROH) on the 'y' axis. We estimated the inbreeding coefficient based on ROH (FROH) as the ratio of SROH to the total length of the autosomal genome covered by SNP positions (herein 2,222,501,653 bp) for each individual. We plotted SROH accounting for different ROH sizes ('short': 100 kb < ROH < 1 Mb; 'intermediate': 1 Mb < ROH < 10 Mb; 'long': ROH > 10 Mb), to highlight the FROH components. Finally, we calculated ROHs coalescence time aiming to define the time at which the ROHs were likely formed. To do so, we used the formula L = 100/2t cM [82] where L is the length of the ROH, cM is the recombination rate and t is the time of coalescence in generations. If demographic analyses revealed some bottleneck, we checked the impact of the ROHs that have been formed during and immediately after the bottleneck (5 generations later) on the population SROH.

2.10 Genetic load analyses

We used publicly available short-read data from two outgroups (OUT), *Canis lupaster* (African wolf, SRA accession: SAMN10199001) and *Canis aureus* (Golden Jackal, SRA accession: SAMN10180427) to polarize SNPs, defining the 'ancestral' and 'derived' states for each variant. To do so, we used est-sfs [83], a tool implementing a maximum likelihood method to infer the unfolded site frequency spectrum (the uSFS) and the ancestral state probabilities for our OUT data. Using a custom script, we obtained the est-sfs input file and extracted the ancestral state with the highest probability for each variant from the output file. If two bases had the same probability of being ancestral, the script randomly picked one of the two. We used the domestic dog genome annotation gtf, cDS and protein files (CanFam3.1; Ensembl release 104 https://may2021.archive.ensembl.org/Canis_lupus_familiaris/Info/Index) to build custom snpEff v.4.3.183 [84] and SIFT4G v.6.084 [85] databases with default settings. We then annotated and predicted the effects of variants using the aforementioned tools. Following annotation, we retained variants where the derived state matched the alternative allele in the dog reference genome. This ensured that the derived alleles had the deleterious effects indicated for the alternative alleles of the dog reference genome. We used classified putatively deleterious variants into three categories: (i) Low-impact (LOW) variants likely to be not deleterious (i.e., synonymous), (ii) Moderate-impact (MODERATE) variants likely to modify protein effectiveness (i.e., nonsynonymous), and (iii) High-impact (HIGH) variants likely to disrupt protein function (i.e. loss of function LoF, stop codons, splice donor variant and splice acceptor, or start codon lost) (56). We used SIFT score to discriminate MODERATE in tolerated nonsynonymous (SIFT score ≥0.05) (MOD-TOL) and putatively deleterious nonsynonymous (SIFT score <0.05) (MOD-DEL) variants. Then we used a custom script to estimate individual allele frequencies and genotype counts of LOW, MOD-TOL, MOD-DEL

and HIGH impact variant derived alleles, for each wolf in the dataset. In particular, the count of heterozygous genotypes represented the 'masked' load, which quantifies the potential loss of fitness due to (partially) recessive deleterious mutations that may become expressed in future generations. The count of homozygous genotypes for the derived alleles represented the 'realized' load, which reduces fitness in the current generation; the sum of 'masked' and 'realized' load represented the 'total' load [12].

3. Results

We sequenced 13 peninsular Italian wolves (WIT) at an average of 15x coverage and recovered 6,636,110 SNPs. Overall, the dataset we used for the analyses ranged 555,575-8,518,995 SNPs, depending on the analysis and the specific populations included in the analyses (Supplementary Table 2). We did not find any highly-related pairs within the sampled population. However, we opted to exclude one individual that displayed two out of three relatedness indexes beyond the threshold along with a reduced genotype quality (Supplementary Fig. 1). We first explored population structure and admixture, aiming to verify the identity of the samples belonging to Italian (WIT), Scandinavian (WSC) and Isle Royale (United States; WUS) populations, and to evaluate if there was dog ancestry in their genomes. The first two components of the explorative PCA explained 54,76% of the genetic variability in the dataset, demonstrating a clear separation between wolves and dogs along the first PC (PC1 - 41.15%). The second PC (PC2 - 13.61%) revealed the distinct clustering among the three wolf populations (Supplementary Fig. 2). The ADMIXTURE clustering mirrored the PCA results. We performed cross-validation (CV) for the admixture analyses (K={2...7}) and found that our dataset has the lowest CV error when estimating four ancestry clusters (Supplementary Fig. 3). With four ancestry clusters, each wolf population formed a discrete cluster, distinct from the dog cluster. Admixture analysis also revealed four WIT, one WSC and 17 DOG individuals as potentially admixed. Re-running the software without the 17 admixed DOG individuals returned the same admixed wolves (Fig. 1A). A 'supervised' ADMIXTURE approach confirmed that two of the four admixed WIT had more than 10% DOG ancestry, while the other two shared a smaller (1-3%) DOG ancestry proportions (Fig. 1B). The only admixed WSC individual had been previously identified as an immigrant wolf [48], suggesting that the dog ancestry proportion can be explained with wolf ancestry not represented in our reference dataset. Both *Canis lupaster* and *Canis aureus* genomes (OUT) did not show dog ancestry (Supplementary Fig. 4), and for this reason, they have been confirmed as valuable outgroups for genetic load analyses. Nevertheless, even though we ran the downstream analyses excluding the admixed individuals, either keeping or removing admixed individuals did not significantly affect the results.

3.1 Genome-wide patterns of variation

We estimated the genome-wide variation for the WIT in comparison with the other highly-inbred wolf populations. Initially, we validated heterozygosity estimation by demonstrating the absence of significant correlation with read depth (p-value = 0.19) (Supplementary Fig 5). Furthermore, when comparing heterozygosity estimates derived from the genotype likelihoods and those from the called genotypes across the three populations, a consistent pattern emerged. WSC had the lowest observed heterozygosity (Ho = 0.00149 ± 0.00025) estimate, followed by WIT and WUS with Ho= 0.00152 (± 0.00018) and Ho=0.00178 (± 0.00019), respectively (Fig. 3). Nucleotide diversity estimate was lowest in WIT (π=0.00082 ± 0.0011) and highest in WSC (π =0.00108 ± 0.00038), with WUS of intermediate value (π =0.00102 ± 0.00043) (Supplementary Fig. 6). Applying the Wilcoxon signed-ranked test, we did not find any difference in Ho and π in the three testes wolf populations.

3.2 Recent historical and contemporary N_e estimates

The demographic analyses carried out with GONE showed that WUS had the smallest estimated N_e, with N_e below 10 during the most recent generations, and N_e 62-71 from 20 to 50 generations before present. WIT and WSC showed similar N_e values (24-32 for WIT, 23-50 for WSC) for the last ten generations, with a severe bottleneck occurring between 12 and 18 generations for WIT and 10-20 for WSC. From 20 to 50 generations before present, both populations showed a constant and stable trend, but significantly different N_e estimates: N_e in WIT ranged between 330 and 344, and in WSC ranged between 508 and 584, although with much statistical uncertainty, as shown in fig. 2A-B. Contemporary N_e estimates (i.e. referring to the last 2-3 generations) obtained with currentNe were equal to 10.4 (90% CI: 7.6-14.3) for WIT, 14.4 (90% CI: 9.3-22.2) for WSC, and 7.4 (90%CI: 5.5-10.0) for WUS (Supplementary Table 3). The exclusion of the highly-related WIT individual from the analyses did not significantly affect recent historical N_e estimates (i.e., last 10 generations: 29-48; last 20-50 generations: 334-386) and contemporary N_e estimates (11.89; 90% CI: 8.4-16.9).

3.3 Runs of homozygosity

WIT had both the highest number of estimated ROHs for each individual (NROH) (mean = 1,374.50) and total length of estimated ROHs for each individual (SROH) (mean = 909.47 Mb), followed by WSC (mean NROH = 1,125.88; mean SROH = 713.03 Mb) and WUS (mean NROH = 420.64; mean SROH = 852.33 Mb). We chose the SROH vs NROH plot based on ROH > 500 kb, given it better represents the demographic history of these samples, after checking different options (ROH > 100 kb, ROH > 1 Mb, ROH > 2 Mb) (Fig. 2C). Individual genomic inbreeding coefficients (FROH) were plotted accounting for all the estimated ROH

(> 100 kb) (Supplementary Fig. 7). WIT had the highest estimated FROH (mean = 0.409), followed by WUS (mean = 0.380), and WSC (mean = 0.321). When plotting SROH accounting for different ROH sizes (Fig. 3), WIT had the highest estimated SROH for 'short' ROHs (mean = 355.43 Mb), followed by WSC (mean = 305.43 Mb) and WUS (mean = 81.12 Mb) (WIT-WUS two-tailed Mann–Whitney U [MWU] test p-value = 1.191e-05; WSC-WUS MWU test p-value = 2.646e-05). Upon considering intermediate size regions, WIT showed the highest estimated SROH (mean = 502.48 Mb), followed by WSC (mean = 401.45 Mb) and WUS (mean = 348.17) (WIT-WUS MWU test p-value = 0.001773). When accounting for long ROHs, WUS clearly separated with a higher estimated SROH (mean = 423.04 Mb) compared to WIT (mean = 11.88 Mb) and WSC (mean = 6.13 Mb) (WUS-WIT MWU test p-value = 0.0001928; WUS-WSC MWU test p-value = 0.0001615). The coalescence time calculation for ROHs revealed that 100 kb ROHs likely formed more than 350 generations ago, 1 Mb ROHs likely formed 37 generations ago, and 10 Mb ROHs likely formed 4 generations ago (Supplementary Fig. 8). During and immediately after the bottleneck, WSC formed ROHs of 1.7-7.4 Mb that contributed to 32.31% of the population SROH; while WIT formed ROHs of 2-5.3 Mb that contributed to 27.14% of the population SROH (Fig. 2D).

3.4 Genetic load

Analysis of derived allele frequencies among individuals (Supplementary Fig. 9) revealed a notable prevalence of likely-not-deleterious alleles (LOW and MOD-TOL) compared to likely deleterious alleles (MOD-DEL and HIGH) across all three populations. Particularly, the most deleterious alleles (HIGH) were found to be the least common. Additionally, genotype count estimations demonstrated a consistent trend across all mutation types, except for WIT, which exhibited the highest average counts for both 'masked' LOW and 'realized' MOD-TOL alleles (Fig. 4). When focusing on deleterious mutations with significant potential impact on individual fitness (MOD-DEL and HIGH), WUS exhibited the highest load, both 'masked' (HIGH: WUS-WIT MWU test p-value = 0.001262; WUS-WSC MWU test p-value = 0.01086) and 'realized' (HIGH: WUS-WSC MWU test p-value = 0.004225). Following WUS, WIT showed a similar estimate of 'realized' load for both MOD-DEL and HIGH variants (WIT-WSC MWU test p-value = 0.01373), while WSC demonstrated a comparable estimate of 'masked' load for MOD-DEL and HIGH variants. The 'total' genetic load assessment positioned WIT as intermediate between WSC and WUS for both MOD-DEL (WIT-WUS MWU test p-value = 0.0132; WSC-WUS MWU test, p-value = 0.01089) and HIGH impact variants (WIT-WSC MWU test p-value = 0.01589; WIT-WUS MWU test p-value = 0.0003185; WSC-WUS MWU test, p-value = 0.0001906).

4. Discussion

Population size (*N*) is a functional parameter to assess numerical population trends [86]. However, to better investigate the conservation status of a population it is necessary to also evaluate genetic and genomic factors, and how they might have been affected by demographic dynamics [1,3]. For example, severe bottlenecks can strongly reduce effective population size (N_e) while facilitating inbreeding. This can determine an increased homozygosity for deleterious alleles, which might be phenotypically expressed and affect individual fitness, therefore potentially leading to inbreeding depression and reducing long-term local adaptive potentials. It has been amply demonstrated that the signatures of genomic erosion associated with past bottlenecks can persist even in populations experiencing a geographic re-expansion and a demographic recovery [87–90]. Therefore, in this research, we focused on the emblematic case study of the peninsular Italian wolf population (WIT), by applying whole-genome approaches to investigate bottleneck legacies on its current genetic diversity.

Comparative genomic variability analyses with other two highly inbred wolf populations (Scandinavian - WSC; Isle Royale, United States - WUS) known to be affected by inbreeding depression, showed that peninsular Italian wolves must be placed among the most genetically eroded wolf populations worldwide. These findings confirm previous studies obtained using traditional genetic markers [23,91], genome-wide SNP panels [42] or a single whole-genome [20].

The recent historical N_e we estimated for WIT showed a significant reduction during 12-18 generations before sampling, that, assuming a generation time of 3-4 years [74,75], occurred between 1938 and 1974. These findings align well with known historical demographic trends of Italian wolves [27,92], indicating that they disappeared from Northern Italy and Sicily immediately after World War II [92], afterwards reaching the minimum population size of about only 100 individuals, in small mountain islets along the Central and Southern Apennines [27]. Our recent historical and contemporary N_e estimates in WIT do not parallel current *N* estimates [33,93] (Fig. 2). This discrepancy might be due to a persisting bottleneck signal in the sampled individuals, suggesting enduring low viability and limited adaptive potential despite the demographic recovery of the population. We cannot exclude, however, that the localized sampling area, compared to the entire Italian peninsula, and the species life-history traits, characterized by few breeders and high juvenile mortality, may contribute to the low N_e estimate and the high *N* estimate. Additionally, a slight downward bias in the N_e estimates may partly be expected due to the inclusion of individuals sampled at different times (2007-2012), which violates the assumption of discrete generations associated with all N_e estimation methods [71,94,95].

The ROH analyses showed that bottleneck and inbreeding deeply affected the genomic make-up of the peninsular Italian wolf population, corroborating previous findings based on a single genome [20]. However, our study also provides a robust population-level

perspective over time, confirming that WIT represents one of the most inbred populations both at the Eurasian scale [96] and at the global scale [20,49]. The SROH vs NROH analyses confirmed a general insight into WIT demographic history. Individuals distribution compared to the intercept line [97] suggested that bottleneck and inbreeding affected the population. The SROH composition, in terms of ROH length, and ROHs coalescence time could provide us information about inbreeding and its magnitude [97]. WIT exhibited the highest SROH estimate for both the shortest ROHs (100kb), which are less likely to significantly impact individual fitness compared to more recent autozygous segments [98], and also for intermediate ROHs (1-10 Mb long), formed between 4-37 generations ago and suggesting a strong bottleneck and post-bottleneck inbreeding.

Our genetic load analyses showed that WIT has a slightly lower realized load compared to deeply-inbred WUS, for both highly (HIGH) and moderately deleterious (MOD-DEL) impact variants, indicating that negative impacts are still acting on individual fitness. This result suggests that purifying selection hasn't strongly worked, and gene flow couldn't act as well, given the ongoing isolation of the peninsular Italian population. In fact, while genetic connectivity between the Italian wolves and wolves from the Dinaric-Balkan-Pindus population has been only recently reinstated in the eastern Alps [32], no such connectivity has been documented in peninsular Italy [31,32]. The post-bottleneck masked load represents the residual load accumulated by the ancestral population and it may be proportional to the historical N_e [12,99]. For this reason, we expected WIT to show the lowest masked load values, since WSC had a recent historical N_e estimate higher than WIT, while WUS, originating from the Great Lakes and Ontario wolf population exhibited a higher historical N_e estimate than WIT in a previous study [100]. Also in this context, the composition of WSC and WUS (i.e., each population including a few individuals sampled over 25 years) could bias N_e estimation (mostly likely downward) [71], but such potential bias should not prevent us from observing general demographic patterns. Considering that WSC masked load estimates could be inflated by immigrants' genomic variation as well [50], the lack of difference between them and WIT suggests that the latter should not be underestimated. For these reasons, the current WIT genetic load could represent a previously unaccounted but potentially serious conservation issue in case of local fragmentation and reduction. Further genomic investigation, however, should include pre-bottleneck WIT samples to have a proper comparison of genomic variation through time.

Finally, our genomic admixture analyses showed 4 out of 13 of the randomly collected WIT individuals exhibiting >1% of dog ancestry, with 2 of them exceeding 10%. These findings are consistent with the hypothesis that pervasive wolf-dog hybridization in the peninsular Italian wolf population might have occurred during the recent wolf population expansion, when wolf encounters and interbreeding with dogs might have been more likely [43]. Although wolf-dog hybridization has been already documented in Italy using genetic and genome-wide markers [34–40,43,44,53], our study provides the first evidence based on complete whole genomes. Despite excluding admixed individuals did not significantly affect

the results obtained in this study, a larger dataset of whole-genome data will be fundamental to disentangle the effects of this phenomenon on the standing genetic variation. Additionally, further research should explore the evolutionary and ecological implications that the introgression of domestic genes might have on the long-term survival of such a morphologically and genetically unique wolf population.

Our study provides insights into the peninsular Italian wolf population through comprehensive population genomics analyses. Although the individuals we typed were sampled in the core of the historical range in the central Apennines, further studies should enlarge the geographic scope of sampling throughout the whole subspecies distribution in the peninsula. Nevertheless, it has to be considered that unrelated individuals might belong to different packs and regions, and no substructure has been revealed in the peninsular Italian wolf population [30]. For these reasons, we can consider our samples as representative for the peninsular Italian wolf population.

Despite demographic growth and range expansions, wild populations could still be susceptible to eventual geographic fragmentations, numerical reductions and natural or human-induced environmental changes. Our findings suggest that, although inbreeding depression has not been directly documented in the peninsular Italian wolf population, the reduced genetic diversity, significant inbreeding signatures, and the non-negligible genetic load that we disclosed could lead to fitness decreasing in case of other future demographic contractions or geographic fragmentations. Moreover, given the difficulties in detecting inbreeding depression in natural populations, we cannot exclude that it may already be occurring to some extent at the local scale. In this context, we maintain that, from a conservation-wise perspective, the protection level of wolves should not be defined only considering demographic data but especially their genomic make-up. Including information about genetic variability, effective population size, inbreeding, genetic load, and hybridization rates among criteria to assess population conservation status would be relevant even to design specific management measures from a wide to a local scale [101–104]. Accordingly, we also maintain that preserving genetic diversity, after receiving full recognition in the Kunming-Montreal Global Biodiversity Framework [2], should be more explicitly and specifically stated among the conservation goals in relevant international (i.e., Bern Convention), European (i.e., Habitats Directive) and national conservation treaties and laws.

Funding

D.B. was supported by a PhD grant of the University of Rome “La Sapienza” & MUR (grant no. DOT1326JZS-17; CUP B85F21005360001). C.V. contribution was funded by Research and Innovation Centre-Fondazione Edmund Mach. P.C. contribution was funded by the European Union - NextGenerationEU National Biodiversity Future Center.

Acknowledgements

We are very grateful to the volunteers that participated in the opportunistic samples collection. We are thankful to Robin Waples and an anonymous reviewer: their comments on a first draft allowed us to significantly improve the paper.

Data availability

All the data and codes processed and produced in this work have been submitted to Dryad (DOI:10.5061/dryad.ngf1vhj2f; https://datadryad.org/stash/share/Ql2eeqaemPKckr_5jYxocEvAPAekaVrWMO7hZy3rlgA).

Raw DNA sequence reads of whole-genomes generated in this study are going to be submitted to the National Center for Biotechnology Information (NCBI) upon article acceptance.

References

1. DeWoody JA, Harder AM, Mathur S, Willoughby JR. 2021 The long-standing significance of genetic diversity in conservation. *Molecular Ecology* **30**, 4147–4154. (doi:10.1111/mec.16051)

2. Ma K. 2023 Kunming-Montreal Global Biodiversity Framework: An important global agenda for biodiversity conservation. *Biodiversity Science* **31**, 23133. (doi:10.17520/biods.2023133)

3. Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, Tallmon DA, Funk WC. 2021 The crucial role of genome-wide genetic variation in conservation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104642118. (doi:10.1073/pnas.2104642118)

4. Waples RS. 2022 What Is *N* e, Anyway? *Journal of Heredity* **113**, 371–379. (doi:10.1093/jhered/esac023)

5. Charlesworth B. 2009 Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* **10**, 195–205. (doi:10.1038/nrg2526)

6. Keller L, Reeve HK. 1994 Partitioning of reproduction in animal societies. *Trends in Ecology & Evolution* **9**, 98–102. (doi:10.1016/0169-5347(94)90204-6)

7. Hedrick PW, Kalinowski ST. 2000 Inbreeding Depression in Conservation Biology. *Annu. Rev. Ecol. Syst.* **31**, 139–162. (doi:10.1146/annurev.ecolsys.31.1.139)

8. Clutton-Brock T. 2021 Social evolution in mammals. *Science* **373**, eabc9699. (doi:10.1126/science.abc9699)

9. Coyne JA, Barton NH, Turelli M. 1997 Perspective: A Critique of Sewall Wright's Shifting Balance Theory of Evolution. *Evolution* **51**, 643. (doi:10.2307/2411143)

10. Wade MJ, Goodnight CJ. 1998 PERSPECTIVE: THE THEORIES OF FISHER AND WRIGHT IN THE CONTEXT OF METAPOPULATIONS: WHEN NATURE DOES MANY SMALL EXPERIMENTS. *Evolution* **52**, 1537–1553. (doi:10.1111/j.1558-5646.1998.tb02235.x)

11. Charlesworth D, Willis JH. 2009 The genetics of inbreeding depression. *Nat Rev Genet* **10**, 783–796. (doi:10.1038/nrg2664)

12. Bertorelle G, Raffini F, Bosse M, Bortoluzzi C, Iannucci A, Trucchi E, Morales HE, Van Oosterhout C. 2022 Genetic load: genomic estimates and applications in non-model animals. *Nat Rev Genet* **23**, 492–503. (doi:10.1038/s41576-022-00448-x)

13. Kardos M, Taylor HR, Ellegren H, Luikart G, Allendorf FW. 2016 Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications* **9**, 1205–1218. (doi:10.1111/eva.12414)

14. Exposito-Alonso M *et al.* 2022 Genetic diversity loss in the Anthropocene. *Science* **377**, 1431–1435. (doi:10.1126/science.abn5642)

15. Boitani L, Linnell JDC. 2015 Bringing Large Mammals Back: Large Carnivores in Europe. In *Rewilding European Landscapes* (eds HM Pereira, LM Navarro), pp. 67–84. Cham: Springer International Publishing. (doi:10.1007/978-3-319-12039-3_4)

16. Chapron G *et al.* 2014 Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* **346**, 1517–1519. (doi:10.1126/science.1257553)

17. Ciucci P, Reggioni W, Maiorano L, Boitani L. 2009 Long-Distance Dispersal of a Rescued Wolf From the Northern Apennines to the Western Alps. *J Wildl Manag* **73**, 1300–1306. (doi:10.2193/2008-510)

18. Hindrikson M *et al.* 2017 Wolf population genetics in E urope: a systematic review, meta-analysis and suggestions for conservation and management. *Biological Reviews* **92**, 1601–1629. (doi:10.1111/brv.12298)

19. Mech D, Boitani L. 2003 *Wolves: Behavior, Ecology, and Conservation*. Chicago: University of Chicago Press. See https://press.uchicago.edu/ucp/books/book/chicago/W/bo3641392.html.

20. Fan Z *et al.* 2016 Worldwide patterns of genomic variation and admixture in gray wolves. *Genome Res.* **26**, 163–173. (doi:10.1101/gr.197517.115)

21. Lucchini V, Galov A, Randi E. 2004 Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. *Molecular Ecology* **13**, 523–536. (doi:10.1046/j.1365-294X.2004.02077.x)

22. Stronen AV *et al.* 2013 North-South Differentiation and a Region of High Diversity in European Wolves (Canis lupus). *PLoS ONE* **8**, e76454. (doi:10.1371/journal.pone.0076454)

23. Pilot M, Greco C, vonHoldt BM, Jędrzejewska B, Randi E, Jędrzejewski W, Sidorovich VE, Ostrander EA, Wayne RK. 2014 Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity* **112**, 428–442. (doi:10.1038/hdy.2013.122)

24. Randi E, Lucchini V, Christensen MF, Mucci N, Funk SM, Dolf G, Loeschcke V. 2000 Mitochondrial DNA Variability in Italian and East European Wolves: Detecting the Consequences of Small Population Size and Hybridization. *Conservation Biology* **14**, 464–473. (doi:10.1046/j.1523-1739.2000.98280.x)

25. Nowak RM, Federoff NE (Brusco). 2002 The systematic status of the Italian wolfCanis lupus. *Acta Theriol* **47**, 333–338. (doi:10.1007/BF03194151)

26. Montana L, Caniglia R, Galaverni M, Fabbri E, Randi E. 2017 A new mitochondrial haplotype confirms the distinctiveness of the Italian wolf (Canis lupus) population. *Mammalian Biology* **84**, 30–34. (doi:10.1016/j.mambio.2017.01.005)

27. Zimen E, Boitani L. 1975 Number and distribution of wolves in Italy. *Zeitschrift Saeugetierkunde 40,.* , 102–112.

28. Fabbri E *et al.* 2007 From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. *Molecular Ecology* **16**, 1661–1671. (doi:10.1111/j.1365-294X.2007.03262.x)

29. Lucchini V, Fabbri E, Marucco F, Ricci S, Boitani L, Randi E. 2002 Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. *Molecular Ecology* **11**, 857–868. (doi:10.1046/j.1365-294X.2002.01489.x)

30. Fabbri E, Caniglia R, Kusak J, Galov A, Gomerčić T, Arbanasić H, Huber D, Randi E. 2014 Genetic structure of expanding wolf (Canis lupus) populations in Italy and Croatia, and the early steps of the recolonization of the Eastern Alps. *Mammalian Biology* **79**, 138–148. (doi:10.1016/j.mambio.2013.10.002)

31. Linnell J, Salvatori V, Boitani L. 2007 Guidelines for Population Level Management Plans for Large Carnivores. A Large Carnivore Initiative for Europe report prepared for the European Commission. *Guidelines for Population Level Management Plans for Large Carnivores. A Large Carnivore Initiative for Europe report prepared for the European Commission.* See https://rm.coe.int/1680746791.

32. Marucco F *et al.* 2023 Transboundary Monitoring of the Wolf Alpine Population over 21 Years and Seven Countries. *Animals* **13**, 3551. (doi:10.3390/ani13223551)

33. Gervasi V *et al.* 2024 Estimating distribution and abundance of wide-ranging species with integrated spatial models: Opportunities revealed by the first wolf assessment in south-central Italy. *Ecology and Evolution* **14**, e11285. (doi:10.1002/ece3.11285)

34. Harmoinen J *et al.* 2021 Reliable wolf-dog hybrid detection in Europe using a reduced SNP panel developed for non-invasively collected samples. *BMC Genomics* **22**, 473. (doi:10.1186/s12864-021-07761-5)

35. Randi E, Lucchini V. 2002 Detecting rare introgression of domestic dog genes into wild wolf (Canis lupus) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics* **3**, 29–43. (doi:10.1023/A:1014229610646)

36. Verardi A, Lucchini V, Randi E. 2006 Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molecular Ecology* **15**, 2845–2855. (doi:10.1111/j.1365-294X.2006.02995.x)

37. Randi E *et al.* 2014 Multilocus Detection of Wolf x Dog Hybridization in Italy, and Guidelines for Marker Selection. *PLoS ONE* **9**, e86409. (doi:10.1371/journal.pone.0086409)

38. Pilot M, Greco C, vonHoldt BM, Randi E, Jędrzejewski W, Sidorovich VE, Konopiński MK, Ostrander EA, Wayne RK. 2018 Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids. *Evolutionary Applications* **11**, 662–680. (doi:10.1111/eva.12595)

39. Salvatori V, Godinho R, Braschi C, Boitani L, Ciucci P. 2019 High levels of recent wolf × dog introgressive hybridization in agricultural landscapes of central Italy. *Eur J Wildl Res* **65**, 73. (doi:10.1007/s10344-019-1313-3)

40. Salvatori V *et al.* 2020 European agreements for nature conservation need to explicitly address wolf-dog hybridisation. *Biological Conservation* **248**, 108525. (doi:10.1016/j.biocon.2020.108525)

41. Santostasi NL, Gimenez O, Caniglia R, Fabbri E, Molinari L, Reggioni W, Ciucci P. 2021 Estimating Admixture at the Population Scale: Taking Imperfect Detectability and Uncertainty in Hybrid Classification Seriously. *J Wildl Manag* **85**, 1031–1046. (doi:10.1002/jwmg.22038)

42. Stronen AV *et al.* 2022 A reduced SNP panel to trace gene flow across southern European wolf populations and detect hybridization with other Canis taxa. *Sci Rep* **12**, 4195. (doi:10.1038/s41598-022-08132-0)

43. Galaverni M, Caniglia R, Pagani L, Fabbri E, Boattini A, Randi E. 2017 Disentangling Timing of Admixture, Patterns of Introgression, and Phenotypic Indicators in a Hybridizing Wolf Population. *Molecular Biology and Evolution* **34**, 2324–2339. (doi:10.1093/molbev/msx169)

44. Caniglia R, Fabbri E, Greco C, Galaverni M, Manghi L, Boitani L, Sforzi A, Randi E. 2013 Black coats in an admixed wolf × dog pack is melanism an indicator of hybridization in wolves? *Eur J Wildl Res* **59**, 543–555. (doi:10.1007/s10344-013-0703-1)

45. Silva P, Galaverni M, Ortega-Del Vecchyo D, Fan Z, Caniglia R, Fabbri E, Randi E, Wayne R, Godinho R. 2020 Genomic evidence for the Old divergence of Southern European wolf populations. *Proc. R. Soc. B.* **287**, 20201206. (doi:10.1098/rspb.2020.1206)

46. Peterson RO, Page RE. 1988 The Rise and Fall of Isle Royale Wolves, 1975-1986. *Journal of Mammalogy* **69**, 89–99. (doi:10.2307/1381751)

47. Vilà C *et al.* 2003 Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proc. R. Soc. Lond. B* **270**, 91–97. (doi:10.1098/rspb.2002.2184)

48. Kardos M *et al.* 2017 Genomic consequences of intensive inbreeding in an isolated wolf population. *Nat Ecol Evol* **2**, 124–131. (doi:10.1038/s41559-017-0375-4)

49. Robinson JA, Räikkönen J, Vucetich LM, Vucetich JA, Peterson RO, Lohmueller KE, Wayne RK. 2019 Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci. Adv.* **5**, eaau0757. (doi:10.1126/sciadv.aau0757)

50. Smeds L, Ellegren H. 2023 From high masked to high realized genetic load in inbred Scandinavian wolves. *Molecular Ecology* **32**, 1567–1580. (doi:10.1111/mec.16802)

51. Mak SST *et al.* 2017 Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. *GigaScience* **6**. (doi:10.1093/gigascience/gix049)

52. Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP. 2018 Single-tube library preparation for degraded DNA. *Methods Ecol Evol* **9**, 410–419. (doi:10.1111/2041-210X.12871)

53. Ciucani MM *et al.* 2023 The extinct Sicilian wolf shows a complex history of isolation and admixture with ancient dogs. *iScience* **26**, 107307. (doi:10.1016/j.isci.2023.107307)

54. Andrews S. 2010 FastQC: A Quality Control Tool for High Throughput Sequence Data. See http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

55. Ewels P, Magnusson M, Lundin S, Käller M. 2016 MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048. (doi:10.1093/bioinformatics/btw354)

56. Gopalakrishnan S *et al.* 2017 The wolf reference genome sequence (Canis lupus lupus) and its implications for Canis spp. population genomics. *BMC Genomics* **18**, 495. (doi:10.1186/s12864-017-3883-3)

57. Lindblad-Toh K *et al.* 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819. (doi:10.1038/nature04338)

58. Schubert M *et al.* 2014 Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat Protoc* **9**, 1056–1082. (doi:10.1038/nprot.2014.063)

59. Danecek P *et al.* 2021 Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008. (doi:10.1093/gigascience/giab008)

60. Lindgreen S. 2012 AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* **5**, 337. (doi:10.1186/1756-0500-5-337)

61. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760. (doi:10.1093/bioinformatics/btp324)

62. Van Der Auwera GA *et al.* 2013 From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *CP in Bioinformatics* **43**. (doi:10.1002/0471250953.bi1110s43)

63. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)

64. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015 Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaSci* **4**, 7. (doi:10.1186/s13742-015-0047-8)

65. Hanghøj K, Moltke I, Andersen PA, Manica A, Korneliussen TS. 2019 Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. *GigaScience* **8**. (doi:10.1093/gigascience/giz034)

66. Waples RK, Albrechtsen A, Moltke I. 2019 Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. *Molecular Ecology* **28**, 35–48. (doi:10.1111/mec.14954)

67. Wickham H. 2016 ggplot2: Elegant Graphics for Data Analysis.

68. Alexander DH, Novembre J, Lange K. 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. (doi:10.1101/gr.094052.109)

69. Stefanović M *et al.* 2024 Range-wide phylogeography of the golden jackals (Canis aureus) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. *Biological Conservation* **290**, 110448. (doi:10.1016/j.biocon.2024.110448)

70. Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356. (doi:10.1186/s12859-014-0356-4)

71. Kardos M, Waples RS. 2024 Low-coverage sequencing and Wahlund effect severely bias estimates of inbreeding, heterozygosity and effective population size in North American wolves. *Molecular Ecology* , e17415. (doi:10.1111/mec.17415)

72. Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A. 2020 Recent Demographic History Inferred by High-Resolution Analysis of Linkage Disequilibrium. *Molecular Biology and Evolution* **37**, 3642–3653. (doi:10.1093/molbev/msaa169)

73. Auton A *et al.* 2013 Genetic Recombination Is Targeted towards Gene Promoter Regions in Dogs. *PLoS Genet* **9**, e1003984. (doi:10.1371/journal.pgen.1003984)

74. Gray MM, Granka JM, Bustamante CD, Sutter NB, Boyko AR, Zhu L, Ostrander EA, Wayne RK. 2009 Linkage Disequilibrium and Demographic History of Wild and Domestic Canids. *Genetics* **181**, 1493–1505. (doi:10.1534/genetics.108.098830)

75. Mech LD, Barber-Meyer SM, Erb J. 2016 Wolf (Canis lupus) Generation Time and Proportion of Current Breeding Females by Age. *PLoS ONE* **11**, e0156682. (doi:10.1371/journal.pone.0156682)

76. Santiago E, Caballero A, Köpke C, Novo I. 2024 Estimation of the contemporary effective population size from SNP data while accounting for mating structure. *Molecular Ecology Resources* **24**, e13890. (doi:10.1111/1755-0998.13890)

77. Waples RS, Anderson EC. 2017 Purging putative siblings from population genetic data sets: a cautionary view. *Molecular Ecology* **26**, 1211–1224. (doi:10.1111/mec.14022)

78. Waples RS. 2024 Practical application of the linkage disequilibrium method for estimating contemporary effective population size: A review. *Molecular Ecology Resources* **24**, e13879. (doi:10.1111/1755-0998.13879)

79. Sánchez-Barreiro F *et al.* 2021 Historical population declines prompted significant genomic erosion in the northern and southern white rhinoceros (*Ceratotherium simum*). *Molecular Ecology* **30**, 6355–6369. (doi:10.1111/mec.16043)

80. Hasselgren M, Dussex N, Von Seth J, Angerbjörn A, Olsen R, Dalén L, Norén K. 2021 Genomic and fitness consequences of inbreeding in an endangered carnivore. *Molecular Ecology* **30**, 2790–2799. (doi:10.1111/mec.15943)

81. Ceballos FC, Hazelhurst S, Ramsay M. 2018 Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics* **19**, 106. (doi:10.1186/s12864-018-4489-0)

82. Thompson EA. 2013 Identity by Descent: Variation in Meiosis, Across Genomes, and in Populations. *Genetics* **194**, 301–326. (doi:10.1534/genetics.112.148825)

83. Keightley PD, Jackson BC. 2018 Inferring the Probability of the Derived *vs.* the Ancestral Allelic State at a Polymorphic Site. *Genetics* **209**, 897–906. (doi:10.1534/genetics.118.301120)

84. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w ^{1118} ; iso-2; iso-3. *Fly* **6**, 80–92. (doi:10.4161/fly.19695)

85. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. 2016 SIFT missense predictions for genomes. *Nat Protoc* **11**, 1–9. (doi:10.1038/nprot.2015.123)

86. Hoffmann AA, Sgrò CM, Kristensen TN. 2017 Revisiting Adaptive Potential, Population Size, and Conservation. *Trends in Ecology & Evolution* **32**, 506–517. (doi:10.1016/j.tree.2017.03.012)

87. Abascal F *et al.* 2016 Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol* **17**, 251. (doi:10.1186/s13059-016-1090-1)

88. Abadía-Cardoso A, Freimer NB, Deiner K, Garza JC. 2017 Molecular Population Genetics of the Northern Elephant Seal Mirounga angustirostris. *Journal of Heredity* **108**, 618–627. (doi:10.1093/jhered/esx053)

89. Ochoa A, Onorato DP, Roelke-Parker ME, Culver M, Fitak RR. 2022 Give and take: Effects of genetic admixture on mutation load in endangered Florida panthers. *Journal of Heredity* **113**, 491–499. (doi:10.1093/jhered/esac037)

90. Adams NE, Edmands S. 2023 Genomic recovery lags behind demographic recovery in bottlenecked populations of the Channel Island fox, *Urocyon littoralis*. *Molecular Ecology* **32**, 4151–4164. (doi:10.1111/mec.17025)

91. Jan M *et al.* 2023 Wolf genetic diversity compared across Europe using the yardstick method. *Sci Rep* **13**, 13727. (doi:10.1038/s41598-023-40834-x)

92. Cagnolaro L. 1974 *Inchiesta sulla distribuzione del lupo (Canis lupus L.) in Italia e nei Cantoni Ticino e Grigioni (Svizzera)*. See https://books.google.it/books/about/Inchiesta_sulla_distribuzione_del_lupo_C.html?id=TDwMHAAACAAJ&redir_esc=y.

93. Galaverni M, Caniglia R, Fabbri E, Milanesi P, Randi E. 2016 One, no one, or one hundred thousand: how many wolves are there currently in Italy? *Mamm Res* **61**, 13–24. (doi:10.1007/s13364-015-0247-8)

94. Nunney L. 1993 THE INFLUENCE OF MATING SYSTEM AND OVERLAPPING GENERATIONS ON EFFECTIVE POPULATION SIZE. *Evolution* **47**, 1329–1341. (doi:10.1111/j.1558-5646.1993.tb02158.x)

95. Waples RS. 2016 Making sense of genetic estimates of effective population size. *Molecular Ecology* **25**, 4689–4691. (doi:10.1111/mec.13814)

96. Salado I, Preick M, Lupiáñez-Corpas N, Fernández-Gil A, Vilà C, Hofreiter M, Leonard JA. 2023 Large variance in inbreeding within the Iberian wolf population. *Journal of Heredity* , esad071. (doi:10.1093/jhered/esad071)

97. Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF. 2018 Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* **19**, 220–234. (doi:10.1038/nrg.2017.109)

98. Stoffel MA, Johnston SE, Pilkington JG, Pemberton JM. 2021 Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nat Commun* **12**, 2972. (doi:10.1038/s41467-021-23222-9)

99. Van Oosterhout C. 2020 Mutation load is the spectre of species conservation. *Nat Ecol Evol* **4**, 1004–1006. (doi:10.1038/s41559-020-1204-8)

100. vonHoldt BM *et al.* 2023 Demographic history shapes North American gray wolf genomic diversity and informs species' conservation. *Molecular Ecology* , mec.17231. (doi:10.1111/mec.17231)

101. Brüniche-Olsen A, Kellner KF, Anderson CJ, DeWoody JA. 2018 Runs of homozygosity have utility in mammalian conservation and evolutionary studies. *Conserv Genet* **19**, 1295–1307. (doi:10.1007/s10592-018-1099-y)

102. Garner BA, Hoban S, Luikart G. 2020 IUCN Red List and the value of integrating genetics. *Conserv Genet* **21**, 795–801. (doi:10.1007/s10592-020-01301-6)

103. Thurfjell H, Laikre L, Ekblom R, Hoban S, Sjögren-Gulve P. 2022 Practical application of indicators for genetic diversity in CBD post-2020 global biodiversity framework implementation. *Ecological Indicators* **142**, 109167. (doi:10.1016/j.ecolind.2022.109167)

104. Hoban S *et al.* 2024 Too simple, too complex, or just right? Advantages, challenges, and guidance for indicators of genetic diversity. *BioScience* **74**, 269–280. (doi:10.1093/biosci/biae006)

FIGURES & TABLES CAPTIONS

Fig. 1 – ADMIXTURE clustering analyses showing: (A) unsupervised ADMIXTURE results for modern dogs and each studied wolf population, and potential admixed individuals, assuming 4 ancestry components (K=4); (B) supervised ADMIXTURE results conducted to validate potentially admixed WIT (HYIT = Italian wolf-dog hybrid) using non-admixed WIT individuals, and non-admixed DOG individuals as populations with known ancestry. Vertical bars correspond to different samples, and the colors represent the estimated ancestry components and proportions.

Fig. 2 – (A) N_e estimates across 20 replicated datasets up to 50 generations ago, each one representing a random sample of approximately 50,000 SNPs per chromosome. Lines represent geometric mean values across replicates, shades are 95% confidence intervals. (B) N_e estimates zoomed for the past 10 generations. (C) Number of ROH (NROH) ('y' axis) compared to the sum of the length of ROH (SROH) ('x' axis) across the autosomes. Each individual is represented by a colored shape, while meanings of colors and shapes are specified in the legend on the right. The black dotted line represents the intercept, with a slope determined by the ratio of the difference in the range of the NROH variable to the difference in the range of the SROH variable. (D) ROH coalescence time up to 50 generations ago in the Italian wolf population WIT), expressed as a function of ROH length ('y' axis) and generations before present ('x' axis). The blue dashed line represents the generation window when the bottleneck occurred, while the red dashed line represents the 5 generations immediately after the bottleneck. The dotted gray line and bar represent the ROH size generated during the aforementioned generation window. 27.14% of SROH belongs to ROHs generated during this time frame.

Fig. 3 – Barplot representing observed heterozygosity estimates (Ho), on the left, and SROH estimates, on the right, for Italian wolves (WIT; in green), Scandinavian wolves (WSC; in purple), and North American wolves from Isle Royale (WUS; in orange). The SROH of different ROH sizes is highlighted with a gray to black palette.

Fig. 4 – Boxplots representing genetic load estimates based on genotype counts ('x' axis) of heterozygotes ('masked load') and homozygous genotypes for deleterious alleles ('realized load') for different degrees impact variants (LOW = low impact variants, likely not deleterious; MOD-TOL = moderate but tolerated impact variants, likely not deleterious; MOD-DEL = moderately deleterious impact variants, likely to modify protein effectiveness; HIGH = high impact variants, likely to disrupt protein function) in each wolf population ('y' axis). The 'total load' represents the sum of 'masked' and 'realized' load.

Figure 1

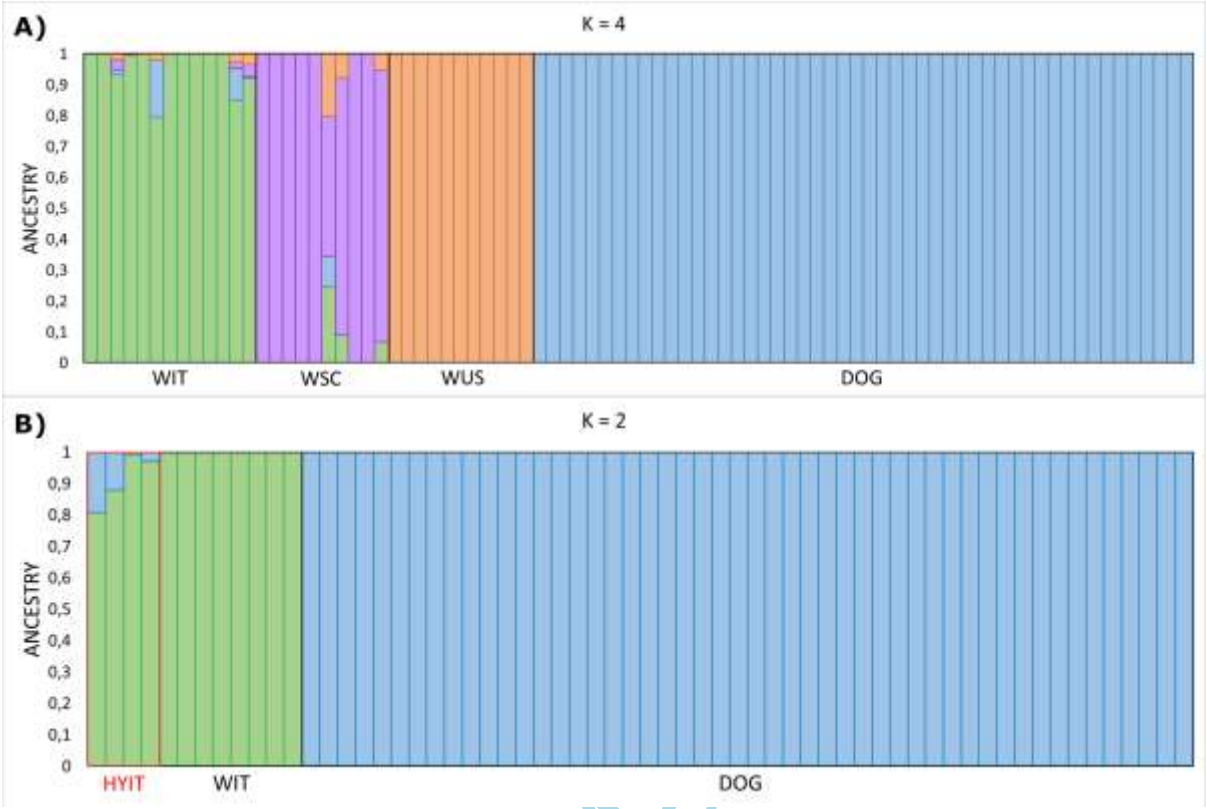


Figure 2

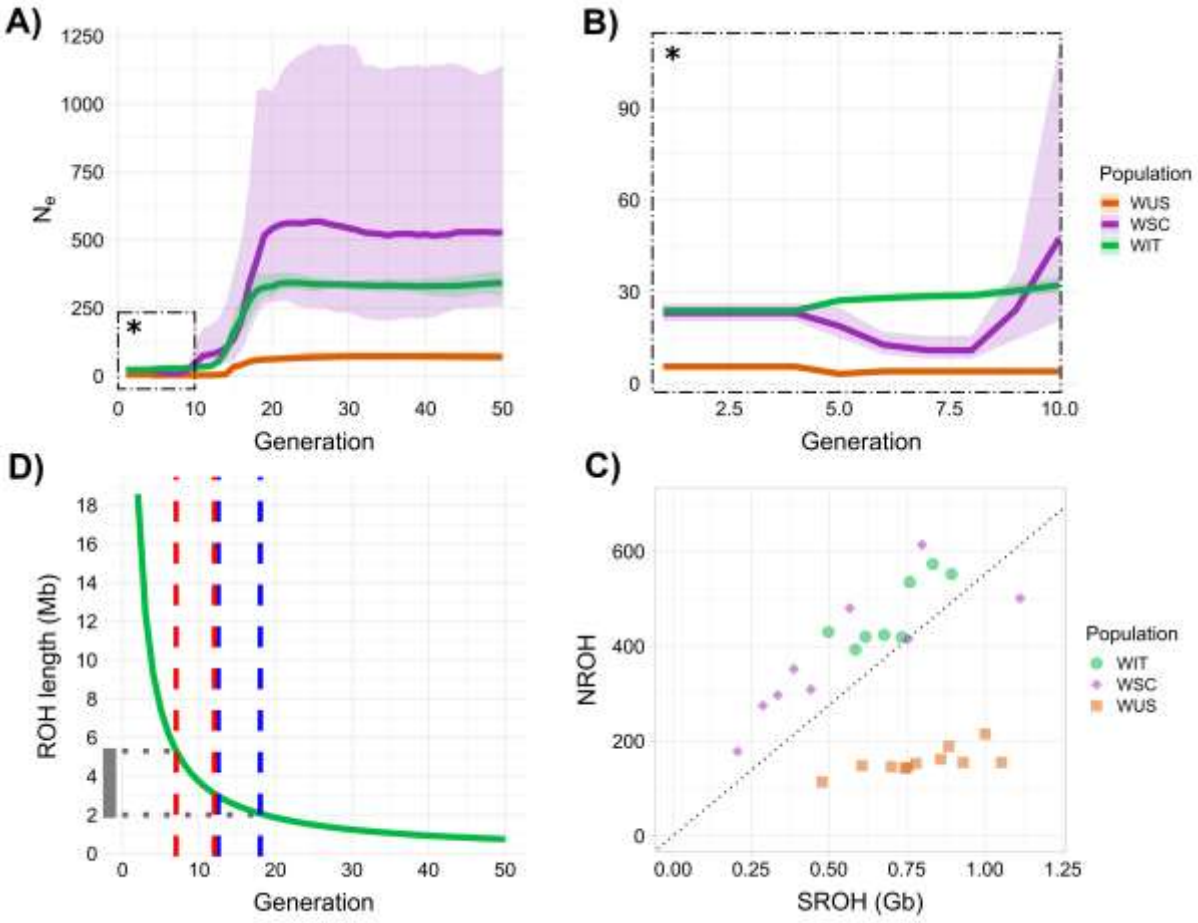


Figure 3

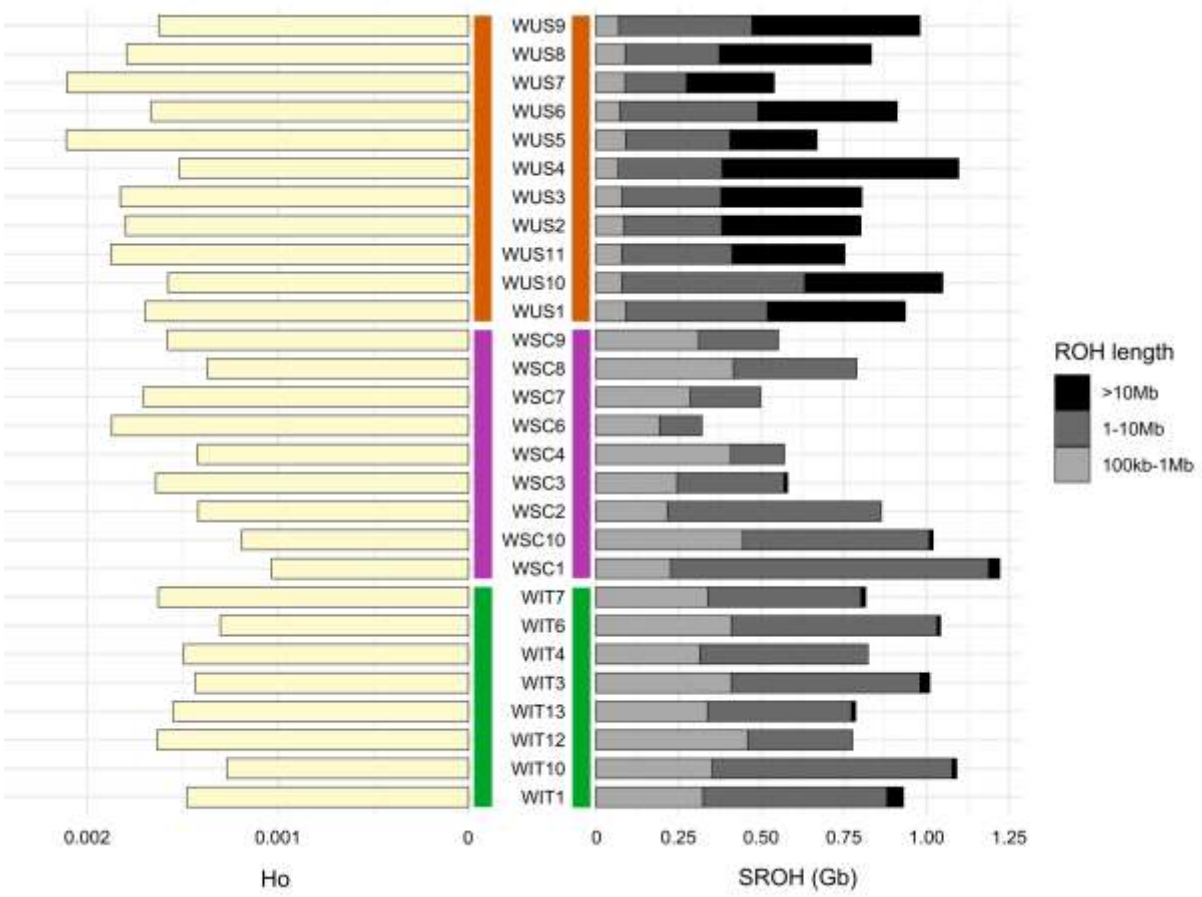


Figure 4