Characterisation of the historic demographic decline of the British European polecat population

R. Shaw [1,2], J. MacPherson [3], A, C. Kitchener[4], G, J. Etherington[1], W. Haerty [1,2]

1 Earlham Institute, Norwich Research Park, Colney Lane, Norwich, NR4 7UZ
rebecca.shaw@earlham.ac.uk

2 University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ

3 Vincent Wildlife Trust, Ledbury, Herefordshire, HR8 1EP

4 Department of Natural Sciences, National Museums Scotland, Edinburgh EH1 1JF, UKUK and School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

Keywords: population genomics, demographic history, range expansion

Abstract

The European polecat (*Mustela putorius*) has a widespread distribution across many countries of mainland Europe, but is documented to be declining within these ranges. In Britain, direct persecution led to a severe decline of the polecat population during the 19th century. Unlike European mainland populations, it is now recovering across much of its former range. The genomic and conservation implications of such a severe demographic decline, followed by the current recovery, have still to be characterised in the European polecat in Britain. Here we carry out population-level whole-genome analyses of 65 polecats from Britain (Wales, England) and the European mainland. Welsh polecats show a unique genetic structure compared to English and European populations. We also reconstructed the demographic history of the Welsh polecat to quantify the magnitude of the bottleneck. Our analyses confirmed the drastic decline of the Welsh polecat's effective population size, with a severe genetic bottleneck around 40 generations ago (1854-1872). We investigated whether genomic signatures reflected this demographic event and found that Welsh polecats had significantly less genetic diversity than English ones, but not European polecats. Runs of homozygosity and the genetic load present in Welsh polecat genomes also indicated recent historic inbreeding. We show that the increase of the British polecat population could be due to admixture events, but also that the Welsh polecat serves as a distinct genetic population, which could be crucial for the overall conservation of European polecats.

1.0 Introduction

Human-mediated actions are considered some of the most serious threats to the survival of small carnivore populations. This includes active persecution and habitat fragmentation reducing connectivity between individuals, thereby limiting gene flow and further contributing to reducing the effective population size (N_e). The ensuing decline in genetic diversity further decreases future resilience to arising challenges including environmental changes. Inbreeding depression is another consequence of a population reduction and can limit the species' ability to purge deleterious mutations, increasing their probability of fixation. These weakly deleterious mutations can negatively affect the population and lead to an increased risk of extinction.

Similarly, populations that have recently undergone rapid expansion often display genetic signatures of low effective population size, partly due to allele surfing, where certain gene variants increase in frequency among individuals at the forefront of the expansion. Likewise, the 'founder effect' can leave a distinct genetic imprint by reducing genetic diversity, often resulting in greater genetic differentiation between populations. Our understanding of these signals is often attributed to more historic (post-glacial) range expansions, or by contemporary (often human-mediated) introductions of species. However, with the increasing impact of climate change affecting land use, wild populations are more likely to be displaced (through human-mediated actions like persecution) and face reductions in their natural habitat range and loss of connectivity (Beyer & Manica, 2020). Therefore, there is a need to investigate these signals and understand how populations adapt to and recover from environmental change and the associated genetic forces that play a consequential role. The temporal fluctuations of genomic changes in a recovering population undergoing range expansion are also important to capture, in order to ensure populations do not remain genetically fragmented (Thomas et al., 2022).

The European polecat (*Mustela putorius*) is a medium-sized carnivore and belongs to the family Mustelidae, which is the largest family in the order Carnivora. The Carnivora displays a diverse adaptive radiation, occupying a wide range of niches within different climates and habitats. The polecat inhabits a wide geographical range across mainland Europe, but in several of these countries their populations are listed or suspected to be declining (Croose et al., 2018). Habitat fragmentation, poaching, poisoning, and persecution are the main threats to the polecat in continental Europe. Hybridisation has been known to occur between the European polecat and other closely related mustelid species, such as the Steppe polecat (*Mustela eversmanii*) and the severely endangered European mink (*Mustela lutreola*) (Cabria et al., 2011; Cserkész et al., 2021).

Polecat population in Britain

In Britain, the European polecat has undergone an extensive population reduction and range contraction, followed by a subsequent rapid range and population expansion over the last 70 years. Although the species is currently protected under both the Wildlife and Countryside Act, 1981 and listed as a priority species on the UK Biodiversity Action Plan (BAP) (BRIG, 2007) since 2007, it has experienced a severe decline in numbers since the late 19th century as the species was persecuted heavily by gamekeepers as pests impacting gamebird populations. This led to a contraction in the species range, with small pockets of populations in Cumbria and Scotland persisting until the early 20th century and a stronghold in central, west Wales and the neighbouring English border counties of Herefordshire and Shropshire (Harris et al., 1995; Langley & Yalden, 1977).

The existing stronghold in Wales started to recover in population size after the end of the First World War, and the banning of gin traps coincided with the subsequent recovery of rabbit populations in the post-myxomatosis period (Langley & Yalden, 1977). As the polecat population began to expand eastwards across its former range, it came into contact with populations of feral domestic ferrets (*Mustela furo*). Ferrets are thought to be derived from the European polecat through domestication that occurred over 2000 years ago. Ferrets have long been used in Britain for hunting purposes and are known to have been released into the wild and accidentally introduced. Owing to the short evolutionary distance between the two species, the European polecat and the domestic ferret readily hybridise and this has been well-documented to occur in Britain (Costa et al., 2013; Etherington et al., 2022).

Previous work on the British population of polecats has focussed on assessing the genome introgression between European polecats and feral domestic ferrets (Etherington et al., 2022) or has aimed to model their demographic history using microsatellites or mitochondrial genes (Costa et al., 2013). Here, we characterise the demographic history of the British European polecat population using whole genome resequencing. We also mine the genome for signatures of the population bottleneck, documenting what impact the decline and rapid expansion had on the population and assessing the conservation implications for British polecats in comparison to mainland European polecat populations.

2.0 Methods

2.1 Sample Collection

For this study, 31 roadkill British polecat samples were resourced from the Vincent Wildlife Trust 2014-2015 survey (Croose, 2016) (Figure 1b). Carcasses were collected by members of the public who were requested to freeze the carcass where possible. The Centre for Ecology and Hydrology (CEH) sent a pre-paid postage box with appropriate packaging material to the recorder who then posted the carcass to CEH via next-day delivery, and the carcasses were frozen upon receipt. We further increased our sample size through the inclusion of publicly available data consisting of eight domestic ferrets (*M. furo*) genomes, 19 British polecats, and 15 European mainland polecat samples (Etherington et al., 2022). Samples were selected based on geographical location from areas around Wales and from a trajectory across England from west to east to capture any diversity across the predicted range expansion of the polecat in Britain (Figure 1c), (Supplementary material S2). The home ranges of polecats in Britain are known to vary seasonally, and male polecats tend to disperse further than females (Birks & Kitchener, 1999), with males travelling as far as 8 km between dens (Blandford, 1987). To avoid incorporating familial bias into our selection through potentially selecting related individuals, we selected individual carcasses that were located greater than 10 km apart.

2.2 DNA Extraction

DNA extraction from tissue samples was carried out using a Qiagen™ DNEasy Blood & Tissue kit according to the manufacturer's protocol. Samples were quality checked using a Nanodrop Spectrophore and DNA quantity was checked using a Qubit Fluorometer. Samples with a low yield 260/280- and 260/230-absorbance ratios outside the desired range (260/280 <1.80, 260/230 <2.0) were further cleaned up using Beckman Coulter™ AMPureXP Beads. A total of 31 samples were selected for library preparation and sent for sequencing. Figure 1C shows the distribution of the samples sent for sequencing and the location of British polecats already sequenced.

2.3 Library preparation & sequencing

Samples were sequenced by The Earlham Institute Genomic Pipelines with LITE library preparation (Perez-Sepulveda et al., 2021) and sequenced using an Illumina NovaSeq 6000 S generating paired-end (2 x 150bp) reads. Samples were run on one lane with a predicted coverage of 10x.

2.4 Sequencing QC, Read mapping & SNP calling

Raw read and sequence quality for all British and European samples were first assessed using fastQC (version 0.11.9) (Andrews, 2010), and any reads below the threshold (Q20) were trimmed using TRIMMOMATIC (version 0.39) (Bolger et al., 2014). Raw reads were mapped to the domestic ferret genome (**GCA_920103865.1**) using BWA MEM (version 0.7.17) (Li & Durbin, 2009). Duplicate reads were then marked using PicardTools (version 2.1.1) ("Picard Toolkit," 2019). Single nucleotide polymorphisms (SNPs) were then called using GATK (McKenna et al., 2010) (version 4.5.0) following the best practice workflow for SNP discovery (van der Auwera & O'Connor, 2020). Briefly, BaseRecalibrator was run on each sample bam file to detect and correct patterns of systematic errors in the base quality scores. The recalibration metric was then applied to each sample bam file, and HaplotypeCaller was run for each sample before consolidating all of them together using GATK GenomicsDBImport. For downstream analysis, BCFtools (Danecek et al., 2021)(version 1.15.1) was used to select only biallelic SNPs with a quality score (QUAL) of at least Q20. A further step of pruning variants over a linkage disequilibrium (LD) threshold of 0.6 was done using the *bcftools prune* command. The resulting SNP dataset was used for all downstream analyses unless stated.

2.5 Relatedness

We assessed relatedness between samples using the KING algorithm in PLINK (version 1.9) (Chang et al., 2015; Purcell et al., 2007). KING kinship coefficients determine first-degree relations (parent-child, full siblings) as values between 0.17-0.35, second-degree relations correspond to 0.08-0.17 and third-degree relations as between 0.04-0.08. One pairing between S17 and S0012_VWT665 had a kinship coefficient of 0.07, suggesting a third-degree level of relatedness. Both of these individuals were English polecats but were found in very different geographical locations (~291km apart). We removed S17 from all of the downstream analysis as this sample was found in Waterrow, England, and was therefore outside of the west-to-east trajectory that we were hypothesising was the route of the Welsh polecat expansion.

2.6 Population Structure

We first generated a principal component analysis (PCA) using the whole-genome SNP dataset of 65 polecats and 8 domestic ferrets in PLINK to determine the British population structure. For these analyses SNPs with a minor allele count of at least 3 and sites with less than 50% missing sites were extracted. A Spearman correlation test was implemented in R, using values from PC1 and sample longitude coordinates, to ascertain whether population structure in Britain could be correlated with geographical location. We used ADMIXTURE (version 1.3) (Alexander et al., 2009) to perform ancestry runs on the SNP dataset for *K* values 1-5, and a CV error was used to compare all *K* values. Pairwise *Fst* values between populations were computed using VCFtools (version 0.1.16) (Danecek et al., 2011).

To further detect structure within and between populations, we used RAXML-ng (version 0.9.0) (Stamatakis, 2014) to infer a phylogenetic tree from our whole-genome SNP dataset. The phylogeny was generated using the GTRGAMMA model of evolution with the Lewis bias ascertainment correction and 100 rapid bootstraps.

2.7 Genomic Diversity

To assess the genetic diversity between populations, we used VCFtools to detect nucleotide diversity in 20 Mb windows across all samples. We tested for significant differences in the genetic diversity estimates between populations by using an analysis of variance test (ANOVA) implemented in R (version 4.1.1). SNPs that were private to one population, or that were shared across populations, were identified using VCFtools. The observed proportion of heterozygous sites and the inbreeding coefficient (f) for each individual per population were also calculated using VCFtools.

2.8 Detection of historic bottleneck: Runs of homozygosity and identification of deleterious mutations

To detect signatures associated with a genetic bottleneck, we calculated the distribution of runs of homozygosity (ROH) using PLINK. Since calling ROH can be sensitive to low sequencing depth in areas of the genome, which can introduce genotype errors (Ceballos et al., 2018), we only calculated ROH for a subsample of six individuals per population that had an average read depth of 5.0 or higher. We then categorised ROH into three length depths; 300-500kb, 500 kb-1Mb, and >1Mb.

SnpEff (version 5.1) (Cingolani et al., 2012) was used to detect the functional impacts of the biallelic SNPs on the genome of the European polecat. SnpEff compares a set of variants to a known annotation database and compares the derived alleles within genic regions to predict the impact of that variant. In this case, we used the domestic ferret annotation as the input database. We assessed whether mutational load differed between populations and this was calculated based on homozygous and heterozygous mutational load within missense and loss-of-function (LoF) categories.

2.9 Historic effective population size

Recent historic effective population size (Ne) was estimated using GONE (Santiago et al., 2020). GONE estimates recent effective population size by calculating LD between pairs of SNPs and finds the Ne that best fits the observed LD spectrum. GONE assumes that the population being tested is closed, with no recent admixture to allow for no biased estimates in the calculation of effective population size. Therefore, we ran GONE for Welsh, English, and European populations separately, removing an individual (S0024_VWT1661) from the Welsh population due to apparent recent hybridisation identified by PCA and ADMIXTURE analyses. For each population, we simulated 2000 generations and calculated 400 bins. GONE takes up to 200 chromosomes as input data, so we removed scaffolds that overlapped with the annotated X chromosome of *Mustela nigripes* and used the largest 100 autosomal scaffolds as input, which constitute 78% of the whole genome assembly. GONE calculated LD for up to 10,000 SNPs per scaffold. Each simulation was run 100 times, with a new subsample of 10,000 SNPs for each run. For each population, we plotted the mean Ne for the 100 runs.

3.0 Results

3.1 Whole Genome Resequencing

We generated whole-genome sequence data for 31 European polecats from Wales and England. Samples had been chosen from the location of the known refugium (Figure 1a), across a trajectory of the historic range expansion following the population's recovery from a genetic bottleneck. These were combined with an already published dataset of 34 European polecats from Italy (5), Germany (2), Austria (3), France (3), Spain (2), Wales (3) and England (16), along with 8 domestic ferrets. Sequence reads were aligned to the domestic Ferret genome. Higher-coverage samples were downsampled to account for coverage biases, and SNPs were then called after downsampling (Supplementary material S1). We identified a total of 2,941,545 SNPs.

3.2 Population Structure Using Genomic Data

Principal component analysis (PCA) of whole-genome SNPs revealed distinct clustering of samples of polecats from mainland Europe, Britain, and domestic ferrets (Figure 2a). PCs 1 (26%) & 2 (15.8%) revealed the English and Welsh samples clustered together and away from mainland European samples. To assess structure within the British population, we repeated the PCA only on the 49 British samples, which revealed a geographical separation of samples from West (Welsh) and East (English) (Figure 2b). Longitudinal coordinates strongly correlated with PC1 ($r_s = -0.754$, $p = <0.01$). Outliers to this relationship that clustered with the Eastern individuals are from northwestern and southern England, suggesting that Welsh samples were distinct from English samples.

ADMIXTURE analyses were best supported by the cross-validation error of *K=3*, indicating the existence of three genetically distinct populations across all of the sampled populations; domestic ferret, British (encompassing English and Welsh populations) and European polecats, thus resolving the British population as a homogenous group distinct from the European mainland population (Figure 2c). Under these parameters, the Welsh population demonstrated a lower degree of admixture with the domestic ferret than English individuals (Figure 2c) (Supplementary material S3), with the exception of one individual, which is likely a polecat-ferret hybrid (S0024_VWT1661). These results suggest repeated hybridisation and continuing introgression between English polecats and domestic ferrets.

Pairwise *Fst* values were computed to detect the genetic distinction between populations. *Fst* values can range from 0-1, with 0 indicating complete sharing of genetic material and a value of 1 indicating no genetic sharing. Our results showed a comparison of English and Welsh populations having an *Fst* value of 0.03, compared to 0.114 between English and European and 0.148 between Welsh and European (Supplementary data figure 1). *Fst* values were not significantly different between population comparisons ($H = 1.25$, $df = 2$, $p = 0.53$).

3.3 Genomic Diversity in Britain in Comparison to Mainland Europe

As the British populations have undergone a severe population contraction prior to their current recovery, we investigated the impact of these demographic events on the genetic diversity of the different populations. European populations demonstrated a higher number of private SNPs (2,096,516) in comparison to English (513,454) and Welsh (1,029,620)

(Figure 3a). This could be attributed to population structure and also demonstrates the genetic distinction between British and European mainland polecats. Welsh individuals showed a higher number of private SNPs in comparison to English individuals, demonstrating unique genetic diversity that still resides within the stronghold population.

Nucleotide diversity estimates (Tajima's π) (Supplementary material S4) revealed that the English population had significantly higher diversity when compared to the Welsh population and European population (H = 7.15, df = 2, p = 0.03). However, there was no significant difference between the Welsh and European populations (Figure 3b). The English population also showed a greater proportion of heterozygous sites in comparison to both Welsh and European populations ($F_{2,60}$ = 17.74, p = <0.001) (Figure 3c).

Individuals belonging to the European population showed a higher inbreeding coefficient ($F_{2,60}$ = 17.19, p = <0.001) (Figure 3d), and a lower proportion of heterozygous sites in comparison to Welsh and English populations overall. However, we did not account for population structure in the grouping of individuals due to the total sample size. French and Austrian populations had higher proportions of heterozygous sites and a lower inbreeding coefficient than German, Italian, and Spanish populations.

3.4 Runs of Homozygosity & Deleterious Mutations

To assess whether the Welsh population had undergone prolonged periods of inbreeding due to the bottleneck effect, we investigated the presence of ROH within the polecat genomes. We separated ROH lengths into different classes to visualise short (300-500kb), medium (500-1Mb), and long (>1Mb) ROH across the genomes of individuals belonging to the Welsh, European, and English populations (Figure 4a) (Supplementary material S5). Longer ROH are indicative of recent inbreeding events, with shorter ROH being indicative of historic events. Individuals belonging to the Welsh population demonstrate the lowest number of ROH segments in all length classes. Individuals belonging to the English population demonstrate a greater number of ROH segments (H = 5.80, df = 2, p = 0.05) in comparison to the European and Welsh populations overall (Figure 4b).

We investigated whether the mutation landscape differed among individuals belonging to European, Welsh, and English populations. However, we found that the pattern of functional impacts, as annotated by SnpEff, was similar across populations (Supplementary material S6). Therefore, we decided to look at mutational load across two functional categories as reported by SnpEff: missense and loss of function (LoF). With homozygous mutations in both categories expected to be the most deleterious, we also considered the load of heterozygous mutations.

The English population demonstrated a significantly higher masked load within LoF mutations in comparison to Welsh individuals ($F_{2,52}$ = 3.169, p = 0.03) (Figure 5b). There was no significant difference amongst populations for LoF homozygous mutations (Figure 5a).

Missense homozygous mutations were significantly higher in Welsh individuals ($F_{2,52}$ = 4.367, p = 0.01) than in the English population (Figure 5c), but the English population showed a significantly higher masked load of missense mutations ($F_{2,52}$ = 6.192, p = <0.01) (Figure 5d).

3.4 Historic effective population size (N_e)

We used GONE to estimate the recent historic effective population size for the British population of European polecats. Our estimates using GONE correlate with a demographic decline in population size occurring around 30-40 generations ago - which using an estimate for generation time of the polecat of 4 years (Di Marco et al., 2013), corresponds to the mid-late 19th century (1854-1872) (Figure 6). Recovery of the polecat population in Wales based on GONE estimates fits with previous estimates of 10-12 generations ago (1966-1974) (Supplementary material S7).

4.0 Discussion

Here, we provide the first evidence of a genetic bottleneck having occurred within the British population of polecats. Our analyses indicate that the population drastically decreased in effective population size from the late 1800s and persisted for ca. 100 years. Our analyses also indicate that the population began recovering from 1960-70 onwards. This coincides with a prolonged period of relaxed persecution from gamekeepers, the recovery of lagomorph populations following myxomatosis, the banning of gin traps, and the introduction of protective legislation. Our results demonstrate that the British polecat population is distinct from European mainland populations, and is recovering from the genetic bottleneck. Conservation efforts should focus on increased protection measures and monitoring of the Welsh population, as it represents unique genetic integrity in comparison to admixed, English populations. Our study also highlights the importance of genomic data in being able to detect signatures within populations to fully detect and monitor the success of their recovery.

4.1 Comparison of British to Mainland European populations

Our analyses indicate that the British polecat population is genetically distinct from European mainland populations. This is highlighted in pairwise *Fst* estimates per population where a higher *Fst* value (closer to 1.0) suggests a population has low levels of breeding and therefore is more genetically distinct from populations compared to lower *Fst* values. Our results indicate that English and Welsh populations have a low fixation index (0.03) and are therefore likely to share genetic variation. *Fst* values between English and European (0.114) and European and Welsh populations (0.148) are slightly higher, suggesting a greater degree of genetic distinction between mainland European and British polecats.

This is also evident from our ADMIXTURE analyses and further corroborates the genetic divergence seen between Welsh and European polecats (Etherington et al., 2022). British and European polecats are thought to have been separated for ca. 8kya (Montgomery et al., 2014), since the flooding of the Doggerland bridge after the Last Glacial Maximum (LGM). Several widespread mammal populations that are native to Britain demonstrate genetic distinction due to isolation over time from mainland European populations. For example, Irish and British populations of red fox (*Vulpes vulpes*) demonstrate unique genetic diversity that distinguishes them from mainland European populations (McDevitt et al., 2022). Likewise, native British populations of roe deer (*Capreolous capreolus*) demonstrate genetic differentiation from mainland European populations, most likely due to colonisation and isolation events (de Jong et al., 2020). Populations of British red deer (*Cervus elaphus*) also demonstrate genetic distinction from mainland European populations, but show a slightly lower genetic distance to Iberian populations, potentially suggesting retention of genetic diversity within populations from similar migration waves during postglacial colonisation

(Carranza et al., 2024). Therefore, the genetic distinction we observe between Welsh and European mainland polecats likely results from their recent isolation, mirroring the pattern of genetic divergence seen across other British populations of wide-ranging mammals. These studies highlight the potential for specific conservation programmes to conserve genetic diversity within widespread Palaearctic species.

4.2 Population Structure

Our analyses indicate population structuring based on geographic location within polecats in Britain (Figure 2b). Ancestry estimates also demonstrate a higher incidence of introgression in polecats that are found at the furthest edges of the polecat range from the Welsh, refugium population (Figure 2c). We also observe a similar relationship in our maximum likelihood phylogeny, whereby English individuals that are found at the more eastern and southern edge of the polecat's range are clustered more closely to domestic ferrets than to the Welsh individuals (Supplementary Figure S4). In agreement with (Costa et al., 2013), we find the least introgression of domestic ferret genes within individuals found in Wales and English border counties. This observed pattern fits the narrative that as the polecat population expanded eastwards, individuals hybridised with feral domestic ferrets within those regions. However, the presence of introgressed polecats within these regions may also have been exacerbated by covert reintroductions of individuals from hybrid stocks (Birks & Kitchener, 1999).

4.3 Demographic Analyses

Using our generated whole genome sequence data, it was possible to detect and analyse historic trends in effective population in the British population of the European polecat. The British polecat bottleneck was thought to have occurred between the late 1800s to the early 1900s, with populations showing recovery since the 1970s (Langley & Yalden, 1977). The most recent historic genetic bottleneck was evident in our results using GONE, with a sharp decrease in N_e occurring in the Welsh population around 30-40 generations (120-160 years ago), from the sampling time of 2014-2015. This coincides with the known history of persecution of polecats from gamekeepers, as with other populations of terrestrial carnivorans in Britain, which subsided for a period during and following World War I, and again after legal protections came into place (Langley & Yalden, 1977; Sainsbury et al., 2019). Alongside legal protections, it is thought that increasing rabbit (*Oryctolagus cuniculus*) populations, an important prey for polecats (Sainsbury et al., 2020), which had declined due to myxomatosis and began recovering from the mid-20th century, are also closely linked with the recovery of the polecat (Langley & Yalden, 1977; Sainsbury et al., 2019; Sumption & Flowerdew, 1985).

We observed that historic N_e estimates were higher for the English population than the Welsh one (Figure 6, Supplementary material S8). This could be explained both by the rapid expansion of individuals from Wales to England and by the hybridisation and subsequent introgression of those individuals with populations of feral domestic ferrets. Our outputs also highlight that modelling demographic histories can be impacted by hybridisation and
those individuals with populations of feral domestic ferrets. Our outputs also highlight that modelling demographic histories can be impacted by hybridisation and population structure. Previous work (Costa et al., 2013), using microsatellites in British polecats could not detect any evidence of a genetic bottleneck. This could be due to the lower resolution from the smaller representation of genetic markers used, but could also be due to the admixture from

domestic ferret populations as suggested by the authors. We observe this pattern with the English population having greater levels of introgression than the Welsh population and therefore observe not only a greater historic effective population size but also a change in the timing of the genetic bottleneck when just modelled using the English population. We see that population structure can obscure the recent historic effective population size when trying to model the European population's demographic histories. Modelling these as a whole population is difficult due to the different environmental pressures that each population faces. GONE advises not to model demographic histories with less than 5 individuals, of which only the Italian population was viable from our European samples. However, our model reconstruction demonstrated the unreliability of this with low power and possible population structure(Supplementary figure S5).

Census data in 2018 estimated that the polecat population in Britain was currently 88,300 (Mathews et al., 2018). Our estimates corroborate survey data from the Vincent Wildlife Trust that the Welsh population is expanding, with more 'true polecat' sightings in Welsh counties and border counties such as Herefordshire and Shropshire (Croose, 2016), which are areas that were thought to contribute to the historic stronghold of the polecat (Langley & Yalden, 1977), and these areas also having the individuals with the least admixture (Figure 2c).

4.4 Genomics signatures associated with the population bottleneck

We hypothesised that we would see genomic signatures within the population associated with the prolonged genetic bottleneck period. Our results demonstrate the presence of potentially harmful deleterious mutations within both Welsh and English individuals. It also shows the presence of heterozygous, potentially masked deleterious variants being more prevalent in admixed English individuals, and a higher unmasked load of missense variants in the Welsh population. This is as expected, as periods of inbreeding in Welsh populations would lead to a higher presence of homozygous deleterious variants within the population. English individuals demonstrate additive genetic diversity, due to the introgression of domestic ferret genes. This is also confirmed in the higher nucleotide diversity across the genomes of English individuals. However, no significant difference amongst populations was found in mutations that were categorised as homozygous LoF, suggesting these mutations may have been purged from the Welsh population due to their expansion and purifying selection acting upon them. Further investigation into historic genomic signatures of the Welsh population would be needed to fully resolve this.

Individuals belonging to the English population had greater overall ROH in their genomes in comparison to those of both European and Welsh individuals (Figure 4a). This could be due to the long-term, low effective population size within the Welsh refugium, followed by the expansion of some individuals to low areas of genetic connectivity. It could also be related to the introgression with feral domestic ferrets that have been subjected to years of selective, human-mediated breeding. Longer ROH lengths can indicate recent inbreeding events, as recombination has not yet acted on breaking apart these segments, whereas shorter ROH blocks are indicative of more ancestral inbreeding events (Ceballos et al., 2018). Despite having a lower historic effective population size from GONE estimates, the Welsh population does not show a greater proportion of short ROH in comparison to the other populations. Despite having a lower historic effective population size from GONE estimates, the Welsh

population does not show a greater proportion of short ROH in comparison to the other populations.

European individuals that were chosen for the ROH analysis, due to high coverage across the genome, show a varied ROH landscape. This could be due to population-level pressures persisting in their environment at the time and may indicate the necessity for similar genomic studies in polecat populations in regions of Europe where the population is declining.

4.5 Conservation Implications

Here, we provide the first evidence of the extent to which the population bottleneck experienced by the Welsh polecat population from the late 1800s to the mid-1950s. We demonstrate that the population started to expand when persecution pressures were lessened and after protective legislation came into action, thus supporting legal protection as a conservation management tool.

We present a unique scenario of population recovery in a once heavily persecuted mammal in Britain. The recovery can largely be attributed to protective legislation and the resurgence of rabbit populations following the myxomatosis outbreak, as well as to the introgression with domestic ferrets, which has facilitated the eastern range expansion of the population. Genomic introgression with a domesticated species is not a typical conservation management tool but has been witnessed to assist in the recovery of other species, such as the American bison (*Bison bison*) (Stroupe et al., 2022), and Iberian ibex (*Capra pyrenaica*) (Münger et al., 2024). However, the decline of the wildcat (*Felis silvestris*) population in Scotland has been severely hampered by hybridisation and subsequent introgression with the domestic cat (Howard-McCombe et al., 2023). It could be that due to the relatively recent nature of the hybridisation of the European polecat and domestic ferret in Britain, the harmful consequences are yet to surface. However, (Costa et al., 2013) demonstrated that the lack of F1 hybrids observed within the British population suggests that these hybridisation events are rare. Feral ferret populations are hard to distinguish in mainland Britain but are known to thrive on surrounding islands where there is a lack of or reduction in competition from polecats and other predatory mammals (Kitchener & Birks, 2008). This suggests that continued hybridisation events in wild settings could be extremely rare, but would still pose a threat to the genetic integrity of the recovering population.

Our analyses also demonstrate that the Welsh population holds unique genetic diversity that could be important for the resilience of the population. This could have important conservation implications for reintroducing polecats from Wales to other parts of the country. Additionally, conservation concerns are underscored by the unmasked missense mutational load persisting in the genome as a result of the prolonged low effective population size (Robinson et al., 2023). The monitoring of this elusive species is hard, due to its low population density and nocturnal habits. Monitoring through non-invasive methods, such as the Vincent Wildlife Trust survey, where sightings and suspected roadkill are reported, has proven successful and has helped to document the polecat's recovery within Wales and England. However, these efforts are not matched in some parts of mainland Europe and therefore its status in some countries needs to be clarified. In a study, that aimed to document the status of the polecat across its known range, of the 34 countries where data were found, 20 countries reported that the polecat population was declining (Croose et al., 2018).

Only two countries reported the polecat population as increasing, i.e., Estonia and Britain. Therefore, conservation priorities should be to continue monitoring the Welsh population to protect its genetic legacy and ensure the population's continued expansion. Given the reported declines in other known polecat ranges, conserving the Welsh population is crucial for supporting the global population. Unfortunately, polecats appear to not only be threatened by hybridisation with feral domestic ferrets but road traffic collisions have been recorded as negatively impacting the dispersal of polecats during peak mobility periods (Barg et al., 2022). The focus should be to avoid habitat fragmentation and the frequency of polecats having to disperse across primary roads. It should be noted that the genetic recovery of a population can often lag behind its demographic recovery. In the case of the British polecat population, we do see genetic structuring due to admixture across parts of England. Another British mustelid, the Eurasian Otter (*Lutra lutra*), has subsequently undergone a recent historic genetic bottleneck and is recovering across its previous ranges (du Plessis et al., 2023). However, barriers, such as landscape changes, have potentially contributed to the fragmentation and low genetic connectivity between populations. The authors recommend that continued monitoring be complemented with continued analysis of genetic data, to ensure that populations in Wales do not become genetically isolated due to landscape change and continue to recover both demographically and genetically (Thomas et al., 2022).

This study examined the recovery of polecats across much of Great Britain, however it did not include data on the population in Scotland. According to survey records, it was thought that by 1915, polecat populations were eradicated within Scotland. Covert re-introductions of polecat were thought to have occurred across some areas in southern England, Cumbria and Argyll (Birks & Kitchener, 1999; Davison et al., 1999). Previous work has highlighted the more ‘ferret-like’ ancestry of polecats found within regions of Scotland (Costa et al., 2013), with feral ferret populations being found on offshore Scottish islands. The 2014-15 survey conducted by the Vincent Wildlife Trust recorded only 29 sightings of polecats, representing just 2% of the total records from Britain (Croose 2016). However, some of these sightings occurred in a newly identified distribution area in Dumfries, which may indicate a range expansion of the Cumbrian populations. Additional sampling and genomic analyses would provide further insights into the population structure and conservation status of the remaining population within Scotland.

Acknowledgments
RS was supported by a NERC PhD studentship through the ARIES DTP (NE/S007334/1, 2465312).WH and GE acknowledge the support of the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation; Earlham Institute Strategic Programme Grants Cellular Genomics BBX011070/1, Decoding Biodiversity BBX011089/1, Delivering Sustainable Wheat BBX011003/1 and their constituent work packages (BBS/E/ER/230001A, BBS/E/ER/230001B, BBS/E/ER/230001C, BBS/E/ER/230002A, BBS/E/ER/230002B, BBS/E/ER/230003A, BBS/E/ER/230003C). The author(s) acknowledge support from the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation, Core Capability Grant BB/CCG1720/1 and the National Capability BBS/E/T/000PR9816, and the work delivered via the Scientific Computing group, as well as support for the physical HPC infrastructure and datacentre delivered via the NBI Computing infrastructure for Science (CiS) group. ACK

thanks the Negaunee Foundation for its continued support of a curatorial preparator who sampled many of the polecats used in this study. We are also grateful to Federica Di Palma for her early input to the project.

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Andrews, S. (2010). *FastQC: Quality Control Tool High Throughput Sequence Data*.

Barg, A., MacPherson, J., & Caravaggi, A. (2022). Spatial and temporal trends in western polecat road mortality in Wales. *PeerJ*, *10*(e14291), e14291.

Beyer, R. M., & Manica, A. (2020). Historical and projected future range sizes of the world's mammals, birds, and amphibians. *Nature Communications*, *11*(1), 5633.

Birks, J. D. S., & Kitchener, A. C. (1999). *The Distribution and Status of the Polecat Mustela Putorius in Britain in the 1990s*. Vincent Wildlife Trust.

Blandford, P. R. S. (1987). Biology of the Polecat Mustela putorius: a literature review. *Mammal Review*, *17*(4), 155–198.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120.

BRIG. (2007). *Report on the Species and Habitat Review*. https://data.jncc.gov.uk/data/bdd8ad64-c247-4b69-ab33-19c2e0d63736/UKBAP-Species-HabitatsReview-2007.pdf

Cabria, M. T., Michaux, J. R., Gómez-Moliner, B. J., Skumatov, D., Maran, T., Fournier, P., de Luzuriaga, J. L., & Zardoya, R. (2011). Bayesian analysis of hybridization and introgression between the endangered european mink (Mustela lutreola) and the polecat (Mustela putorius): HYBRID IDENTIFICATION BETWEEN TWO MUSTELID SPECIES. *Molecular Ecology*, *20*(6), 1176–1190.

Carranza, J., Pérez-González, J., Anaya, G., de Jong, M., Broggini, C., Zachos, F. E., McDevitt, A. D., Niedziałkowska, M., Sykut, M., Csányi, S., Bleier, N., Csirke, L., Røed, K., Saint-Andrieux, C., Barboiron, A., Gort-Esteve, A., Ruiz-Olmo, J., Seoane, J. M.,

Godoy, J. A., … Membrillo, A. (2024). Genome-wide SNP assessment of contemporary European red deer genetic structure highlights the distinction of peripheral populations and the main admixture zones in Europe. *Molecular Ecology*, e17508.

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews. Genetics*, *19*(4), 220–234.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80–92.

Costa, M., Fernandes, C., Birks, J. D. S., Kitchener, A. C., Santos-Reis, M., & Bruford, M. W. (2013). The genetic legacy of the 19th-century decline of the British polecat: evidence for extensive introgression from feral ferrets. *Molecular Ecology*, *22*(20), 5130–5147.

Croose, E. (2016). *The Distribution and Status of the Polecat (Mustela putorius) in Britain 2014-2015*. https://www.vwt.org.uk/wp-content/uploads/2016/04/Polecat-Report-2016.pdf

Croose, E., Duckworth, J. W., Ruette, S., Skumatov, D. V., Kolesnikov, V. V., & Saveljev, A. P. (2018). A review of the status of the Western polecat*Mustela putorius*: a neglected and declining species? *Mammalia*, *82*(6), 550–564.

Cserkész, T., Kiss, C., Barkaszi, Z., Görföl, T., Zagorodniuk, I., Sramkó, G., & Csorba, G. (2021). Intra- and interspecific morphological variation in sympatric and allopatric populations of Mustela putorius and M. eversmanii (Carnivora: Mustelidae) and detection of potential hybrids. *Mammal Research*, *66*(1), 103–114.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker,

R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, *27*(15), 2156–2158.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008.

Davison, A., Birks, J. D. S., Griffiths, H. I., Kitchener, A. C., Biggins, D., & Butlin, R. K. (1999). Hybridization and the phylogenetic relationship between polecats and domestic ferrets in Britain. *Biological Conservation*, *87*(2), 155–161.

de Jong, M. J., Li, Z., Qin, Y., Quéméré, E., Baker, K., Wang, W., & Hoelzel, A. R. (2020). Demography and adaptation promoting evolutionary transitions in a mammalian genus that diversified during the Pleistocene. *Molecular Ecology*, *29*(15), 2777–2792.

Di Marco, M., Pacifici, M., Santini, L., Baisero, D., Francucci, L., Grottolo Marasini, G., Visconti, P., & Rondinini, C. (2013). Generation length for mammals. *Nature Conservation*, *5*, 89–94.

du Plessis, S. J., Blaxter, M., Koepfli, K.-P., Chadwick, E. A., & Hailer, F. (2023). Genomics Reveals Complex Population History and Unexpected Diversity of Eurasian Otters (Lutra lutra) in Britain Relative to Genetic Methods. *Molecular Biology and Evolution*, *40*(11). https://doi.org/10.1093/molbev/msad207

Etherington, G. J., Ciezarek, A., Shaw, R., Michaux, J., Croose, E., Haerty, W., & Di Palma, F. (2022). Extensive genome introgression between domestic ferret and European polecat during population recovery in Great Britain. *The Journal of Heredity*, *113*(5), 500–515.

Harris, S., Morris, P., Wray, S., & Yalden, D. (1995). *A Review of British Mammals: Population Estimates and Conservation Status of British Mammals Other Than Cetaceans*. https://www.researchgate.net/publication/242583185_A_Review_of_British_Mammals_Population_Estimates_and_Conservation_Status_of_British_Mammals_Other_Than_C

etaceans

Howard-McCombe, J., Jamieson, A., Carmagnini, A., Russo, I.-R. M., Ghazali, M., Campbell, R., Driscoll, C., Murphy, W. J., Nowak, C., O'Connor, T., Tomsett, L., Lyons, L. A., Muñoz-Fuentes, V., Bruford, M. W., Kitchener, A. C., Larson, G., Frantz, L., Senn, H., Lawson, D. J., & Beaumont, M. A. (2023). Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease. *Current Biology: CB*, *33*(21), 4761–4769.e5.

Kitchener, A. C., & Birks, J. D. S. (2008). *Feral ferret Mustela furo*. https://scholar.google.com/citations?user=b6MOxGsAAAAJ&hl=en&oi=sra

Langley, P. J. W., & Yalden, D. W. (1977). The decline of the rarer carnivores in Great Britain during the nineteenth century. *Mammal Review*, *7*(3-4), 95–116.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760.

Mathews, F., Kubasiewicz, L., Gurnell, J., Harrower, C., Mcdonald, R., & Shore, R. (2018). *A review of the population and conservation status of British mammals*. Unpublished. https://doi.org/10.13140/RG.2.2.28059.31521

McDevitt, A. D., Coscia, I., Browett, S. S., Ruiz-González, A., Statham, M. J., Ruczyńska, I., Roberts, L., Stojak, J., Frantz, A. C., Norén, K., Ågren, E. O., Learmount, J., Basto, M., Fernandes, C., Stuart, P., Tosh, D. G., Sindicic, M., Andreanszky, T., Isomursu, M., … Wójcik, J. M. (2022). Next-generation phylogeography resolves post-glacial colonization patterns in a widespread carnivore, the red fox (Vulpes vulpes), in Europe. *Molecular Ecology*, *31*(3), 993–1006.

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.

Montgomery, W. I., Provan, J., McCabe, A. M., & Yalden, D. W. (2014). Origin of British and Irish mammals: disparate post-glacial colonisation and species introductions.

Quaternary Science Reviews, *98*, 144–165.

Münger, X., Robin, M., Dalén, L., & Grossen, C. (2024). Facilitated introgression from domestic goat into Alpine ibex at immune loci. *Molecular Ecology*, *33*(14), e17429.

Perez-Sepulveda, B. M., Heavens, D., Pulford, C. V., Predeus, A. V., Low, R., Webster, H., Dykes, G. F., Schudoma, C., Rowe, W., Lipscombe, J., Watkins, C., Kumwenda, B., Shearer, N., Costigan, K., Baker, K. S., Feasey, N. A., Hinton, J. C. D., Hall, N., & 10KSG consortium. (2021). An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes. *Genome Biology*, *22*(1), 349.

Picard toolkit. (2019). In *Broad Institute, GitHub repository*. Broad Institute. https://broadinstitute.github.io/picard/

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.

Robinson, J., Kyriazis, C. C., Yuan, S. C., & Lohmueller, K. E. (2023). Deleterious Variation in Natural Populations and Implications for Conservation Genetics. *Annual Review of Animal Biosciences*, *11*, 93–114.

Sainsbury, K. A., Shore, R. F., Schofield, H., Croose, E., Campbell, R. D., & Mcdonald, R. A. (2019). Recent history, current status, conservation and management of native mammalian carnivore species in Great Britain. *Mammal Review*, *49*(2), 171–188.

Sainsbury, K. A., Shore, R. F., Schofield, H., Croose, E., Hantke, G., Kitchener, A. C., & McDonald, R. A. (2020). Diets of European polecat Mustela putorius in Great Britain during fifty years of population recovery. *Mammal Research*, *65*(2), 181–190.

Santiago, E., Novo, I., Pardiñas, A. F., Saura, M., Wang, J., & Caballero, A. (2020). Recent Demographic History Inferred by High-Resolution Analysis of Linkage Disequilibrium. *Molecular Biology and Evolution*, *37*(12), 3642–3653.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*, *30*(9), 1312–1313.

Stroupe, S., Forgacs, D., Harris, A., Derr, J. N., & Davis, B. W. (2022). Genomic evaluation of hybridization in historic and modern North American Bison (Bison bison). *Scientific Reports*, *12*(1), 6397.

Sumption, K. J., & Flowerdew, J. R. (1985). The ecological effects of the decline in Rabbits *(Oryctolagus cuniculus L.)* due to myxomatosis. *Mammal Review*, *15*(4), 151–186.

Thomas, N. E., Hailer, F., Bruford, M. W., & Chadwick, E. A. (2022). Country-wide genetic monitoring over 21 years reveals lag in genetic recovery despite spatial connectivity in an expanding carnivore (Eurasian otter, Lutra lutra) population. *Evolutionary Applications*, *15*(12), 2125–2141.

van der Auwera, G., & O'Connor, B. D. (2020). *Genomics in the cloud: Using docker, GATK, and WDL in Terra*. O'Reilly Media.

Author Contribution

The language used to recognise author contributions is described by CRediT Taxonomy (https://credit.niso.org/).
Conceptualization: GE, WH; formal analysis: RS; funding acquisition: GE, WH, FDP; investigation: RS; methodology: RS, GE, WH; project administration: RS, GE, WH; resources: JM, ACK, GE, WH; software: RS; supervision: JM, GE, WH; visualisation: RS; writing - original draft: RS; writing review & editing: RS, JM, ACK, GE, WH

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

All raw reads for the sequences generated during this study have been uploaded to ENA with the project number PRJEB76468.

Figures

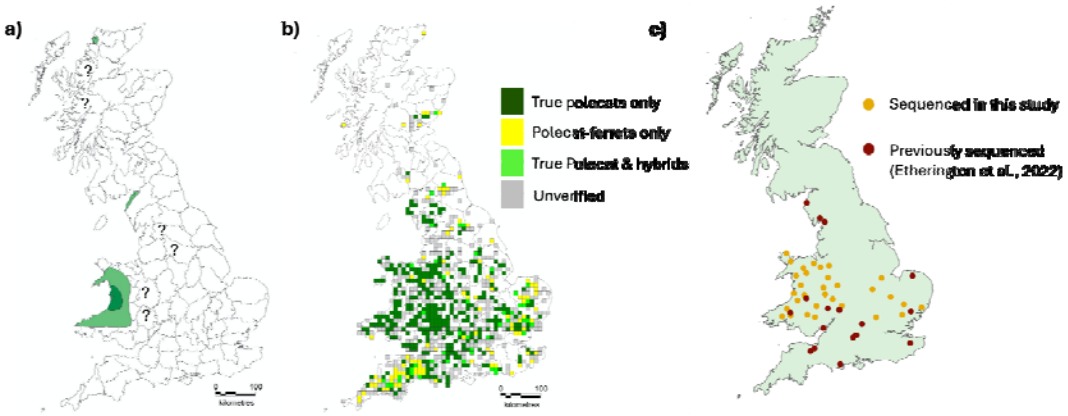

Figure 1: (a) The range of the European polecat in Britain in 1915. Dark green areas indicate the population stronghold, light green indicates known occurrences, and “?” refers to uncertain occurrences of polecats in those areas (Langley & Yalden, 1977). (b) Records received during the Vincent Wildlife Trust 2014-15 survey of verifiable sightings, based on their phenotype, of true polecats (dark-green), polecat-ferret hybrids (yellow), both true polecats and hybrids (lime green), and unverifiable (grey) (Croose, 2016). (c) Samples selected to be sequenced in this study (dark yellow) and (red) samples that were previously sequenced by (Etherington et al., 2022). Figures 1(a) and 1(b) were adapted from Croose, 2016.

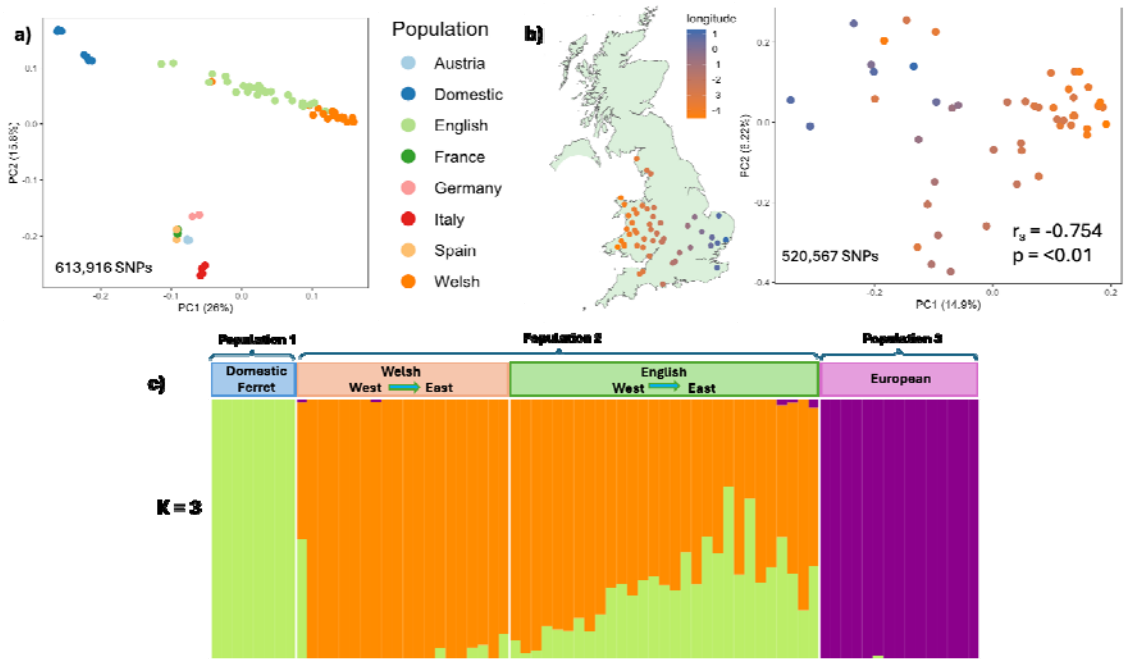

Figure 2. Population structure of the European polecat in Great Britain and across mainland Europe. (a) PCA based on whole genome SNPs for all European polecats and domestic ferrets. (b) Map of Great Britain demonstrating longitude coordinates of samples and PCA based on whole genome SNPs for all British polecats, which is supported by a significant correlation between PC1 and longitude coordinates. (c) Admixture results for whole-genome SNPs between the four populations.

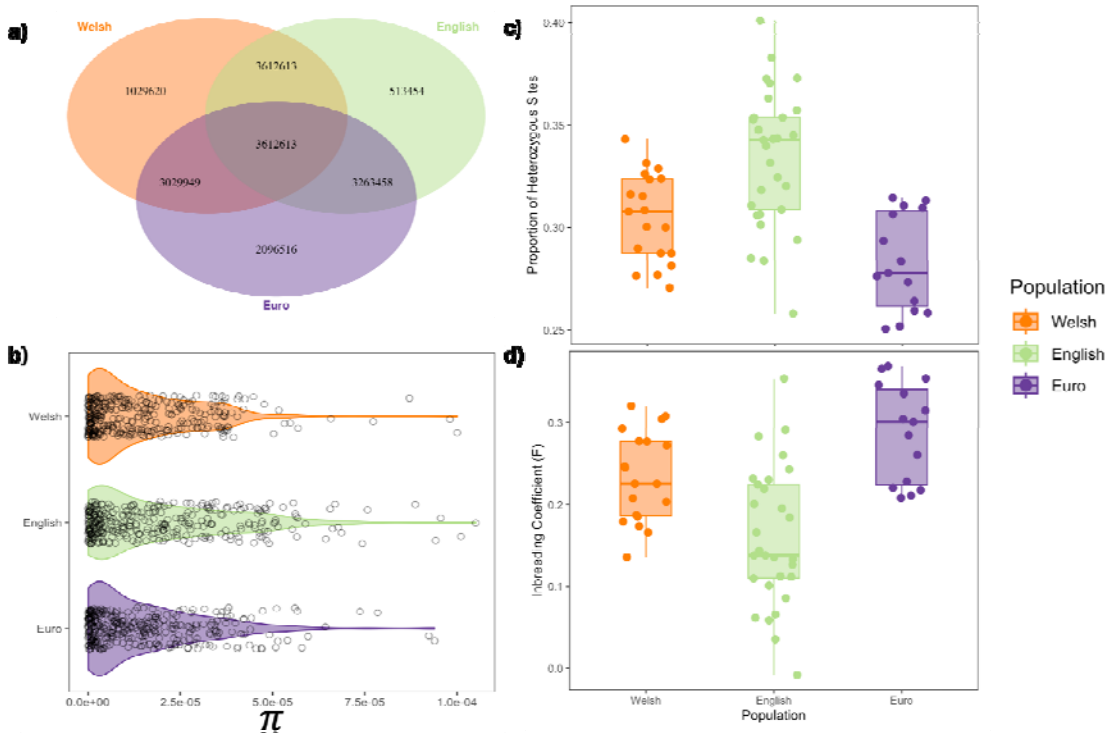

Figure 3. Genetic diversity estimates and comparisons across British and European populations (a) Shared and private SNPs between populations. (b) Population nucleotide diversity values (π), calculated in 20Mb windows across the genome. (c) Population-level inbreeding coefficient values and (d) proportion of heterozygous sites calculated using VCFtools.

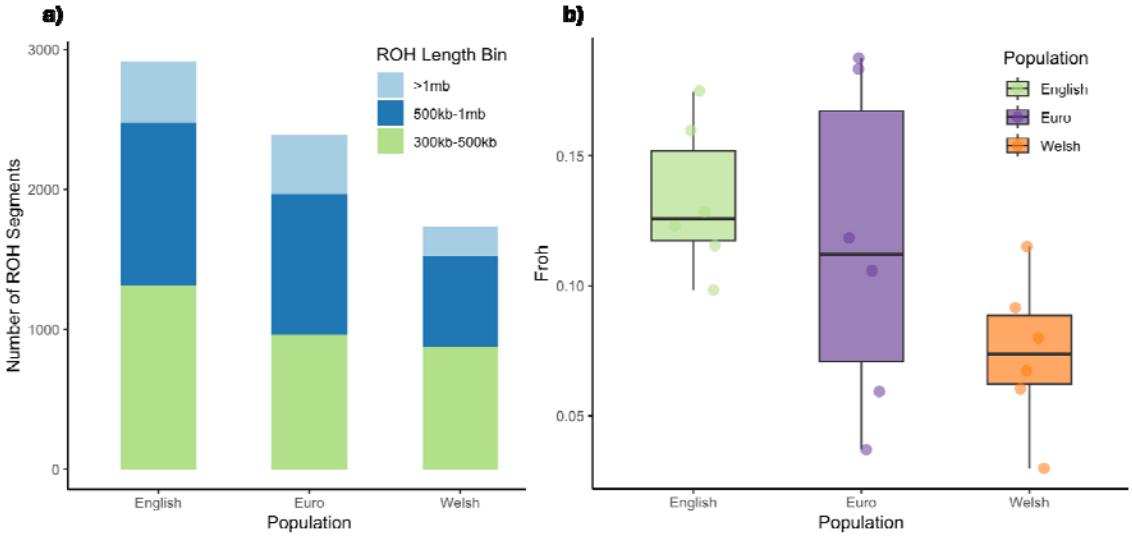

Figure 4. Population estimates of (a) the number of ROH segments categorised by length of ROH for English, European (Euro), and Welsh polecats, and (b) the calculated proportion of the genome in ROH (F_{ROH}) for the same populations

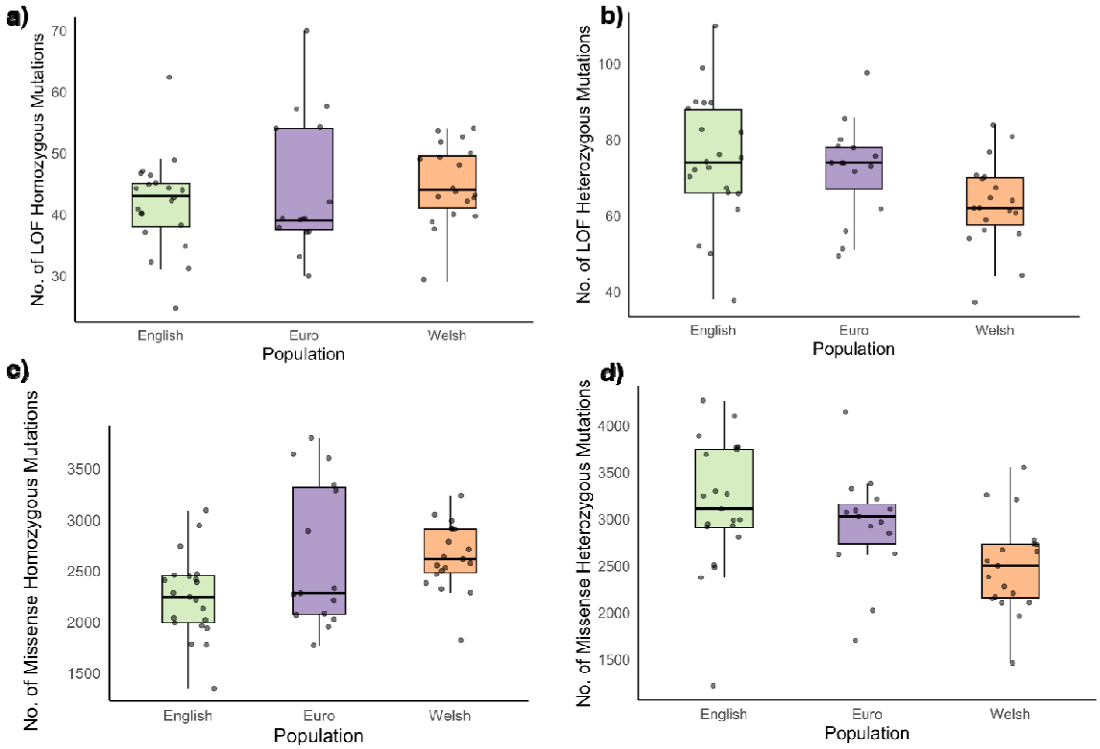

Figure 5. Mutation load landscape across populations within different functional categories. Loss of function (LoF) (a) homozygous mutational load per population and (b) heterozygous mutational load and missense (c) homozygous mutational load and (d) heterozygous mutational load.

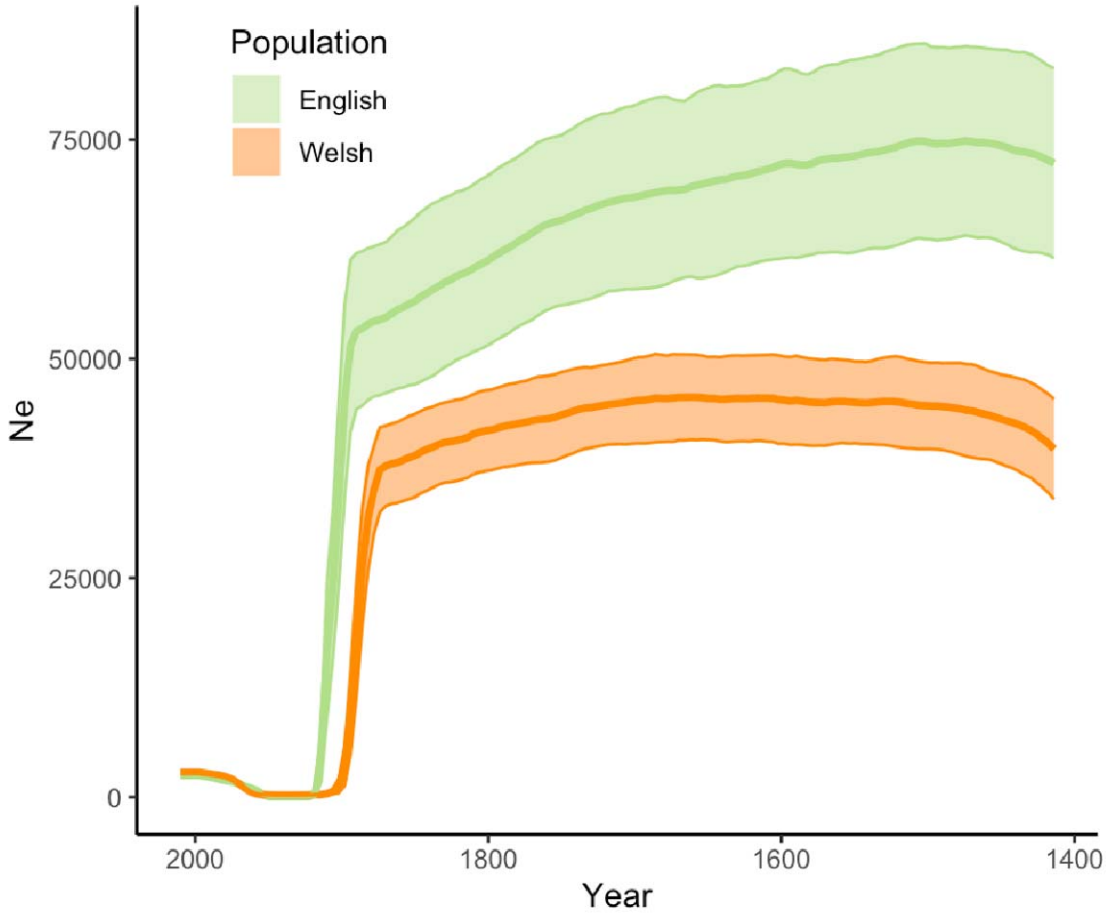

Figure 6: Estimated mean effective population size (solid line) with 95% confidence intervals (shaded area) for English (green) and Welsh (orange) populations across 200 generations (generations = 4 years) based on whole genome SNP data computed in GONE.