Current Biology

Report

Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease

Highlights

- The onset of significant hybridization is estimated to begin from the late 1950s
- Hybridization arises from expansion after a bottleneck estimated at 100 years ago
- Immune-related genes, including the MHC, show excess domestic cat introgression
- All but 0.13% (3.2 Mb) of the Scottish wildcat genome is present in our sample

Authors

Jo Howard-McCombe, Alexandra Jamieson, Alberto Carmagnini, ..., Helen Senn, Daniel J. Lawson, Mark A. Beaumont

Correspondence

jhmccombe@rzss.org.uk (J.H.-M.), hsenn@rzss.org.uk (H.S.), dan.lawson@bristol.ac.uk (D.J.L.), m.beaumont@bristol.ac.uk (M.A.B.)

In brief

The Scottish wildcat has long been held as a typical conservation exemplar following anthropogenic hybridization. Analysis of whole genomes from modern, museum, and archaeological samples by Howard-McCombe et al. demonstrates that hybridization with domestic cats has been recent and potentially mediated by pathogens.

Howard-McCombe et al., 2023, Current Biology *33*, 4761–4769
November 6, 2023 © 2023 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.cub.2023.10.026

CellPress

Current Biology

CellPress
OPEN ACCESS

Report

Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease

Jo Howard-McCombe,[1,6,21,*] Alexandra Jamieson,[2,3] Alberto Carmagnini,[3,4] Isa-Rita M. Russo,[5] Muhammad Ghazali,[6] Ruairidh Campbell,[7,8] Carlos Driscoll,[9] William J. Murphy,[10] Carsten Nowak,[11] Terry O'Connor,[12] Louise Tomsett,[13] Leslie A. Lyons,[14] Violeta Muñoz-Fuentes,[15] Michael W. Bruford,[5,19] Andrew C. Kitchener,[16,17] Greger Larson,[2] Laurent Frantz,[3,4] Helen Senn,[6,20,*] Daniel J. Lawson,[18,20,*] and Mark A. Beaumont[1,20,*]

[1]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK
[2]The Palaeogenomics & Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford OX1 3QY, UK
[3]Palaeogenomics Group, Department of Veterinary Sciences, Ludwig-Maximilians-Universität, Munich, Germany
[4]School of Biological and Behavioural Sciences, Queen Mary University of London, London E1 4NS, UK
[5]School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK
[6]RZSS WildGenes Laboratory, Conservation Department, Royal Zoological Society of Scotland, Edinburgh EH12 6TS, UK
[7]Wildlife Conservation Research Unit, Department of Zoology, University of Oxford Recanati-Kaplan Centre, Tubney House, Abingdon Road, Tubney OX13 5QL, UK
[8]NatureScot, Great Glen House, Leachkin Road, Inverness IV3 8NW, UK
[9]InnerCat, 221 Center Street, Frederick, MD 21701, USA
[10]Texas A&M University, Veterinary Integrative Biosciences, College Station, TX 77843, USA
[11]Senckenberg Research Institute and Natural History Museum, Center for Wildlife Genetics, 63571 Weimar, Germany
[12]BioArCh, Department of Archaeology, University of York, York YO10 5NG, UK
[13]Mammal Section, Science Department, Natural History Museum, London SW7 5BD, UK
[14]Department of Veterinary Medicine & Surgery, College of Veterinary Medicine, University of Missouri, Columbia, MO 65211, USA
[15]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK
[16]Department of Natural Sciences, National Museums Scotland, Edinburgh EH1 1JF, UK
[17]School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK
[18]School of Mathematics, University of Bristol, Bristol BS8 1UG, UK
[19]Deceased
[20]Senior author
[21]Lead contact
*Correspondence: jhmccombe@rzss.org.uk (J.H.-M.), hsenn@rzss.org.uk (H.S.), dan.lawson@bristol.ac.uk (D.J.L.), m.beaumont@bristol.ac.uk (M.A.B.)
https://doi.org/10.1016/j.cub.2023.10.026

SUMMARY

The European wildcat population in Scotland is considered critically endangered as a result of hybridization with introduced domestic cats,[1,2] though the time frame over which this gene flow has taken place is unknown. Here, using genome data from modern, museum, and ancient samples, we reconstructed the trajectory and dated the decline of the local wildcat population from viable to severely hybridized. We demonstrate that although domestic cats have been present in Britain for over 2,000 years,[3] the onset of hybridization was only within the last 70 years. Our analyses reveal that the domestic ancestry present in modern wildcats is markedly over-represented in many parts of the genome, including the major histocompatibility complex (MHC). We hypothesize that introgression provides wildcats with protection against diseases harbored and introduced by domestic cats, and that this selection contributes to maladaptive genetic swamping through linkage drag. Using the case of the Scottish wildcat, we demonstrate the importance of local ancestry estimates to both understand the impacts of hybridization in wild populations and support conservation efforts to mitigate the consequences of anthropogenic and environmental change.

RESULTS AND DISCUSSION

Understanding the causes and consequences of hybridization in threatened species is a challenging priority for conservation.[4–7] For example, hybridization affects the operational definition of species from a legislative perspective[6]; engenders discussion of species identity, integrity, and genetic uniqueness[7]; and informs discussions pertaining to whether preservation of genetic

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

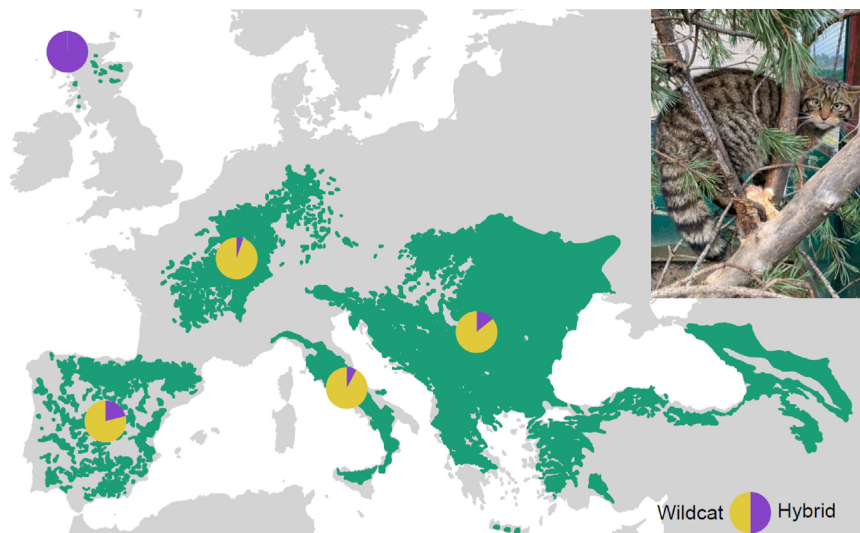

Figure 1. Current range of European wildcats and estimated hybridization rates
Median proportion of hybrid individuals (i.e., any individual with detectable domestic cat ancestry) reported across the four main metapopulations and Scottish population of European wildcat[1] (for the full list of reported values, see Data S1). The image in the top right shows a captive wildcat individual from Scotland. Photo credit: Saving Wildcats.

diversity or preservation of habitats and associated ecosystem function should be the goal.[8] Many wild populations are now subject to anthropogenic stressors[9] and environmental change.[10,11] Thus, within conservation biology, a key question is how to balance the input of potentially adaptive variation from introgressive hybridization against the loss of biodiversity that arises from genetic swamping.[5]

Hybridization between domestic cats and European wildcats exemplifies many of these issues. The European wildcat (*Felis silvestris*), historically distributed across western Eurasia from Turkey to Britain[12] (Figure 1), can hybridize with domestic cats (*F. catus*) and produce fertile offspring. *F. catus* derived from the African-Asian wildcat *F. lybica*,[13] whose range extends from southern Africa to China.[12] *F. lybica* and *F. catus* are genomically divergent from *F. silvestris* (Table S1) and phenotypically distinguishable using a suite of morphological characters.[1,14] Hybridization between domestic cat and wildcat populations in continental Europe has been previously reported (Data S1), and median hybridization rate (i.e., proportion of hybrids) ranges from 5.25% in western and central Europe to 21% on the Iberian Peninsula (Figure 1; Data S1).

As a consequence of combined pressures from persecution and habitat loss,[1] European wildcats have experienced dramatic population declines and range contraction over the last few centuries. In Britain, wildcats persisted in England and Wales until the late 19th century[15] but are now locally extinct. The remaining extant population is fragmented across the Scottish Highlands. Following the extinction of the lynx (*Lynx lynx*), which likely took place in the 7th century CE,[16,17] wildcats are now the only extant wild felid species in Britain and the most endangered mammalian carnivore.[18] Habitat-based estimates of wildcat population size suggest that between 30 and 430 individuals remain in the wild.[18] In Scotland, hybridization with domestic cats is now far greater than in other European populations (Figure 1).[19,20] Currently, no wild-living or captive Scottish wildcat is measurably free of domestic cat ancestry.[20] Genetic studies show that the wild population can be described as a hybrid swarm[20,21] and is therefore at serious risk of extinction via genetic swamping.[2]

It is important to understand how the uniquely acute anthropogenic hybridization in Scotland has arisen in the event that other European wildcat populations experience the same circumstances. Paleogenomic and zooarchaeological studies have demonstrated that domestic cats have been present in Britain since at least ∼350 BCE.[3] For most of their shared evolutionary history (i.e., at least until the 16th century CE), hybridization between wildcats and domestic cats was limited.[3] It is unclear, however, when wildcats in Britain transitioned from a population with occasional hybridization to one at risk of imminent extirpation[1,2] as a direct result of introgression.

Here, we generated and analyzed whole-genome sequence data to date the onset of significant hybridization in Britain and assess potential consequences for wildcats carrying introgressed domestic cat DNA. Our data comprised 22 domestic cats, 30 wild-living individuals from Scotland sampled across the hybrid swarm,[20] and six wildcats from the UK captive population (founded on wild animals from Scotland[22]). Seven additional wildcat samples were obtained from Germany and Portugal. We also made use of low-coverage, whole-genome sequence data from historic and archaeological samples, specifically four putative Scottish wildcats (museum specimens sampled 1906–1939, 0.3–4.7×) and two archaeological samples from Medieval (16th century, 0.9×) and Mesolithic Britain (8,459–8,272 cal. BP, 0.2×).[3] Additional low-coverage genomic data (0.02–0.07×) were obtained from 20 museum samples of putative Scottish wildcats.

Domestic cat introgression in the Scottish wildcat

The continuum of genetic ancestry in the hybrid swarm is illustrated by a principal component analysis (PCA) of 65 putative wildcats, hybrids, and domestic cats (Figure 2; sample information provided in Data S1). Additional low-coverage whole-genome data from historic and ancient samples are also projected onto the PCA (for sample information, see Data S1). This shows that ancient and historic wildcat samples from Britain clustered with modern wildcats from continental Europe and with an archaeological (8,459–8,272 cal. BP) wildcat from Scotland that pre-dates the introduction of domestic cats in the region. Given the confirmed absence of introgression in more recent archaeological samples,[3] this suggests a low degree of introgression following the introduction of domestic cats ∼350 BCE.[3]

To quantify domestic cat introgression in the Scottish wildcat population, we computed F-statistics[23] for both the historic and

CellPress
OPEN ACCESS

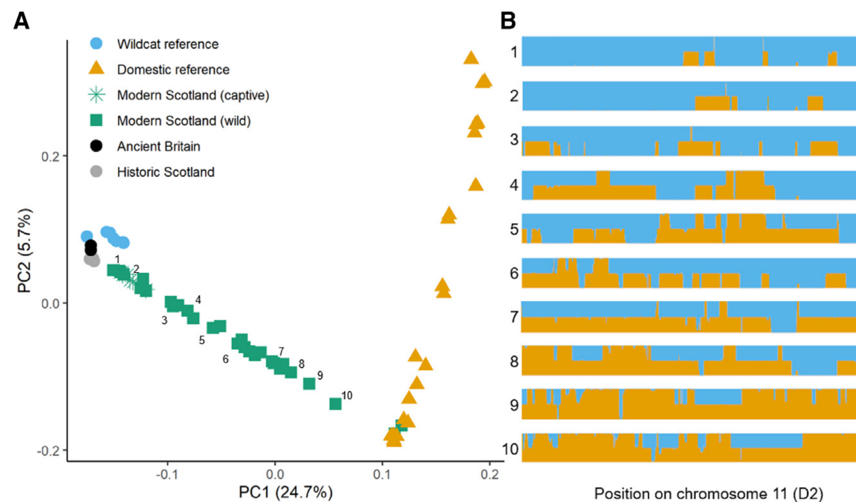

Figure 2. Genetic clustering and observed introgression across the sample set

(A) Principal component analysis shows strong differentiation of wildcats from mainland Europe (blue) and domestic cats (orange) across PC1, with Scottish individuals (green) distributed between these two groups. The clustering of museum and archaeological samples from Britain (gray and black) indicates this was not the historic situation in Scotland.

(B) Introgressed domestic cat DNA was identified in Scottish hybrids using MOSAIC. Horizontal bars show example inferred ancestries for ten individuals sampled across the hybrid swarm (shown on the PCA) for both copies of chromosome 11 (D2). Orange, probable domestic ancestry; blue, wildcat.

See also Data S2.

modern populations from Scotland (Data S1). A strong correlation ($R = 0.88$, $p < 2.2e^{-16}$) was observed between ancestry estimates computed using F_4 ratios and those obtained from STRUCTURE analyses of 35 SNPs routinely used to assay low-quality DNA samples for conservation monitoring.[20] A low-to-zero rate of hybridization was evident in the historic samples from Scotland, including data from the additional 20 individuals sampled between 1903 and 1985. Evidence of statistically significant domestic cat introgression was observed in five individuals, with estimated introgression ranging from 2.3% to 6.6% (Data S1). Scottish wildcat individuals only exceeded more than ~5% domestic cat ancestry after 1956. All modern individuals (1997–2018) showed evidence of introgression from domestic cats, and the proportion of domestic ancestry ranged between 11% and 74%. A captive wildcat population was established in the UK during the 1960s, initially from a small number of founders and with no hybrid testing.[22] Though this population has not avoided introgression, the mean proportion of domestic cat ancestry (18%) is significantly lower than that observed in the modern wild population (Data S1).

Dating the onset of introgression and historical population size

In order to date the onset of hybridization in Scotland, we used MOSAIC[24] to identify and quantify local ancestry (wildcat versus domestic cat) across the hybrid genomes (examples shown for chromosome D2, Figure 2; full results in Data S2), analyzing each putative wildcat individual separately to account for potentially continuous admixture. Local ancestry tract length can be used to deduce the timing of admixture events.[23,25] Based on the analysis of 36 individuals from Scotland, the mean estimate for the onset of hybridization was 8.6 (95% CI 8.3–9.8) generations before present. The distribution of hybridization events from individual coancestry curves ranged between 4.7 and 17.9 generations (Table S2). Observed patterns of exponential decay support the conclusion that admixture was rare prior to ~20 generations before the present[25] (Data S3).

Using a generation time of 3 years (STAR Methods) and accounting for sampling date (where known), our estimate for the onset of hybridization is between the late 1950s and mid-1990s (Figure 3). This time frame is consistent with results from PCA and AdmixTools (Figure 2; Data S1), highlighting low levels of introgression in early 20th century individuals, and confirms results from demographic modeling of the wildcat population in Scotland.[21] This result is also congruent with the breakdown in the consistency of pelage markings (scored at seven key morphological features on an ordinal scale of 1, 2, and 3 for domestic, hybrid, or wildcat features, respectively) that began in the late 1950s[14] (Figure S1).

Using MOSAIC local ancestry, we obtained ancestry-specific estimates of recent effective population size with GONE[26] (Figure 3), which uses linkage-disequilibrium (LD) information to infer demographic history. For the wildcat population, results were broadly consistent with the pattern of wildcat population decline and recent expansion observed in historical records.[15] We observed a decline in the effective population size beginning in the mid-19th century that accelerated in the 20th century, with a population minimum 22 generations before present, around 1950. This is somewhat more recent than reported by previous studies that, based on historical records of wildcat sightings, proposed a population low for Scottish wildcats in the 1910s.[15] Domestic ancestry from hybrid cats and five Scottish domestic cat reference samples shows effective population size for domestic cats increasing steadily over the last 150 generations. The dramatic decline ~5 generations before present may be due to the effect of current Population Structure, which will increase LD. By contrast, using the admixed data, a rapid population decline was observed from ~11 generations before the present, likely due to mixture LD[27] and consistent with the recent onset of admixture shown in previous analyses (Figures 2 and 3).

Signatures of selection in the hybrid population

Genomic regions observed in hybrid individuals that contain excess domestic or wildcat ancestry are candidates for selection. We counted excess local ancestry (STAR Methods) across the combined (admixed) wild and captive sample, accounting for each individual's overall ancestry, which is indicative of selection post-dating admixture.[28] Our analyses detected positive

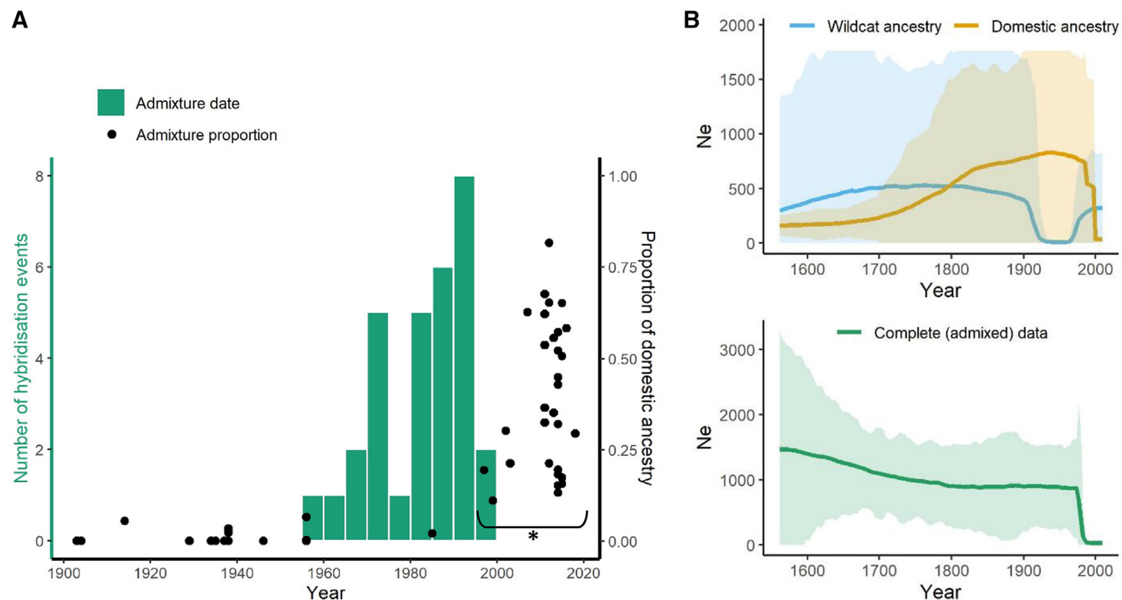

Figure 3. Predicted pattern of 20th-century admixture in Scottish wildcats

(A) A histogram shows the distribution of estimated admixture events (green). Superimposed are the predicted proportions of domestic cat ancestry (AdmixTools F_4 ratio test) in historic and modern Scottish samples (modern samples fall within the bracket denoted with an asterisk). Modern samples were used for MOSAIC analyses (generating the observed distribution of hybridization events).

(B) Estimates of recent effective population size (N_e); solid line indicates mean estimate, shaded area 95% confidence intervals. Ancestry-specific N_e (top) recovers historic decline and expansion of Scottish wildcats (blue) in the context of increasing domestic cat (orange) ownership in Britain. Analysis of the complete data for the hybrid population (bottom) is confounded by admixture LD consistent with recent admixture in Scotland, which causes an artificial population crash 11 generations before present.

See also Figure S1, Table S2, and Data S1 and S3.

selection on domestic cat ancestry in 43 regions spanning ~26 Mb across 10 chromosomes, including one region that spanned a portion of the major histocompatibility complex (MHC) locus on chromosome B2 (Figure 4). We also identified a region of purely domestic ancestry on chromosome C2 (~3.2 Mb). Using Wright-Fisher simulations of drift, assuming known ancestry, we demonstrate that the overall distribution of excess domestic genes at the MHC locus is not predicted by genetic drift alone (Figure S2). Further, exclusion of the MHC region from MOSAIC analysis showed that other regions of excess domestic ancestry could not be explained by unusual patterns of LD within the MHC itself (Figure S2).

Within regions of excess domestic ancestry, 273 genes were identified based on the Ensembl gene annotation (v106)[29] for domestic cats. Using G-profiler,[30] 212 functional categories were found to be significantly enriched for this set of genes ($p < 0.005$, Benjamini-Hochberg correction). The strongest hit (GO:0042611, $p < 5.0e^{-24}$) was associated with the MHC protein complex, and of the 66 categories with a corrected $p < 1.0e^{-4}$ (for complete list, see Data S1), 55 were directly linked to immune function. While many immune genes are found in the regions of excess domestic ancestry (Data S1), the MHC contains a large number of genes associated with numerous GO terms. Furthermore, it is close to the centromere on chromosome B2 and thus in a region of reduced recombination.[31] Since each immune-related GO term is associated with many MHC genes, it is likely that the MHC region as a whole contributes strongly to any functional enrichment in this category. Indeed, a G-profiler analysis excluding the MHC shows significant enrichment after multiple-testing correction only for "severe combined immunodeficiency." Strikingly, selection in hybrids was only identified within regions of domestic cat ancestry, and no wildcat ancestry passed the detection threshold. It should be emphasized that enrichment analysis should be viewed only as indicative under recent admixture, as it assumes independence between genetic regions, which are correlated due to admixture, and therefore may overstate the strength of the evidence.

Domestic cats and feral domestic cats are recognized as a reservoir of infectious diseases that pose a threat to all wild felids.[32] Even in apparently healthy cats, the prevalence of feline immunodeficiency virus (FIV) and feline leukemia virus in the UK domestic cat population is predicted to be ~6% and 5%, respectively.[33] A recent survey of 125 feral and hybrid cats in Scotland identified the presence of many common pathogens including FIV (7.3%), feline calicivirus (20%), and hemoplasma species (*Mycoplasma haemofelis*, 4.8%; *Mycoplasma haemominutum*, 23.4%).[32] Therefore, disease transmission is an important potential driver of selection in wildcat populations, and hybridization provides a mechanism for the transfer of pathogen resistance (including MHC variants[34,35]) from domestic cats.

Perspectives on anthropogenic hybridization

Our results demonstrate that though domestic cats have lived sympatrically with wildcats in Britain for >2,000 years, gene flow between the two populations was minimal until the second half of the 20th century. The abrupt ending of this genetic

CellPress
OPEN ACCESS

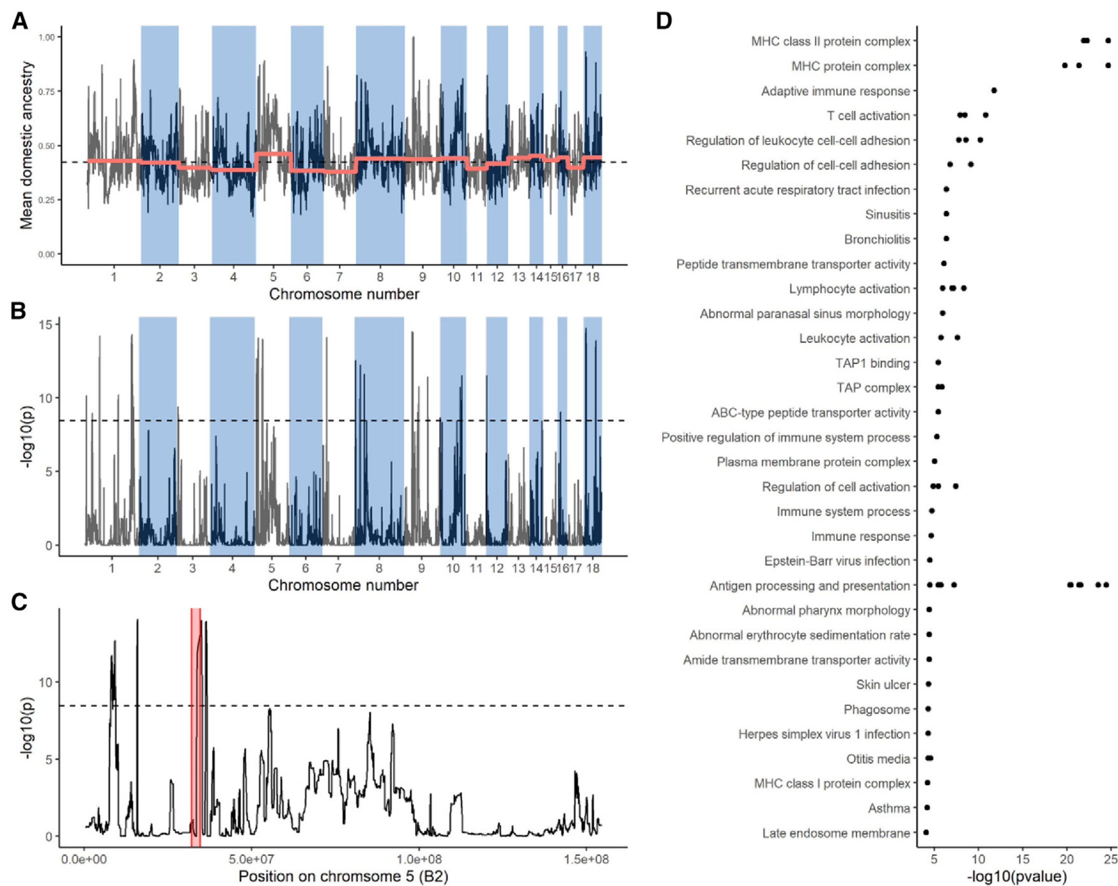

Figure 4. Evidence of selection for domestic cat genes in the hybrid population

(A) The mean proportion of domestic cat ancestry in the hybrid population varied across the genome. The dashed line indicates overall mean domestic ancestry; the red bar shows the mean per chromosome.

(B) Several regions showed a significant excess of domestic ancestry (dashed line indicates the significance threshold used).

(C) This included the MHC locus on chromosome 5 (B2), where domestic variation may confer disease resistance.

(D) A summary of the top 66 functional categories ($p < 1.0e^{-4}$), identified by G-profiler (each point represents a single term; terms with similar functions have been grouped).

See also Figures S2–S4 and Data S1.

segregation then led to the genetic swamping and near extirpation of the wildcat population in Britain.[1,2] The high level of introgression poses a challenge for the successful management of this population. As with many jurisdictions,[6] while wildcats are a protected species in the UK, legal protection from persecution extends neither to "hybrids" (colloquially taken to be F1 individuals) nor to introgressed individuals that have different degrees of backcrossing.[2]

The wildcat demographic decline in Britain pre-dates the onset of significant hybridization (Figure 3). Our evidence supports the view that the population of wildcats in Scotland is exceedingly low and has been affected by a strong recent bottleneck (70 individuals [CI 46–94], 33 generations before present) (Figure 3). There has been some, limited, recovery following this bottleneck, and the wildcat range has expanded over the 20th century from a small area in the northwest Highlands into central Scotland.[2] This range expansion is a proposed driver of introgression in continental wildcat populations,[36] since there would have been a low population density of conspecific mates at the leading edge of the expansion. A key isolating mechanism between wildcats and domestic cats is ecological and spatial separation,[37] which likely breaks down due to habitat degradation and the spread of wildcats in anthropogenic environments[38] where domestic cats predominate. Low wildcat population density across a fragmented habitat has persisted over the 20th century in the face of prey fluctuation, persecution, slow habitat recovery (specifically afforestation), urbanization, and land-use change, including intensive agriculture.[2]

In addition to the demographic factors noted here, our analysis suggests that genetic swamping of wildcat ancestry has potentially been exacerbated by asymmetrical selection favoring alleles acquired from domestic cats in immune genes such as MHC (Figure 4). High levels of genetic diversity (including, e.g., sequence diversification or gene duplication) make the MHC a difficult genomic region to analyze.[39] Sequence alignment is technically challenging, particularly for non-model organisms, and tens of thousands of samples are required to obtain completely reliable haplotypes[40]; strong balancing selection[41]

acts to strongly reduce genetic separation between species, leading to incomplete lineage sorting. However, our results do not require perfect alignment (observed variant quality; Figure S3) and our interpretation is appropriate as long as the *excess* of domestic cat ancestry, as inferred by MOSAIC, is real. Beyond the predicted and observed genetic drag of domestic cat ancestry into chromosome B2 (Figure 4), we ensured that unadmixed wildcat and domestic cat samples are sufficiently separated to allow a clear separation in a neighbor-joining tree (Figure S4), implying that ancestry calls are correct on average. Given the short timescale of hybridization that has led to the observed long sections of introgressed ancestry (Figure 2), it is likely that any putative signals of selection associated with these patterns are strong. It should be noted, however, that on longer timescales we may expect relationships between parental ancestry and local recombination rates due to weakly deleterious variants and Bateson-Dobzhansky-Muller incompatibilities.[42] Intriguingly, and prompting the hypothesis that contact with a large population harboring novel pathogens may induce selection on the MHC, an analogous signal of excess local ancestry is observed in humans at this locus,[43] based on analysis of ancient DNA. Modern Europeans were observed to possess excess Steppe pastoralist ancestry at the MHC (over genome-wide averages), disproportionately replacing the MHC genotype of their contemporary hunter-gatherers. As with domestic cats and wildcats, the pastoralists are thought to have formed a larger population and to have been exposed to a wider range of diseases in comparison to hunter-gatherers.

A limitation of gene enrichment analysis for immunity is that each immune-related GO term is shared by many MHC-associated loci, which means that the region as a whole has high leverage in statistical tests. However, there are many immune-related genes identified across the genome in regions of excess domestic cat introgression. Therefore, it is reasonable to hypothesize that contact between domestic cats and wildcats brings a disease burden that is at least partially offset by introgression of domestic immune genes, likely also dragging linked maladaptive variants from the domestic genome into the wildcat population. Given the asymmetry in population and disease density, this general phenomenon may have relevance for the conservation and rewilding of species with domestic relatives, including gray wolves,[44,45] polecats,[46] and ibexes.[35]

There is a potential advantage of hybridization in genetically rescuing the small Scottish wildcat population from inbreeding depression,[47] but recent analyses[48] show that there is a danger of replacing a significant fraction of native wildcat diversity with that of the phylogenetically, ecologically, and behaviorally distinct domestic cat. A more compelling argument favoring domestic diversity in the wildcat is that we have identified infectious disease as the primary selected phenotype in hybrids. Maintaining domestic cat variation at the MHC, and potentially at other immune-related loci, may be important for population viability in the presence of domestic cats, but at the risk of dragging linked variants maladapted to the wildcat niche. Selective introgression of the MHC has been previously reported both in the case of anthropogenic hybridization[35] and in a natural hybrid zone.[49]

Our results indicate that the modern Scottish wildcat population is heavily admixed. In this regard it is technically "genomically extinct" by the definition of Allendorf et al.[4] because we have found no unadmixed individuals. In such cases there may be evidence of strong selection against non-native alleles,[50] whereas our evidence suggests that genome segments from domestic cats are disproportionately favored. However, our data provide sufficient information to demonstrate that an attempt at reconstruction of the ancestral wildcat genome using a de-introgression strategy,[51] based on the estimated proportion of wildcat ancestry per individual (hybrid index, HI), is theoretically feasible. To combat inbreeding depression and maintain diversity, the inclusion of continental European wildcat genetic variation may be important in the future. An alternative approach[6,8] would be to use HI within a program of habitat restoration and exclusion of domestic cats and individuals with a low HI, so that selection can favor wildcat traits in their natural environment. In conclusion, our evidence demonstrates that the captive Scottish wildcat population and wild-living hybrids retain almost the full ancestral genome required to reconstruct the original species, and these results will inform future conservation strategies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Data generation
  - Sequence alignment and variant calling
  - Phasing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Population structure
  - Quantifying introgression
  - Dating admixture
  - Masking domestic ancestry in hybrid individuals
  - Recent effective population size
  - Tests for selection

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2023.10.026.

ACKNOWLEDGMENTS

We are grateful for the assistance of the wildcat captive holding community in the UK and to David Barclay for discussions relating to the studbook. We thank Danielle Gunn-Moore at the University of Edinburgh for access to domestic cat reference samples. We also thank staff, volunteers, and collaborators of the National Heritage Lottery Funded Scottish Wildcat Action project for their participation in sampling during this project. Some wild study samples collected prior to Scottish Wildcat Action were provided by Ruairidh Campbell and Jenny Bryce in association with the NatureScot commissioned survey[52] with additional support from WildCru and PTES. A selection of wild study samples from Scotland are archived in the CRYOARKS biobank (https://www.

cryoarks.org), and samples from captive individuals are archived in the EAZA biobank supported via BBSRC grant BB/R015260/1 and the Royal Zoological Society of Scotland. We thank Hugo AndersonWhymark and Alison Sheridan for providing samples. Harry Gordon, Holly Scott, and Hollie Smith are thanked for their preliminary work on post-MOSAIC analyses. Generation of the whole-genome data was funded by the People's Trust for Endangered Species and the Royal Zoological Society of Scotland, with sequencing of historic and ancient samples funded by NERC-Oxford DTP in Environmental Research, the ERC (grant ERC-2013-StG-337574UNDEAD). This research used the Leibniz-Rechenzentrum High Performance Computing facility. J.H.-M. is supported by the NERC Doctoral Training Partnership, with additional funding from the People's Trust for Endangered Species. H.S. and M.G.'s work at the RZSS WildGenes lab was supported by the Players of the People's Postcode Lottery. Preparation of cat specimens at National Museums Scotland was supported by the Negaunee Foundation. The authors would like to thank everyone who provided wildcat samples to National Museums Scotland. This paper is dedicated to the memory of the late Mike Bruford, a close colleague and mentor, who, together with Elizabeth Barratt, initiated the first genetic survey of the Scottish wildcat.

AUTHOR CONTRIBUTIONS

Conceptualization, D.J.L., J.H.-M., and M.A.B.; methodology, A.C., A.J., D.J.L., J.H.-M., and I.-R.M.R.; resources, A.C.K., C.D., C.N., H.S., L.A.L., L.T., M.G., R.C., T.O., V.M.-F., and W.J.M.; writing, D.J.L., J.H.-M., and M.A.B.; supervision, D.J.L., H.S., M.A.B., M.W.B., A.C.K., G.L., and L.F.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: July 31, 2023
Revised: September 24, 2023
Accepted: October 17, 2023
Published: November 6, 2023

REFERENCES

1. Gerngross, P., Ambarli, H., Angelici, F.M., Anile, S., Campbell, R., Ferreras de Andres, P., Gil-Sanchez, J.M., Götz, M., Jerosch, S., Mengüllüoglu, D., et al. (2022). Felis silvestris. The IUCN Red List of Threatened Species 2022. https://doi.org/10.2305/IUCN.UK.2022-1.RLTS.T181049859A181050999.en.
2. Breitenmoser, U., Lanz, T., and Breitenmoser-Würsten, C. (2019). Conservation of the wildcat (Felis silvestris) in Scotland: review of the conservation status and assessment of conservation activities. Wildcat in Scotland. https://www.nature.scot/sites/default/files/2019-02/Wildcat%20in%20Scotland%20-%20Review%20of%20conservation%20status%20and%20activities_1.pdf.
3. Jamieson, A., Carmagnini, A., Howard-McCombe, J., Doherty, S., Hirons, A., Dimopolous, E., Lin, A.T., Allen, R., Anderson-Whymark, H., Barnett, R., et al. (2023). Limited historical admixture between European wildcats and domestic cats. Curr. Biol. *33*. https://doi.org/10.1016/j.cub.2023.08.031.
4. Allendorf, F.W., Leary, R.F., Spruell, P., and Wenburg, J.K. (2001). The problems with hybrids: setting conservation guidelines. Trends Ecol. Evol. *16*, 613–622. https://doi.org/10.1016/S0169-5347(01)02290-X.
5. Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G., and Rieseberg, L.H. (2016). Hybridization and extinction. Evol. Appl. *9*, 892–908. https://doi.org/10.1111/eva.12367.
6. Chan, W.Y., Hoffmann, A.A., and van Oppen, M.J.H. (2019). Hybridization as a conservation management tool. Conserv. Lett. *12*, e12652, https://doi.org/10.1111/conl.12652.
7. Quilodrán, C.S., Montoya-Burgos, J.I., and Currat, M. (2020). Harmonizing hybridization dissonance in conservation. Commun. Biol. *3*, 1–10. https://doi.org/10.1038/s42003-020-1116-9.
8. Wayne, R.K., and Shaffer, H.B. (2016). Hybridization and endangered species protection in the molecular era. Mol. Ecol. *25*, 2680–2689. https://doi.org/10.1111/mec.13642.
9. Ottenburghs, J. (2021). The genic view of hybridization in the Anthropocene. Evol. Appl. *14*, 2342–2360. https://doi.org/10.1111/eva.13223.
10. Cahill, A.E., Aiello-Lammens, M.E., Fisher-Reid, M.C., Hua, X., Karanewsky, C.J., Ryu, H.Y., Sbeglia, G.C., Spagnolo, F., Waldron, J.B., Warsi, O., and Wiens, J.J. (2013). How does climate change cause extinction? Proc. Biol. Sci. *280*, 20121890. https://doi.org/10.1098/rspb.2012.1890.
11. Brauer, C.J., and Beheregaray, L.B. (2020). Recent and rapid anthropogenic habitat fragmentation increases extinction risk for freshwater biodiversity. Evol. Appl. *13*, 2857–2869. https://doi.org/10.1111/eva.13128.
12. Kitchener, A.C., Breitenmoser-Würsten, C., Eizirik, E., Gentry, A., Werdelin, L., Wilting, A., Yamaguchi, N., Abramov, A.V., Christiansen, P., Driscoll, C., et al. (2017). A revised taxonomy of Felidae. The final report of the Cat Classification Task Force of the IUCN/SSC Cat Specialist Group. Cat News Spec *11*.
13. Driscoll, C.A., Menotti-Raymond, M., Roca, A.L., Hupe, K., Johnson, W.E., Geffen, E., Harley, E.H., Delibes, M., Pontier, D., Kitchener, A.C., et al. (2007). The Near Eastern origin of cat domestication. Science *317*, 519–523.
14. Kitchener, A.C., Yamaguchi, N., Ward, J.M., and Macdonald, D.W. (2005). A diagnosis for the Scottish wildcat (Felis silvestris): a tool for conservation action for a critically-endangered felid. Anim. Conserv. *8*, 223–237. https://doi.org/10.1017/S1367943005002301.
15. Langley, P.J.W., and Yalden, D.W. (1977). The decline of the rarer carnivores in Great Britain during the nineteenth century. Mamm Rev. *7*, 95–116. https://doi.org/10.1111/j.1365-2907.1977.tb00363.x.
16. Hetherington, D.A., Lord, T.C., and Jacobi, R.M. (2006). New evidence for the occurrence of Eurasian lynx (Lynx lynx) in medieval Britain. J. Quat. Sci. *21*, 3–8. https://doi.org/10.1002/jqs.960.
17. Raye, L. (2021). An 18th century reference to a Eurasian lynx (Lynx lynx) in Scotland. Mamm. Commun. *7*, 47–52.
18. Mathews, F., Kubasiewicz, L.M., Gurnell, J., Harrower, C.A., McDonald, R.A., and Shore, R.F. (2018). A Review of the Population and Conservation Status of British Mammals: Technical Summary (Natural England).
19. Beaumont, M., Barratt, E.M., Gottelli, D., Kitchener, A.C., Daniels, M.J., Pritchard, J.K., and Bruford, M.W. (2001). Genetic diversity and introgression in the Scottish wildcat. Mol. Ecol. *10*, 319–336. https://doi.org/10.1046/j.1365-294X.2001.01196.x.
20. Senn, H.V., Ghazali, M., Kaden, J., Barclay, D., Harrower, B., Campbell, R.D., Macdonald, D.W., and Kitchener, A.C. (2019). Distinguishing the victim from the threat: SNP-based methods reveal the extent of introgressive hybridization between wildcats and domestic cats in Scotland and inform future in situ and ex situ management options for species restoration. Evol. Appl. *12*, 399–414. https://doi.org/10.1111/eva.12720.
21. Howard-McCombe, J., Ward, D., Kitchener, A.C., Lawson, D., Senn, H.V., and Beaumont, M. (2021). On the use of genome-wide data to model and date the time of anthropogenic hybridization: an example from the Scottish wildcat. Mol. Ecol. *30*, 3688–3702. https://doi.org/10.1111/mec.16000.
22. Barclay, D. (2019). Scottish Wildcat Studbook, Species360 Zoological Information Management System (ZIMS). zims.Species360.org.
23. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture

in human history. Genetics *192*, 1065–1093. https://doi.org/10.1534/genetics.112.145037.

24. Salter-Townshend, M., and Myers, S. (2019). Fine-scale inference of ancestry segments without prior knowledge of admixing groups. Genetics *212*, 869–889. https://doi.org/10.1534/genetics.119.302139.
25. Hellenthal, G., Busby, G.B.J., Band, G., Wilson, J.F., Capelli, C., Falush, D., and Myers, S. (2014). A genetic atlas of human admixture history. Science *343*, 747–751. https://doi.org/10.1126/science.1243518.
26. Santiago, E., Novo, I., Pardiñas, A.F., Saura, M., Wang, J., and Caballero, A. (2020). Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. Mol. Biol. Evol. *37*, 3642–3653. https://doi.org/10.1093/molbev/msaa169.
27. Waples, R.S., and England, P.R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. Genetics *189*, 633–644. https://doi.org/10.1534/genetics.111.132233.
28. Nelson, R.M., Wallberg, A., Simões, Z.L.P., Lawson, D.J., and Webster, M.T. (2017). Genomewide analysis of admixture and adaptation in the Africanized honeybee. Mol. Ecol. *26*, 3603–3617. https://doi.org/10.1111/mec.14122.
29. Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., et al. (2021). Ensembl 2021. Nucleic Acids Res. *49*, D884–D891. https://doi.org/10.1093/nar/gkaa942.
30. Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. *44*, W83–W89. https://doi.org/10.1093/NAR/GKW199.
31. Li, G., Hillier, L.W., Grahn, R.A., Zimin, A.V., David, V.A., Menotti-Raymond, M., Middleton, R., Hannah, S., Hendrickson, S., Makunin, A., et al. (2016). A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. G3 (Bethesda) *6*, 1607–1616. https://doi.org/10.1534/g3.116.028746.
32. Meredith, A., Bacon, A., Allan, B., Kitchener, A., Senn, H., Brooks, S., Kortland, K., Hetherington, D., and Davies, S. (2018). Domestic cat neutering to preserve the Scottish wildcat. Vet. Rec. *183*, 27–28. https://doi.org/10.1136/vr.k2905.
33. Hosie, M.J., Robertson, C., and Jarrett, O. (1989). Prevalence of feline leukaemia virus and antibodies to feline immunodeficiency virus in cats in the United Kingdom. Vet. Rec. *125*, 293–297.
34. Hedrick, P.W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. Mol. Ecol. *22*, 4606–4618. https://doi.org/10.1111/mec.12415.
35. Grossen, C., Keller, L., Biebach, I., International Goat Genome Consortium, Croll, D., Ajmone, P., Amills, M., Boitard, S., Chen, W., Cheng, S., et al. (2014). Introgression from domestic goat generated variation at the major histocompatibility complex of alpine ibex. PLoS Genet. *10*, e1004438, https://doi.org/10.1371/journal.pgen.1004438.
36. Nussberger, B., Currat, M., Quilodran, C.S., Ponta, N., and Keller, L.F. (2018). Range expansion as an explanation for introgression in European wildcats. Biol. Conserv. *218*, 49–56. https://doi.org/10.1016/j.biocon.2017.12.009.
37. Gil-Sánchez, J.M., Jaramillo, J., and Barea-Azcón, J.M. (2015). Strong spatial segregation between wildcats and domestic cats may explain low hybridization rates on the Iberian Peninsula. Zoology *118*, 377–385. https://doi.org/10.1016/j.zool.2015.08.001.
38. Kilshaw, K., Montgomery, R.A., Campbell, R.D., Hetherington, D.A., Johnson, P.J., Kitchener, A.C., Macdonald, D.W., and Millspaugh, J.J. (2016). Mapping the spatial configuration of hybridization risk for an endangered population of the European wildcat (Felis silvestris silvestris) in Scotland. Mamm. Res. *61*, 1–11. https://doi.org/10.1007/s13364-015-0253-x.
39. Dilthey, A., Cox, C., Iqbal, Z., Nelson, M.R., and McVean, G. (2015). Improved genome inference in the MHC using a population reference graph. Nat. Genet. *47*, 682–688. https://doi.org/10.1038/ng.3257.
40. Luo, Y., Kanai, M., Choi, W., Li, X., Sakaue, S., Yamamoto, K., Ogawa, K., Gutierrez-Arcelus, M., Gregersen, P.K., Stuart, P.E., et al. (2021). A high-resolution HLA reference panel capturing global population diversity enables multi-ancestry fine-mapping in HIV host response. Nat. Genet. *53*, 1504–1516. https://doi.org/10.1038/s41588-021-00935-7.
41. Alter, I., Gragert, L., Fingerson, S., Maiers, M., and Louzoun, Y. (2017). HLA class I haplotype diversity is consistent with selection for frequent existing haplotypes. PLoS Comput. Biol. *13*, e1005693, https://doi.org/10.1371/journal.pcbi.1005693.
42. Schumer, M., Xu, C., Powell, D.L., Durvasula, A., Skov, L., Holland, C., Blazier, J.C., and Sankararaman, S. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. Science *660*, 656–660. https://doi.org/10.1126/science.aar3684.
43. Barrie, W., Yang, Y., Attfield, K.E., Irving-Pease, E., Scorrano, G., Jensen, L.T., Armen, A.P., Dimopoulos, E.A., Stern, A., Refoyo-Martinez, A., et al. (2022). Genetic risk for multiple sclerosis originated in Pastoralist Steppe populations. Preprint at bioRxiv. https://doi.org/10.1101/2022.09.23.509097.
44. Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., and Randi, E. (2017). Disentangling timing of admixture, patterns of introgression, and phenotypic indicators in a hybridizing wolf population. Mol. Biol. Evol. *34*, 2324–2339. https://doi.org/10.1093/molbev/msx169.
45. Muñoz-Fuentes, V., Darimont, C.T., Paquet, P.C., and Leonard, J.A. (2010). The genetic legacy of extirpation and re-colonization in Vancouver Island wolves. Conserv. Genet. *11*, 547–556. https://doi.org/10.1007/s10592-009-9974-1.
46. Costa, M., Fernandes, C., Birks, J.D.S., Kitchener, A.C., Santos-Reis, M., and Bruford, M.W. (2013). The genetic legacy of the 19th-century decline of the British polecat: evidence for extensive introgression from feral ferrets. Mol. Ecol. *22*, 5130–5147. https://doi.org/10.1111/mec.12456.
47. Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol. Ecol. *16*, 463–475. https://doi.org/10.1111/j.1365-294X.2006.03148.x.
48. Harris, K., Zhang, Y., and Nielsen, R. (2019). Genetic rescue and the maintenance of native ancestry. Conserv. Genet. *20*, 59–64. https://doi.org/10.1007/s10592-018-1132-1.
49. Dudek, K., Gaczorek, T.S., Zieliński, P., and Babik, W. (2019). Massive introgression of major histocompatibility complex (MHC) genes in newt hybrid zones. Mol. Ecol. *28*, 4798–4810. https://doi.org/10.1111/mec.15254.
50. Kovach, R.P., Hand, B.K., Hohenlohe, P.A., Cosart, T.F., Boyer, M.C., Neville, H.H., Muhlfeld, C.C., Amish, S.J., Carim, K., Narum, S.R., et al. (2016). Vive la résistance: genome-wide selection against introduced alleles in invasive hybrid zones. Proc. Biol. Sci. *283*, 20161380, https://doi.org/10.1098/rspb.2016.1380.
51. Amador, C., Hayes, B.J., and Daetwyler, H.D. (2014). Genomic selection for recovery of original genetic background from hybrids of endangered and common breeds. Evol. Appl. *7*, 227–237. https://doi.org/10.1111/eva.12113.
52. Littlewood, N.A., Campbell, R.D., Dinnie, L., Gilbert, L., Hooper, R., Iason, G., Irvine, J., Kilshaw, K., Kitchener, A., Lackova, P., et al. (2014). Survey and scoping of wildcat priority areas. Scottish Natural Heritage Commissioned Report No. 768.
53. Buckley, R.M., Davis, B.W., Brashear, W.A., Farias, F.H.G., Kuroki, K., Graves, T., Hillier, L.W., Kremitzki, M., Li, G., Middleton, R.P., et al. (2020). A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. PLoS Genet. *16*, e1008926–e1008928. https://doi.org/10.1371/journal.pgen.1008926.
54. Yu, H., Xing, Y.T., Meng, H., He, B., Li, W.J., Qi, X.Z., Zhao, J.Y., Zhuang, Y., Xu, X., Yamaguchi, N., et al. (2021). Genomic evidence for the Chinese

CellPress
OPEN ACCESS

mountain cat as a wildcat conspecific (Felis silvestris bieti) and its introgression to domestic cats. Sci. Adv. *7*, eabg0221. https://doi.org/10.1126/sciadv.abg0221.

55. Li, G., Davis, B.W., Eizirik, E., and Murphy, W.J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Res. *26*, 1–11. https://doi.org/10.1101/gr.186668.114.
56. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
57. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.
58. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110.20.
59. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
60. 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., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
61. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab0084. https://doi.org/10.1093/gigascience/giab008.
62. 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., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.
63. Browning, S.R., and Browning, B.L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am. J. Hum. Genet. *81*, 1084–1097. https://doi.org/10.1086/521987.
64. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909. https://doi.org/10.1038/ng1847.
65. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190–e2093. https://doi.org/10.1371/journal.pgen.0020190.
66. Petr, M. (2020), admixr: an interface for running "ADMIXTOOLS" analyses. R package version 0.9.1. https://CRAN.R-project.org/package=admixr.
67. Knaus, B.J., and Grünwald, N.J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. *17*, 44–53. https://doi.org/10.1111/1755-0998.12549.
68. Paradis, E. (2010). Pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics *26*, 419–420. https://doi.org/10.1093/bioinformatics/btp696.
69. Paradis, E., and Schliep, K. (2019). Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
70. Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of population structure using dense haplotype data. PLoS Genet. *8*, 10024533–e1002517. https://doi.org/10.1371/journal.pgen.1002453.
71. Hudson, R.R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics *18*, 337–338. https://doi.org/10.1093/bioinformatics/18.2.337.
72. Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., et al. (2011). Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database *2011*. bar030-9. https://doi.org/10.1093/database/bar030.
73. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763. https://doi.org/10.1073/pnas.1314445110.
74. Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., and Allentoft, M.E. (2015). Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. *5*, 11184–11212. https://doi.org/10.1038/srep11184.
75. Gansauge, M.T., and Meyer, M. (2014). Selective enrichment of damaged DNA molecules for ancient genome sequencing. Genome Res. *24*, 1543–1549. https://doi.org/10.1101/gr.174201.114.24.
76. Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S.S.T., Sinding, M.H.S., Samaniego, J.A., Wales, N., Sicheritz-Pontén, T., and Gilbert, M.T.P. (2018). Single-tube library preparation for degraded DNA. Methods Ecol. Evol. *9*, 410–419. https://doi.org/10.1111/2041-210X.12871.
77. Charlesworth, B. (1994). Evolution in Age-Structured Populations, 2nd Ed. (Cambridge University Press).
78. O'Ryan, C., Harley, E.H., Bruford, M.W., Beaumont, M., Wayne, R.K., and Cherry, M.I. (1998). Microsatellite analysis of genetic diversity in fragmented South African buffalo populations. Anim. Conserv. *1*, 85–94. https://doi.org/10.1017/S136794309800002X.
79. Mattucci, F., Galaverni, M., Lyons, L.A., Alves, P.C., Randi, E., Velli, E., Pagani, L., and Caniglia, R. (2019). Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. Sci. Rep. *9*, 11612, https://doi.org/10.1038/s41598-019-48002-w.
80. Quilodrán, C.S., Nussberger, B., Macdonald, D.W., Montoya-Burgos, J.I., and Currat, M. (2020). Projecting introgression from domestic cats into European wildcats in the Swiss Jura. Evol. Appl. *13*, 1–12. https://doi.org/10.1111/eva.12968.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
57 wildcat/hybrid/domestic cat blood or tissue samples	This paper	Table S3
24 wildcat/hybrid skin samples (historic)	This paper	Table S3
Deposited data		
Raw sequencing reads	This paper	NCBI BioProject: PRJNA1030997
Domestic cat genome assembly	[53]	GenBank: GCA_000181335.3
Domestic cat sequencing reads	NCBI	NCBI BioProject: PRJNA343389; PRJNA512113
Domestic cat, *F. bieti, F. lybica ornata* sequencing reads	[54]	NCBI BioProject: PRJNA478778
F. margarita	[55]	NCBI BioProject: PRJNA286909
Putative British wildcat, ancient genomic data	[3]	Table S3
Software and algorithms		
FastQC v0.11.2	Babraham Institute	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/; RRID: SCR_014583
Trimmomatic v0.39	[56]	http://www.usadellab.org/cms/index.php?page=trimmomatic; RRID: SCR_011848
Bowtie2 v2.2.9	[57]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; RRID: SCR_016368
GATK Toolkit v4.0.8.1	[58]	https://software.broadinstitute.org/gatk/; RRID: SCR_001876
Samtools v1.9	[59]	https://www.htslib.org/; RRID: SCR_002105
VCFtools v0.1.17.0	[60]	https://vcftools.github.io/index.html; RRID: SCR_001235
BCFtools v1.8	[61]	http://samtools.sourceforge.net/mpileup.shtml; RRID: SCR_005227
Plink	[62]	https://www.cog-genomics.org/plink/
BEAGLE v5.1	[63]	http://faculty.washington.edu/browning/beagle/beagle.html; RRID: SCR_001789
Eigensoft v7.2.1	[64,65]	https://reich.hms.harvard.edu/software; RRID: SCR_004965
AdmixTools v7.0.2	[23]	https://github.com/DReichLab/AdmixTools; RRID: SCR_018495
Admixr	[66]	https://github.com/bodkan/admixr
MOSAIC	[24]	https://csgitlab.ucd.ie/mst/mosaic
vcfR	[67]	https://CRAN.R-project.org/package=vcfR; RRID: SCR_023453
pegas	[68]	https://cran.r-project.org/web/packages/pegas/index.html
Ape	[69]	https://CRAN.R-project.org/package=ape; RRID: SCR_017343
GONE	[26]	https://github.com/esrud/GONE
fineSTRUCTURE v4.4.1	[70]	http://paintmychromosomes.com; RRID: SCR_018170
Ms	[71]	http://home.uchicago.edu/~rhudson1/source/mksamples.html
Ensembl v106	[29]	http://www.ensembl.org/; RRID: SCR_002344
BioMart MartView	[72]	http://www.ensembl.org/info/data/biomart/index.html; RRID: SCR_010714
Gprofiler	[30]	http://biit.cs.ut.ee/gprofiler/; RRID: SCR_006809

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Jo Howard-McCombe (jhmccombe@rzss.org.uk).

CellPress
OPEN ACCESS

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw sequencing reads generated by this study are available at NCBI Short Read Archive under the BioProject: PRJNA1030997.
- Code used to process and analyse raw data is available at GitHub (https://github.com/johowardmcc/Howard_McCombe_et_al_CurrBiol_2023; https://github.com/danjlawson/localselection). Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

For this study whole-genome resequencing data were generated for 57 modern wildcats (*F. silvestris, F. bieti, F. lybica, F. margarita*) and domestic cats. This included 40 putative wildcat/hybrid individuals and 5 domestic cats from Scotland, 6 German and 1 Portuguese wildcat (*F. silvestris*), 1 sample of *F. bieti,* 1 sample of *F. margarita* and 3 samples of *F. lybica ornata* (see Data S1). Additionally, publicly available data were retrieved for 17 domestic cat samples, 1 sample of *F. margarita*, 1 sample of *F. lybica ornata* and 4 samples of *F. bieti* (for accession numbers see Data S1).

Low-coverage genetic data, including 4 whole-genome sequences, were generated for 24 historical samples of Scottish wildcats. Whole-genome sequence data for two archaeological samples of British wildcats were provided by Jamieson et al.[3]

METHOD DETAILS

Data generation

Modern samples

Whole-genome sequence data were generated for 57 modern samples (Data S1). Raw read data from 23 additional samples were obtained from public databases and downloaded in FASTQ format (NCBI accession numbers provided in Data S1).

Wild individuals from Scotland were selected with the aim of sampling a wide geographical distribution and range of genetic backgrounds from across the hybrid swarm.[20] DNA was extracted by the WildGenes laboratory, Royal Zoological Society of Scotland, as described by Senn et al.[20] European wildcat data, outside of Scotland, included six German wildcat individuals and one Portuguese wildcat. DNA extraction for the German wildcat samples was carried out using the Qiagen DNeasy Tissue kit. For the Portuguese sample DNA and library build was conducted by BioServe.

Sequencing of the Scottish samples was performed by the Beijing Genomics Institute (BGISEQ-500), generating paired-end read data at 15x coverage. German wildcat samples were sequenced by Edinburgh Genomics, whole-genome sequencing PCR-free libraries on the HiSeq X platform, generating paired-end data at 30x coverage. The Portuguese wildcat sample was sequenced using the Illumina HiSeq 2500 platform and aligned to the domestic cat reference genome v9.0[53] (accession: GCA_000181335.3) (mapped read data, BAM format, were provided to the authors of this study).

Eleven individuals from this dataset were sampled from additional Felis spp. (specifically, *F. margarita, F. bieti* and *F. lybica ornata*, using the taxonomy described in Kitchener et al.[12]) (Data S1). These samples were exclusively used to calculate F-statistics (see 'Quantifying Introgression') and were excluded from additional downstream analyses. DNA samples from one *F. bieti*, one *F. margarita* and one *F. lybica ornata* were generated by BioServe and sequenced using the Illumina HiSeq platform. Two *F. lybica ornata* samples were extracted using the Qiagen Puregene Kit and sequenced using the BGISEQ platform. An additional four *F. bieti*, one *F. lybica ornata* and one *F. margarita* samples were available on public databases.

Historic samples

Genetic data were obtained from historic and ancient wildcat samples (n = 26, Data S1) from Britain as described in Jamieson et al.[3] (see also below). These data comprised two datasets: whole-genome sequencing of four historic samples from Scotland and two ancient samples from Britain, and low-coverage data (not whole-genome sequencing) from 20 additional historic samples from Scotland.

Low-coverage data were generated using the extraction protocol described in Senn et al.[20]

Whole-genome sequence data were generated from four historic specimens from Scotland, sampled between 1906 and 1939 (Data S1). DNA was extracted from these samples using the method described in Senn et al.[20] Whole-genome sequence data from two archaeological samples were processed as described in Jamieson et al.[3] In brief, bone samples were cut using a Dremel 3000 electric hand-drillto between 50 and 200 mg. The surface of the bone was removed using a circular cutting disk to eliminate any surface contamination of modern DNA. The bones were powdered using a Retsch MM400 micro-dismembrator. The DNA was extracted using a protocol based on Dabney et al.[73] with modifications from Damgaard et al.[74]

Samples were sequenced as part of a larger palaeogenomic study, as described in Jamieson et al.[3]: Illumina libraries were built following either Gansauge and Meyer[75] or Carøe et al.,[76] but with the addition of a six basepair barcode added to the IS1_adapter.P5 and IS3_adapter.P5+P7 adapter pair. The libraries were then amplified on an Applied Biosystems StepOnePlus Real-Time PCR

system to check that library building was successful, and to determine the minimum number of cycles to use during the indexing amplification PCR reaction. A six base-pair barcode was used during the indexing amplification reaction resulting in each library being double-barcoded with an "internal adapter" directing adjacent to the ancient DNA strand and which would be the first bases sequenced, and a traditional external barcode that would be sequenced during Illumina barcode sequencing. The number of PCR cycles was determined by performing qPCR (Quantitative PCR) using the StepOne Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions. The amplified libraries were then pooled together in three batches and purified using the QIAGEN MinElute columns following the manufacturer's instructions. This was followed by size selection using SPRI beads (Beckman Coulter), 32 μL of DNA extract using Carøe et al.'s method.[76] SPRI beads (Beckman Coulter) were used for purification of the libraries as instructed in the protocol. The prepared libraries were then assessed for the optimal number of cycles for PCR and quality control conducted on a TapeStation 2200 (Agilent Technologies) prior to being sent for sequencing. The first batch was sequenced on a single lane of a HiSeq 4000 instrument at the Crick Institute, London. The second and third batches were sequenced on a single lane each of a HiSeq4000 instrument at Novogene, Sacramento.

Libraries with >15% endogenous DNA were selected for deeper sequencing. These libraries were sequenced in two batches, the first consisted of five indexed libraries sequenced alongside another five indexed libraries from another project, the second batch consisted of six indexed libraries sequenced alongside another ten indexed libraries from another project. Both were sequenced on one lane each of a NovaSeq 6000 at Novogene, Sacramento. For a summary per sample see Table S3.

Sequence alignment and variant calling

Modern samples

All bioinformatic processing used the BlueCrystal Phase 3 platform. Poor-quality read data were removed from FASTQ files using Trimmomatic.[56] A sliding window approach removed bases from the 3′ end of any sequence of four positions with a mean quality score less than 20. Bases at the leading or trailing ends of forward or reverse reads with a quality score less than three were removed. Trimmed read data were aligned to the domestic cat reference genome v9.0[53] (accession: GCA_000181335.3), using Bowtie,[57] and the default parameters for paired-end data. Samtools[59] was used to sort and index BAM files per sample. Alignment rate appeared to be consistent across domestic cat, wildcat, and hybrid samples (mean alignment rate of 96.1%, 97.6% and 98.3% for domestic cats, putative wildcats and known hybrid samples, respectively).

Read-group information was added and duplicate reads were identified and flagged for each BAM file using GATK AddOrReplaceReadGroups and GATK MarkDuplicates[58] (GATK v4.0.8.1).

GATK was used to call variants per sample, using HaplotypeCaller in GVCF mode. GATK GenomicsDBImport was used to aggregate per-sample GVCFs for joint-genotyping (per chromosome) with GenotypeGVCFs.

An existing set of high-quality reference data is not available for *Felis silvestris*, so a hard-filtering approach was taken instead of variant recalibration with GATK. Two rounds of filtering were carried out. Firstly, using GATK SelectVariants, filtering out sites with low quality by read depth (QD < 2), poor mapping quality (MQ < 40,MQRankSum< -12.5, ReadPosRankSumTest< -8), or strand bias (FS > 60, SOR>3). Read depth, variant quality and SNP density were then assessed per chromosome to inform a second round of filtering using VCFtools[60] and BCFtools,[61] removing sites with low-quality calls (QUAL<50) or excessive read depths (DP > 2000). Four closely related individuals in the captive population (kinship >0.125, as identified using pedigree information[22]) were removed from the dataset. Multi-allelic sites and sites with a genotyping rate of less than 100% were discarded. After this, 11,863,892 biallelic SNPs remained for analyses.

Historic samples

Historic and ancient whole-genome data were genotyped as described in Jamieson et al.[3] For the low-coverage historical data, Samtools[59] v.1.9 was used to coordinate sort and index the resulting bam files. Bcftools[61] (v1.8) mpileup (–adjust-MQ 50) and bcftools call (-vmO z) were used to generate genotype likelihoods and call variants across all samples. Subsequent filtering removed one individual (WCQ1030) with large proportions of missing data. Hard-filtering of variants retained bi-allelic loci of moderate quality (QUAL≥20). Of these, only sites also genotyped in the modern data were retained for downstream analyses (72,389 SNPs).

Phasing

Beagle v.5.2[63] was used to phase modern samples genotyped at 11,863,892 SNPs. A genetic linkage map has been generated for domestic cats.[31] This was modified to phase the wildcat dataset, pruning the original dataset to remove non-contiguous SNPs, and imposing a minimum recombination rate of 5×10^{-7}. The final map used for phasing included 5,860 markers. Beagle was run using 100 iterations and a burn-in of 50.

Phased data were subsequently thinned; a minor allele count of at least three was imposed, and variants thinned at random to one SNP per 2kb. The total number of SNPs remaining for analyses was 1,011,786.

QUANTIFICATION AND STATISTICAL ANALYSIS

Population structure

Principal component analysis (PCA) of 71 modern, historic and archaeological samples (65 modern, 4 historic and 2 archaeological) used Eigensoft's smartpca.[64,65] Prior to PCA, low coverage (historic and ancient) samples were filtered to include only bi-allelic sites with a genotyping quality score of at least 20 (QUAL≥20). Modern data were used to compute the first ten principal components

(PCs), projecting low-coverage samples onto these axes (lsqproject = YES). Outlier removal was disabled (outliermode = 2). To minimise missing-ness across the dataset, PCs were computed using those SNPs genotyped in at least one lowcoverage sample, thinned to one SNP per 1kb. The final number of markers used for PCA was 862,730. Missing data in the historic and archaeological samples varied between 1.5% (WCQ1008) and 78.1% (AJ419).

Population pairwise F_{ST} (Table S1) was calculated between domestic cats, continental wildcats, captive Scottish wildcats and wild-living Scottish samples using VCFtools[60] –weir-fst-pop.

Quantifying introgression

F4 statistics were calculated using AdmixTools[23] and its R interface, admixr.[66] Ancestry proportions were estimated for all putative Scottish wildcats: historic (whole genome sequences and additional low coverage data, n = 23), modern wild-living (n = 30) and modern captive (n = 6) using the F4 ratio test and qpAdm. Models tested are summarised in Data S1. Scenarios, including the historic screening data, were based on the analysis of 72,389 SNPs.

(QUAL$\geq$20, genotyped in at least one historic sample) and the rest used 862,730 SNPs, as described above for PCA.

Dating admixture

MOSAIC[24] analysis used 65 modern samples, genotyped at 1,011,786 SNPs (see Population Structure'). Three reference panels were provided to MOSAIC: wildcats from mainland Europe (n = 7), Scottish domestic (n = 5) and non-Scottish domestic cats (n = 17). We specified a model of two-way admixture (A = 2). Given the observed introgression in the captive Scottish wildcat population (Data S1), these samples (n = 6) were treated as admixed and included in the target population.

Coancestry curves were fitted for the target population as a whole and for each sample individually (Data S3), allowing for the fact that individuals within a population may experience an admixture event at different points in their history.[25] Sampling date was accounted for, where possible, to give an estimated date of admixture per sample (Table S2). The population mean estimate for the onset of admixture was taken from the three coancestry curves generated for the target population as a whole. Confidence intervals for the inferred population mean were generated using 100 bootstrap samples.

To determine a calendar date from these data, we employed the standard life-table generation time, i.e., the mean age of the parents of an individual when it is born, averaged across all individuals in the population.[77,78] An accurate generation time remains uncertain for wildcats, as is the case for many elusive species for which it is challenging to obtain detailed life tables. Additionally, the impact of hybridization on reproductive rate is not understood. Breeding age is usually between 1 and 8 years in captive wildcats[1] and estimates of generation time, including those previously used to date hybridization between wildcats and domestic cats, range between two[79] to three[36,80] years.

An exploratory test, including the captive Scottish wildcat population as an additional reference panel, showed MOSAIC to be robust to admixed panels, a finding that accords with Salter-Townshend & Myers.[24]

Masking domestic ancestry in hybrid individuals

Using local ancestry estimates from MOSAIC,[24] regions of putative introgressed domestic cat ancestry (visualised in Data S2) were removed (or 'masked') from individual genomes to obtain ancestry-specific measures. For each chromosome ancestry estimates were first mapped back to the original phasing (pre-MOSAIC phase hunter), using MOSAIC's phase_local(localanc, final.flips), and then mapped to the corresponding bp position using the MOSAIC function grid_to_pos() (Salter-Townshend, personal communication, November 2021). Sites with >0.8 probability of wildcat ancestry across both haplotypes (per individual) were retained for downstream analyses. The remaining sites were coded as missing data ("."). Genotype information was modified in R using vcfR.[67] VCFs generated per chromosome, using this method, were combined using bcftools[61] concat.

Once masked, strong negative correlation was observed between proportion of missing data and hybrid index ($R = -0.96$, $p < 2.2e-16$). Two individuals (WCQ099 and WCQ165) were removed due to a high proportion of missing data (imiss>0.99), supporting these individuals as putative feral domestic cats. Subsequent PCA using the masked data eliminated the hybrid swarm observed in Figure 1, and clustered Scottish wildcats with wildcats from mainland Europe and ancient and historic samples from Britain. Principal components were computed in Eigensoft[64,65] with seven wildcats from mainland Europe and 22 domestic cats genotyped at 959,539 SNPs, projecting the remaining samples (ancient, historic, or with masked domestic cat ancestry).

Recent effective population size

Recent effective population size was estimated using GONE.[26] GONE[26] was run using the default settings and exploiting phasing information (PHASE = 1) and linkage information from the modified domestic cat recombination map (generated with fineSTRUCTURE's[70] convertrecfile.pl). All modern individuals of putative Scottish wildcat ancestry were included in the analyses (captive and wild-living), genotyped at the thinned set of markers described above (see 'Population Structure'), autosomes only (nSNPs = 972,731).

GONE analyses also included ancestry specific estimates of effective population size. For this, haploid data were generated, removing domestic cat or wildcat ancestry (as identified by MOSAIC[24]) per haplotype, using the method as described above ('Masking domestic ancestry in hybrid individuals'). GONE was run treating these data as pseudohaploid (PHASE = 0). For wildcat ancestry this analysis was run using 34 individuals (68 haploids, excluding putative feral domestic cats WCQ0099 and WCQ0165) genotyped

at 204,637 SNPs (removing monomorphic loci). For domestic ancestry this analysis was run using 41 individuals (82 haploids, including WCQ0099, WCQ0165 and the five domestic cat reference samples from Scotland), genotyped at 362,348 SNPs.

These data were not dense enough to generate confidence intervals by resampling subsets of SNPs, as described by Santiago et al.[26] Although the trajectory was consistent, N_e appeared to be dependent on sample size using this method. Instead, we applied a leave-one-out jack-knife approach, running each analysis 18 times, leaving out data from one chromosome for each run. Jack-knife bias appeared to be small using this method (mean of -551.61 and 251.02 for wildcat and domestic cat ancestry-specific estimates, respectively).

Tests for selection

The method described by Nelson et al.[28] was used to identify regions of excess wildcat or domestic ancestry in the hybrid population, based on local ancestry estimates from.

MOSAIC[24] (see 'Dating Admixture'). Briefly, this method assumes that the local ancestry of each SNP l for haplotype i is binomially distributed with probability p_i, given by the genome-wide ancestry of individual i. By treating each haplotype as independent, the total local ancestry is Poisson-binomial distributed and the p value is computed using a two-tailed test accounting for multiple testing. These p values were shown to be well-calibrated for a simulated population (using ms[71]) of 1000 individuals that admixed 30 generations ago. Because there is no modelling of SNPs, the model does not depend on LD except for multiple-testing correction to which it is conservative, and is therefore robust to genetic drift.

To further empirically verify that the MHC signal cannot be explained by genetic drift, we implemented a diploid Wright-Fisher model with crossover recombination via meiosis, with recombination occurring uniformly at random at rate 1 per Morgan of a 3 Morgan genome. We initialised the simulation to have the mean domestic ancestry observed (0.42) in our data. We followed local ancestry explicitly and ran the simulation a fixed number of generations G with a fixed population size N, before sampling 36 individuals to match our sample. The count of ancestry as modelled in Figure 4 is compared in Figure S2. For no values of G between 5 and 20 and N between 400 and 1200 do we see the extreme domestic ancestry observed in the hybridised wildcats, with a clear net excess genome-wide, and the distribution on chromosome B2 is an outlier. Code is at https://github.com/danjlawson/localsancestrybreeding in the 'geneticdrift' folder.

Domestic cat genes within these regions were identified using the Ensembl gene annotation v106[29] (http://www.ensembl.org/biomart/martview/). Tests for functional enrichment across this set of genes were carried out using g-profiler,[30] using all available data sources for the domestic cat (Gene Ontology, KEGG and Human Phenotype Ontology), a significance threshold of p = 0.05, following a Benjamini-Hochberg correction, and otherwise default settings.