

Population genomics of Corsican wildcats: Paving the way toward a new subspecies within the *Felis silvestris* spp. complex?

Elodie Portanier¹  | H  l  ne Henri¹ | Pierre Benedetti² | Fr  d  ric Sanchis² | Corinne R  gis¹ | Pascale Chevret¹  | Marco Zedda³ | Adil El Filali¹ | Sandrine Ruet  te⁴ | S  bastien Devillard¹

¹Laboratoire de Biom  trie et Biologie Evolutive, University of Lyon, Universit   Claude Bernard Lyon 1, CNRS, Villeurbanne, France

²Office Fran  ais de la Biodiversit   (OFB), Direction R  gionale PACA-Corse, Cellule Technique Corse, Hameau de Mezzana, Moltifao, France

³Dipartimento di Medicina Veterinaria, Universit   degli Studi di Sassari, Sassari, Italy

⁴Office Fran  ais de la Biodiversit   (OFB), Direction de la Recherche et de l'Appui Scientifique, UPAD, Montfort, Birieux, France

Correspondence

Elodie Portanier, Sorbonne Universit  , CNRS, UMR 7144, Station Biologique de Roscoff, Roscoff, France.
Email: elodie.portanier@gmail.com

Funding information

Direction R  gionale de l'Environnement, de l'Am  nagement et du Logement Corse; Laboratoire de Biom  trie et Biologie Evolutive; Office de l'Environnement de la Corse; Office Fran  ais de la Biodiversit  ; Universit   de Lyon

Handling Editor: David Coltman

Abstract

In the context of the current extinction crisis, identifying new conservation units is pivotal to the development of sound conservation measures, especially in highly threatened taxa such as felids. Corsican wildcats are known by Corsican people since a very long time but have been little studied. Meaningful information about their phylogenetic position is lacking. We used ddRADseq to genotype phenotypically homogenous Corsican wildcats at 3671 genome-wide SNPs and reported for the first time their genetic identity. We compared this genomic information to domestic cats *Felis silvestris catus* from Corsica and mainland France, European wildcats *F. s. silvestris* and Sardinian wildcats *F. s. lybica*. Our premise was that if the Corsican wildcat, as a phenotypic entity, also represents a genetic entity, it deserves conservation measures and to be recognized as a conservation unit. Corsican wildcats appeared highly genetically differentiated from European wildcats and genetically closer to Sardinian wildcats than to domestic cats. Domestic cats from Corsica and mainland France were closer to each other and Sardinian wildcats were intermediate between Corsican wildcats and domestic cats. This suggested that Corsican wildcats do not belong to the *F. s. silvestris* or *catus* lineages. The inclusion of more high-quality Sardinian samples and Near-Eastern mainland *F. s. lybica* would constitute the next step toward assessing the status of Corsican wildcat as a subspecies and/or evolutionarily significant unit and tracing back wildcat introduction history of in Corsica.

KEYWORDS

conservation genetics, Corsica, double-digest RAD sequencing, *Felis reyi*, *Felis silvestris*, population genomics

1 | INTRODUCTION

Because of human activities, biodiversity is currently facing an extinction crisis (Brummitt et al., 2015; Ceballos et al., 2015;

Cowie et al., 2017, 2022; Hallmann et al., 2017; Thomas et al., 2004). This crisis represents a threat to global biodiversity but also to humanity and its activities (Ceballos et al., 2015, 2020; Diaz et al., 2019) making biodiversity conservation, besides a moral obligation, an

Sandrine Ruet  te and S  bastien Devillard share equal responsibilities.

urgent need to ensure a viable environment for future human generations. Biodiversity is defined according to three levels (genetic, species and ecosystem diversities) which, to be adequately protected, need to be characterized and their functioning understood. For instance, among the information necessary for the design of suitable conservation and management strategies, species, subspecies or evolutionarily significant unit delineation and biogeography characterization are part of the first steps (Coates et al., 2018). Identifying populations of the same species and/or unique gene pools, indeed, helps to assign a legal status to this genetic diversity that can then be considered in conservation plans, favouring the protection of both the species and the global biodiversity (Coates et al., 2018). Genetic diversity is crucial to preserve since it is linked to the fitness of individuals, the short-term maintenance of populations, but also to the adaptive potential of populations, and thus, to the ability of species to cope with environmental changes (DeWoody et al., 2021; Frankham et al., 2004; Hoban et al., 2021 and references therein). In addition to this conservation significance, identifying new and previously ignored specific entities is also a positive and enthusiastic message sent to both the conservation biology community and governments, underlying that work performed in investigating biodiversity might be successful. Defining conservation units is nevertheless challenging, although the advance of population genomics approaches greatly improved our understanding of the species-population continuum and facilitated the identification of relevant units (Coates et al., 2018; Funk et al., 2012; Hohenlohe et al., 2021).

In the case of wild species having domestic relatives, it is even more challenging. Indeed, many of the wild ancestors of domestic species are understudied and have uncertain status due to the frequent lack of clear phenotypical demarcation and the hybridization with domestic relatives which may participate to obscure delineation signals (Smith et al., 2022). One example of such species of which the conservation is challenged partly due to the existence of domestic population are wildcats *Felis silvestris*. Although felids are one of the most intensively studied group and the target of large conservation efforts, wildcats do not benefit from the same level of concern as the largest and most charismatic species of this group (Albert et al., 2018; Anile & Devillard, 2020). Despite being considered as "Least Concern" by the IUCN, *Felis silvestris* is experiencing a population decline and is considered an endangered taxon, at least locally, in some countries, where it is protected by European laws (Yamaguchi et al., 2015). Threats to wildcats include human activities (e.g., persecution, road kills), habitat loss and fragmentation and hybridization with feral domestic cats *Felis silvestris catus* (Yamaguchi et al., 2015). The *Felis silvestris* lineage in Eurasia is actually subdivided into several subspecies, according to their geographical ranges: the African-Near Eastern wildcat (*F. s. lybica*), the Asian wildcat (*F. s. ornata*), the European wildcat (*F. s. silvestris*), the Southern African wildcat (*F. s. cafra*) and the Chinese cat (*F. s. bieti*) (Driscoll et al., 2007). *F. s. silvestris* diverged first, 230,000 years before present (BP) while *F. s. cafra*, *F. s. ornata* and *F. s. lybica* diverged from each other more recently (173,000 years BP, Driscoll et al., 2007). The African wildcat *F. s. lybica* is now recognized as

the ancestor of domestic cats (Driscoll et al., 2007, 2009; Ottoni et al., 2017). The domestication process was initiated in the Near East during the Neolithic, about 10,000 years ago, and *F. s. lybica* then most probably followed human migration toward Europe as a commensal species (Baca et al., 2018; Faure & Kitchener, 2009; Nilson et al., 2022; Ottoni et al., 2017; Vigne et al., 2004).

During the migration from the Fertile Crescent to Europe, ancestral representatives of several mammalian species were introduced on Mediterranean islands, such as mouflons (*Ovis gmelini musimon*), with individuals being first introduced in Cyprus around 10,500 years BP and reaching Sardinia and Corsica 3000–4000 years later (Poplin, 1979; Vigne, 1992; Zeder, 2008) but also of red foxes (*Vulpes vulpes*), house mice (*Mus musculus*), wild boars (*Sus scrofa*) or weasels (*Mustela nivalis*) and probably wildcats (Cucchi et al., 2020; Faure & Kitchener, 2009; Lebarbenchon et al., 2010; Ottoni et al., 2013). Some of these later species, such as sheep (*Ovis aries*) and pigs (*Sus domesticus*), were then domesticated and are thus thought to have been introduced on these islands as early domestic animals, while others, such as the house mouse and most probably cats, migrated as commensal species (Cucchi et al., 2020; Faure & Kitchener, 2009). The domestication process of animals at those times must nevertheless have remained very primitive (Rezaei, 2007; Zeder, 2008, 2012), and populations that since inhabit these islands might then represent unique gene pools needing to be recognized as such taxonomically and to benefit from conservation strategies based on scientific knowledge, as evidenced for mouflons or weasels (Guerrini et al., 2015; Lebarbenchon et al., 2010; Portanier, Chevret, et al., 2022). This is particularly relevant to *Felis* species. Indeed, *F. s. lybica* remains that dated back to around 9500 years BP were discovered in Cyprus (Vigne et al., 2004). In addition, Sardinian wildcats are genetically closer to *F. s. lybica* than to *F. s. silvestris* (Mattucci et al., 2016; Randi et al., 2001) and show some morphological specificities that led to their consideration as a specific variety *F. s. lybica var. sarda* (Mura et al., 2013). This highlights the singularity of the Sardinian wildcat gene pool, which might be one of the unique representatives of ancestral *F. s. lybica* introduced to the Western Mediterranean. Island populations, indeed, often evolved with low levels of competition due to both geographic and genetic isolation, which results in a high level of endemism and the creation of biodiversity hotspots, which are high-priority conservation areas (Kier et al., 2008; Loso & Ricklefs, 2009; Myers et al., 2000; Whittaker & Fernández-Palacios, 2007).

Through time, Sardinia and Corsica have shared a large part of their history, being one single landmass during the last glaciation but also sharing similar history of colonization by humans and genetic ancestry (Grimaldi et al., 2001; Tamm et al., 2019 and references therein). A high phylogenetic proximity has also been evidenced for several species they host (e.g., mouflon, Portanier, Chevret, et al., 2022; weasel *Mustela nivalis*, Lebarbenchon et al., 2010; Corsican red deer *Cervus elaphus corsicanus*, Doan et al., 2017). While the presence of wildcats in Corsica has been known for a long time, with Lavauden describing them for the first time in 1929 as "*Felis reyi*" and Arrighi and Salotti mentioning them in a scientific

publication in 1988 (Arrighi & Salotti, 1988; Lavauden, 1929; see also Vigne, 1992), no further scientific studies have been performed on these animals. Based on fur characteristics, Arrighi and Salotti (1988) classified Corsican wildcats as *Felis silvestris lybica*, while Vigne (1988) classified other specimens as domestic cats returned to wildlife. Corsican people, nevertheless, have always distinguished “u ghjattu-volpe” (Corsican “cat-fox”), representing wildcats, and “u ghjattu insalvaticu”, representing feral cats. In accordance, the IUCN Cat Specialist Group reports Corsican wildcats as *Felis lybica reyi* (Kitchener et al., 2017). It is thus still unclear if these cats represent feral domestic cats or “true” wildcats and, in the latter case, if they are representatives of *Felis silvestris silvestris* or of *Felis silvestris lybica*, as in Sardinia. A major link is thus lacking in cat history in the Western Mediterranean and Europe. The main reason for the lack of recent studies on Corsican wildcats was that none have been detected in Corsica from the early 1980s to the 2000s. However, in 2008, a shepherd accidentally captured an animal that was probably a Corsican wildcat since it showed a similar phenotype to the one historically described (Arrighi & Salotti, 1988; Lavauden, 1929). This detection revived a monitoring program on the Corsican wildcat and during the 2010s, several noninvasive detections occurred through camera trapping. It allowed us to hypothesize that the Corsican wildcat was still present and to notice that these individuals were phenotypically similar across locations. Their phenotypes differed from domestic cat phenotypes and, as previously reported (Arrighi & Salotti, 1988; Lavauden, 1929), from European wildcat phenotypes. The Corsican phenotype was described as close to the Sardinian and African wildcat phenotype (see Arrighi & Salotti, 1988 and *Materials and Methods* below, but see Lavauden, 1929). Since then, multiple individuals that would phenotypically match the description of the Corsican wildcats have been detected and captured. It thus appeared urgent to resolve the taxonomic position of this Corsican wildcat phenotype.

Taxonomic uncertainties such as this one might indeed have deleterious effects on species conservation. On one hand, for the Corsican wildcat itself, since it would need to be officially recognized as a wild conservation unit to benefit from European conservation measures (Yamaguchi et al., 2015). On the other hand, it is necessary to confirm that Corsican wildcat are not feral cats since these may have significant detrimental effects on some vulnerable species, especially in island ecosystems (Nogales et al., 2013). If confirmed as a wild entity, Corsican wildcat conservation would also favour the preservation of numerous other species and habitat types since, like many carnivores, wildcats are an umbrella species (Jerosch et al., 2009, 2018; Noss et al., 1996; Virgos et al., 2002). It is a critical point since islands of the Mediterranean Basin are biodiversity hotspots hosting many endemic species (e.g., Escoriza & Hernandez, 2019; Grill et al., 2007; Jeanmonod et al., 2015) and particularly threatened by climate change (Courchamp et al., 2014; Ducrocq, 2016; Giorgi, 2006; Myers et al., 2000).

In the present study, the first one performed at such a scale on Corsican wildcats, we thus aimed to determine if Corsican wildcats are genetically distinct from domestic cats, Sardinian wildcats and

European wildcats using population genomics approaches. We were particularly interested in evaluating if the Corsican wildcat should be considered as a management unit (MU) with a unique gene pool (Moritz, 1994; Palsbøll et al., 2007), a result that would favour a better conservation of *Felis silvestris* and open discussions about the definition of a new *F. silvestris* subspecies. To achieve this goal, we used double-digest restriction-site-associated DNA sequencing (ddRADseq) to genotype samples of wild and domestic cats from both Corsica and mainland France as well as Sardinian wildcats at genome-wide SNPs. Our premise was that if Corsican wildcats, as a phenotypic entity, also represent a distinct molecular entity, then the Corsican wildcat deserves conservation measures and conservation unit delineation.

2 | MATERIALS AND METHODS

2.1 | Study sites, capture protocol and phenotypic data

Overall, 72 samples were available for DNA sequencing, among which 10 and 12 were domestic cats from Corsica and mainland France, respectively; 20 and 23 were wildcats from Corsica and mainland France, respectively; and seven were Sardinian wildcats (Tables 1, S1). Corsican wildcat samples were obtained from 13 locations on the island (Figure 1, Table S1). Samples were collected from individuals captured in box-traps baited with fish, road-killed or old dead animals examined by Dr. M. Salotti and Dr B. Condé (Natural History Museum of Nancy, France) (see Tables 1, S1). All of these samples showed homogeneous phenotypes (Figure 2) and corresponded to the criteria described by Dr M. Salotti (Salotti, 1992). The coat of the Corsican wildcat indeed differs from that of the European wildcat and the domestic tabby cat but is close to that of the Sardinian wildcat. It is composed of two phases, a fawn grey and a fawn brown, on which are present characteristic drawings of darker tones or even black. The ventral surface of the forelegs and hind legs are black, and black lines are drawn on the forehead, the cheeks, the occipital and cervical regions, the neck, the forelegs and the hind legs. A darker mediodorsal stripe is present. The tail ends in three well-formed black rings, which are wider on their dorsal part, and in two or three sketches of rings less and less visible toward the trunk of the animal (Figure 2).

For captured animals, box-traps were placed in habitats known to be used by Corsican wildcats thanks to a monitoring that occurred in autumn 2016 using camera traps. When captured, animals were anaesthetized using Domitor (0.1 mg/kg; Orion Pharma) and revived with Antisédan (0.2 mg/kg; Orion Pharma) to ensure a rapid reversal of sedation. Before release, ear punches and hairs were sampled for molecular analyses and photographs were taken for phenotypic description. Trapping and handling were performed by trained French Biodiversity Agency (OFB) officers according to the appropriate national laws governing animal welfare, following the ethical conditions detailed in the specific accreditations delivered by the Ministry in

TABLE 1 Sampling information for domestic and wildcats available for sequencing (see Table S1 for details of each sample).

Population	Subspecies	N	Sampling mode	Sampling years	Sample type
Domestic Corsican	<i>F. s. catus</i>	1	Road killed animal	2014	Ear punches in ethanol
		9	Lived veterinarian animals	2014	Hair
Domestic mainland	<i>F. s. catus</i>	12	Road killed animals	2010–2016	Ear punches in ethanol
Wild Corsican	<i>F. s. lybica?</i>	9	Capture	2016–2017	Ear punches in ethanol or dried, hairs, tail tissue, dried skin
		2	Dead animals	2010–2018	
		6	Dead animals	1986–1992	
		3	Museum specimen	1964–1966	
Wild mainland	<i>F. s. silvestris</i>	23	Road killed animals	1996–2016	Ear punches in ethanol
Wild Sardinia	<i>F. s. lybica</i>	7	Museum specimen, road-killed animals	NA	Dried skin, Ear punch or liver in ethanol

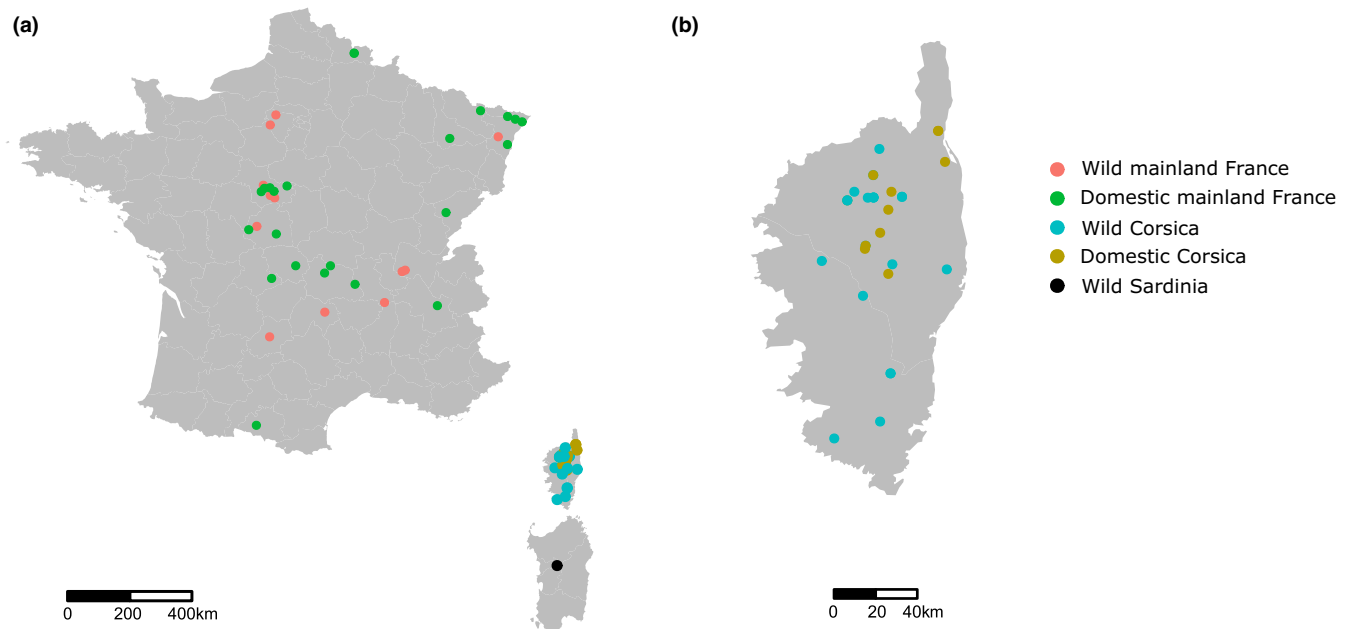


FIGURE 1 Maps of sampling locations in (a) mainland France and (b) in Corsica.

charge of ecology in agreement with the French environmental code (Art. L. 411–1 and L. 421–1 and R. 411–1). Domestic Corsican cats were sampled thanks to collaborations with veterinarians (Figure 1, Table 1). Mainland domestic and wildcat samples were obtained from road-killed animals monitored by the French Biodiversity Agency (Figure 1a). Assignment to either European wildcats *Felis silvestris silvestris* or domestic cats was confirmed genetically (following O'Brien et al., 2009; Devillard et al., 2014). Finally, Sardinian wildcats ($n = 7$, Table 1) were obtained from museum specimen and road-killed animals as a courtesy of M. Zedda.

2.2 | DNA extractions, ddRADseq library preparation and sequencing

Genomic DNA was extracted using either NucleoSpin Tissue (Macherey-Nagel) or Blood and Tissue (Qiagen) kits (see Table S1),

following the manufacturers' recommendations. Samples were thus lysed overnight at 56°C using proteinase K, and DNA was then purified and isolated using purification columns. The double-stranded DNA concentration of each sample was then quantified using a Qubit dsDNA HS Assay kit (Life Technologies), and purity was measured using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). ddRADseq libraries were prepared using 2–1200 ng of DNA according to sample quality (i.e., 1200 ng when possible, less otherwise, Table S1). The libraries were prepared based on the protocol of Peterson et al. (2012) but with modifications, in particular concerning the preparation of the adapters (see Henri et al., 2015). The detailed protocol can be found in Data S1. Briefly, it consisted in the double digestion of the DNA in each sample using *SbfI* and *MseI* enzymes (New England Biolabs), followed by ligation of P1-adaptors (including unique barcodes 8 and 10 bp long) and a common P2-adaptor. All samples were then pooled into a single sequencing library. Size selection of fragments (between 350 and 450 bp) was performed

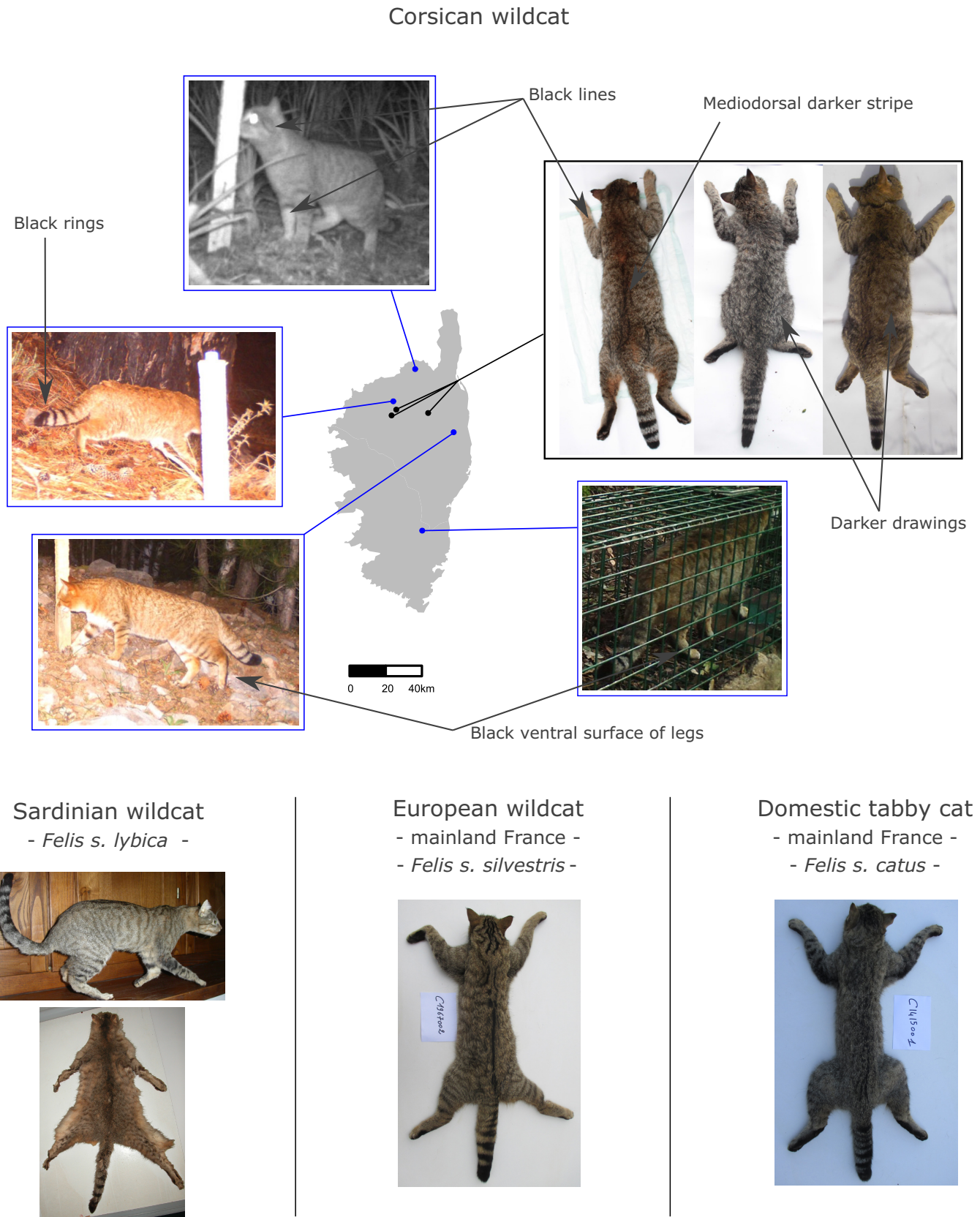


FIGURE 2 Corsican wildcat phenotype illustration. Pictures were obtained using camera traps (blue points) or during capture sessions (black point, anaesthetized animals). It is noteworthy that coat patterns do not always allow subspecies discrimination, especially when considering hybrid individuals (Devillard et al., 2014; O'Brien et al., 2009).

using Blue Pippin technology (Sage Science), and the Agilent 2100 Bioanalyser system (Agilent Technologies) was used to verify fragment sizes. The enrichment step was performed using 15 cycles of polymerase chain reaction (PCR) with the Library Amplification kit (Kapa Biosystems), and the enriched library was subjected to a final purification and quantification step (Qubit and a quantitative PCR, Library quantification kit for Illumina platform, Kapa Biosystems). The pooled library was sequenced on lanes of HiSeq 3000 and NovaSeq 6000 Illumina sequencers by the GeT-PlaGe sequencing platform (Genotoul) and allowed the procurement of 150-bp reads. As a first assessment of sequencing success, a single-end sequencing run was performed on 23 samples (five Corsican wildcats, nine mainland domestic cats, six mainland wildcats and three Sardinian wildcats). All these individuals, except the three Sardinian for which all available DNA was included in the first run (i.e., no more available DNA), were sequenced again using paired-end sequencing. This second sequencing run also included all other available samples (i.e., total of 20+49 samples). We thus obtained reads for the 72 previously mentioned individuals, among which 20 were sequenced twice (once single- and once paired-end). These 20 individuals served as sequencing duplicates to calculate sequencing error rate (see below). Two negative controls were also included in the paired-end run to control for contaminations.

2.3 | SNP calling and filtering

The quality of the raw reads was checked using FastQC version 0.11.5 (Andrews, 2010). Reads from each sample (unique individuals and sequencing duplicates, $n = 92$) were demultiplexed using the *process_radtags* module of the Stacks version 2.52 pipeline (Catchen et al., 2011; Rochette et al., 2019) using the *-c*, *-q* and *-r* options to remove any read with an uncalled base, discard reads with low-quality scores and rescue barcodes and RAD-tags. The adapter sequence search was also activated with two mismatches allowed. After demultiplexing, the total numbers of retained reads were 108,368,724 and 686,123,726 reads for single-end and paired-end sequencing, respectively. Reads (paired or single-end) were then mapped using the Burrow-Wheeler Aligner version 0.7.15 (BWA, Li & Durbin, 2009) and the *mem* algorithm with default parameters values against the reference genome of domestic cat *Felis_catus_9.0* (GenBank accession number GCA_000181335.4, Buckley et al., 2020, downloaded from Ensembl Genome Browser http://mart.ensembl.org/Felis_catus/Info/Index) previously indexed (*bwa -index*). Output files were then converted to BAM format and sorted using samtools version 1.9 (Li et al., 2009), which was also used to discard reads with a mapping score < 30.

The Stacks *refmap.pl* module was then used to build loci and identify and genotype SNPs for each sample. This first run allowed us to identify 23 samples poorly genotyped due to a low number of reads (all < 1.4 M with 20 individuals having < 100,000 reads). These individuals also showed a *gstacks* effective coverage < 10 and a high proportion of missing data (around 100% for most samples).

Unfortunately, all samples from Sardinia were among these low-quality samples. We thus reran *refmap.pl* as well as the *populations* module without the low-quality samples using two filtering settings: $-r = 0.80$, meaning that at least 80% of the individuals in a population are required to process a locus for that population, and $-p = 3$, meaning that a locus must be present in at least three populations to be processed. These parameters were chosen as a result of several runs of the *populations* module, varying *-r*, *-p* and *-R* (the minimum percentage of individuals across populations required to process a locus) to obtain a good compromise between the number of SNPs and the missing data quantity (per individuals and SNP). Only the first SNP per locus was kept, avoiding linkage disequilibrium issues in population genetics analyses (*write-single-snp* option). Using *populations* we also filtered the data set to exclude SNPs showing a heterozygosity higher than 70% (to avoid paralogous loci) and a minimal allelic frequency higher than 0.05 (i.e., an allele must be present at least X times to be kept, with $X = 0.05 \times 2N$, and $N =$ number of individuals; with 51 individuals, an allele must be present five times). SNPs were then additionally filtered using *vcftools* version 0.1.14 (Danecek et al., 2011) to keep only biallelic SNPs, alleles with a minor allele frequency > 0.05 (since few SNPs that were kept by Stacks were actually eliminated by the *Vcftools* filter for minor allele frequency), a maximal percentage of missing genotypes of 25% and genotypes with a minimal coverage of 10 and a maximal coverage of 780 (i.e., twice the mean depth). *Vcftools* was also used to calculate the mean depth per individual and the proportion of missing SNPs per individual (Table S1). After filtering, we used duplicates to calculate a sequencing error rate. To this end, we compared genotypes at each SNP for each pair of duplicates and calculated the error rate as the ratio of the number of SNP that had different genotypes and the total number of SNPs. We then excluded one of the two duplicates, keeping in the final data set the one showing the least missing data.

Using the same setups and filters, except for the *populations* module for which we parameterized $p = 5$, we aimed at obtaining a reduced data set in which some Sardinian individuals could be kept. We thus excluded three Sardinian individuals having < 5000 reads and proceeded with the four others (between $\approx 16,000$ and 85,000 reads). Using $-p = 5$ allowed us to keep only SNP sampled in all populations, including the Sardinian one, and we obtained a data set of 249 SNP in which all individuals had less than 15% of missing SNP (and 0% for the four Sardinian remaining individuals). This reduced data set was then processed, as the "large data set", as follows. Using the R version 4.1.0 (R Core Team, 2021) packages *vcfR* version 1.12.0 (Knaus & Grünwald, 2017), *pegas* version 1.0-1 (Paradis, 2010) and *q-value* version 2.24.0 (Storey et al., 2021), we excluded SNPs showing departure from Hardy-Weinberg (HW) equilibrium (q -value < 0.05) in at least one of the populations considered (i.e., tests were performed within each population since departure from Hardy-Weinberg equilibrium is expected due to possible genetic differentiation between European, Corsican, Sardinian wildcats and domestic cats). Finally, using the R package *pcadapt* version 4.3.3 (Luu et al., 2017; Privé et al., 2020), we filtered for outlier SNPs. The number of principal components (PCs) retained was chosen as

the one before the straight line in the scree-plot (i.e., subsequent PCs only accounted for random variation, Cattell's rule, as recommended by Luu et al., 2017). To be conservative, all SNPs showing p -values $< .05$ after adjustment through q -values, the false discovery rate (Benjamini & Hochberg, 1995) or Bonferroni correction, were excluded from further analyses.

2.4 | Population genetics analyses

Population genetic structure was investigated using both large and reduced data sets through several approaches. We first used a principal component analysis (PCA) performed using the R package *adeigenet* version 2.1.3 (Jombart, 2008; Jombart & Ahmed, 2011). To investigate the genetic structure more in depth we then used a discriminant analysis of principal components (DAPC, *adeigenet*), which allowed us to identify the optimal number of genetic clusters (K) using a k-means algorithm (*find.clusters*) and assign individuals to clusters (the DAPC itself). We ran *find.clusters* for K varying from 1 to 10, using 10,000 iterations and 1000 different starting points. The optimal number of clusters was identified as the one minimizing the BIC values, or, for competing solutions, as the smallest value for which there was a clear break with preceding values. A cross-validation procedure was then used to identify the optimal number of principal components to retain to perform the DAPC, which was subsequently applied to assign individuals to a cluster and calculate membership probabilities.

The Bayesian clustering algorithm implemented in the STRUCTURE software version 2.3.4 (Pritchard et al., 2000) was also used to cluster individuals without a priori information about their population of origin and determine their membership coefficients. We used the admixture and correlated allele frequency models for a varying number of clusters (K, from 1–5), with 10 independent repetitions for each K value and an MCMC length of 1,000,000 iterations (burnin: 100,000). The optimal number of clusters was determined using both the likelihood of each K ($\ln \Pr(X|K)$) and the method described by Evanno et al. (2005) as implemented in STRUCTURE HARVESTER version 0.6.94 (Earl & vonHodt, 2012). Independent runs for the optimal K were combined using CLUMPP version 1.1.2 (Jakobsson & Rosenberg, 2007) as implemented in CLUMPAK (Kopelman et al., 2015). We also applied the maximum likelihood approach implemented in ADMIXTURE version 1.3 (Alexander et al., 2009; Zhou et al., 2011) to estimate individual ancestries. Data sets were formatted using PLINK version 1.9 (Chang et al., 2015; Purcell et al., 2007). ADMIXTURE was run 10 times for each value of K from 1 to 5 and a 10-fold cross validation (CV) was applied to then determine the optimal K value. All other parameters were default parameters. Optimal K value was identified as the one minimizing the average CV-errors value across the 10 runs. For the optimal K value, independent runs were combined using CLUMPP as implemented in CLUMPAK.

Finally, based on a priori population assignments, global and between-populations genetic differentiation indices (F_{ST} , theta

estimator, Weir & Cockerham, 1984) were calculated using *StAMPP* R package (Pembleton et al., 2013). Confidence intervals and significance of F_{ST} values were obtained using bootstrap over loci ($n = 1000$ repetitions). The package *hierfstat* version 0.5–7 (Goudet & Jombart, 2020) was used to determine genetic diversity indices (observed and expected heterozygosity, allelic richness and F_{IS}), while nucleotide diversity (π) was obtained by running the *Stacks populations* module using only the SNPs that passed all previously described filters (*white-list*).

3 | RESULTS

3.1 | Sequencing and data filtering

The Fastqc quality check revealed no sequencing issues. The per base sequence quality was >28 all along reads 1 for paired-end sequencing and all along reads for single-end sequencing. Reads 2 showed a slightly lower quality (22–26) of the first three bases of reads. After demultiplexing, the mean number of retained reads per sample was 8,833,359 for paired-end sequencing and 4,152,531 for single-end sequencing. Negative controls obtained only few reads (of a total 1144 and 56,160, only 542 and 5696 were retained after *process_radtags* filtering; most of the reads were discarded because no Rad tags were detected, confirming that no DNA was present in the negative controls). When the 23 low quality samples were excluded, the mean *gstacks* effective coverage was 60x. Most (95%) of the primary alignments produced by BWA were kept. A total of 903,417 loci of an average of 260 bp were built, but most were discarded by *populations* filters. A final number of 11,884 present in at least three populations and in 80% of individuals within each population were obtained, leading to 5446 SNPs genotyped before SNP filtering.

After filtering using *vcftools*, 4249 SNPs remained, showing a mean coverage of 330x and a mean of 10.4% missing data. The genotyping error rate per individual, calculated thanks to the 20 duplicated individuals, was low (mean of 3.7%). After removing sequencing duplicates from the data set, 228 SNPs appeared to depart from HW equilibrium in at least one of the four populations, and 350 were detected as outliers by *pcadapt* (keeping four PCs, see Figure S1). These SNPs were thus excluded to obtain a final data set of 3671 genotyped in 51 unique individuals (9 wild Corsican, 9 domestic Corsican, 12 domestic mainland and 21 wild mainland cats) with moderated missing data proportions (2.3 and 3.8% for mainland wild and domestic cats, respectively, and 7.1 and 34.5% for Corsican wild and domestic cats, respectively). Domestic Corsican samples were of low quality (from hairs with low DNA quantity, see Table S1), which may explain such higher missing data rate. In the reduced data set (i.e., the one including four Sardinian individuals), 15 SNPs were excluded for departing from HW equilibrium in at least one of the five populations and 32 were detected as outliers by *pcadapt* (keeping three PCs). The final reduced data set included 202 SNP and 60 individuals (11 Corsican wildcats, 10 Corsican domestics, 12

mainland domestics, 23 mainland wildcats and 4 Sardinian wildcats). All individuals had <15% of missing data (mean = 4.01%, 2.33%, 3.25%, 1.97% and 0% for wild Corsican, domestic Corsican, domestic mainland, wild mainland and Sardinian wildcats, respectively).

3.2 | Population genetic structure including mainland European wildcats

The PCA performed to explore the population genetic structure revealed a clear genetic distinction between all four populations (Figure 3a). The first axis explained 27.2% of the variance contained in the data and separated wild mainland cats from all other individuals. The second axis explained 4.9% of the variation and showed a distinction between Corsican wildcats and domestic cats from both Corsica and mainland France. The pattern revealed by the PCA first axis was in accordance with patterns revealed by clustering approaches. Indeed, DAPC k-means and Bayesian STRUCTURE algorithms both pointed to the presence of two genetic clusters (see Figures S2 and S3; BIC was minimal for $K = 2$, while $\ln \Pr(X|K)$ started to plateau after $K = 2$). The DAPC, performed using one discriminant function and 20 PCs (cross-validation procedure), as well as the STRUCTURE approach evidenced a strong genetic distinction between European wildcats from mainland France and all

other individuals. In both analyses, one cluster grouped all mainland wildcats, while the second one was composed of all other individuals (domestics from mainland France and Corsica, wildcats from Corsica, Figure 3b,c). In the DAPC, the assignment success was one (i.e., assignment of k-means algorithm and DAPC itself were exactly the same), and posterior membership probabilities were also equal to one for each individual in its genetic cluster, indicating strong support for this genetic structure. In STRUCTURE, all individuals also had high posterior probability of membership: all were equal to 1 for cluster 1 (European wildcats) and ranged from 0.99 to 1 for cluster 2 except for one individual (Figure 3c). Similar results were observed using ADMIXTURE (Figures S4, S5).

Finally, F_{ST} values also confirmed the marked genetic differentiation between European wildcats from mainland France and all other populations and highlighted a significant genetic differentiation between Corsican wild and domestic cats as well as domestic cats from mainland France (global $F_{ST} = 0.40(0.39-0.41)_{IC95\%}$, see Table 2 for pairwise values). The lowest F_{ST} value was obtained between domestic cats from mainland France and from Corsica. Interestingly, European wildcats from mainland France exhibited the lowest genetic diversity (Table 3). Similar results were obtained while performing the PCA and calculating F_{ST} on the reduced data set including four Sardinian individuals (Figure 3d, Table 2). Indeed, wild mainland individuals exhibited large differentiation with all

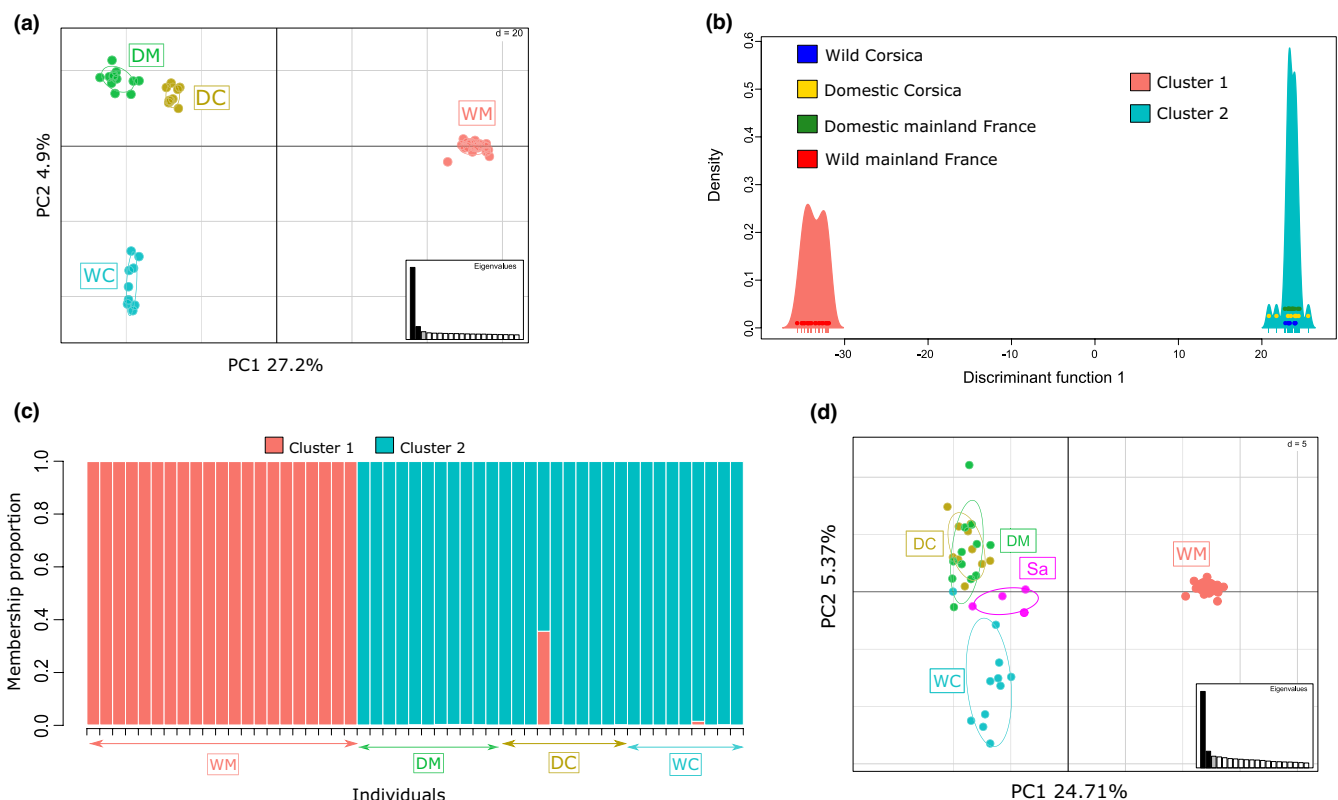


FIGURE 3 Genetic structure of mainland, Corsican and Sardinian wild and domestic cats investigated through (a) principal component analysis (PCA), (b) discriminant analysis of principal components (DAPC) and (c) STRUCTURE software. (d) PCA performed on the reduced data set including four Sardinian individuals. DM, domestics from mainland France; DC, domestics from Corsica; WC, wildcats from Corsica; Sa, wildcats from Sardinia; WM, wildcats from mainland France.

TABLE 2 Pairwise F_{ST} values and associated 95% confidence intervals and significance calculated between Corsican and mainland wild and domestic cats.

	WC	DC	DM	Sa ^a
DC	0.069 (0.061–0.077)***			
DM	0.067 (0.061–0.073)***	0.015 (0.011–0.019)***		
Sa ^a	0.043 (0.018–0.069)**	0.050 (0.016–0.92)**	0.033 (0.007–0.061)*	
WM	0.518 (0.505–0.530)***	0.547 (0.531–0.561)***	0.507 (0.494–0.519)***	0.693 (0.635–0.742)***

Abbreviations: DC, domestics from Corsica; DM, domestics from mainland France; Sa, wildcats from Sardinia; WC, wildcats from Corsica; WM, wildcats from mainland France.

*** $p < .001$, ** $p < .01$, * $p < .05$ after Bonferroni correction.

^aFor Sardinian individuals, F_{ST} values were calculated using the reduced data set (202 SNP).

TABLE 3 Mean \pm SD of genetic diversity indices calculated over loci for each population of either domestic or wild cats.

Population	n	H_o	H_e	A_r	Allele number	π	F_{IS}
WC	9	0.24 \pm 0.20	0.27 \pm 0.19	1.77 \pm 0.39	6282	0.23 \pm 0.003	0.08 \pm 0.35
DC	9	0.25 \pm 0.19	0.28 \pm 0.18	1.80 \pm 0.37	4474	0.22 \pm 0.003	0.09 \pm 0.34
DM	12	0.25 \pm 0.17	0.28 \pm 0.17	1.80 \pm 0.34	6850	0.24 \pm 0.003	0.08 \pm 0.29
Sa ^a	4	0.11 \pm 0.17	0.16 \pm 0.21	1.43 \pm 0.50	289	0.16 \pm 0.02	0.23 \pm 0.45
WM	21	0.11 \pm 0.14	0.12 \pm 0.15	1.43 \pm 0.39	5956	0.10 \pm 0.002	0.07 \pm 0.21

Abbreviations: A_r , allelic richness (rarefaction method, El Mousadik & Petit, 1996); H_e , expected heterozygosity; H_o , observed heterozygosity; π , nucleotide diversity calculated on variant positions, F_{IS} , averaged inbreeding coefficient overall loci. Population: DC, domestics from Corsica; DM, domestics from mainland France; Sa, wildcats from Sardinia; WC, wildcats from Corsica; WM, wildcats from mainland France.

^aFor Sardinian individuals, values were calculated using the reduced data set (202 SNP).

individuals (including Sardinian ones) and were separated from all other individuals in the first axis of the PCA (Figure 3d). The second axis separated Corsican wildcats from domestic individuals and Sardinian individuals were intermediate between these two groups (Figure 3d). Using ADMIXTURE, wild mainland individuals were also isolated from all other individuals (Figures S4, S12).

3.3 | Population genetic structure excluding mainland European wildcats

Since we were interested in determining if Corsican wildcats were genetically close to domestic cats, Sardinian wildcats or represented a unique gene pool, we repeated all the previous analyses but excluding European wildcats from mainland France. The *populations* module outputted a total of 6846 loci present in all three populations and in 80% of individuals within each population, leading to 3494 SNPs. After filtering using *vcftools* (maximal depth = 960), 3378 SNPs remained. Among these SNPs, 116 departed from HW equilibrium in at least one population, and 149 were detected as outliers by *pcadapt* (keeping three PCs, see Figure S6). The final data set consisted of 3113 SNPs genotyped in 30 unique individuals (9 wild Corsican, 9 domestic Corsican, 12 domestic mainland). It is noteworthy that, following removal of mainland European wildcats, the missing data rate was drastically reduced in Corsican domestic cats (2.05%). Among the 3113 SNPs genotyped, 156 exhibited private alleles for Corsican wildcats; 129 and 222 for Corsican and mainland

domestic cats, respectively. Using the reduced data set while excluding wildcats from mainland France, 258 SNP were genotyped by the *populations* module, among which 17 were excluded for departing from HW equilibrium and one was detected as outliers by *pcadapt* (keeping two PCs). A total of 240 SNPs was thus genotyped in 37 unique individuals (11 wild Corsican, 10 domestic Corsican, 12 domestic mainland and 4 Sardinian wild individuals).

For the large data set, the first axis of the PCA explained 7.08% of the variance contained in the data and separated wild Corsican cats from all domestic individuals (Figure 4a). The second axis explained 4.54% of the variation and showed some distinction between domestic Corsican cats and mainland domestic cats. The results of the DAPC k-means algorithm were less clear. Indeed, the BIC was minimal for $K = 1$, although a break in slope was visible for $K = 2$ (Figure S7). We nevertheless looked at clustering results when choosing $K = 2$ to evidence, if present, genetic differentiation between Corsican wildcats and domestic cats. In doing so, the DAPC (performed using 18 PCs [cross-validation procedure] and one discriminant function) revealed patterns in accordance with the one observed on the PCA. Indeed, one cluster grouped all Corsican wildcats, while the other grouped Corsican and mainland domestic cats (Figure 4b). Although it was not the optimal clustering solution, analyses performed using $K = 2$ received strong support since, in DAPC, the assignment success and the posterior membership probabilities were equal to one. It is noteworthy that, when performed using a priori assignments (i.e., based on sampling and not on k-means algorithm), the DAPC led to similar results and discriminated between Corsican wildcats

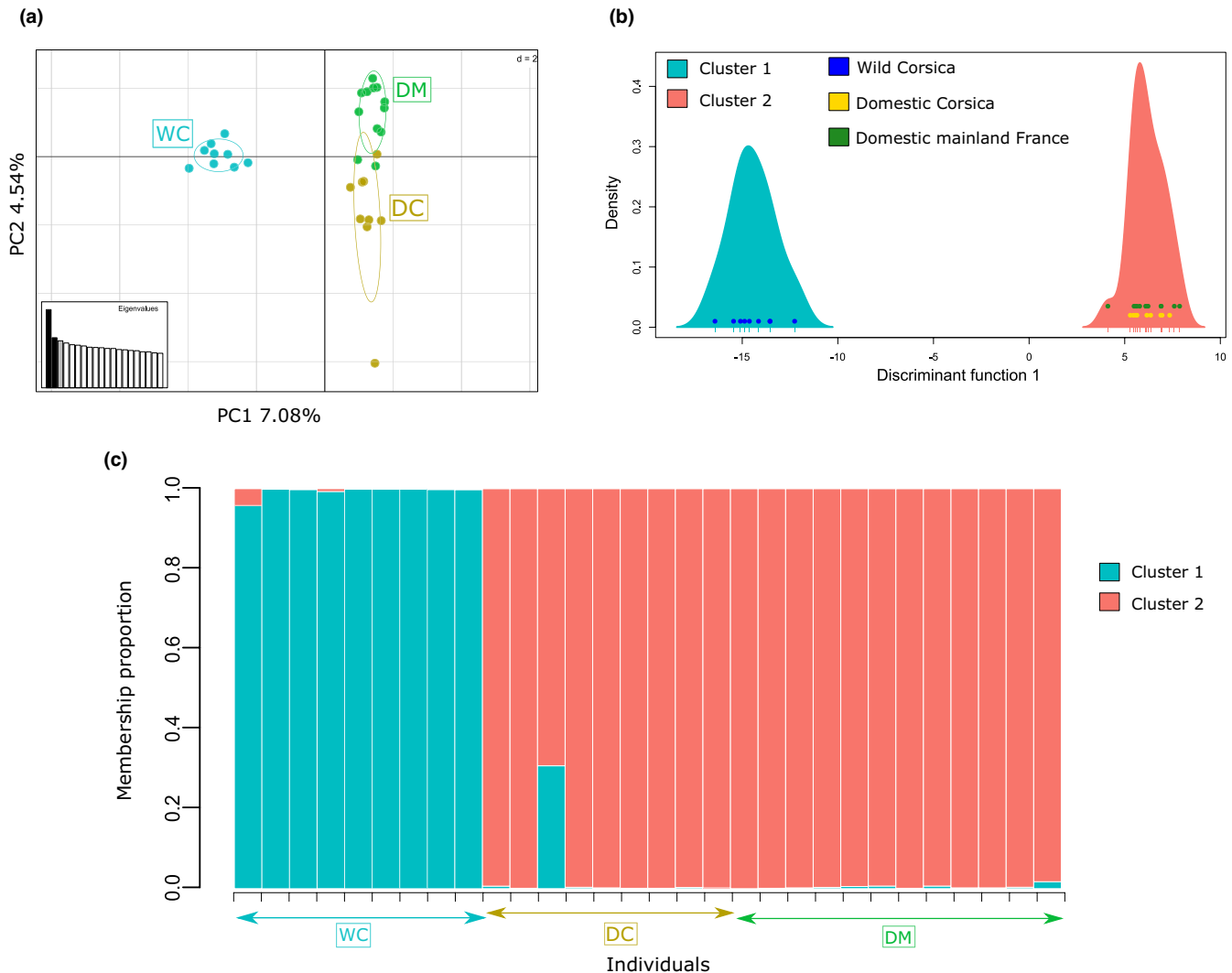


FIGURE 4 Genetic structure of wild Corsican cats, Corsican and mainland domestic cats investigated through (a) a principal component analysis (PCA), (b) a discriminant analysis of principal components (DAPC) and (c) STRUCTURE software. DM, domestics from mainland France; DC, domestics from Corsica; WC, wildcats from Corsica.

and domestic cats. Using the STRUCTURE Bayesian clustering approach, both the Evanno method and the $\ln \Pr(X|K)$ suggested that $K = 2$ was the optimal number of clusters (Figure S8). As previously, the two clusters separated Corsican wildcats and domestic cats from mainland France and Corsica, and membership proportions were very high in each cluster, ranging between 0.95 and 1 except for one domestic individual from Corsica (Figure 4c). Although using ADMIXTURE the most supported K value was $K = 1$, as with the DAPC, when considering $K = 2$ the two clusters separated Corsican wildcats and domestic cats with high membership probabilities (Figures S4, S5). Finally, F_{ST} values corresponded to previous observations, since global F_{ST} was $0.05(0.04-0.50)_{IC95\%}$, and pairwise F_{ST} values were also of this order of magnitude, with the lowest F_{ST} value obtained between domestic cat populations (Table 4).

When considering the reduced data set including four Sardinian individuals, the PCA confirmed the intermediate position of Sardinian wildcats between Corsican wildcats and domestic individuals as well as the clear separation of Corsican wildcats (Figure 5a).

TABLE 4 Pairwise F_{ST} values and associated 95% confidence intervals and significance calculated between wild and domestic Corsican cats and mainland domestic cats.

	WC	DC	DM
DC	0.065 (0.058–0.072)***		
DM	0.062 (0.056–0.068)***	0.014 (0.010–0.018)***	
Sa ^a	0.037 (0.013–0.059)**	0.044 (0.015–0.081)*	0.035 (0.009–0.062)*

Abbreviations: DC, domestics from Corsica; DM, domestics from mainland France; Sa, wildcats from Sardinia; WC, wildcats from Corsica. *** $p < .001$, ** $p < .01$, * $p < .05$ after Bonferroni correction.

^aFor Sardinian individuals, F_{ST} values were calculated using the reduced data set (240 SNP).

As for the large data set, the DAPC results were less clear since the BIC was minimal for $K = 1$ but a break in slope could be detected for $K = 2$ (Figure S9). Looking at clustering results with $K = 2$, Corsican

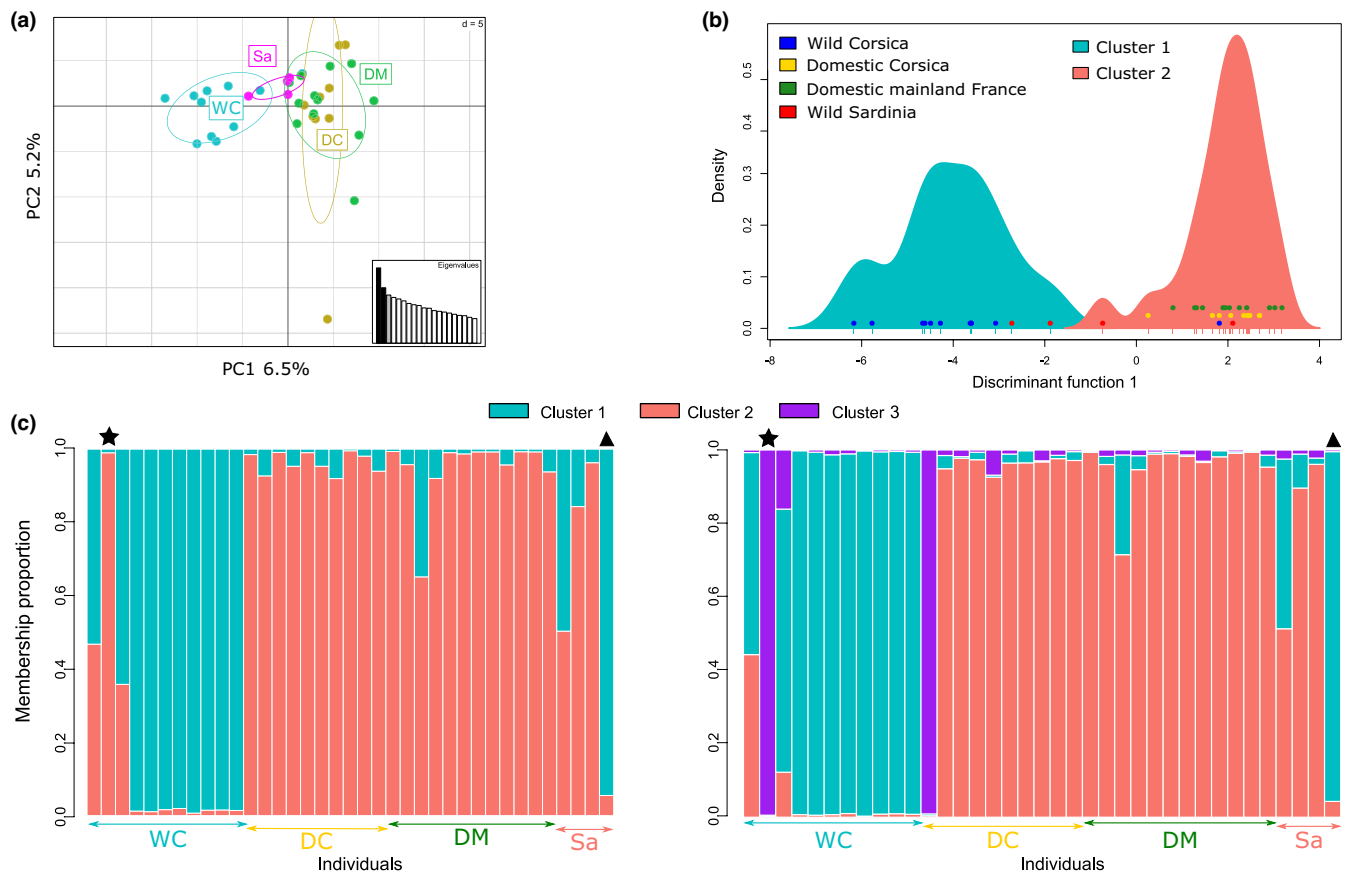


FIGURE 5 Genetic structure of mainland France, Corsican and Sardinian wild and domestic cats investigated through a (a) principal component analysis (PCA), (b) discriminant analysis of principal components (DAPC) and (c) STRUCTURE software. Analyses were performed on the reduced data set including four Sardinian individuals. DM, domestic from mainland France; DC, domestic from Corsica; WC, wildcats from Corsica; Sa, wildcats from Sardinia. ★ individual C1820001 and ▲ individual SAR005 (see main text).

wildcats grouped in one cluster while domestic cats grouped in the second one (Figure 5b). Two of the Sardinian individuals (SAR004 and SAR005, Table S1) grouped with Corsican wildcats and two grouped with domestic cats. One Corsican wildcat (C1820001, not present in the large data set) grouped with domestic cats. STRUCTURE also detected the presence of several genetic groups since Evanno's method pointed to $K = 4$ as the optimal number of cluster while $\ln \text{Pr}(X|K)$ suggested $K = 3$ (Figure S10).

We thus examined membership proportions for both $K = 2$ (as supported with the large data set), $K = 3$ and $K = 4$. Both $K = 2$ and $K = 3$ clustering solutions isolated Corsican wildcats in one cluster and grouped domestic cats in another one (Figure 5c). The same Corsican wildcat as previously was assigned to the domestic cluster (C1820001, star on Figure 5c) and two additional individuals showed intermediate membership proportions. Sardinian, while all except one (SAR005, triangle on Figure 5c) assigned to the domestic cluster, also showed some degree of admixture with Corsican wildcats. When considering $K = 3$, similar results were observed except that the third genetic group clustered one Corsican wildcat (C1820001) and one domestic cat from Corsica (Figure 5c). Membership probabilities when considering $K = 4$ showed similar results except that domestic cats were shared between clusters 2 and 4 (Figure S11).

ADMIXTURE analyses confirmed all these results, although $K = 1$ received the most support as the optimal the number of clusters (Figure S4) and considering $K = 3$ led to inconclusive results (Figure S12). Overall, ancestry coefficients were more intermediate using ADMIXTURE than using STRUCTURE (Figure S12). F_{ST} values agreed with an intermediate position of Sardinian wildcats between wild Corsican and domestics since the lowest F_{ST} value was observed between Sardinian and Corsican wildcats, followed by F_{ST} between Sardinian and domestics and the highest value was observed between Corsican wildcats and domestics (Table 4).

4 | DISCUSSION

By investigating, for the first time, Corsican wildcats using ddRAD sequencing and population genetics approaches, we evidenced their genetic uniqueness by revealing a very strong genetic differentiation with European wildcats and a non-negligible genetic differentiation with domestic cats from both Corsica and mainland France as well as with *F. s. lybica* from Sardinia. Domestic cats were genetically closer to each other than to Corsican wildcats, a result that, in conjunction with phenotypic particularities (coat patterns) and genetic

differentiation with Sardinian wildcats, opened the discussion about the necessity of recognizing the Corsican wildcat as a management unit in need of conservation measures and further ecological studies. Because of the low DNA quality of Sardinian samples, we only relied on a reduced data set including *F. s. lybica* but it suggested a higher genetic proximity of Corsican wildcats with Sardinian wildcats than with domestic cats. This called for further comprehensive studies including more Sardinian high-quality samples but also Sardinian domestic cats and mainland Near-Eastern *F. s. lybica* that will allow the precise repositioning of Corsican wildcats in the *Felis* phylogeny and to characterize the divergence between Sardinian and Corsican wildcats and original *F. s. lybica*.

4.1 | Genetic structure of Corsican wildcats

Although they have been known by the Corsican people since ancient times, Corsican wildcats have been little studied and never using genetic approaches. It was thus unknown if these individuals were feral domestic cats or “true” wildcats, as assumed given their phenotypic particularities reminiscent of those of wild *Felis silvestris* spp., and if these “ghjattu-volpe” diverged from European wildcats *Felis silvestris silvestris*. Here, both multivariate and genetic clustering approaches, as well as genetic differentiation measures, revealed a strong genetic divergence between European and Corsican wildcats ($F_{ST} = 0.50$). This confirmed that Corsican wildcats do not belong to the *F. s. silvestris* subspecies. The absence of mainland *F. s. lybica* in our data set prevented us from making direct conclusions about the genetic proximity between Corsican wildcats and *F. s. lybica*. However, the inclusion of few Sardinian individuals on a smaller number of SNPs revealed that Corsican wildcats were genetically closer to Sardinian wildcats than to domestic cats. Since Sardinian wildcats are described as *F. s. lybica*, and although that would need to be ascertained in future studies, this may argue in favour of a non-negligible genetic proximity between Corsican wildcats and *F. s. lybica*. In addition, both Corsican and Sardinian wildcats genetic proximity with domestic cats suggested that they might share a common ancestry with domestic cats and thus be wild representatives of *F. s. lybica*. Indeed, since domestic cats diverged from *F. s. lybica*, a genomic proximity between domestic cats and *F. s. lybica* has been observed (Driscoll et al., 2007).

It is even more true since, in cats, the domestication process is likely to have been less intensive than in other species and to have had less strong impacts on behavioural, morphological and genomic characteristics (Faure & Kitchener, 2009; Mattucci et al., 2019; Montague et al., 2014). Preliminary analyses using mitochondrial DNA additionally confirmed the genetic proximity of *F. s. lybica* (including one individual from Sardinia), domestic cats and Corsican wildcats (Data S2, Figure S13, Tables S2–S3). This corresponded to the intermediate position we observed for Sardinian individuals between domestic cats and Corsican wildcats (Figure 5). Alternatively, this intermediate position, the assignment of some individuals to the domestic cluster and others to the Corsican clusters and the presence

of admixed individuals (Figure 5) may suggest that Sardinian wildcats actually derived from Corsican wildcats and are nowadays victim of domestic introgression. In addition, increasing the number of genetic clusters did not separate Sardinian wildcats but instead created spurious results (Figures 5, S11, S12). However, in a previous study using microsatellite markers, Mattucci et al. (2013) detected no admixture between domestic cats and Sardinian wildcats. The intermediate position we observed in the present study may thus result from a sampling bias toward hybrid individuals. A comprehensive sampling scheme including more Sardinian wildcats and also Sardinian domestic cats as well as mainland *F. s. lybica* would allow to clarify the relations between Sardinian and Corsican wild and domestic cats.

The non-negligible level of genetic differentiation existing between Corsican wildcats and domestic cats from both the mainland and Corsica, and more importantly the closer proximity of domestics from the mainland and Corsica, suggested that Corsican wildcats are not domestic feral cats. Indeed, if the Corsican wildcat was feral, we would have expected a higher genetic proximity between wild and domestics from Corsica than between domestics from mainland and Corsica. This close proximity of domestics from distant geographic areas might be in line with the differing histories of domestic and wild cats. Cat domestication started in the Near East around 10,000 years ago, and these commensal organisms then probably colonized Central Europe by following Neolithic farmers (Baca et al., 2018). Phylogenetic analyses nevertheless also revealed that the genetic make-up of contemporary domestic cats actually resulted from both Near Eastern and Egyptian cats, the latter colonizing Europe much later, during the Roman Empire (around 2000 years BP, Ottoni et al., 2017; Baca et al., 2018). This genetic history of domestic cats might explain the higher genetic proximity observed between Corsican and mainland domestic cats, arguing in favour of a different origin of Corsican wild and domestic cats. The former could be a representative of Neolithic commensal *F. s. lybica*, and the latter a representative of contemporary domestic cats introduced much later on the island.

The degree of genetic differentiation we observed between domestic and Corsican wildcats ($F_{ST} = 0.06–0.07$, Table 4) was in accordance with values observed between *F. s. lybica* from Sardinia and North Africa and domestic cats, although differences in the genetic markers used makes the comparison challenging (Mattucci et al., 2016). It is, in addition, noteworthy that, in the SNP panels investigated here the genome representation might be biased. Indeed, by requiring that loci should be present in at least 80% of individuals in all populations ($r = 0.8$, $p = 3$ or $p = 4$ in Stacks analyses), while we considered probably divergent populations, we might have sampled the SNPs genotyped in the most conserved genome regions, leading to an underestimation of the genetic differentiation existing between populations. In addition, we considered SNP panels excluding SNP identified as outliers and thus showing the largest F_{ST} values (*pcadapt*). The loci may be outliers for different reasons (e.g., being under selection) but may also reflect the genetic divergence between populations. Along with the reduction of statistical power due to the small number of individuals involved in our analyses, such

conservative genome representation might explain why the DAPC pointed to $K = 1$ as the optimal number of clusters. This conservative genome representation may also explain why European wildcats showed the lowest genetic diversity (Table 3). Indeed, since this population was the most divergent, sampling SNP in the most conserved regions of the genome may have led to sample less variable portions of the European wildcat genome. The relatively low genetic diversity observed here do not correspond to what was previously observed in French European wildcats (Portanier, Léger, et al., 2022).

4.2 | Conservation implications

The genetic distinction of Corsican wildcats from domestic, European and Sardinian wildcats calls for conservation considerations. The Mediterranean basin is one of the most threatened biodiversity hotspots (Ducrocq, 2016; Giorgi, 2006; Myers et al., 2000), and identifying conservation units on Mediterranean islands might favour its preservation by allowing its biodiversity to be recognized at a legislative level (Coates et al., 2018). Based on our assessment of Corsican wildcat population genetic structure, we argue that this possibly unique organism should be considered as a conservation unit within the *Felis silvestris* species complex. Several denominations exist on the population-species continuum when speaking of conservation units, such as evolutionarily significant units (ESUs) and management units (MUs) (see Moritz, 1994; Palsbøll et al., 2007; Funk et al., 2012; Coates et al., 2018 for reviews). These concepts of infraspecific categories have definitions that may still be under debate and that evolve through time, but some consensus seems to emerge from the scientific literature. It is nowadays mostly accepted that ESUs should be defined as “populations that have substantial reproductive isolation, which has led to adaptive differences so that the population represents a significant evolutionary component of the species” (Palsbøll et al., 2007). While the ESUs concept includes notions of adaptive divergence and reproductive isolation, the MUs concept relies on the demographic independence of populations since they are defined as “populations of conspecific individuals among which the degree of connectivity is sufficiently low so that each population should be monitored and managed separately” (Palsbøll et al., 2007). To be recognized as an ESU, Corsican wildcats would need to be replaced more precisely in the *Felis silvestris* spp. phylogeny, and a study on adaptive genetic markers would be necessary to evidence adaptive divergence between this population and other ones, such as the Sardinian one and mainland Near-Eastern *F. s. lybica*. While the present study does not allow to make conclusions about ESUs, we nevertheless suggest that Corsican wildcats constitute a MU, since they inhabit an island and, to our knowledge, no translocations have occurred from Sardinia or other wildcat populations for many thousand years. It is noteworthy that the genetic differentiation detected between Sardinian and Corsican wildcat may result from this long-term isolation, which may have led to adaptive divergence but also favoured strong genetic drift if both populations have low population sizes (no data available at present).

5 | CONCLUSION AND PERSPECTIVES

Investigating, for the first time, Corsican cats that have been described as wild for a very long time by the Corsican people, and using genome-wide SNPs obtained through next-generation sequencing, we evidenced that Corsican wildcats, “ghjattu-volpe”, might represent a new unique gene pool in Corsica and a Management Unit within the *F. silvestris* spp. complex. Although the sample size and data set composition prevented us from making conclusions about the definition of the Corsican wildcat as an ESU or a subspecies, our results open the way to further studies. Description of the Corsican wildcat as a subspecies would aid immensely in the conservation of this MU, since species and subspecies are better recognized legislatively than other infra-specific categories, especially in the European laws (Coates et al., 2018). Our study also illustrated how using relatively low-cost sequencing strategies and reduced genome representation can be useful for conservation purpose. It illustrated, for instance, how these approaches can reveal cryptic diversity previously overlooked using less powerful genetic markers and bring enthusiastic news in the current context of extinction crisis. It also emphasized the importance of focusing more on understudied Corsican fauna, and more generally on understudied insular fauna, which can be made up of numerous undiscovered MUs or ESUs.

Future works on the Corsican wildcat should focus on repositioning it more precisely in the *Felis silvestris* phylogeny and on evolutionary history by including more *F. s. lybica* samples from mainland Near East and from Sardinia and investigating adaptive divergence as well as demographic history, but also on gaining knowledge of their ecology. Implementing comprehensive biological monitoring would allow the effective conservation of the Corsican wildcat by bringing knowledge on their ecological niche, habitat suitability, spatial distribution and spatial ecology, the dynamics of their populations, their population genetic structure at finer spatial scales, including both landscape genetic connectivity and hybridization with domestic cat investigations. All these knowledges would help to protect Corsican wildcats but also bring important information about their interaction with other Corsican species, which would in turn improve discussions and debates about the conservation of anthropochorous species (Gipoliti & Amori, 2002, 2006). In addition to the investigation of demographic history, an urgent pursuit of the present study regards hybridization since hybridization between mainland wild and domestic cats has been reported (Beugin et al., 2016, 2020; Say et al., 2012) and might also occur between Corsican wild and domestic cats. Further studies involving higher sample sizes and sampling wild and domestic individuals living in close proximity would allow to determine if hybridization with domestic cats does not threaten Corsican wildcats as it does for European ones (Yamaguchi et al., 2015). For this purpose, the ddRAD sequencing data generated in the present study could allow to derive ancestry informative markers for Corsican wildcat and SNP particularly suitable to investigate introgression, as previously done

for European wildcats and domestic cats (Mattucci et al., 2019; Nussberger et al., 2014). If future studies reveal an absence or a low rate of hybridization with domestic cats, Corsican wildcats could be seen as living fossils of African-Near Eastern wildcats introduced several thousands of years ago, offering astonishing opportunities to study cat domestication.

AUTHOR CONTRIBUTIONS

E.P., P.C., P.B., F.S., S.R. and S.D. conceptualized and designed the research. P.B. and F.S. supervised the fieldwork in Corsica. M.Z. provided samples from Sardinia. H.H. and C.R. conducted laboratory steps. E.P. performed data analyses with the contribution of H.H., A.E.F. and P.C. during the first steps of the project. E.P. wrote the manuscript with feedback from S.D., S.R. and H.H. and all the authors. All authors contributed to interpreting the results and approving the manuscript.

ACKNOWLEDGEMENTS

We warmly thank all the professionals from the Office Français de la Biodiversité (OFB) for their technical support in sampling the different populations, especially Charles-Antoine Cecchini, Valérie Grisoni, Laurence Henry and François Léger. We also gratefully acknowledge the CC Laboratoire de Biométrie et Biologie Evolutive/ Pôle Rhône-Alpes de Bioinformatique (PRABI) for providing computer resources, Fanny Dens as a student for her field and laboratory work and Jeremy Larroque for helpful discussions about SNP filtering and data analyses. We are grateful to veterinarians Dr Marc Memmi, Dr Bernard Fabrily, Dr Anne Guiard-Marigny, Dr Nathalie Kilburg, Dr Marion Terrazoni, Dr Claud d'Angeli and Dr Pascal Jugnet for collecting Corsican domestic cat samples and to the Museum of Natural History of Nancy, the Museum of Natural History of Strasbourg and Dr M. Salotti (Corsican University) who collected samples from furs. We also thank the sequencing GeT-PlaGe sequencing platform (Genotoul, Castanet-Tolosan, France). This research was funded by the Office de l'Environnement de la Corse (OEC), the OFB, the Direction Régionale de l'Environnement, de l'Aménagement et du Logement Corse (DREAL), the University of Lyon and the Laboratoire de Biométrie et Biologie Evolutive. We are also grateful to the two anonymous reviewers for their useful comments that helped us to improve the quality of this study.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

Raw sequence reads and related metadata have been deposited in the SRA BioProject PRJNA811807. Sample accession numbers are reported in Table S1. The final SNP data sets are available from the FigShare repository (<https://doi.org/10.6084/m9.figshare.19298864.v1>). The 65 new mitochondrial sequences are available from European Nucleotide Archive with accession nos. OW373342-OW373406.

BENEFIT-SHARING STATEMENT

Benefits generated: A research collaboration was developed with scientists from the countries that provided genetic samples. All collaborators are included as coauthors. The results of research have been shared with the provider communities and the broader scientific community (see above) and the research addresses a priority concern, in this case, the conservation of organisms being studied. More broadly, our group is committed to international scientific partnerships, as well as institutional capacity building.

ORCID

Elodie Portanier  <https://orcid.org/0000-0002-6566-6263>

Pascale Chevret  <https://orcid.org/0000-0002-4186-875X>

REFERENCES

- Albert, C., Luque, G. M., & Courchamp, F. (2018). The twenty most charismatic species. *PLoS One*, *13*, e0199149.
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Anile, S., & Devillard, S. (2020). Spatial variance-mass allometry of population density in felids from camera-trapping studies worldwide. *Scientific Reports*, *10*, 14814.
- Arrighi, J., & Salotti, M. (1988). Le chat sauvage (*Felis silvestris*) en Corse: une certitude. *Mammalia*, *52*, 123–125.
- Baca, M., Popovic, D., Panagiotopoulou, H., Marciszak, A., Krajcarz, M., Krajcarz, M. T., Makowiecki, D., Weglenski, P., & Nadachowski, A. (2018). Human-mediated dispersal of cats in the Neolithic Central Europe. *Heredity*, *121*, 557–563.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*, 289–300.
- Beugin, M. P., Leblanc, G., Queney, G., Natoli, E., & Pontier, D. (2016). Female in the inside, male in the outside: Insights into the spatial organization of a European wildcat population. *Conservation Genetics*, *17*, 1405–1415.
- Beugin, M. P., Salvador, O., Leblanc, G., Queney, G., Natoli, E., & Pontier, D. (2020). Hybridization between *Felis silvestris silvestris* and *Felis silvestris catus* in two contrasted environments in France. *Ecology and Evolution*, *10*, 263–276.
- Brummitt, N. A., Bachman, S. P., Griffiths-Lee, J., Lutz, M., Moat, J. F., & Farjon, A. (2015). Green plants in the red: A baseline global assessment for the IUCN sampled red list index for plants. *PLoS One*, *10*(8), Article e0135152.
- 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., Minx, P., Tomlinson, C., Lyons, L. A., Murphy, W. J., & Warren, W. C. (2020). A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *PLoS Genetics*, *16*, Article e1008926.
- Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182.
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., Garcia, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*, Article e1400253.

- Ceballos, G., Ehrlich, P. R., & Raven, P. H. (2020). Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *PNAS*, *117*, 13596–13602.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, s13742-015-0047-8.
- Coates, D. J., Byrne, M., & Moritz, C. (2018). Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*, *6*, Article 165.
- Courchamp, F., Hoffman, B. D., Russel, J. C., Leclerc, C., & Bellard, C. (2014). Climate change, sea-level rise, and conservation: Keeping Island biodiversity afloat. *Trends in Ecology & Evolution*, *29*, 127–131.
- Cowie, R. H., Bouchet, P., & Fontaine, B. (2022). The sixth mass extinction: Fact, fiction or speculation? *Biological Reviews*, *97*, 640–663.
- Cowie, R. H., Régnier, C., Fontaine, B., & Bouchet, P. (2017). Measuring the sixth extinction: What do mollusks tell us? *The Nautilus*, *131*, 3–41.
- Cucchi, T., Papayianni, K., Cersoy, S., Aznar-Cormano, L., Zazzo, A., Debruyne, R., Bălăşescu, A., Simmons, A., Valla, F., Hamilakis, Y., Mavridis, F., Mashour, M., Darvish, J., Siaharsari, R., Biglari, F., Petrie, C. A., Weeks, L., Sardari, A., Maziari, S., ... Vigne, J. D. (2020). Tracking the near eastern origins and European dispersal of the western house mouse. *Scientific Reports*, *10*, Article 8276.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R., Lunter, G., Marth, G., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158.
- Devillard, S., Jombart, T., Léger, F., Pontier, D., Say, L., & Ruetter, S. (2014). How reliable are morphological and anatomical characters to distinguish European wildcats, domestic cats and their hybrids in France? *Journal of Zoological Systematics and Evolutionary Research*, *52*, 154–162.
- DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, *30*, 4147–4154.
- Diaz, S., Settele, J., Brondizio, E., et al. (2019). Pervasive human-driven decline of life on earth points to the need for transformative change. *Science*, *366*, eaax3100.
- Doan, K., Zachos, F. E., Wilkens, B., Vigne, J. D., Piotrowska, N. N., Stankovic, A., Jedrzejska, B., Stefaniak, K., & Niedzialkowska, M. (2017). Phylogeography of the Tyrrhenian red deer (*Cervus elaphus corsicanus*) resolved using ancient DNA of radiocarbon-dated subfossils. *Scientific Reports*, *7*, 2331.
- Driscoll, C. A., Macdonald, D. W., & O'Brien, S. J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *PNAS*, *106*, 9971–9978.
- 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., Yamaguchi, N., O'Brien, S. J., & Macdonald, D. W. (2007). The near eastern origin of cat domestication. *Science*, *317*, 519–523.
- Ducrocq, V. (2016). Climate change in the Mediterranean region. In J.-P. Moatti & S. Thiébaud (Eds.), *The Mediterranean region under climate change: A scientific update*. IRD Editions.
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361.
- El Mousadik, A., & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, *92*, 832–839.
- Escoriza, D., & Hernandez, A. (2019). Using hierarchical spatial models to assess the occurrence of an Island endemism: The case of *Salamandra corsica*. *Ecological Processes*, *8*, 15.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*, 2611–2620.
- Faure, E., & Kitchener, A. C. (2009). An archaeological and historical review of the relationships between felids and people. *Anthrozoös*, *22*, 221–238.
- Frankham, R., Ballou, J. D., & Briscoe, D. A. (2004). *A primer of conservation genetics*. Cambridge University Press.
- Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, *27*, 489–496.
- Giorgi, F. (2006). Climate change hot-spots. *Geophysical Research Letters*, *33*, L08707.
- Gippoliti, S., & Amori, G. (2002). Anthropochorous wild mammal taxa and conservation lists. *Conservation Biology*, *16*, 1162–1164.
- Gippoliti, S., & Amori, G. (2006). Ancient introductions of mammals in Mediterranean Basin and their implications for conservation. *Mammal Review*, *36*, 37–48.
- Goudet, J., Jombart, T. (2020). Hierfstat: Estimation and tests of hierarchical F-statistics. R Package Version 0.5-7. <https://CRAN.R-project.org/package=hierfstat>
- Grill, A., Casula, P., Lecis, R., & Menken, S. (2007). Endemism in Sardinia. In S. Weiss & N. Ferrand (Eds.), *Phylogeography of southern European refugia*. Springer.
- Grimaldi, M. C., Crouau-Roy, B., Amorós, J. P., Cambon-Thomsen, A., Carcassi, C., Orru, S., Viader, C., & Contu, L. (2001). West Mediterranean islands (Corsica, Balearic Islands, Sardinia) and the Basque population: Contribution of HLA class I molecular markers to their evolutionary history. *Tissue Antigens*, *58*, 281–292.
- Guerrini, M., Forcina, G., Panayides, P., Lorenzini, R., Garel, M., Anayiotos, P., Kassinis, N., & Barbanera, F. (2015). Molecular DNA identity of the mouflon of Cyprus (*Ovis orientalis ophion*, Bovidae): Near eastern origin and divergence from Western Mediterranean conspecific populations. *Systematics and Biodiversity*, *13*, 472–483.
- Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörrén, T., Goulson, D., & de Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One*, *12*(10), Article e0185809.
- Henri, H., Cariou, M., Terraz, G., Martinez, S., El Filali, A., Veysiere, M., Duret, L., & Charlat, S. (2015). Optimization of multiplexed RADseq libraries using low-cost adaptors. *Genetica*, *143*, 139–143.
- Hoban, S., Bruford, M. V., Funk, W. C., et al. (2021). Global commitments to conserving and monitoring genetic diversity are now necessary and feasible. *Bioscience*, *71*, 964–976.
- Hohenlohe, P., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*, 62–82.
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806.
- Jeanmonod, D., Naciri, Y., Schlüssel, A., & Gamisans, J. (2015). Floristic analyses of the Corsican flora: Biogeographical origin and endemism. *Candollea*, *70*, 21–41.
- Jerosch, S., Götz, M., Klar, N., & Roth, M. (2009). Characteristics of diurnal resting sites of the endangered European wildcat (*Felis silvestris silvestris*): Implications for its conservation. *Journal for Nature Conservation*, *18*, 45–54.
- Jerosch, S., Kramer-Schadt, S., Götz, M., & Roth, M. (2018). The importance of small-scale structures in an agriculturally dominated landscape for the European wildcat (*Felis silvestris silvestris*) in Central Europe and implications for its conservation. *Journal for Nature Conservation*, *41*, 88–96.
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.
- Jombart, T., & Ahmed, I. (2011). ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

- Kier, G., Kreft, H., Ming Lee, T., Jetz, W., Ibish, P. L., Nowicki, C., Mutke, J., & Barthlott, W. (2008). A global assessment of endemism and species richness across Island and mainland regions. *PNAS*, *106*, 9322–9327.
- Kitchener, A. C., Breitenmoser-Würsten, C., Eizirik, E., Gentry, A., Werdelin, L., Wilting, A., Yamaguchi, N., Abramov, A. V., Christiansen, P., Driscoll, C., Duckworth, J. W., Johnson, W., Luo, S.-J., Meijaard, E., O'Donoghue, P., Sanderson, J., Seymour, K., Bruford, M., Groves, C., ... Tobe, S. (2017). A revised taxonomy of the Felidae. The final report of the cat classification task force of the IUCN/SSC cat specialist group. *Cat News Special Issue*, *11*, 80.
- Knaus, B. J., & Grünwald, N. J. (2017). VCFR: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*, 44–53.
- Kopelman, N. M., Mayze, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191.
- Lavauden, M. L. (1929) Sur le chat sauvage de la Corse. Comptes rendus hebdomadaires des séances de l'Académie des sciences, séance du 2 décembre 1929, 1923-1924. Accessed on Gallica - Bibliothèque nationale de France. <https://gallica.bnf.fr/ark:/12148/bpt6k3142j/f1027.image>.
- Lebarbenchon, C., Poitevin, F., Arnal, V., & Montgelard, C. (2010). Phylogeography of the weasel (*Mustela nivalis*) in the western-Palaeartic region: Combined effects of glacial events and human movements. *Heredity*, *105*, 449–462.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with BurrowsWheeler transform. *Bioinformatics*, *25*, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, *25*, 2078–2079.
- Loso, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, *457*, 830–836.
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*, 67–77.
- Mattucci, F., Galaverni, M., Lyons, L. A., Alves, P. C., Randi, E., Velli, E., Pagani, L., & Caniglia, R. (2019). Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Scientific Reports*, *9*, 11612.
- Mattucci, F., Oliveira, R., Bizzarri, L., Vercillo, F., Anile, S., Ragni, B., Lapini, L., Sforzi, A., Alves, P. C., Lyons, L. A., & Randi, E. (2013). Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecology and Evolution*, *3*, 2443–2458.
- Mattucci, F., Oliveira, R., Lyons, L. A., Alves, P. C., & Randi, E. (2016). European wildcat populations are subdivided into five main biogeographic groups: Consequences of Pleistocene climate changes or recent anthropogenic fragmentation? *Ecology and Evolution*, *6*, 3–22.
- Montague, M. J., Li, G., Gandolfi, B., Khan, R., Aken, B. L., Searle, S. M. J., Minx, P., Hillier, L. W., Koboldt, D. C., Davis, B. W., Driscoll, C. A., Barr, C. S., Blackstone, K., Quilez, J., Lorente-Galdos, B., Marques-Bonet, T., Alkan, C., Thomas, G. W. C., Hahn, M. W., ... Warren, W. C. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *PNAS*, *111*, 17230–17235.
- Moritz, C. (1994). Defining 'evolutionary significant units' for conservation. *Trends in Ecology & Evolution*, *9*, 373–375.
- Mura, A., Gadau, S., & Zedda, M. (2013). Morphological and morphometrical features of Sardinian wild cat (*Felis silvestris lybica* var. *sarda*). *Journal of Biological Research*, *86*, 35–37.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–858.
- Nilson, S. M., Gandolfi, B., Grahn, R. A., Kurushima, J. D., Lipinski, M. J., Randi, E., Waly, N. E., Driscoll, C., Murua Escobar, H., Schuster, R. K., Maruyama, S., Labarthe, N., Chomel, B. B., Ghosh, S. K., Ozpinar, H., Rah, H. C., Millán, J., Mendes-de-Almeida, F., Levy, J. K., ... Lyons, L. A. (2022). Genetics of randomly bred cats support the cradle of cat domestication being in the near east. *Heredity*, *129*, 346–355.
- Nogales, M., Vidal, E., Medina, F. M., Bonnaud, E., Tershy, B. R., Campbell, K. J., & Zavaleta, E. S. (2013). Feral cats and biodiversity conservation: The urgent prioritization of Island management. *Bioscience*, *63*, 804–810.
- Noss, R. F., Quigley, H. B., Hornocker, M. G., Merrill, T., & Paquet, P. C. (1996). Conservation biology and carnivore conservation in the Rocky Mountains. *Conservation Biology*, *10*, 949–963.
- Nussberger, B., Wandeler, P., & Camenisch, G. (2014). A SNP chip to detect introgression in wildcats allows accurate genotyping of single hairs. *European Journal of Wildlife Research*, *60*, 405–410.
- O'Brien, J., Devillard, S., Say, L., Vanthomme, H., Leger, F., Ruetten, S., & Pontier, D. (2009). Preserving genetic integrity in a hybridising world: Are European wildcats (*Felis silvestris silvestris*) in eastern France distinct from sympatric feral domestic cats? *Biodiversity and Conservation*, *18*, 2351–2360.
- Otoni, C., Girdland, F. L., Evin, A., Georg, C., De Cupere, B., & Van Neer, W. (2013). Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Molecular Biology and Evolution*, *30*, 824–832.
- Otoni, C., Van Neer, W., De Cupere, B., Daligault, J., Guimaraes, S., Peters, J., Spassov, N., Prendergast, M. E., Boivin, N., Morales-Muñiz, A., Bălăşescu, A., Becker, C., Benecke, N., Boroneant, A., Buitenhuis, H., Chahoud, J., Crowther, A., Llorente, L., Manaseryan, N., ... Geigl, E. M. (2017). The palaeogenetics of cat dispersal in the ancient world. *Nature Ecology and Evolution*, *1*, 0139.
- Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology & Evolution*, *22*, 11–16.
- Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*, 419–420.
- Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*(5), 946–952.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), Article e37135.
- Poplin, F. (1979). Origine du mouflon de Corse dans une nouvelle perspective paléontologique, par marronnage. *Annales de génétique et de sélection Animale*, *11*, 133–143.
- Portanier, E., Chevret, P., Gélén, P., Benedetti, P., Sanchis, F., Barnabera, F., Kaerle, C., Queney, G., Bourgoïn, G., Devillard, S., & Garel, M. (2022). New insights into the past and recent evolutionary history of the Corsican mouflon (*Ovis gmelini musimon*) to inform its conservation. *Conservation Genetics*, *23*, 91–107.
- Portanier, E., Léger, F., Henry, L., Gayet, T., Queney, G., Ruetten, S., & Devillard, S. (2022). Landscape genetic connectivity in European wildcat (*Felis silvestris silvestris*): A matter of food, shelters and demographic status of populations. *Conservation Genetics*, *23*, 653–668.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.
- Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. B. (2020). Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*, *37*, 2153–2154.
- 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 toolset for whole-genome association

- and population-based linkage analysis. *American Journal of Human Genetics*, 81, 559–575.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing <https://www.R-project.org/>
- Randi, E., Pierpaoli, M., Beaumont, M., Ragni, B., & Sforzi, A. (2001). Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution*, 18, 1679–1693.
- Rezaei, H. (2007). Phylogénie moléculaire du genre *Ovis* (Mouton et Mouflons)—Implications pour la conservation du genre et pour l'origine de l'espèce domestique. Doctoral thesis, Université Joseph Fourier, Grenoble.
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754.
- Salotti, M. (1992). Le Chat sauvage (*Felis silvestris lybica*) en Corse: une population en danger d'extinction. Faculté des Sciences et techniques de Corte, Corte. 4p.
- Say, L., Devillard, S., Léger, D., Pontier, D., & Ruetten, S. (2012). Distribution and spatial genetic structure of European wildcat in France. *Animal Conservation*, 15, 18–27.
- Smith, W. J., Quilodran, C. S., Jezierski, M. T., Sendell-Price, A. T., & Clegg, S. M. (2022). The wild ancestors of domestic animals as a neglected and threatened component of biodiversity. *Conservation Biology*, 36, e13867.
- Storey, J. D., Bass, A. J., Dabney, A., & Robinson, D. (2021). Q value: Q-value estimation for false discovery rate control. R Package Version 2.26.0.
- Tamm, E., Di Cristofaro, J., Mazières, S., Pennarun, E., Kushniarevich, A., Raveane, A., Semino, O., Chiaroni, J., Pereira, L., Metspalu, M., & Montinari, F. (2019). Genome-wide analysis of Corsican population reveals a close affinity with northern and Central Italy. *Scientific Reports*, 9, 13581.
- Thomas, J. A., Telfer, M.G., Roy, D. B., Preston, C. D., Greenwood, J. J. D., Asher, J., Fox, R., Clarke, R. T., & Lawton, J. H. (2004). Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science*, 303, 1879–1881.
- Vigne, J. D. (1988). *Paléontologie, archéologie et mammalogie*. Xième Colloque Francophone de Mammalogie.
- Vigne, J. D. (1992). Zooarchaeology and the biogeographical history of the mammals of Corsica and Sardinia since the last ice age. *Mammal Review*, 22, 87–96.
- Vigne, J. D., Guilaine, J., Debue, K., Haye, L., & Gérard, P. (2004). Early taming of the cat in Cyprus. *Science*, 304, 259.
- Virgos, E., Telleria, J. L., & Santos, T. (2002). A comparison on the response to forest fragmentation by medium-sized Iberian carnivores in Central Spain. *Biodiversity and Conservation*, 11, 1063–1079.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.
- Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island biogeography* (2nd ed.). Oxford University Press.
- Yamaguchi, N., Kitchener, A., Driscoll, C., & Nussberger, B. (2015). *Felis silvestris*. The IUCN red list of threatened species 2015: E.T60354712A50652361. <https://doi.org/10.2305/IUCN.UK.2015-2.RLTS.T60354712A50652361.en>
- Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *PNAS*, 105, 11597–11604.
- Zeder, M. A. (2012). Pathways to animal domestication. In P. Gepts, T. R. Famula, R. L. Bettinger, S. B. Brush, A. B. Damania, P. E. McGuire, & C. O. Qualset (Eds.), *Biodiversity in agriculture – domestication, evolution and sustainability*. Cambridge University Press.
- Zhou, H., Alexander, D. H., & Lange, K. (2011). A quasi-Newton method for accelerating the convergence of iterative optimization algorithms. *Statistics and Computing*, 21, 261–273.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Portanier, E., Henri, H., Benedetti, P., Sanchis, F., Régis, C., Chevret, P., Zedda, M., El Filali, A., Ruetten, S., & Devillard, S. (2023). Population genomics of Corsican wildcats: Paving the way toward a new subspecies within the *Felis silvestris* spp. complex? *Molecular Ecology*, 00, 1–17. <https://doi.org/10.1111/mec.16856>