Alma Mater Studiorum – Università di Bologna

Settore Concorsuale di afferenza: 05/11

Settore Scientifico disciplinare: BIO/18

Dottorato di Ricerca in

Biodiversità ed Evoluzione

Conservation genetics of European wildcat (*Felis silvestris silvestris*): a wide and integrating analysis protocol for admixture inferences and population structure

Presentata da:
Federica Mattucci

Coordinatore Dottorato
Barbara Mantovani

Relatore
Ettore Randi

Ciclo XXV

Esame finale anno 2014

TABLE OF CONTENTS

CHAPTER 1 – General Introduction
1.1. Cat domestication: a complex evolutionary process 5
1.1.1. Domestication effects on wildcats populations 9
1.2. Wildcats in Europe 11
1.3. Conservation genetics of European wildcats 14
1.4. Genetic tools for cat studies: assessing population structure and admixture occurrence 19
1.4.1. Toward an implemented set of informative autosomal (SNPs and STRs) and uniparental (mtDNA and Y) molecular markers 22
1.5. Objectives and thesis outline 26
1.6. References 31

CHAPTER 2 – European wildcats in Italian Peninsula
Paper I- Population genetic structure of wildcats (*Felis silvestris*) in Italy 51
2.1 References 73
2.2 Supporting Information 81
2.3 List of Tables 86
2.4 List of Figures 87

CHAPTER 3 – European wildcat population structure
Paper II - Population genetic structure of wildcats (*Felis silvestris*) in Europe 91
3.1 References 119
3.2 Supporting Information 127
3.3 List of Tables 131
3.4 List of Figures 133

CHAPTER 4 – Improving the molecular toolbox
Paper III - Towards a wide genetic approach for the European wildcat (Felis silvestris) Conservation 135
4.1 References 163
4.2 Supporting Information 169
4.3 List of Tables 174
4.4 List of Figures 175

CHAPTER 5 – General Discussion
5.1. Genetic structure of European wildcats in Italy 178
5.2. Genetic structure of European wildcat populations in Europe 182
5.3. Improving the molecular tools in wildcat studies for assessing admixture 186
5.3.1. Detection of most distinctiveness SNPs between European wild and domestic cats 186
5.3.2. Combining use of most informative autosomal SNPs with uniparental markers (mtDNA and Y-chromosome) for the assessment of hybridization in European wildcats (*Felis silvestris*) 192
5.4. Future perspectives and final remarks of the PhD work 194
5.5. References 200

ACKNOWLEDGEMENTS 211

Chapter 1

General Introduction

1.1. Cat domestication: a complex evolutionary process

The transition from hunting-gathering to farming and herding is a significant threshold in human history (Benecke 1987; Tchernov & Valla 1997), resulting in a radical restructuring of human societies, worldwide alterations in biodiversity, and significant changes in the Earth's landforms and its atmosphere.

It is believed that, during the Upper Paleolithic, Homo sapiens had already initiated the process of domestication (Vigne 2011) but only in the beginning of the Neolithic period civilization began to better recognize and well explore animal's qualities for their benefit. Around at 10,000 - 9,500 B.P.†, 1,000 years after the domestication of crop plants in the southern Levant (Uerpmann 1979; Bar-Yosef & Meadow 1995), humans brought animals under their care mainly to produce food, provide protection and help with servile labor.

Three principal areas of livestock domestication have been identified; in particular: southwest Asia (Fig 1; the FERTILE CRESCENT, extending from the Mesopotamian plains, through the Taurus mountains, the Mediterranean coast to the Levant, and its eastern margin, towards the Indus valley region), east Asia (China and countries south of China), and the Andean chain of South America (Bruford *et al*. 2003).

However, recent studies (Vigne *et al.* 2011) demonstrated how the geographic range of the earliest cultivators in the Near East was wider than previously suspected (O'Brien & Mayr 1991a; O'Brien & Mayr 1991b), extending beyond the mainland to Cyprus. Most likely, the island's occupation was a consequence of a combined geographic expansion driven by strong demographic growth (Bradshaw *et al.* 1996), with an increase in exchange systems (Elton 1953), already observed in other regions during the transition to farming.

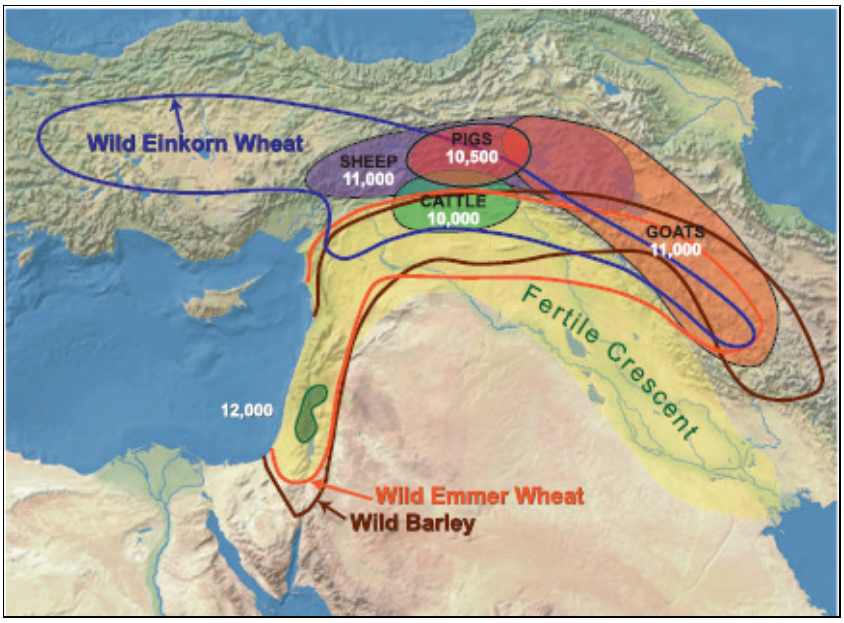


Figure 1. Map of the Near East indicating the FERTILE CRESCENT. Shaded areas indicate the approximate areas of domestication of pig, cattle, sheep, and goats with dates of initial domestication in calibrated years B.P. (Driscoll *et al.* 2009; Zeder 2008). The Fertile Crescent is part of present day Iran, Iraq, Syria and Turkey.

During recent decades, archaeology, archaeozoology, and genetics provided many new data relating on the complex human-mediated evolutionary process of domestication (e.g. Davis 2005; Dobney and Larson 2006; Redding 2005; Zeder 2006a, 2011; Zeder *et al.* 2006), most likely, generated from natural and antropic directed selection in anthropogenic environments (Tchernov 1984; Driscoll *et al.* 2009; Smith 2012; Hu *et al.* 2013).

It is believed that, this mutualistic relationship between humans and animals occurred in a number of independent places and through a surprisingly high number of episodes, as a reaction of numerous different factors (e.g., climate, biogeography, environment, demography, techno-economic practices, diet and health, social structure and mentality), and through a multitude of pathways which varied in length, direction, and travel time (Zeder 2012). For instance, prey pathways have been well documented through culling of ancient goats, sheep, and cattle (Zeder 2012). In this case, the animals were primary preyed for meat and hides for thousands of years and subsequently domesticated by the same hunters. On the contrary, directed pathways had occurred when humans domesticated wild species with desirable resources by taking advantage of all the knowledge gained during the management of already domesticated animals. However, most interesting are the commensally pathways that characterized dogs, pigs and cats route domestication (Zeder 2012).

Wolves (Canis lupus) are the first species known to have been domesticated during the Late Glacial by both European and Asian hunter-gatherers. Although date and location are still debated, it is believed that domestication occurred somewhere around 17–15 kyrs BP (Benecke 1987; Pionnier-Capitan *et al.* 2011) or perhaps earlier, around 20–30 kyrs BP (Germonpré *et al.* 2008) in Central Europe (Klein & Cruz-Uribe 1984; Zeder 2006a). Most likely, the process happened in 2 interwoven phases. First, a founder group of less-fearful wolves would have been pulled toward nomadic encampments to scavenge kills or perhaps salvage wounded escapees from the hunt. Thereafter, these wolves may have found utility as barking sentinels, warning of human and animal invaders approaching at night (Lindsay 2000). Gradually, natural selection and genetic drift resulting from human activities began to differentiate these wolves from the wide autonomous population. Once people had direct interaction with wolves, a subsequent, 'cultural process' would have begun. Suitable 'preselected' wolf pups taken as pets would have been socialized to humans and unconsciously and unintentionally selected for decreased flight behavior and increased sociality (Muller 2002; Driscoll *et al.* 2009). However, this process did not bring major modifications in the way of life for humans, except perhaps small changes in hunting strategies, tactics or techniques (Vigne 2011).

By contrast, cat domesticates arose much later after humans built houses, farms, and settlements. For many years, the Nile Valley in Ancient Egypt was considered the centre of cat's domestication basing on Middle Kingdom Egyptian art dated to ca. 4,000 B.P. (Driscoll *et al*. 2009, Faure & Kitchener 2009). However, the oldest archaeological evidence pushes back the date of cat domestication to 10,600 YA (Vigne *et al.* 2012), when the earliest evidence of a cat-human close relationship was found in Cyprus deposits.

Recent genetic studies have suggested that the origins of cat domestication occurred in the adjacent Near Eastern sites (Driscoll *et al*. 2007; Lipinski *et al*. 2008) as domestic cats have derived mitotypes from regional wildcats and the genetic diversity of modern domestic cats within these regions is highest.

Most likely, this process would have followed two steps: a "commensalization" that attracted cats to the villages because of an high concentrations of mice, whose presence was due to an increase in stocked foodstuffs; and a consecutive protection provided by the villagers to some individuals for the commensalism (Pascal *et al.* 2006; Vigne *et al.* 2011).

Thus, wildcat domestication occurred through a self-selective process in which behavioral reproductive isolation evolved as a correlated character of assortative mating coupled to habitat choice for urban environments. It is, therefore, reasonable to believe that humans did not try to

influence breeding and behavior of the first house cat, and this favored their admixture with local wildcats (Driscoll *et al.* 2009). Evidence of this is furnished by the multiple genetic analyses that traced the maternal origins of cat domestication to at least 5 wildcat lines originating in the Near East.

From the Near East origins of domestication, cats subsequently became common in Europe and Asia by the 10th century, as a consequence of their dispersal throughout Europe by Roman legions, and were ultimately transported around the world on the major land and sea trade routes. By the time of the industrial revolution (late 18^{th} early 19^{th} century), pet cat owners were selectively mating their pet tabbies to produce fancy breeds (Zeder 2012).
Over the past 140 years, a plethora of pedigreed cat varieties has developed due to mankind's imposed artificial selection on the process of cat domestication. Since the first cat showed in London in 1871 with only five breeds, the development of pedigreed cats has increased in popularity. Of the 41 breeds for competition recognized by the Cat Fanciers' Association (CFA), and the accepted 57 breeds of The International Cat Association (TICA, http://www.tica.org), 16 "natural breeds" are thought to be regional variants that experienced some degree of isolation, which resulted in fixation of alleles for distinctive morphological traits of the breed. The remaining breeds were developed over the past 50 years, as single-gene variants derived from the natural breeds (Kurushima *et al.* 2012; Lipinski *et al.* 2011; Menotti-Raymond *et al.* 2008).

It is reasonable believe that true domestication of cats arrived very recently and might still be an ongoing even since no selective breeding was perpetrated for long time. Moreover, overlapping niches between the wildcat progenitors, random-bred feral cats, random-bred house cats and fancy breeds likely produces continual, although limited, horizontal gene flow throughout the domestic cat world. Nonetheless, a recognizable and diagnostic population structure among the domestic cat breeds has been recognized (Lipinski *et al.* 2008; Menotti-Raymond *et al*. 2008). Twenty-seven population clusters were, in fact, identified by Menotti-Raymond *et al.* (2008), 6 of which composed of multiple breeds, mainly organized into four distinct groups: Asia, Mediterranean basin, Western Europe, and East Africa. American cats consistently grouped with cats from Western Europe, suggesting European settlers probably brought cats to the New World and the cat's time in America has been too brief for significant genetic differentiation (Lipinski *et al.* 2008).
While date and location of cat's domestication has been widely debated until recent years, the origin seemed rather clear: archaeological (Kitchener 1991; Yamaguchi *et al*. 2004a), morphologic (Yamaguchi *et al.* 2004a), and genetic (e.g. Randi & Ragni 1991; Johnson & O´Brien 1997;

Driscoll *et al*. 2007) evidences concur that the domestic cat derived, very recently, from the Near East group of the wildcat species *Felis silvestris*: the *Felis silvestris lybica* subspecies.

1.1.1 Domestication effects on wildcats populations

The biological consequences of domestication vary according to the species. However, for vertebrates (including birds and fish) there are a number of constant modifications including a decrease in aggressiveness, increased fertility, sexual dimorphism (and often body size) decrease, shortening of the face, decrease of the braincase volume, appearance of a new coat colors, soft ears (carnivores) and voice changing (barking of dogs; Clutton-Brock 1981; Davis 1987; Bokonyi 1974; Zohary *et al.* 1998). Some of these modifications result from hormonal changes, due to environmental conditions and the stress of captivity (Zohary *et al.* 1998; Arbuckle 2005), whereas others result from epigenetic/developmental changes or are purely genetic mutations possibly selected by humans (Vigne *et al.* 2011).

From a genetic point of view, domestication through the artificial selection has caused selective sweeps at multiple loci in domestic species (Zeder *et al.* 2006a; Wiener & Wilkinson 2011). During the various phases of this process, different portions of the genome are sculpted by artificial selection, from the early steps of initial taming until the devoted formation and improvement of breeds (Wiener & Wilkinson 2011). Comparing to the little variation in body morphology and coat colour patterns of European wildcats, domestic cats show a variety of distinct coat colour, fur type, behaviour, breeding cycle, and morphologic variants (see Lyons 2010 and Lyons 2012 for a summary of most relevant phenotype/genotype associations among domestic cat breeds). Moreover, specific mutations determining variable patterns in domestic cats are predictably absent in natural wildcat populations, and, thus, they might be strong diagnostic candidate genetic variants for the distinction of wild and domestic relatives (Cruz *et al*. 2008).

However, this variability is not necessarily beneficial. The domestication process, in fact, generates two genome-wide bottlenecks at specific loci through breeds' development: i) during the beginning, because of the small demographic size of the initiating domestic population, ii) and subsequently, because of the constant improvement (through strong selection) of purebred individuals with specific desired phenotypic traits (Bruford *et al.* 2003; Zeder 2006a). Hence, selection at linked sites decreases locus-specific effective population size, which increases the probability of deleterious mutations becoming fixed (Hill and Robertson 1966; Cameron *et al.* 2008). For instance, a comparison of the d_N/d_S ratio (ω) in dog and wolf lineages indicated that dogs appear to

have been accumulating nonsynonymous mutations in mitochondrial DNA (mtDNA) genes at a greater rate than wolves (Björnerfeldt *et al.* 2006). The same pattern has been recognized for wildcat, where over 20 deleterious genetic disorders have been recognized in modern cats, and all have been identified in pure breeds (Lipinski *et al*. 2008; Cruz *et al.* 2008).

Therefore, domestication causes, in one hand, genome wide reductions of genetic diversity (resulting from the early stages of domestication and from breeds improvement) at both neutral loci because of demographic events, and selected specific loci because of positive selection of advantageous alleles. On the other hand, following the strong bottleneck effects, relaxation of selective pressures may increase domesticates' genetic diversity through the increment of post-domestication (de novo) mutational events (Doebley *et al.* 2006). Thus, it appears needed understand what forces play roles in shaping patterns of genetic variation, decrease in diversity, and target of selection in domesticated species.

In spite of their rapid spread, cats have remained amazingly similar to their wild felid ancestors (*Felis silvestris* subspp.) in form and function (Dobney & Larson 2006; Clutton-Brock 1999). The progenitor species of domestic cats remained compatible with human agriculture. Gene flow between feral and tame modern cats, and between modern cats and their wild subspecies (Beaumont *et al.* 2001; Wiseman *et al.* 2000), has not negatively impacted the role of cats as the principle small carnivore in human dominated ecosystems.

Therefore, the impetus to change cats to suit certain human needs and the resulting manipulation of domesticates gene pools (Lipinski *et al.* 2008), have caused tremendous phenotypic changes and created a variety of cat breeds. Hence, diversifying selection is mostly related with morphologic aesthetic traits, related with pelage types appreciated by humans as opposed to complex behaviors and qualities, such as hunting skills and meat or milk production in dogs or in other livestock species. Many of the cat's phenotypic attributes, even those that affect body and appendage morphologies, are traits with basic Mendelian inheritance patterns (Kurushima *et al.* 2012). The immense phenotypic diversity that commonly segregates in domestic species provides exceptional opportunities to establish specific genotype/phenotype associations and to study the general mechanisms by which genetic variation governs biological functions.

A number of cat studies focused on detection of mutation related with hair length (e.g. four different mutations at FGF5 (Drogemuller *et al*. 2007, Kehler *et al*. 2007), and SNPs at KRT71 (Gandolfi *et al*. 2010)), and hair color, in which genes like TYR (Lyons *et al.* 2005a; Schmidt-Küntzel *et al.* 2005), TYRP1 (Lyons *et al.* 2005b; Schmidt-Küntzel *et al.* 2005), MLPH (Ishida *et al.* 2006) or ASIP (Eizirik *et al.* 2003) are determinants. Recently, it has been suspected that coat color genes might have pleiotropic effects upon morphology and behavior (Pontier *et al.* 2009). Moreover, the

same candidate gene proved also to be extremely useful for identifying mutations related with inherited diseases (see Lyons 2012 for details). The recent and rapid genesis of cat breeds from a limited number of individuals suggests that, in many cases, a small number of genes of large effect are responsible for breed characteristics (Pollinger *et al.* 2005). Today, approximately 33 genes containing 50 mutations are known to cause feline health problems or alterations in the cat's appearance (Lyons 2010).

1.2. Wildcats in Europe

Modern felid species descend from relatively recent (G11 million years ago) divergence and speciation events. A highly resolved molecular phylogeny defines eight principal lineages produced through at least 10 intercontinental migrations facilitated by sea-level fluctuations. The most recently derived group was the domestic cat lineage originated 6,2 MYA and differentiated with 4 species at Ethiopian and Paleartic distribution: *Felis chaus* (jungle cat), *Felis nigripes* (black-footed cat), *Felis margarita* (desert cat) and about 2,5 MYA *Felis silvestris* (Johnson *et al.* 2006).

Although several studies were conducted, aiming at extricating the complex phylogeny of *Felis silvestris*, there is still no clear consensus to relate geographical variation in the morphology and genetics of the globally widespread wildcats to its taxonomy and systematic (Kitchener & Rees 2009). The latest phylogeographical analysis (Driscoll *et al.* 2007; Macdonald *et al.* in press) classified *Felis silvestris* as a polytypic wild species composed of five distinct inter-fertile subspecies plus a domesticated form (as shown in Figure 2; Drischoll *et al.* 2007)

- *F. s. silvestris* in Europe (Schreber 1775)
- *F. s. lybica* in the north of Africa and south-west Asia (Forster 1780)
- *F. s. ornata* in the Middle East and central Asia (Gray 1830)
- *F. s. cafra* in the south of Saharan Africa (Desmarest 1822)
- *F. s. bieti* in the North East corner of the Tibetan Plateau (Milne-Edwards 1872)
- *F. s. catus* the domesticated form derived from wildcats in the Middle East or Egypt (Clutton-Brock 1999; Nowell & Jackson 1996; Pocock 1907; Randi & Ragni 1991)

While an alternative taxonomic treatment considers *F. bieti, F. silvestris*, and *F. lybica* (including *ornata* and *cafra*) as three recently radiated phylogenetic species (Kitchener & Rees 2009, Macdonald *et al.* in press).

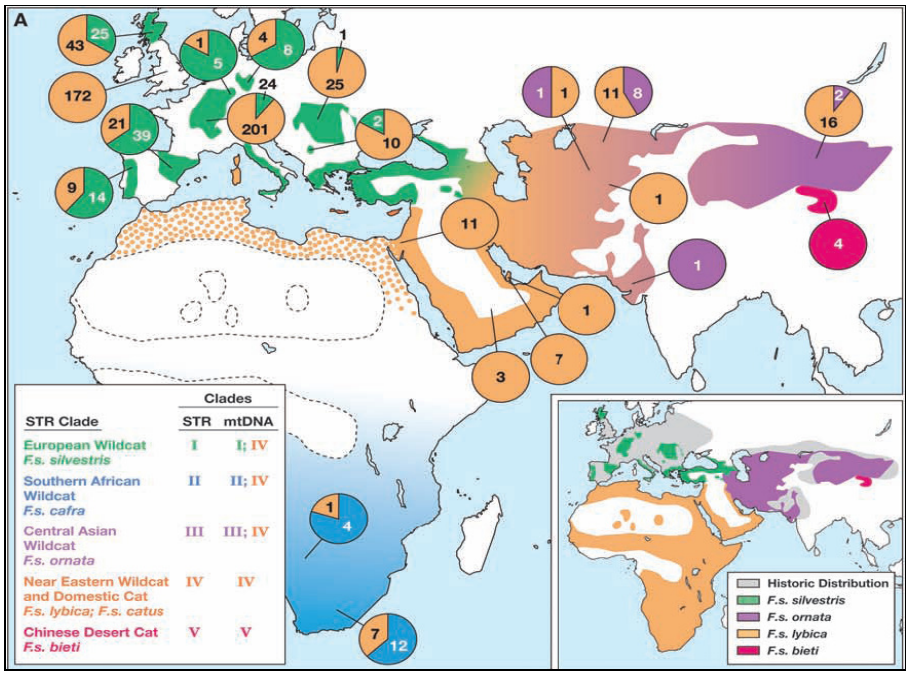


Figure 2. The current range of *Felis silvestris* and areas of sample collection are shown. The colored regions reflect the location of capture of individuals with different STR and mitochondrial DNA (mtDNA) clade genotypes (defined in the bottom left). In pie chart are reported frequencies and number of specimens carrying each mtDNA haplotype clade. The inset on the right shows the current and historic range of *F. silvestris* subspecies on the basis of traditional morphology-based taxonomy. Adapted from Driscoll *et al.* 2007.

Felis silvestris appeared for the first time in the fossil deposits of the Holsteinian Interglacial (300,000 - 400,000 YA; Wolsan 1993) and probably descended from Martelli's wild cat (*F.* (*s.*) *lunensis;* Martelli 1906). Archeological remains founded in the Palestine region and South Africa (Kurtén 1965a; Klein 1986) suggest that wildcat may have expanded its range suddenly in the last *c.* 50,000 YA; timing that well coincides, on a geological time scale, with the supposed divergence at *c.* 20,000 YA between European wildcats and African wildcats (Randi & Ragni 1991). Therefore, the evolution of the modern wildcat probably consisted of at least three different range expansions punctuated by two differentiation events, during which wildcat moved out of Europe to colonize the Middle East and subsequently spread quickly eastward to Asia and southward to Africa (Kitchener & Rees 2009).

Despite the wide geographical distributions ranging from western Europe throughout most of Africa, and from Arabia and south-west Asia to China (Nowell & Jackson 1996; Sunquist & Sunquist 2002), wildcats were classified as "Least Concern" in the IUCN Red List, because of habitat destruction and population decline. Among the five wild subspecies the European wildcat is

probably the most endangered and, thus, several molecular and archeological studies were conducted in the last few decades. Between the late 1700s and mid 1900s, in fact, severe declines and local extirpations occurred in Europe resulting in a fragmented relict distribution (Stahl & Artois 1991; Nowell & Jackson 1996; Peichocki 2001). Despite the Italian population that is spreading northwards into Austria (Lapini & Molinari 2006), European wildcat is considered extinct in the Netherlands (Nowell & Jackson 1996), Austria (Spitzenberger 2005) and Czech Republic (IUCN 2007).

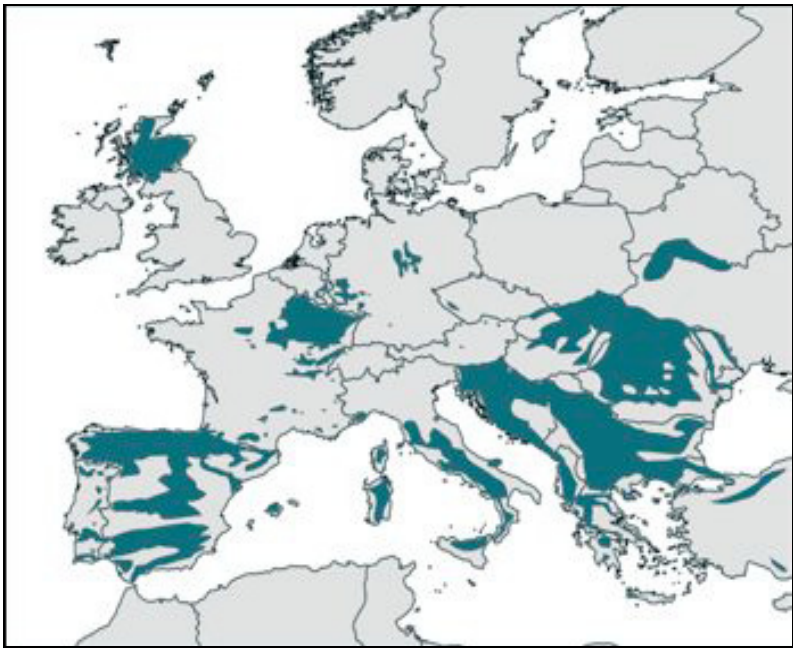


Figure 3. Fragmented distribution of the European wildcat (*F. s. silvestris*) in Europe. Adapted from Grabe & Worel (2001)

Current patterns of European wildcats distribution (Fig 3) and genetic variability are most probably a reflection of both natural and anthropogenic events. The vast majority of wildcat's endemic populations are today considerably small and fragmented (e.g. France, O'Brien *et al.* 2009; Germany, Germain *et al*. 2008), and many of them are known to be critically decreasing (e.g. Poland, Wolsan & Okarma 2001; Scotland, Yamaguchi *et al*. 2004b). The Iberian subpopulation, for instance, is suspected to have decreased at a rate of >30% over three generations and consequently the European wildcat is listed as 'vulnerable' (IUCN 2007). In Scotland, the population has been reduced to approximately 400 individuals (IUCN 2007). Contrastingly, small increments on the species range have been detected in Switzerland and Belgium (IUCN 2007), and although previously extinct in the Netherlands (Nowell & Jackson 1996), wildcats may be recolonizing from German populations in the Eifel or Ardennes forests (Canters *et al.* 2005).

Different studies in central Europe also report that the species distribution seems stable or even expanding in the later years, due to some increment in forest cover as a result of massive agriculture abandonment (Raimer 2006; Hertwig *et al.* 2009; O'Brien *et al*. 2009). Among the Mediterranean isles, Sicily is the only one populated by European wildcats, both Sardinia and Corsica isles, in fact, host populations of i) feral domestic cats introduced by Neolithic navigators, about 8,000–6,000 YA (at an early stage of domestication (Gippoliti & Amori 2006, Kitchener *et al.* 2010); ii) or *Felis silvestris lybica* introduced by the same navigators coming from North Africa (Ragni 1981; Pierpaoli *et al*. 2003).

Wildcats occurred in a wide variety of habitats (Stahl & Leger 1992), displaying an individual and seasonal variation in habitat selection (Wittmer 2001). The species is distributed from sea level to 2,250 m in the Pyrenees (Palomo & Gisbert 2002), from deserts and scrub grassland in the Mediterranean area (Lonzano *et al.* 2013) to dry and mixed forest in the temperate bioclimatic areas (Germain *et al.* 2008) and resulted lacking only from the rainforest and coniferous forest. In particular, European wildcats avoided areas of intensive cultivation, preferring forest landscape with low densities of humans or Mediterranean maquis scrubland and riparian forest. African wildcats are found everywhere outside tropical rainforest, although thinly distributed in true desert (Nowell & Jackson 1996). The Chinese alpine steppe cat is restricted to the alpine meadow habitats in the eastern edge of the Tibetan plateau at 2,500-5,000 m (He *et al.* 2004). Asiatic wildcats have a different elevation range (up to 2,000-3,000 m) and are most typically associated with scrub desert (Nowell & Jackson 1996). Despite being generally considered a species specialized in consuming rodents (Nowell & Jackson 1996), in Mediterranean ecosystems it preferred rabbits as food source (Lozano *et al.* 2006).

1.3. Conservation genetics of European wildcats

In Europe, the wildcat has disappeared from much of its original distribution area, resulting in severe fragmentation of its populations. The lacking of sufficient biological knowledge (Stahl & Artois 1991) and the current disrupted distribution have leaded to the inclusion of the species on the EU Habitats and Species Directive (Annex IV; 12/05/1992) and the Bern Convention (Appendix II; 19/04/1979). Wildcat is fully protected across most of its range in Europe and Asia, but only some of its African range (Nowell & Jackson 1996), and is classed as "Least Concern or Vulnerable" at the national level in many European range states (IUCN 2007).

Hundreds of years of intensive hunting, road kills (Nowell & Jackson 1996, Lüps *et al.* 2002, Schulenberg 2005), human persecution for pest control (e.g. Scotland; Macdonald *et al.* 2004), and

trapping for fur (e.g. Asia; Nowell & Jackson 1996), has played a major role in the species extirpation in many places (Langley & Yalden 1977; Duarte & Vargas 2001).
Pressure of persecution decreased since the termination of eradication programs, which caused the species to disappear from large areas of its range (see Langley & Yalden 1977; Pierpaoli *et al.* 2003), but new researches show that a significant number of individuals are still killed in predator control programs (Duarte & Vargas 2001; Herranz 2001; Spanish Environment Ministry, unpublished data). Moreover, the extensive road network probably acts as a handicap to dispersal, thus limiting the gene flow and ultimately resulting in a hidden genetic structure within the European wildcat populations (Eckert 2003; Mölich 2006; summary in Simon 2006). In addition, predator control measures in a number of European countries may result in a drastic demographic decline of the species (e.g. Iberian Peninsula (e.g. Lozano *et al.* 2007; Monterroso *et al.* 2009)).

Two main extinction risks are underlined by European Council (Stahl & Artois 1991) and considered as essential for the survival and conservation of the species in the near future: i) habitat destruction that leads to fragmentation and isolation of populations; and ii) hybridization between wildcats and domestic cats (*Felis s. catus*), which may compromise the species genetic purity. Thus, guidelines for conservation need to evaluate the biological and ecological effects of this threat and promptly indentify management tools to delete or mitigate their effects aiming at preserving the species in Europe (Stahl & Artois 1991; McOrist & Kitchener 1994; Lozano *et al.* 2007).

i) habitat loss and fragmentation

Fragmentation of natural habitats is a major challenge in conservation biology and one of the top threats to biodiversity (Hanski 1999; Fahrig 2003; Henle *et al.* 2004). The negative effects result from the decrease of habitat availability and changes in quality and spatial configuration of habitats (Fahrig 2003; Ezard & Travis 2006).
At a population level, habitat fragmentation first erode neutral and adaptive genetic diversity, reducing the effective population size and the inter-population connectivity (Johansson *et al.* 2007). Subsequently, all remnant isolated and small populations experience a genetic drift that increase their inbreeding level and limit their gene flow from surrounding populations. The evolutionary consequences of this genetic bottleneck induce, on one hand, to reduce the genetic variability and enlarge the differentiation among remnant populations; on the other hand, to impoverish the evolutionary potential of these small populations increasing their risk of extinction (Avise *et al.* 1987; Young *et al.* 1996; Saccheri *et al.* 1998; Reed & Frankham 2003).

For many years, deforestation (Nowell & Jackson 1996; Krüger *et al.* 2009) and loss of Mediterranean scrubland for fire control (Lozano *et al.* 2003), concomitantly with urbanization, extensive road networks and intensive agriculture (Easterbee *et al.* 1991; Stahl & Artois 1991; McOrist & Kitchener 1994; Nowell & Jackson 1996; Klar *et al.* 2008, 2009), have simultaneously depleted and isolated natural European wildcat populations, and some of these threats persist in current days. Thus, preservation of wildcat habitat, through the maintenance of both connectivity via habitat corridors (Harris 1984; Bennett 1990; Saunders *et al.* 1991) and the main preys survival, is regarded as a guarantee for the long-term conservation of the species.

ii) introgression and hybridization

Natural hybridization has been defined as 'Successful matings in nature between individuals from two populations, or groups of populations that are distinguishable on the basis of one or more heritable characters' (Arnold & Burke 2004).

The role of hybridization in the evolution of living organisms has been extensively discussed among evolutionary ecologists (e.g. Arnold 1992; Dowling & Secor 1997; Barton 2001; Fitzpatrick 2004). On one hand, interspecies hybridization can facilitate evolutionary diversification in both plants and animals, including the origin of new species (Rieseberg 1997; Arnold 2004; Grant *et al*. 2005) so that hybridization can lead to evolutionary innovation and even speciation (Anderson & Stebbins 1954; Barton & Hewitt 1985; Allendorf *et al.* 2001; Rieseberg *et al*. 2003). On the other hand, when hybridization is driven by anthropogenic changes (e.g. invasive species, domestication, habitat loss and fragmentation), it might become a conservation concern (Allendorf *et al.* 2001). Moreover, when the process of hybridization induces a flow of genes between genetically differentiated species or populations it became introgressive hybridization.

In a conservation oriented-view, introgression became stronger when it occurred between translocated or invasive organisms and local wild individuals because it can produce a deep impact on the structure of local communities, loss of local adaptation generated during the evolutionary process and eventually local extinctions (Rhymer & Simberloff 1996; Allendorf *et al*. 2001).

Under this scenario, domestication and the consequently worldwide spread of breeds, have assumed a critical role on negatively affect natural populations through predation, resource competition (Corbett 1979; Birò *et al*. 2004-2005; Germain *et al.* 2008) and disease transmission (e.g. the Feline leukaemia virus found in French wildcats may increase the mortality of individuals in the wild

influencing the growth rate of the populations; Fromont *et al.* 2000; in Great Britain see McOrist 1992).

The infusion of domestic genes via interbreeding (Rhymer & Simberloff 1996; McGinnity *et al.* 2003) might have introduced genes favored under artificial selection that are maladaptive in the natural environment, and disrupted locally adapted gene complexes. Both of these processes can reduce the fitness of wild populations (Rhymer & Simberloff 1996; Allendorf *et al*. 2001; McGinnity *et al*. 2003; Hutchings & Fraser 2008).

Among the endangered European carnivores, a number of species are affected by artificially mediated hybridization where domesticated forms are leading protagonist (Gittleman *et al.* 2001), including the *gray wolf* (Hope 1994; Vilà & Wayne 1999; Randi 2008; Godinho *et al.* 2011), polecats *Mustela putorius* (Lodé *et al.* 2005), American mink *Neovison vison* (Kidd *et al*. 2009; Tamlin *et al*. 2009), red fox *Vulpes vulpes* (Sacks *et al.* 2011) or wild arctic fox *Alopex lagopus* (Noren *et al.* 2005).

Several factors may have promoted the reproductive interaction between wild and domestic cats in Europe. Most likely, the fragmented and demographic decline of wildcats in comparison to the increasing distribution of free-ranging domestic cats may have been played a crucial role, intensifying their contact.

Observations of wildcats' spatial activities suggest that most of the admixture events are probably occurring between male wildcats and domestic females (Birò *et al*. 2004), while observations in Ardennes Mountains in France indicate that hybridization might not involve domestic females, but probably the low densities of wild males promote the crossbreeding between wild females and domestic males (Germain *et al.* 2008).

The daily activities and feeding habits overlapping (77-88%; Birò *et al*. 2005) among wildcat and domestic cat found in France (Germain *et al.* 2008) and Hungary (Birò *et al.* 2005) probably promote their interaction. However, the lacking of any concordance between the space use patterns of wild and domestic cats led to think that probably hybridization may occur because of rare excursions of cat outside their home range (Germain *et al.* 2008). In particular, some observations in Ardennes Mountains in France indicate that crossbreeding probably is favored by the low densities of wild males that induce wild females to mate with domestic males (Germain *et al*. 2008).

Moreover, hybrids may play a key role in hybridization by behaving as wildcats and by sharing at least a part of their range with them as well as with domestic cats. Barriers between them and wildcats may not exist because of their similarity in morphology and spatial behavior (Birò *et al.* 2004; Germain *et al.* 2009).

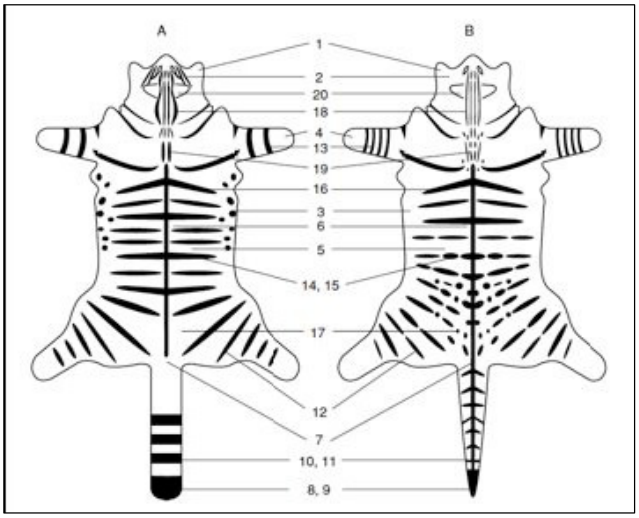
The wide trophic nice discovered in hybrid individuals, intermediate between wild and domestic cats and totally overlapping the wildcat's one, suggest that competition in the use of food resources' may exist, and that wildcat survival in case of food limitations may be negatively affected, on the contrary of hybrids that can find food also close to human settlements (Birò *et al.* 2005). Admixed animals may also be less susceptible to habitats changes and more proficient in colonizing new habitats (Germain *et al*. 2009).

In cases of long-term sympatry, it has been hypothesized that, wildcats may exhibit an antagonistic behavior towards domestic cats (Hubbard *et al.* 1992), and thus averting hybridization to a certain extend (Easterbee *et al*. 1991). Under the scenario of long-lasting simpatry in Europe, mating between wild and domestic cats should occur only occasionally. If this is the case, after crossbreeding takes place, hybrids – rather than domestic forms – may be protagonists in maintaining and expanding hybridization (Germain *et al.* 2009). This reality poses the question if in places where putative wildcat populations have been apparently stable or even expanding that is a sign of populations "health" or, conversely, the result of frequent incidence of feral domestic and hybrid cats. At least in France, evidences show that wildcat's expansion should not be an artefact of hybridization since crossbreeding was detected over the entire range of the species and does not congregate on newly occupied areas (Germain *et al.* 2009).

The absent of differential selective pressures for wild and domestic cat combining with the absent of behavioral barriers between hybrids and wildcats make the hybrid identification in wildcat population arduous and uncertain. Moreover, the fertility of admixed individuals and the potential mating with parental groups may lead to the disappearance of their 'pure' parental genomes (Mallet 2005) and the formation of a hybrid swarm (Allendorf *et al.* 2001). Thus, the ability to detect 'pure' or admixed individuals became the basic requisite to understand the conservation status of the populations and develop adequate management plans.

The detection of hybrid individuals relied upon morphological characteristics until the mid-1960s (Allendorf *et al*. 2001). First morphological studies in Italy (Ragni & Possenti 1996) have shown that it is possible to differentiate domestic cats and European wildcats reliably using only pelage variables. In addition, Puzachenko (2002) investigated differences between European and African wildcats using 42 cranial characters, 11 of which are diagnostic for the subspecies identification. While Yamaguchi *et al.* (2004b) used 31 measurements of the skull and mandible and five derived indices to identify possible differences in the three recent subspecies of *F. silvestris*. Most recently, Kitchener *et al.* (2005) studied 20 pelage variables in 135 specimens of presumed wild-living cats from Scotland in various collections to develop and test a reliable definition of the Scottish wildcat (Fig 4). The morphological identification of individuals with hybrid origin, however, remains

difficult because of the long-term sympatry and interbreeding of domestic and wild cats and is usually based on the presence of a mixture of characters assumed to be typical for both forms. Moreover, morphological criteria frequently do not allow to identify hybrids beyond the F1 generation, as they are often morphologically indistinguishable from the parental species (Barbour *et al*. 2007; Krüger *et al*. 2009; Seiler *et al*. 2009; Ostberg *et al*. 2011). Thus, the develop of molecular markers it became needed for assessing admixture occurrence in wildcat populations.



(1) White on chin
(2) Stripes on cheek
(3) Dark spots underside
(4) White on paw
(5) White on flank
(6) White on back
(7) Extent of dorsal line
(8) Shape of tail tip
(9) Color of tail tip
(10) Distinctness of tail bands
(11) Alignment of tail bands
(12)* Stripes on hind leg
(13)* Bands encircling foreleg
(14) Tabby coat patterns
(15) Broken stripes on flanks & hindquarters
(16)* Stripes on body
(17) Spots on flanks & hindquarters
(18) Stripes on nape
(19) Stripes on shoulder
(20) Color of the back of ear

Figure 4. Diagram highlighting 20 diagnostic morphologic characters used to score cats as wild, domestic or hybrid (adapted from Kitchener *et al.* 2005). A=European wildcat; B=domestic cat

1.4. Genetic tools for cat studies: assessing population structure and admixture occurrence

Until now, a number of several studies have been performed on European wildcat populations providing first insights on wildcat's genetics (Pierpaoli *et al*. 2003) and the evolutionary history of *Felis silvestris silvestris* (Driscoll *et al*. 2007). Nevertheless, various drawbacks were disclosed, hindering the possibility to assess the occurrence of admixture events, the conservation status of this endangered subspecies and, thus, correctly define European-wide management plans.

In particular, the uneven or small sampling analyzed have often provided a very low geographic coverage preventing, in this way, the chance to generalized the specific case-study results. In addition, the small congruence between morphologic and molecular data and between data sets produced in different labs have created misunderstandings and difficulties. However, the more limiting aspect of the previous studies is related with type and numbers of molecular markers used, as they influenced the correct detection of "pure" reference specimens and hybrids.

In the light of wildcats' conservation, hybridization and genetic fragmentation of populations represent the main complex and controversial issues and, thus, implementing the resolution and robustness of hybrids detection and disclosing the genetic structure of European wildcat populations appeared to be the main challenge.

i) molecular advances in admixture analysis

The problematic definition of morphological criteria allowing unambiguous distinction between wild, domestic cats and cryptic hybrids forms, along with the particularly challenging identification of hybrids beyond the first generation (Daniels *et al.* 1998; Allendorf *et al.* 2001), prompted the initiation of genetic studies into diagnostic molecular traits. Recent advances over the past two decades in the development of molecular markers and bio-statistical techniques have led to a better recognition of hybridization and admixture patterns in wildcat populations.

The inability of allozymes variability to unequivocally differentiate wildcats and their domestic counterparts in the first molecular studies (Randi & Ragni 1991), led to involve the use of maternally inherited mitochondrial sequences aiming at properly investigating traces of historical splitting events and gene transfer between genetic lineages (Gottelli *et al*. 1994; Ward *et al*. 1999; Adams *et al*. 2003). However, the weak resolved phylogenetic signals, the poor differentiation among lineages and the low-resolution networks disclosed, led to think that probably an occasional gene flow between wildcats and domestic cats might be occurred during their long history of coexistence in Europe (see also Driscoll *et al.* 2007). Hence, the utility of mtDNA to accurately identify hybrids and assign individuals of unknown origin to wild or domestic populations has been limited (Randi *et al.* 2001; Eckert *et al.* 2009; Hertwig *et al*. 2009).

Only the combination of highly polymorphic microsatellite markers (e.g. Lecis *et al.* 2006; Pierpaoli *et al*. 2003; Randi & Lucchini 2002; Wiseman *et al*. 2000) with sophisticated Bayesian clustering methods proved to be reliable for identifying population's structure, individual assignment and admixture analyses, and deep investigating all the complex evolutionary processes involving the subspecies (e.g. Randi *et al.* 2001; Oliveira *et al.* 2007, 2008; Vila *et al.* 2003).

Over the last decades, a number of studies tried to develop and optimize methods to correctly asses individual assignment and admixture proportion, by taking advantage of Bayesian clustering procedure described by Pritchard *et al.* (2000) and recently updated in STRUCTURE (Falush *et al.* 2003), and Bayesian-method developed by Anderson & Thompson (2002) and performed in NEWHYBRIDS. In particular, STRUCTURE appeared to be useful to primary identify the number of

genetically distinct clusters that maximize the likelihood of the data, and subsequently to assign the individuals to these clusters.

The power of admixture analyses to detect F1, F2 hybrids and first-generation backcrosses, given the number of markers and the level of genetic differentiation between parental populations (estimated by F_{ST}), proved to be reliable assessed by simulations with HYBRIDLAB (Nielsen *et al.* 2006). While the ancestry detection of admixed genotypes proved to be reliable assessed with NEWHYBRIDS, by inferring the posterior probability assignment (Q) of each sampled individual to different genotype frequency classes, including parental and first generation hybrid categories (F1, F2 and Backcross I).

Previous studies evidenced the presence of diverse degrees rates of domestic genes introgression among wildcats populations in Europe, suggesting that hybridization does not constitute a uniform threat throughout the entire range of the species. However, the predominant factors determining such a geographical variability of hybridization still remain unclear (Beaumont *et al.* 2001; Pierpaoli *et al.* 2003; Kitchener *et al.* 2005). In particular, wide admixture rate has been detected in Hungary (31% of admixed cats; Lecis *et al.* 2006) and Scotland (the main hybrid swarm; Hubbard *et al*. 1992; Beaumont *et al.* 2001; Daniels *et al.* 2001; Pierpaoli *et al.* 2003; McDonald *et al.* 2004; Kitchener *et al.* 2005), while low admixture rate has been disclosed in Italy (between 2-8 %; Randi *et al*. 2001; Lecis *et al.* 2006) and Iberian Peninsula (7%; Oliveira *et al.* 2007,2008). High level of crossbreeding occurred also in France (23,8 % ; O' Brien 2009). In Germany, contrasting results has been reported with introgression level ranging from 18,4% (Hertwig *et al.* 2009) to 3% (Eckert *et al.* 2009; Pierpaoli *et al.* 2003). This variability among studies remarkably highlight the need to carefully evaluate all estimations of hybridization, and suggests that it might be promising face this complex issue with a spatial and multilocus integrated approach.

ii) molecular advances in population structure

In a conservation oriented view, providing a correct detection and quantification of genetic diversity pattern among populations it appears crucial to reconstruct their evolution and conservation status and, to properly direct their management tools.

Hence, in addition to the traditional population genetic analyses dealing with the distribution of allele frequencies between and within populations (for example Wright *F* estimation; Wright 1931), different methods based on maximum-likelihood assignment procedures (Paetkau *et al.* 1995; Rannala & Mountain 1997; Cornuet *et al.* 1999) and Bayesian clustering models (Pritchard *et al.*

2000) have lately emerged as prominent computational tools to infer population structure in molecular ecology (Beaumont & Rannala 2004).

Recently, the integration of both geographic and genetic information into spatial and landscape genetic models, have provided a useful tool to identify weakly differentiated populations, spatial clusters locations and genetic barriers (Storfer *et al.* 2007). In particular, the Spatial Principal Component Analysis (sPCA, Jombart *et al.* 2008) implemented in R, and the Bayesian Poisson-Voronoi tessellation of the sampling areas, obtained using GENELAND (Guillot *et al*. 2005b), ensured that both genetic diversity and spatial patterns were taken into account to delineate the spatial organization of genetic populations.

Several studies focused on wildcat distribution tried to prevent the potentially detrimental effects of fragmentation on small populations gene pool, via the identification of definable phylogenetic units (i.e. taxonomic groups, evolutionary significant units (ESU's, Ryder 1986) or isolated populations) to implement in the effective conservation management plans.

First Pierpaoli *et al.* (2003) have suggested the presence of European wildcat subpopulations and a sharp separation between wildcats of southern and central Europe. In particular, a genetic connection has been disclosed among populations in northwest France, Belgium and Luxembourg (Hille *et al*. 2000) with the German southwest one (Eifel, Hunsrück and Pfälzerwald, southwest of the Rhine-valley). By contrast, the German northern wildcats (Harz and Solling) proved to have suffered dramatic demographic decline in the recent past (Haltenorth 1957), and are still geographically isolated from all the other populations (Pierpaoli *et al.* 2003).

All the other studies on wildcat population structure have been regional, specifically delimited to the Iberian Peninsula (Oliveira *et al*. 2008), Germany (Eckert *et al.* 2009; Hertwing *et al.* 2009) and France (Say-L *et al.* 2012). However, the lacking of connected sampling over the entire wide home range distribution of the subspecies, have prevented the possibility to assess the genetic signatures of past and recent bottlenecks, the phylogeographical origin and evolution of current populations and, thus, the identification of significant conservation units needed for the subspecies survival.

1.4.1. Toward an implemented set of informative autosomal (SNPs and STRs) and uniparental (mtDNA and Y) molecular markers

Recent model-based Bayesian statistical techniques combined with highly polymorphic molecular markers such as microsatellites (Pritchard *et al*. 2000; Anderson & Thompson 2002) proved to be useful for identifying potential hybridization or hybrid individuals in a wide range of scenarios

(Beaumont *et al*. 2001; Randi *et al*. 2001; Hansen 2002; Randi & Lucchini 2002; Pierpaoli *et al*. 2003) and for defining hybrid zones (Nielsen *et al*. 2003) or introgression (Susnik *et al*. 2004). Nevertheless, detection of 'pure wildcat' and second generation backcrosses (Oliveira *et al*. 2008) are still uncertain because of the low number and low resolution of genetic markers used until now (Vähä & Primmer 2006; O'Brien *et al.* 2009; Hertwig *et al.* 2009). An efficient F_1 hybrid detection, in fact, might be achieved with a minimum number of 12 or 24 loci with pairwise F_{ST} between hybridizing parental populations of 0.21 or 0.12, respectively, while the backcrosses identification might require at least 48 loci (Vaha & Primmer 2006). However, all previous cat studies have been based on less than 20 unlinked microsatellites.

Together with the quality and quantity of molecular markers, great relevance on hybrid detection have been provided by the use of simulations and reference pure or hybrid genotypes to set the thresholds of individual *qi* values for identifying admixed individuals, and to estimate the range of their confidence intervals (CI) (Randi 2008). On the other hand, the advantages of using physically-linked loci (Falush *et al*. 2003; Lecis *et al.* 2006;Vähä & Primmer 2006; Oliveira *et al.* 2008) and *a priori* population information (Lecis *et al.* 2006; O' Brien *et al*. 2009; Randi *et al.* 2001) in admixture analyses are still matter of some controversy.

Great advances are expected by using uniparental markers to statistically improve the detection of pure and admixed individuals and unearthing ancient gene flow (Hertwing *et al*. 2009). The selection of specific mtDNA region including diagnostic mutations distinctiveness for wild and domestic cat, for example, might be useful to detect past (maternal) introgression events (Randi 2008). Recently, Driscoll *et al.* (2011) has developed a multi locus protocol combining 36 unlinked microsatellites with two mitochondrial genes, containing 7 diagnostic mutations between the two subspecies (Menotti-Raymond *et al.* 2003), that proved to be reliable for hybridization detection. Moreover, extremely suitable appeared the combining use of slow (e.g., SNPs) and more rapid (e.g., microsatellites) evolving polymorphic markers on the Y chromosome to solve recent population ancestry in humans (de Knijff 2000; Hill *et al.* 2000). For instance, sequences from *SMCY*, *SRY*, *UBE1Y*, *ZFY* genes and SINEs within the Y chromosome have been recently used to elucidate the Felidae phylogeny, providing a valuable patrilineal perspective of cat evolution (Pecon-Slattery & O'Brien 1998; Pecon-Slattery *et al.* 2000a; Pecon-Slattery *et al.* 2000b, Pecon-Slattery *et al.* 2004).

However, the most promising outcomes to improve wildlife conservation-oriented studies and address questions in evolutionary biology and ecology are expected by scanning the entire genome of this endangered species. Over the past few decades, in fact, several important biological and

informatics resources, required for genetic investigations in cats, were developed (O' Brien 2009). The great affection toward cats raised after the discovery of about 250 feline infectious agents analogous to human disorders. Thus, in few times, cats began powerful natural models to study and cure human diseases (O' Brien 2009). A number of genome projects were developed and increased the coverage of domestic cat genome (Pointius *et al.* 2007; Mullikin *et al.* 2010), along with the latest autosomal (Menotti-Raymond *et al.* 2009) and X genetic linkage maps (Schmidt-Küntzel *et al.* 2009), providing a large number of single nucleotide polymorphisms (SNPs) and various previously undiscovered nuclear mitochondrial DNA (Numt) sequences distributed across cat chromosomes (O' Brien 2008). Furthermore, Davis *et al.* (2009) recently reported a high-resolution cat radiation hybrid (RH) map that constitutes a comprehensive framework for identifying genes controlling feline phenotypes of interest. In the last few years, ~9.55 million SNPs from the three combined genome sequencing efforts were submitted to Illumina to produce an ~63K array. The SNPs panel was assembled avoiding rare SNPs, SNPs near repeats or within duplication sites, and SNPs only found in Cinnamon, the Abyssinian cat used for genome sequencing. The Illumina Infinium iSelect 63K Cat DNA Array was tested using 288 cats from different 12 breeds, 10 wildcats, 10 western random bred cats and 10 eastern random bred cats, in addition with five trios, the Abyssinian cat (Cinnamon) and 6 cats from the Hill's SNP discovery project. The preliminary discovery of ~5,000 wildcat SNPs has opened the way to largely increase the number of informative loci available for wildcat research.

The most straightforward contribution of genomics to conservation is the enormously increase of genetic markers that have been intractable until now to answer many important questions. Among them, single nucleotide polymorphisms (SNPs) have been recognized as the new frontier for population studies (Allendorf *et al.* 2010; Ouborg *et al*. 2010). SNPs are, in fact, the most promising genetic markers for the study of both artificial selection and neutral variability in genome-wide analyses (Luikart *et al*. 2003; Wayne & Morin 2004) because i) are abundant and widespread in many species' genomes (coding and non-coding regions); ii) evolve in a manner well described by simple mutation models, such as the infinite sites model (Vignal *et al.* 2002); iii) their genotype code can be standardized in public databases allowing a direct comparison among studies, in contrast with microsatellite loci (LaHood *et al.* 2002); iv) provide an alternative to the maternally inherited and linked mitochondrial markers (whose application is moreover affected by the existence of pseudogenes in the nuclear genome, e.g Lopez *et al.* 1996; Antunes *et al*. 2007) and to the high homoplasic microsatellite loci, whose analyses are fraught with difficulties.; v) nuclear

SNPs are measured on the same mutational scale as mitochondrial SNPs, making intergenomic comparisons easier (Brumfield *et al*. 2003).

However, since the mainly bi-allelic transmission might reduce the power of SNPs for parentage analyses (Prodohl *et al.* 1998) or for detecting fine-scale geographic structure, increasing the number of loci is anyway recommended offsetting this drawback (Giordano *et al.* 1999). In species where large numbers of SNPs have been screened (e.g. humans), it has been shown that they can successfully work (Rosenberg *et al*. 2003; Turakulov & Easteal 2003; Lao *et al*. 2006; Paschou *et al*. 2007), and sometimes outperforming microsatellites (Liu *et al*. 2005). In general, the required number of loci is difficult to assess a priori because each study has a different evolutionary context (e.g. Mariette *et al.* 2002; Pritchard *et al.* 2000), and simulation studies are needed to further elucidate SNP numbers and characteristics for population genetics studies (Morin *et al.* 2004). For instance, linkage studies required approximately three times as many SNPs in comparison to microsatellites (Kruglyak 1997). While the relative number of SNPs needed to estimate population genetic parameters with statistical confidence is likely to be the same (Brumfield *et al.* 2003). Nonetheless, the above examples suggest that at least two to six times more SNPs might be necessary to achieve the same resolution as achieved by microsatellite loci (Morin *et al.* 2004).

Until now, only few studies have compared microsatellite and single-nucleotide polymorphism (SNP) diversity within the same samples, achieving equivocal results. For example, a positive correlation has been found in Atlantic salmon (Ryynänen *et al*. 2007), in contrast with the absent of any correlation in wolves and North American coyotes studies (Väli *et al.* 2008; Ouborg *et al.* 2010). On the contrary, Chinook salmon offered a powerful example of the highly performing ability of 41 SNPs to detect population structure in comparison with nine microsatellites and 22 allozymes (Smith *et al*. 2007). In particular, when sample sizes are large and/or F_{ST} is expected to be > 0.01 (Morin *et al.* 2009), SNPs resulted to be more efficient than microsatellites. In addition to this, further SNPs advantages has been detected regarding the lower error rates associated, the presence of a simple mutation model with low homoplasy and the high genotyping efficiency. Since noninvasive techniques are becoming one of the most popular ways of studying critically threatened species, high advantages were also offered by SNPs improving the possibility of efficiently genotyping molecular marker on degraded samples (Morin & McCarthy 2007).

So that, SNPs are becoming efficient tools for wildlife conservation-oriented studies. In particular, the most promising application of SNPs-based genomic approach might provide opportunities to assess differential rates of introgression across different genomic regions. For example, many

diagnostic SNP markers have been used to detect introgression in hybridizing fish taxa (Finger *et al.* 2009; Simmons *et al.* 2009; Hohenlohe *et al.* 2011; Amish *et al*. 2012). In wildcat studies, recently Nussemberg *et al.* (2013) developed a set of 48 diagnostic SNP markers ($F_{ST} > 0.8$) for identifying wildcats, domestic cats, as well as their hybrids and backcrosses. First and second generation hybrids were reliably recognized with this set of SNP markers, but the reduced and regional sampling used did not guarantee their application to the others wildcat populations of Europe. Furthermore, the SNPs power for hybrid identification might have been overestimated by using individuals too much differentiated than average for the simulations. So that, further studies are recommended to increase SNPs availability for admixture detection.

1.5 Objectives and thesis outline

Crossbreeding with widespread free ranging domestic cats and fragmentation of populations resulting from both natural and anthropogenic events, critically threatened the survival of wildcats in Europe (Stahl & Artois 1994). Thus, a long-term conservation management plans need to: i) monitor the distribution and demographic trend of current wildcat populations by analyzing their variability; ii) detect the genetic-geographic structure of European wildcats aiming at understanding the genetic consequences of fragmentation and identifying eventual evolutionary or conservation units; iii) deeply investigate hybridization and its effects; and finally iv) evaluate the mortality due to illegal hunting or road kills (Stahl & Artois 1991).

For 25 years, wildcat hybridization pattern with feral domestic cats had been assessed through molecular approaches in numerous places within its distribution area (French *et al.* 1988; Hubbard *et al.* 1992; Daniels *et al*. 1998; Beaumont *et al*. 2001; Daniels and Corbett 2003; Pierpaoli *et al.* 2003; Kitchener *et al.* 2005; Lecis *et al.* 2006; Oliveira *et al.* 2008a,b; Hertwig *et al.* 2009; O' Brien *et al*. 2009; Eckert *et al*. 2010). However, several drawbacks were detected in previous studies that still hinder the thoughtful knowledge of wildcat populations in Europe; in particular: the low resolution in hybrids detection and admixture classes discrimination (F1, F2, and backcrosses) due to the type and number of markers chose; the reduced and geographical confined sampling analyzed; and the heterogeneity of data and statistical tools that have made difficult the comparison of hybridization pattern detected. Hence, in the light of wildcats' conservation, it has paved the way to develop a wide genetic study combining a large sampling covering the main wildcats home range distribution with an expanded set of molecular markers and novel bio-statistical techniques.

In this context, it has been enhanced my PhD project, carried out in the conservation genetic molecular laboratory of ISPRA (Ozzano Emilia) of Ettore Randi, and resulted from a strictly long-term collaborations between the same laboratory of ISPRA with the Portuguese research genetic centre of CIBIO of Paulo C. Alves and the Davis Comparative Genetics Laboratory of Dr. Leslie Lyons in California.

The guideline followed during the developing of the thesis led to:

i) identify a reliable genetic analysis protocol able to discriminate hybrids from 'pure' reference, assess their distribution over the European wildcat range and deeply investigate on ancient introgressive events;

ii) and concomitantly, detect the genetic variability and the current distribution patterns of European wildcat populations, and moreover, built evolutionary scenarios to assess timing and origin of wildcats fragmentation.

The thesis has been organized in four chapters. The first Chapter resumes the evolution and main threats of the studied subspecies, aiming at providing an exhaustive overview of the current status of European wildcats. In particular, it has been described the complex process of domestication and the biological and evolutionary consequences that it produced on wildcats via introgressive hybridization. In addition, a summary of the new genetic tools obtained by the advances in molecular biology of the last decades, open the way for a promising improvement of existing, overcoming the previously detected drawbacks.

In the second Chapter, we tried to provide a better understanding of the historical factors that shaped the European wildcat population structure in Italy, and led to the identification of significant conservation units. The results are organized in the *Paper I* already published in SCI journal:

Paper I. Mattucci F, Oliveira R, Bizzarri L, Vercillo F, Anile S, Ragni B, Lapini L, Sforzi A, Alves PC, Lyons LA, Randi E (2013) Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecology and Evolution,* **3(8)**, 2443-2458

A number of molecular studies was conducted on Italian wildcats population (Randi & Ragni 1991; Randi *et al.* 2001; Pierpaoli *et al.* 2003; Lecis *et al.* 2006) aiming at developing a suitable set of markers to both distinguish cat subspecies hosted in the Peninsula (*F.s. silvestris* over the Apennine ridge, in Sicily and North-Eastern Alps; *F. s. catus* widespread everywhere; and *F. s. lybica* in Sardinia) and hybrid individuals. However, all previously works experienced the necessity to

increase the number of molecular marker in order to improve the identification of 'pure' and admixed cat populations. Thus, in *Paper I* we analyzed the genetic variability of 35 autosomal unlinked microsatellites in 346 cat tissues, increasing two or three times the number of markers previously used and the sampling. Combining the use of highly polymorphic loci, Bayesian statistical inferences and landscape analyses tools, we provided for the first time evidences that the European wildcat (*Felis silvestris silvestris*) populations currently distributed in Italy differentiated in, and expanded from two distinct glacial refuges, located in the southern Apennines and at the periphery of the eastern Alps. The main genetic separation between the north-eastern Alps and the peninsular and Sicilian wildcats probably originated during the Last Glacial Maximum; on the other hand, genetic and geographic substructure detected between the eastern (Apennine mountains and hills) and western (Maremma hills and lowlands) sides of the Apennine ridge, resulted to be more recent as a consequence of adaptation to specific ecological conditions of the Mediterranean habitats.

The efficiency of molecular and statistical tools used to detect and investigate the origin of the current geographic-genetic substructure of European wildcats in Italy (as disclosed in *Paper I*), led us to extend the survey area trying to cover the majority home range distribution of wildcats in Europe. In the third Chapter, we tried to overcome one of the main drawbacks of previous studies related to the low number of markers and samples used. In this case, the results allowed the design of a manuscript that is currently under preparation:

***Paper II*.** Mattucci F, Oliveira R, Lyons LA, Alves PC, Randi E Population genetic structure of wildcats (*Felis silvestris*) in Europe. In prep

In *Paper II*, we analyzed the genetic variability of 1114 cat tissues collected from 15 localities of Europe in twelve years (from 1998 to 2010) with 38 autosomal unlinked microsatellites. By taking advantage from descriptive statistics, multivariate and Bayesian analyses we: i) evaluated the extent of genetic diversity within and between wild and domestic cat populations; ii) estimated the extent of population structuring and fragmentation; iii) identified genetic signatures of past and recent bottlenecks. Our findings confirm the previous works on the presence of a hybrid swarm in Hungary and Scotland. Cryptic hybrids were detected, nevertheless, also in other not deeply introgressed populations, suggesting that crossbreeding might negatively affect these populations in case of demographic decline and fragmentation.

European wildcat populations are strongly structured into 5 geographic-genetic macro clusters corresponding to: the Italian peninsular & Sicily; Balkans & north-eastern Italy; Germany eastern; central Europe; and Iberian Peninsula. Estimation of past demographic changes and simulations on divergence times led us to hypothesize that central European population might be the consequence of extra-Mediterranean Würm ice age refuge areas in Europe (Northern Alps, Carpathians, and the Bulgarian mountain systems), while the divergence among and within the southern European populations resulted by the Pleistocene bio geographical framework of Europe, with three southern refugia localized in the Balkans, Italian Peninsula and Iberia Peninsula.

Even though, preliminary admixture analyses were performed in both *Paper I* and *Paper II* aiming at identifying 'pure' wildcat to further examine for landscape and genetic structure investigations, simulations inferences revealed an unsatisfying analytical power of microsatellites used to discriminate backcrosses hybrids among the analysed wildcat populations, prompting the search of new informative molecular markers. Thus, in Chapter 4 we performed a wide genetic investigation of novel loci to establish a set of diagnostic tools suitable for accurately detect levels of introgressive hybridization between European wild and domestic cats. Part of the results were organized in a manuscript that is currently under preparation:

***Paper III*.** Mattucci F, Velli E, Lyons LA, Alves PC, Oliveira R, Randi E Combining use of most informative autosomal SNPs with uniparental markers (mtDNA and Y-chromosome) for the assessment of hybridization in European wildcats (*Felis silvestris*). In prep

In *Paper III*, we selected 151 most informative SNPs from the Illumina Infinium iSelect 63K Cat DNA Array to amplify 150 cat tissues, collected from 15 localities of Europe and previously analyzed with 38 unlinked microsatellites. We estimated levels of genetic variability and differentiation among wild and domestic cat populations, and evaluated the power of SNPs to accurately identify admixture events and discriminate the different hybrid categories that might results from crossbreeding (F1, F2 and backcrosses). Results from Bayesian model-based computations of simulated and empiric genotypes showed that the entire set of 151 markers provided successful estimates of admixture, correctly assigning all parental, F1-F2 hybrids and first and second generation backcrosses. A direct comparison between most informative SNPs and STRs proved the outperforming power of SNPs on admixture detection and inferring admixed ancestries. By taking advantage of a productive collaboration with a PhD student of Roma Tre, dealing with non-invasive monitoring of wildcat populations in northern-Apennines, we further analyzed the

mtDNA with 21 SNPs of the ND5 region and 719 bp of the control-region, and a feline specific Y marker.

Results from the combining use of most informative autosomal SNPs with uniparental markers (mtDNA and Y-linked markers) led us to disclose a total of 11 admixed individuals, while the presence of a number of domestic mitochondrial haplotypes shared with some wild individuals led us to hypnotize the possibility that ancient introgressive events might have occurred and that further investigation should be recommended.

The last Chapter (Chapter 5) resumes the most relevant results obtained in this work and the future perspective that may help to identify under selection traits to improve the identification of 'pure' and admixed individuals in wildcats populations, in order to efficiently support management plans.

1.6 References

Adams KL, Cronn R, Percifield R, Wendel JF (2003) Genes duplicated by polyploidy show unequal contribution to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences USA*, **100**, 4649-4654

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613-622

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11(10)**, 697-709

Amish SJ, Hohenlohe PA, Painter S et al. (2012) RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species-diagnostic SNP assays. *Molecular Ecology Resources*, **12**, 653-660

Anderson E, Stebbins GL (1954) Hybridization as an evolutionary stimulus. *Evolution,* **8,** 378-388

Anderson EC, Thompson EA (2002) A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics,* **160(3)**, 1217-1229

Antunes A, Pontius J, Ramos MJ, O'Brien SJ, Johnson WE (2007) Mitochondrial Introgressions into the Nuclear Genome of the Domestic Cat. *Journal of Heredity*, **98**, 414-420

Arbuckle BS (2005) Experimental animal domestication and its application to the study of animal exploitation in Prehistory, in: J.D. Vigne, D. Helmer, J. Peters (Eds.), New methods for the first steps of the domestication, Oxbow Books, Oxford Pp. 18-33

Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics*, **23**, 237-261

Arnold ML, Burke JM (2004) Natural hybridization. In: *Evolutionary Genetics: Concepts and Case Studies* (eds FoxCW, WolfJB). Oxford University Press, Oxford, in press.

Avise JC, Arnold J, Ball RM (1987) Intraspecific phylogeography: the mitochondrial bridge between population genetics and systematics. *Annual Review of Ecology, Evolution and Systematics*, **18**, 489-522

Bar-Yosef O, Meadow RH (1995) The origins of agriculture in the Near East. *Last Hunters, First Farmers*, eds Price TD, Gebauer, A-B (School of American Research Press, Santa Fe, NM), pp 39-94

Barbour RC, Potts BM, Vaillancourt RE (2007) Gene flow between introduced and native Eucalyptus species: morphological analysis of tri-species and backcross hybrids involving *E. nitens*. *Silvae Genetica*, **56**, 127-133

Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology*, **10**, 551-568

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and* Systematics, **16**, 113-148

Beaumont M, Barratt EM, Gottelli D, Kitchener AC, Daniels MJ, Pritchard JK, Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat.*Molecular Ecology*, **10**, 319-336

Beaumont MA, Rannala B, (2004) The Bayesian devolution in genetics. *Nature Reviews Genetics*, **5(4)**, 251-261

Benecke N (1987) Studies on early dog remains from Northern Europe. *Journal of Archaeological Science*, **14**, 31-49

Bennett AF (1990) Habitat corridors and the conservation of small mammals in a fragmented forest environment. *Landscape Ecology*, **4**, 109-122

Biró Z, Szemethy L, Heltai M (2004) Home range sizes of wildcats (Felis silvestris) and feral domestic cats (Felis silvestris catus) in a hilly region of Hungary. *Mammalian Biology*, **69**, 302-310

Biró Z, Lanszki J, Szemethy L, Heltai M, Randi E (2005) Feeding habits of feral domestic cats (Felis catus), wild cats (Felis silvestris) and their hybrids: trophic niche overlap among cat groups in Hungary. *Journal of Zoology*, **266**, 187-196

Bokonyi S (1974) History of domestic mammals in Central and Eastern Europe, Akadémiai Kiad, Budapest

Bradshaw JWS, Goodwin D, Legrand-Defretin V, Nott HMR (1996) Food selection by the domestic cat, an obligate carnivore. *Comparative Biochemistry and Physiology Part A*: *Physiology*, **114**, 205-209

Bruford MW, Bradley DG, Luikart G (2003) DNA markers reveal the complexity of livestock domestication**.** Nature Reviews Genetics*,* **4**, 900-910

Brumfield R, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249-256

Björnerfeldt S, Webster MT, Vilà C (2006) Relaxation of selective constraint on dog mitochondrial DNA following domestication. *Genome Research*, **16**, 990-994

Cameron EK, Bayne EM, Coltman DW (2008) Genetic structure of invasive earthworms *Dendrobaena octaedra* in the boreal forest of Alberta: insights into introduction mechanisms. *Molecular Ecology*, **17**, 1189-1197

Canters KJ, Thissen JBM, Van Diepenbeek A, Jansman HAH, Goutbeek K (2005) The wildcat (*Felis silvestris*) finally recorded in the Netherlands. *Lutra*, **48**, 67-90

Clutton-Brock J (1981) Domesticated animals, British Museum (Nat. Hist.)

Clutton-Brock J (1999) A Natural History of Domesticated Mammals, Cambridge Univ. Press, New York

Corbett JA (1979) Psychiatric morbidity and mental retardation. In: James F, Snaith R, editors. Psychiatric Illness and Mental Handicap. London: Gaskell Press. Pp 11-25

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New Methods Employing Multilocus Genotypes to Select or Exclude Populations as Origins of Individuals. *Genetics*, **153**, 1989-2000

Cruz F, Vila C, Webster MT (2008) The Legacy of Domestication: Accumulation of Deleterious Mutations in the Dog Genome. *Molecular Biology and Evolution*, **25**, 2331-2336

Daniels MJ, Balharry D, Hirst D, Kitchener AC, Aspinall RJ (1998) Morphological and pelage characteristics of wild living cats in Scotland, implications for defining the wildcat. *Journal of Zoology*, **244**, 231-247

Daniels MJ, Beaumont MA, Johnson PJ, Balharry D, MacDonald DW, Barratt E (2001) Ecology and genetics of wild-living cats in the north-east of Scotland and the implications for the conservation of the wildcat. *Journal of Applied Ecology*, **38**, 146-161

Daniels MJ, Corbett L (2003) Redefining introgressed protected mammals: when is a wildcat a wild cat and a dingo a wild dog? *Wildlife Research,* **30(3)**, 213-218

Davis SJM (1987) The Archaeology of Animals, B. T. Batsford Ltd, London

Davis SJM (2005) Why domesticate food animals? Some zoo-archaeological evidence from the Levant. *Journal of Archaeological Science*, **32(9),** 1408-1416

Davis BW, Raudsepp T, Wilkerson AJP, Agarwala R, Schäffer AA, Houck M, Chowdhary BP, Murphy WJ (2009) A high-resolution cat radiation hybrid and integrated FISH mapping resource for phylogenomic studies across Felidae. *Genomics*, **93**, 299-304

Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. *Journal of Zoology*, **269**, 261-271

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell*, **127**,1309-1321

Dowling TE, Secor CL (1997) The role of hybridization in the evolutionary diversification of animals. *Annual Review of Ecology, Evolution and Systematics,* **28**, 593-619

Driscoll C, Juliet C, Andrew C, O'Brien S (2009) The taming of cats. *Scientific American*, **300**, 68-75

Driscoll C, Yamaguchi N, O'Brien SJ, Macdonald DW (2011) A Suite of Genetic Markers Useful in Assessing Wildcat (*Felis silvestris ssp*.) - Domestic Cat (*Felis silvestris catus*) Admixture. *Journal of Heredity*, **102**, 87-90

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ (2007) The Near Eastern origin of cat domestication. *Science*, **317**, 519-523

Driscoll CA, Macdonald DW, O'Brien SJ (2009) From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, **106(1)**, 9971-9978

Drogemuller C, Rufenacht S, Wichert B, Leeb T (2007) Mutations within the FGF5 gene are associated with hair length in cats. *Animal Genetics*, **38**, 218-221

Duarte J, Vargas JM (2001) Son selectivos los controles de predadores en los cotos de caza? *Galemys*, **13**, 1-9

Easterbee N (1991) The Wildcat. In: The Handbook of British Mammals (eds Corbet GB, Harris S), Pp. 431-437. Blackwell Scientific Publications, Oxford

Eckert J (2003) Predictive values and quality control of techniques for the diagnosis of Echinococcus multilocularis in definitive hosts. *Acta Tropica*, **85**, 157-163

Eckert I, Suchentrunk F, Markov G, Hartl GB (2009) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mammalian Biology*, **75**, 160-174

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ (2003) Molecular genetics and evolution of melanism in the cat family. *Current Biology*, **13**, 448-53

Elton CS (1953) The use of cats in farm rat control. *The British Journal of Animal Behaviour*, **1**, 151-155

Ezard THG, Travis JMJ (2006) The impact of habitat loss and fragmentation on genetic drift and fixation time. *Oikos*, **114**, 367–375

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annual Review *of* Ecology, *Evolution, and Systematics*, **34**, 487–515

Falush D, Stephens M, Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587

Faure E, Kitchener A (2009) An archaeological and historical review of the relationships between felids and people. Anthrozoos, **22(3)**, 221-238

Finger AJ, Stephens MR, Clipperton NW, May B (2009) Six diagnostic single nucleotide polymorphism markers for detecting introgression between cutthroat and rainbow trouts. *Molecular Ecology Resources*, **9(3)**, 759-763

Fitzpatrick BM (2004) Rates of evolution of hybrid inviability in birds and mammals. *Evolution*, **58(8)**, 1865-1870

French DD, Corbett LK, Easterbee N (1988) Morphological discriminant functions of Scottish wildcats (*Felis silvestris*), domestic cats (*F. catus*) and their hybrids. *Journal of Zoology (London)*, **214**, 235-259

Fromont E, Sager A, Léger F, Bourguemestre F, Jouquelet E, Stahl P, Pontier D, Artois M (2000) Prevalence and pathogenicity of retroviruses in wildcats in France. *Veterinary Record*, **146**, 317-9

Gandolfi B, Outerbridge CA, Beresford LG, Myers JA, Pimentel M, Alhaddad H, Grahn JC, Grahn RA, Lyons LA (2010) The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71. *Mammalian Genome*, **21**, 509-515

Germain E, Benhamou S, Poulle ML (2008) Spatio-temporal sharing between the European wildcat, the domestic cat and their hybrids. *Journal of zoology*, **276**, 1 95-203

Germain E, Ruette S, Poulle M-L (2009) Likeness between the food habits of European wildcats, domestic cats and their hybrids in France. *Mammalian Biology Zeitschrift fur Saugetierkunde*, **74**, 412-417

Germonpré M, Sablin MV, Stevens TE, Hedges REM, Hofreiter H, Stiller M, Jaenicke-Desprese V (2008) Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes, *Journal of Archaeological Science*, **36**, 473-490

Giordano M, Oefner PJ, Underhill PA, Luca Cavalli Sforza L, Tosi R, Richiardi PM (1999) Identification by denaturing high-performance liquid chromatography of numerous polymorphisms in a candidate region for multiple sclerosis susceptibility. *Genomics*, **56**, 247-253

Gippoliti S, Amori G (2006) Ancient introductions of mammals in the Mediterranean Basin and their implications for conservation. *Mammal Review*, **36**, 37-48

Gittleman JL, Funk SM, Macdonald DW, Wayne RK (2001) Carnivore conservation (Conservation Biology, Vol. 5). Lavoisier, Cambridge

Godinho R, Llaneza L, Blanco L, Lopes S, Álvares F, García E, Palacios V, Cortés Y, Talegón J and Ferrand N (2011) Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology*, **20**, 5154-5166

Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA, Wayne R (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology*, **3**, 301-312

Grabe H, Worel G (2001) Die Wildkatze. Zurück auf leisen Pfoten, Pp. 14-27. Buch and Kunstverlag Oberpfalz, Amberg

Grant PR, Grant BR, Petren K (2005) Hybridization in the recent past. *American Naturalist*, **166**, 56-67

Guillot G, Mortier F, Estoup A (2005b) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes,* **5(3)**, 712-715

Haltenorth T (1957) Die Wildkatze. Die neue Brehm Bücherei 189. Ziemsen Wittenberg Lutherstadt

Hansen MM (2002) Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. *Molecular Ecology*, **11**, 1003-1015

Hanski I (1999) Metapopulation ecology. Oxford University Press, Oxford, United Kingdom

Harris LD (1984) The Fragmented Forest: Island Biogeography Theory and Preservation of Biotic Diversity. University of Chicago Press, Chicago

He L, Garcia-Perea R, Li M, Wei F (2004) Distribution and conservation status of the endemic Chinese mountain cat *Felis bieti***.** *Oryx*, **38**, 55-61

Henle K, Davies KF, Kleyer M, Margules C, Settele J (2004) Predictors of species sensitivity to fragmentation. In: Henle K, Lindenmayer DB, Margules CR, Saunders DA, Wissel C: Species Survival in Fragmented Landscapes: Where to from now? Special Issue Biodiversity Conservation **13**, 207-251

Herranz J (2001) Efectos de la depredacion y del control de predadores sobre la caza menor en Castilla-La Mancha. Ph.D. Thesis. Universidad Autonoma de Madrid

Hertwig ST, Jungnickel A, Stepanow S, Böhle UR, Fischer MS (2009) Regionally high rates of hybridization and introgression in German wildcat populations (Felis silvestris, Carnivora, Felidae). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 283-297

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. *Genetic Resourses*, **8**, 269-294

Hille A, Pelz O, Trinzen M, Schlegel M, Peters G (2000) Using microsatellite markers for genetic individualization of European wildcats (*Felis silvestris*) and domestic cats. *Bonn. Zool. Beitr*, **49**, 165-176

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, **11**, 117-122

Hope J (1994) The nature of the scrapie agent - the evolution of the virino. *Annals of the New York Academy of Sciences*, **724**, 282-289

Hu Y, Hu S, Wang W, Wu X, Marshall FB, Chen X, Hou L, Wang C (2013) Earliest evidence for commensal processes of cat domestication. *PNAS*, doi: 10.1073/pnas.1311439110

Hubbard AL, McOris S, Jones TW, Boid R, Scott R, Easterbee N (1992) Is survival of European wildcats Felis silvestris in Britain threatened by interbreeding with domestic cats? *Biological Conservation,* **61**, 203-208

Hutchings JA, Fraser DJ (2008) The nature of fisheries and farming induced evolution. *Molecular Ecology,* **17**, 294-313

Ishida Y, David VA, Eizirik E, Schäffer AA, Neelam BA, Roelke ME, Hannah SS, O'brien SJ, Menotti-Raymond M (2006) A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics*, **88**, 698-705

IUCN (2007) *European Mammal Assessment*. IUCN, Gland, Switzerland and Cambridge, UK

Johansson M, Primmer CR, Merila J (2007) Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (*Rana temporaria*). *Molecular Ecology*, **16**, 2693-2700

Johnson WE, O'Brien SJ (1997) Phylogenetic reconstruction of the felidae using 16S rRNA and NADH-5 mitochondrial genes. *Journal of Molecular Evolution*, **44(1)**, 98-116

Johnson WE, Eizirik E, Pecon-Slattery J, Murphy WJ, Antunes A, Teeling E, O'Brien SJ (2006) The Late Miocene Radiation of Modern Felidae: A Genetic Assessment. *Science*, **311 (5757)**, 73-77

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92-103

Kehler JS, David VA, Schaffer AA, Bajema K, Eizirik E, Ryugo DK, Hannah SS, O'Brien SJ, Menotti-Raymond M (2007) Four independent mutations in the feline fibroblast growth factor 5 gene determine the long-haired phenotype in domestic cats. *Journal of Heredity*, **98**, 555-566

Kidd AG, Bowman J, Lesbarrères D, Schulte-hostedde AI (2009) Hybridization between escaped domestic and wild American mink (Neovison vison). *Molecular Ecology*, **18**, 1175-1186

Kitchener AC (1991) The natural history of the wild cats. Cornell university press, New York. Pp. 280

Kitchener AC, Yamaguchi N, Ward JM, Macdonald DW (2005) A diagnosis for the Scottish wildcat (*Felis silvestris*): a tool for conservation action for a critically-endangered felid. *Animal Conservation*, **8**, 223-237

Kitchener AC, Rees EE (2009) Modelling the dynamic biogeography of the wildcat: implications for taxonomy and conservation. *Journal of Zoology*, **279**, 144–155

Kitchener AC, Daniels M, Yamaguchi N, Kilshaw K, Driscoll C (2010) Reversing cryptic extinction: the history, present and future of the Scottish Wildcat. In: D.W. Macdonald and Loveridge A.C. (eds), Biology and conservation of wild felids, Oxford University Press, Oxford, Pp. 471-491

Klar N, Fernández N, Kramer–Schadt S, Herrmann M, Trinzen M, Bu!ttner I, Niemitz C (2008) Habitat selection models for European wildcat conservation. *Biological Conservation*, **141**, 308-319

Klar N Herrmann M, Kramer-Schadt S (2009) Effects and Mitigation of Road Impacts on Individual Movement Behavior of Wildcats. *Journal of Wildlife Management*, **73 (5)**, 631-638

Klein RG, Cruz-Uribe K (1984) The analysis of animal bones from archeological sites, The Univ. Chicago Press, Chicago

Klein RG (1986) The prehistory of Stone Age herders in the Cape Province of South Africa. *South African Archaeological Society Goodwin Series,* **5**, 5-12

Knijff P (2000) Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *The American Journal of Human Genetics*, **67**, 1055-1061

Krüger M, Hertwig ST, Jetschke G, Fischer MS (2009) Evaluation of anatomical characters and the question of hybridization with domestic cats in the wildcat population of Thuringia, Germany. *Journal of Zoological Systematics and Evolutionary Research*, **47**, 268-282

Kruglyak L (1997) The use of a genetic map of biallelic markers in linkage studies. *Nature Genetics*, **17**, 21-24

Kurtén B (1965) The Carnivora of the Palestine Caves. *Acta Zoologica Fennica*, **107**, 1-74

Kurushima JD, Lipinski MJ, Gandolfi B, Froenicke L, Grahn JC, Grahn RA, Lyons LA (2012) Variation of cats under domestication: genetic assignment of domestic cats to breeds and worldwide random-bred populations. *Animal Genetics*, **44(3)**, 311-324

LaHood ES, Moran P, Olsen J, Grant WS, Park LK (2002) Microsatellite allele ladders in two species of Pacific salmon: preparation and field-test results. *Molecular Ecology Notes*, **2**, 187-190

Langley PJW, Yalden DW (1977) The decline of the rarer carnivores in Great Britain during the nineteenth century. *Mammal Review*, **7**, 95-116

Lao O, Van Duijn K, Kersbergen P, Knijff P, Kayser M (2006) Proportioning whole-genome SNP diversity for the identification of geographic population structure and genetic ancestry. *The* American *Journal of* Human Genetics, **78(4)**, 689-90

Lapini L, Molinari P (2006) Nach zehn Jahren taucht in Kärnten die Wildkatze wieder auf. *Carinthia II*, **197(117)**, 59-66

Lecis R, Pierpaoli M, Biro ZS, Szemethy L, Ragni B, Vercillo F, Randi E (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119-131

Lindsay SR (2000) Handbook of Applied Dog Behavior and Training, Volume 1, Adaptation and Learning. Iowa State University Press, Ames, Iowa, p. 12

Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, Pedersen NC, Lyons LA (2008) The ascent of cat breeds: Genetic evaluations of breeds and worldwide random-bred populations. *Genomics*, **91(1)**, 12-21

Liu N, Chen L, Wang S, Oh C, Zhao H (2005) Comparison of singlenucleotide polymorphisms and microsatellites in inference of population structure. *BMC Genetics*, **6(1)**, 26

Lyons L (2010) Feline Genetics: Clinical Applications and Genetic Testing Topics in Companion Animal. *Medicine*, **25**, 203-212

Lyons L (2012) Genetic testing in domestic cats. *Molecular and Cellular Probes*, **26(6)**, 224-30

Lyons LA, Imes DL, Rah HC, Grahn RA (2005a) Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (*Felis catus*). *Animal Genetics*, **36**, 119-126

Lyons LA, Foe IT, Rah HC and Grahn RA (2005b) Chocolate coated cats: TYRP1 mutations from brown color in domestic cats. *Mammalian Genome*, **16**, 356-366

Lodé T, Guiral G, Peltier D (2005) European Mink–Polecat Hybridization Events: Hazards from Natural Process? *Journal of Heredity*, **96**, 89-9

Lopez JV, Cevario S, O'Brien SJ (1996) Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. *Genomics*, **33**, 229-246

Lozano J, Virgós E, Malo AF, Huertas DL, Casanovas JG (2003) Importance of scrub-pastureland mosaics for wild-living cats occurrence in a Mediterranean area: implications for conservation of the wildcat (*Felis silvestris*). *Biodiversity and Conservation*, **12**, 921-935

Lozano J, Moleón M, Virgós E (2006) Biogeographical patterns in the diet of the wildcat, *Felis silvestris* Schreber, in Eurasia: factors affecting the trophic diversity. *Journal of Biogeography,* **33**, 1076-1085

Lozano J, Virgós E, Cabezas–Díaz S, Mangas JG (2007) Increase of large game species in Mediterranean areas: is the European wildcat (*Felis silvestris*) facing a new threat? *Biological Conservation*, **138**, 321-329

Lozano J, Virgós E, Cabezas-Díaz S (2013) Monitoring European wildcat *Felis silvestris* populations using scat surveys in central Spain: are population trends related to wild rabbit dynamics or to landscape features? *Zoological Studies*, **52**, 16

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews in Genetics*, **4**, 981-94

Lüps P, Flückiger PF, Peier D, Schmidt P (2002) Fund einer Waldkatze *Felis silvestris* bei Oberbuchsiten. *Mitteilungen der Naturforschenden Gesellschaft des Kantons Solothurn*, **39**, 41-45

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20 (5)**, 229-237

Mariette S, Le Corre V, Austerlitz F, Kremer A (2002) Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology*, **11**, 1145-1156

McDonald DW, Daniels MJ, Driscoll C, Kitchener A, Yamaguchi N (2004) The Scottish Wildcat. Analyses for Conservation and an Action Plan. Wildlife Conservation Research Unit, Oxford.

McGinnity P, Prodhl P, Ferguson A, Hynes R, Maoilidigh N, Baker N, Cotter D, O'Hea B, Cooke D, Rogan G, Taggart J, Cross T (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 2443-2450

McOrist S (1992) Deseases of the European wildcat (*Felis silvestris* Schreber, 1777) in Great Britain. *Rev Sci Tech*, **11(4)**, 1143-9

McOrist S, Kitchener AC (1994) Current threats to the European wildcat, *Felis silvestris*, in Scotland. *Ambio*, **23**, 243-245

Menotti-Raymond M, David VA, Roelke ME, Chen ZQ, Menotti KA, Sun S, Schäffer AA, Tomlin JF, Agarwala R, O'Brien SJ, Murphy WJ (2003) Secondgeneration integrated genetic linkage/radiation hybrid maps of the domestic cat (*Felis catus*). *Journal of Heredity*, **94**, 95-106

Menotti-Raymond M, David VA, Pflueger SM, Lindblad-Toh K, Wade CM, O'Brien SJ, Johnson WE (2008) Patterns of molecular genetic variation among cat breeds. *Genomics*, **91(1)**, 1-11

Menotti-Raymond M, David VA, Schäffer AA, Tomlin JF, Eizirik E, Phillip C, Wells D, Pontius JU, Hannah SS, O'Brien SJ (2009) An autosomal genetic linkage map of the domestic cat, Felis silvestris catus. *Genomics*, **93**, 305-313

Mölich T (2006) Ein Rettungsnetz für die Wildkatze: Das Drei- Länder-Projekt Hessen, Thüringen und Bayern. In: Kleine K, Grosse R (ed), Naturschutz-Akademie Hessen; Bund für Umweltund Naturschutz Deutschland; Institut für Tierökologie und Naturbildung. NAH Akedemie-Berichte 5, NZH Verlag, Wetzlar, Pp 39-44 (in German).

Monterroso P, Brito JC, Ferreras P, Alves PC (2009) Spatial ecology of the European wildcat in a Mediterranean ecosystem: dealing with small radio–tracking datasets in species conservation. *Journal of Zoology*, **279**, 27-35

Morin PA, Luikart G, Wayne RK, SNP-Workshop Group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**, 208-216

Morin PA, McCarthy M (2007) Highly accurate SNP genotyping from historical and low-quality samples. *Molecular Ecology Notes*, **7**, 937-946

Morin PA, Martien KK, Taylor BL (2009) Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, **9**, 66-73

Muller W (2002) The domestication of the wolf - the inevitable first? In: Vigne JD, Peters J, Helmer D (eds): "The First Steps of Animal Domestication: New Archaeological Approaches". Oxbow Books (Oxford), Pp 34-40

Mullikin JC, Hansen NF, Shen L, Ebling H, Donahue WF, Tao W, Saranga DJ, Brand A, Rubenfield MJ, Young AC, Cruz P, NISC Comparative Sequencing Program, Driscoll C, David V, Al-Murrani SW, Locniskar MF, Abrahamsen MS, O'Brien SJ, Smith DR, Brockman JA (2010) Light whole genome sequence for SNP discovery across domestic cat breeds. *BMC Genomics*, **11**, 406

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Grønkjær P (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea, revealed by individual admixture analysis. *Molecular Ecology*, **12**, 1497-1508

Nielsen EE, Bach LA, Kotlicki P (2006) HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Resources*, **6(4)**, 971-973

Norén K, Dalén L, Kvaløy K, Angerbjörn A (2005) Detection of farm fox and hybrid genotypes among wild arctic foxes in Scandinavia. *Conservation Genetics,* **6**, 885-894

Nowell K and Jackson P (1996) European wildcat, *Felis silvestris*, silvestris group Schreber, 1775. In Wild cats: status survey and conservation action plan: 110–113. Nowell, K. and Jackson, P. (Eds). Gland, Switzerland: IUCN

Nussberger B, Greminger MP, Grossen P, Keller LF, Wandeler P (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources,* **13**, 447-460

O'Brien SJ, Mayr E (1991a) Bureaucratic mischief: Recognizing endangered species and subspecies. *Science,* **251**, 1187-1188

O'Brien SJ, Mayr E (1991b) Response. *Science,* **253**, 251-252

O'Brien SJ, Johnson W, Driscoll C, Pontius J, Pecon-Slattery J, Menotti-Raymond M (2008) State of cat genomics. *Trends in Genetics*, **24**, 268-279

O'Brien J, Devillard S, Say L, Vanthomme H, Léger F, Ruette 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 Conservation*, **18**, 2351-2360

Oliveira R, Godinho R, Randi E, Ferrand N, Alves PC (2007) Molecular analysis of hybridisation between wild and domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**, 1-11

Oliveira R, Godinho R, Randi E, Alves PC (2008) Hybridizations versus conservation: are domestic cats threatening the genetic integrity of wildcats (*Felis silvestris silvestris*) in Iberian Peninsula? *Philosophical Transactions of the* Royal Society *B: Biological*, **363**, 2953-2961

Ostberg CO, Duda JJ, Graham JH, Zhang S, Haywood III KP, Miller B, Lerud TL (2011) Growth, Morphology, and Developmental Instability of Rainbow Trout, Yellowstone Cutthroat Trout, and Four Hybrid Generations. *Transactions of the American Fisheries Society,* **140(2)**, 334-344

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. *Trends in Genetics*, **26**, 177-187

Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, Zaibert V, Thorpe N, Evershed RP (2009) The earliest horse harnessing and milking. *Science*, **323(5919)**, 1332-1335

Paetkau D, Calvert W, Stirling I, and Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347-354

Palomo LJ Gisbert J (2002) Atlas de los Mamıferos Terrestres de Espana. Direccin General de Conservacion de la Natureza-SECEM-SECEMU, Madrid

Pascal M, Lorvelec O, Vigne JD (2006). Invasions biologiques et extinctions: 11,000 ans d'histoire des vertébrés en France. Paris: Belin.

Paschou P, Ziv E, Burchard EG *et al.* (2007) PCA-correlated SNPs for structure identification in worldwide human populations. *PLoS Genetics*, **3**, 1672-1686

Pecon-Slattery J, O'Brien SJ (1998) Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. *Genetics*, **148**, 1245-1255

Pecon-Slattery J, Murphy WJ, O'Brien SJ (2000a) Patterns of diversity among SINE elements isolated from three Y-chromosome genes in carnivores. *Molecular Biology and Evolution*, **17**, 825-829

Pecon-Slattery J, Sanner-Wachter L, O'Brien SJ (2000b) Directed gene conversion occurs between mammalian X-Y homologs outside the pseudoautosomal region. *Proceedings of the National Academy of Sciences of the USA*, **97**, 5307-5312

Pecon-Slattery J, Wilkerson AJ, Murphy WJ, O'Brien SJ (2004) Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. *Molecular Biology* and *Evolution*, **21(12)**, 2299-2309

Peichocki R (2001) Lebensräume. Die Verbreitung der Wildkatze in Europa. In: H. Grabe and G. Worel (eds), Die Wildkatze. Zurück auf leisen Pfoten, Pp. 14-27. Buch and Kunstverlag Oberpfalz, Amberg

Pierpaoli M, Biro ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (Felis silvestris) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585-2598

Pionnier-Capitan M, Bemilli C, Bodu P, Cé lé rier G, Ferrié JG , Fosse P, Garcia M, Vigne JD (2011) New evidence for upper Palaeolithic small domestic dogs in South Western Europe. *Journal of Archaeological Science*, under revision

Pocock RI (1951) Catalogue of the genus Felis. London: British Museum

Pollinger JP, Bustamante CD, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK (2005) Selective sweep mapping of genes with large phenotypic effects. *Genome Resources*, **15**, 1809-1819

Pontier D, Fouchet D, Bahi-Jaber N, Poulet H, Guiserix M, Natoli E, Sauvage F (2009) When domestic cat (*Felis silvestris catus*) population structures interact with their viruses. *Comptes rendus biologies*, **332**, 321-328

Pontius JU, Mullikin JC, Smith DR, Lindblad-Toh K, Gnerre S, et al. (2007) Initial sequence and comparative analysis of the cat genome. *Genome Research*, **17**, 1675-1689

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959

Prodohl PA, Loughry WJ, McDonough CM, Nelson WS, Thompson EA, Avise JC (1998) Genetic maternity and paternity in a local population of armadillos assessed by microsatellite DNA markers and field data. *American Naturalist*, **151**, 7-19

Puzachenko AY (2002) Hybrid syndrome and method for identification of hybrids in museum collections of *Felis silvestris* and *Felis libyca*. *Säugetierkundliche Informatien*, **26**, 234-248

Raimer F (2006) Die Wildkatzenpopulationen in Hessen und Niedersachsen seit dem 18. Jahrhundert. In: Kleine K, Grosse R (ed), Naturschutz-Akademie Hessen; Bund für Umwelt- und Naturschutz Deutschland; Institut für Tierökologie und Naturbildung. NAH Akedemie-Berichte 5, NZH Verlag, Wetzlar, Pp 39-44

Ragni B (1981) Gatto selvatico, *Felis silvestris* Schreber, 1777. Distribuzione e biologia di 22 specie di mammiferi in Italia. Consiglio Nazionale delle Ricerche, Rome

Ragni B, Possenti M (1996) Variability of coat-colour and markings system in *Felis silvestris*. *Italian Journal of Zoology*, **63**, 285-292

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, **17**, 285-293

Randi E, Ragni B (1991) Genetic variability and biochemical systematics of domestic and wild cat populations (Felis silvestris: Felidae). *Journal of Mammalogy*, **72**, 79-88

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

Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analysis of microsatellite variation. *Conservation Genetics,* **3**, 31-45

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *PNAS, Proceedings of the National Academy of Sciences*, **94**, 9197-9201

Redding RW (2005) Breaking the mold, a consideration of variation in the evolution of animal domestication. Pp. 41–8 in J-D Vigne, J Peters, and D Helmer (eds.) The First Steps of Animal Domestication. Oxford: Oxbow Books

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, **17**, 230-237

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics*, **28**, 359-389

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 1211-1216

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109

Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. *The American Journal of Human Genetics*, **73**, 1402-1422

Ryynänen HJ, Tonteri A, Vasemägi A, Primmer CR (2007) A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *Journal of Heredity*, **98**, 692-704

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology & Evolution*, **1**, 9-10

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491-494

Sacks B, Moore M, Statham MJ, Wittmer HU (2011) A restricted hybrid zone between native and introduced red fox (*Vulpes vulpes*) populations suggests reproductive barriers and competitive exclusion, *Molecular Ecology*, **20(2)**, 326-341

Say L, Devillard S, Léger F, Pontier D, Ruette S (2012) Distribution and spatial genetic structure of European wildcat in France. *Animal Conservation*, **15**, 18-27

Saunders DA, Hobbs RJ, Mergules CR (1991) Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, **5**, 18-32

Schmidt-Küntzel A, Eizirik E, O'Brien SJ, Menotti-Raymond M (2005) Tyrosinase and tyrosinase related protein 1 alleles specify domestic cat coat color phenotypes of the albino and brown loci. *Journal of Heredity*, **96**, 289-301

Schmidt-Küntzel A, Nelson G, David VA, Schäffer AA, Eizirik E, Roelke ME, Kehler JS, Hannah SS, O'Brien SJ, Menotti-Raymond M (2009) A domestic cat X chromosome linkage map and the sexlinked orange locus-mapping of orange, multiple origins, and epistasis over non-agouti. *Genetics*, **181**, 1415-1425

Schulenburg J (2005) Säugetiere (Mammalia). In: Gu!nther A., Nigmann U. and Achtziger R. (eds.), Analyse der Gefährdungsursachen planungsrelevanter Tiergruppen in Deutschland zur Ergänzung der bestehenden Roten Listen gefährdeter Tiere. Naturschutz und Biologische Vielfalt 21, Bonn-Bad Godesberg, Pp 70-124

Seiler SM, Keeley ER (2009) Competition between native and introduced salmonid fishes: cutthroat trout have lower growth rate in the presence of cutthroat–rainbow trout hybrids. *Canadian Journal of Fisheries and Aquatic Sciences*, **66(1)**, 133-141

Simon O (2006) Die Wildkatze (*Felis silvestris*) als Leitart für den LebensraumverbundKellerwald-Burgwald-Rothaargebirge. In:Kleine K, Grosse R (ed), Naturschutz-Akademie Hessen; Bund für Umweltund Naturschutz Deutschland; Institut für Tierökologie und Naturbildung. NAH Akedemie-Berichte 5, NZH Verlag, Wetzlar, Pp 39-44 (in German)

Smith B (2012) A cultural niche construction theory of initial domestication. *Biological Theory*, **6(3)**, 260-271

Smith CT, Lagrange A, Ardren W (2007) Genetic analysis of juvenile Chinook salmon collected in White Salmon River. United States Fish and Wildlife Service, Abernathy Fish Technology Center Annual Report.

Spitzenberger F (2005) Rote Liste der Säugetiere Österreichs. In: K. P. Zulka (ed.), *Rote Listen gefährdeter Tiere Österreichs*, Böhlau, Wien, Köln, Weimar

Stahl P, Artois M (1991) Status and conservation of the wildcat (*Felis silvestris*) in Europe and around the Mediterranean rim. Convention on the Conservation of European Wildlife and Natural habitats Standing Committee. Council of Europe Nature and Environment Series, **69**, 1-76

Stahl P, Leger F (1992) Le chat sauvage (*Felis silvestris*, Schreber, 1777). Encyclopédie des Carnivores de France (ed. by M. Artois and H. Maurin), Pp. 1-50. Société Francaise pour l'Etude et la Protection des Mammiféres (SFEPM), Bohallard, Puceul

Stahl P, Artois M (1994) Status and conservation of the wildcat (Felis silvestris) in Europe and around the Mediterranean rim. Nature and Environment Series, No. 69, Council of Europe Press, Strasbourg, Pp. 78

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the landscape in landscape genetics. *Heredity*, **98**, 128-142

Sunquist M, Sunquist F (2002) Wild Cats of the World. The University of Chicago Press, Chicago

Sušnik S, Kozmus P, Poklukar J, Meglic V (2004) Molecular characterisation of indigenous *Apis mellifera carnica* in Slovenia, *Apidologie*, **35**, 623-636

Tamlin AL, Bowman J, Hackett DH (2009) Separating wild from domestic American mink based on skull morphometrics. *Wildlife Biology*, **15**, 266-277

Tchernov E (1984) Commensal animals and human sedentism in the Middle East. Animal and Archaeology eds Clutton-Brock J, Grigson C (British Archaeological Reports, International Series, Oxford), pp 91-115

Tchernov E, Valla FR (1997) Two new dogs, and other Natufian dogs, from the Southern Levant. *Journal of Archaeological Science*, **24**, 65-95

Trut LN (1999) Early canid domestication: The farm-fox experiment. *American Scientist*, **87**, 160-169

Turakulov R, Easteal S (2003) Number of SNP loci needed to detect population structure. *Human Heredity*, **55**, 37-45

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends of Ecology Evolution*, **16**, 330-343

Uerpmann HP (1979) Problems of the Neolithization of the Mediterranean Region (original in German). *Supplements to the Tuebingen Atlas of the Near Orient*, series B. (Ludwig Reichert, Weisbaden)

Ward TJ, Bielawski JP, Davis SK, Templeton SW, Derr JN (1999) Identification of domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA markers and the parametric bootstrap. *Animal Conservation***, 2(1),** 51-57

Wayne RK, Morin PA (2004) Conservation genetics in the new molecular era. *Frontiers in Ecology and the Environment*, **2**, 89-7

Wiener P, Wilkinson S (2011) Deciphering the genetic basis of animal domestication. *Proceedings of the Royal Society B*, **278**, 3161-3170

Wiseman R, O'Ryan C, Harley EH (2000) Microsatellite analysis reveals that domestic cat (*Felis catus*) and southern African wild cat (*F. lybica*) are genetically distinct, *Animal Conservation*, **3**, 221-228

Wittmer HU (2001) Home range size, movements, and habitat utilization of three male European wildcats (*Felis silvestris* Schreber, 1777) in Saarland and Rheinland-Pfalz (Germany). *Zeitschrift für Säugetierkunde - Mammalian Biology*, **66**, 365-370

Wolsan M (1993) *Mustela putorius* Linnaeus, 1758 – Waldiltis, Europäischer Iltis, Iltis. – In: Stubbe M & Krapp F (eds.): Handbuch der Säugetiere Europas, Band 5/I.699–769; Wiesbaden (Aula)

Wolsan M, Okarma H (2001) Felis silvestris Schreber, 1775. In: Glowacinski (ed.), Polish red data book of animals, pp. 94-95, 83-84. Warszawa

Wright S (1931) Evolution in mendelian populations. *Genetics*, **16**, 97-159

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63-72

Väli U, Einarsson A, Waits L, Ellegren H (2008) To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*, **17**, 3808-3817

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, **34**, 275-305

Vigne JD (2011) The origins of animal domestication and husbandry: A major change in the history of humanity and the biosphere. *Comptes Rendus Biologies*, **334(3)**, 171-181

Vigne JD, Carrère I, Briois F, Guilaine J (2011) The Early Process of Mammal Domestication in the Near East. Current Anthropology, Vol. 52, No. S4, The Origins of Agriculture: New Data, New Ideas, Pp. 255-271

Vigne JD, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J (2012) First wave of cultivators spread to Cyprus at least 10,600 y ago, *PNAS***, 109 (22)**, 8445-8449

Vilà C, Wayne R (1999) Hybridization between wolves and dogs. *Conservation Biology*, **13**, 195-198

Vilà C, Walker C, Sundqvist A, Flagstad Ø, Andersone Z, Casulli A, Kojola I, Valdmann H, Halverson J, Ellegren H (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity*, **90**, 17-24

Yamaguchi N, Driscoll CA, Kitchener AC, Ward JM, Macdonald DW (2004a) Craniological differentiation between European wildcats (*Felis silvestris silvestris*), African wildcats (*F. s. lybica*) and Asian wildcats (*F. s. ornata*): implications for their evolution and conservation. *Biological Journal of the Linnean Society*, **83**, 47-63

Yamaguchi N, Kitchener AC, Driscoll CA, Ward JM, Macdonald DW (2004b) Craniological differentiation amongst wild-living cats in Britain and southern Africa: natural variation or the effects of hybridisation? *Animal Conservation*, **7**, 339-351

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, **11**, 413-418

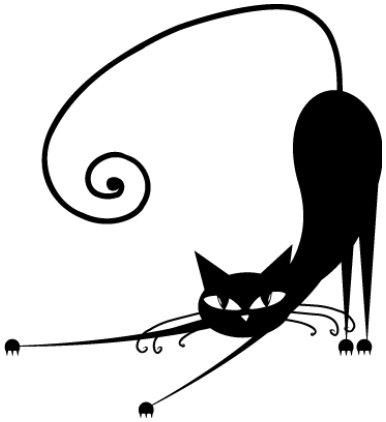
Zeder MA (2006a) Central questions in the domestication of plants and animals. *Evolutionary Anthropology,* **15**, 105-17

Zeder MA (2011) The origins of agriculture in the Near East. *Current Anthropology*, **52(4)**, 221-225

Zeder MA (2012) Pathways to Animal Domestication. In *Harlan II: Biodiversity in Agriculture: Domestication, Evolution, & Sustainability*, edited by Damania A and Gepts P, University of California, Davis, Pp 227-259

Zeder MA, Emshwiller E, Smith BD, Bradley DG, eds (2006) *Documenting Domestication* (Univ of California Press, Berkeley)

Zohary D, Tchernov E, Horwitz LK (1998) The role of unconscious selection in the domestication of sheep and goat, *Journal of Zoology London*, **245**, 129-135



Chapter 2

European wildcats in Italian Peninsula

Genetic structure of wildcat (*Felis silvestris*) populations in Italy

F. Mattucci[1*], R. Oliveira[2], L. Bizzarri[3], F. Vercillo[3], S. Anile[3], B. Ragni[3], L. Lapini[4], A. Sforzi[5], P. C. Alves[2,6], L. A. Lyons[7], E. Randi[1]

[1] Laboratorio di Genetica, ISPRA, Istituto Superiore per la Protezione e la Ricerca Ambientale, Via Cà Fornacetta 9, 40064 Ozzano dell'Emilia (BO), Italy

[2] CIBIO/InBIO Laboratório Associado, Centro de Investigação em Biodiversidade e Recursos Genéticos and Dep. de Biologia da Faculdade de Ciências do Porto, Universidade do Porto, Portugal

[3] Dipartimento di Biologia Cellulare ed Ambientale, Università degli Studi di Perugia, Via Elce di Sotto, 06123 Perugia, Italy

[4] Museo Friulano di Storia Naturale, Via Marangoni 39, 33100 Udine, Italy

[5] Museo di Storia Naturale della Maremma, Strada Corsini 5, 58100 Grosseto, Italy

[6] Wildlife Biology Program, College of Forestry and Conservation, University of Montana, Missoula, USA

[7]School of Veterinary Medicine, Population Health and Reproduction, ,

,

ABSTRACT

Severe climatic changes during the Pleistocene shaped the distributions of temperate-adapted species. These species survived glaciations in classical southern refuges with more temperate climates, as well as in western and eastern peripheral Alpine temperate areas. We hypothesized that the European wildcat (*Felis silvestris silvestris*) populations currently distributed in Italy differentiated in, and expanded from two distinct glacial refuges, located in the southern Apennines and at the periphery of the eastern Alps. This hypothesis was tested by genotyping 235 presumed European wildcats using a panel of 35 domestic cat-derived microsatellites. To provide support and controls for the analyses, 17 know wildcat *x* domestic cat hybrids and 17 Sardinian wildcats (*F. s. libyca*) were included. Results of Bayesian clustering and landscape genetic analyses showed that European wildcats in Italy are genetically subdivided into three well-defined clusters corresponding to populations sampled in: 1) the eastern Alps; 2) the peninsular Apennines; and 3) the island of Sicily. Further, the peninsular cluster is split into two subpopulations distributed on the eastern (Apennine

mountains and hills) and western (Maremma hills and lowlands) sides of the Apennine ridge. Simulations indicated Alpine, peninsular and Sicilian wildcats were isolated during the Last Glacial Maximum. Population subdivision in the peninsula cluster of central Italy arose as consequence of a more recent expansions of historically or ecologically distinct European wildcat sub-populations associated with distinct the Continental or Mediterranean habitats. This study identifies previously unknown European wildcat conservation units and supports a deep phylogeographical history for Italian wildcats.

***Keywords*:** European wildcat, African wildcat, glacial refuges, hybridization, admixture analysis, landscape genetics, conservation genetics

INTRODUCTION

Cyclic glacial/interglacial climatic changes during the Quaternary significantly shaped the biogeography of plant and animal species in the northern Hemisphere. Major glacial refuges have been identified in Iberia, Italy and the southern Balkans, supporting the classical Southern Mediterranean refuge model (Hewitt 2000; Weiss & Ferrand 2006). Two complementary phylogeographic models have been suggested: 1) the existence of cryptic Northern refugia, putatively located in sheltered areas scattered in central Europe north of the Pyrenees and the Alps (Stewart & Lister 2001; Deffontaine *et al.* 2005); and 2) the occurrence of repeated colonization waves of species migrating into continental Europe from eastern European and Asian refugia (Bilton *et al.* 1998; Randi 2007; Fløjgaard *et al.* 2011). In addition, research on paleopollen-reconstructed biomes and species distribution modelling identified a number of plant biodiversity hotspots scattered around the Mediterranean basin that correspond to putative late Pleistocene refugia (Médail & Diadema 2009), driving phylogeographic research to develop more complex hypotheses that include the occurrence of refugia in northern areas and of refugia-within-refugia (Gómez & Lunt 2007). These same refugia hypotheses have been proposed for different organisms. However, clear refugia are still less defined for species with wide geographical ranges, like carnivores (Davison *et al.* 2011; Pilot *et al.* 2010; Johnson *et al.* 2004; Driscoll *et al.* 2007), mainly due to the complex interlinks of past (natural) and recent (mainly anthropogenic) population dynamics. To this end, reconstructing the phylogeography of wildcat (*Felis silvestris*), which have both widespread and fragmented populations, might deepen our knowledge of the carnivores' biogeography in Europe.

Wildcats are widely distributed in three continents (Europe, Africa and Asia), with at least five well-documented subspecies (Driscoll & Nowell 2010). Since the late Pleistocene, and before the dramatic population decline caused in the last centuries by deforestation and human persecution, the European wildcat subspecies (*F. s. silvestris*) was widespread across Europe. The fossil record documents a rapid expansion of the European wildcat throughout most of central and western Europe from the Late Glacial (ca. 15,000 – 9,600 Y B.C.) to the Holocene (Sommer & Benecke 2006). Interestingly, fossils from the Last Glacial Maximum (LGM) were found not only in classical southern Mediterranean refuges in Spain and Italy, but also in western and eastern Alpine peripheral regions (Sommer & Benecke 2006). Mild climatic conditions at low elevation and in coastal areas, plus the buffering effects of the Mediterranean sea, allowed the persistence of temperate forest communities at the periphery of glaciated Alps (Willis & Van Andel 2004). During the LGM, patches of deciduous forest persisted at the tips of the western Alps (the Maritime Alps in

Italy and southern France) and the eastern Alps (in Italy, Slovenia and Istria; Tribsch & Schönswetter 2003; Szövényi *et al.* 2009). Those areas provided refuges for *Fagus*, *Betula*, *Populus* and *Salix* trees (Magri *et al.* 2006; Petit *et al.* 2003; Maliouchenko *et al.* 2007), preserving habitat patches suitable to populations of forest-associated small-mammals (e.g., *Clethrionomis*, *Apodemus* and *Microtus*; Michaux *et al.* 2004; 2005; Fløjgaard *et al.* 2009) and their predators, including the European wildcat (Sommer & Benecke 2006). In the southern Italian peninsula, the LGM drove back European wildcat populations to coastal and mountain southernmost refugia, where rare and scattered fossils have been found in paleontological and archaeological sites (Ragni *et al.* 1994).

Italy currently hosts at least three geographically distinct populations of the European wildcat (Fig. 1), which might represent the living remnants of Pleistocene refugial populations: 1) wildcats in the eastern Italian Alps (Friuli Venezia Giulia and Veneto), which are presumably connected with neighbour populations in Slovenia and Croatia (Lapini 1989; 2006); 2) a widespread population network that is distributed across the central and southern Italian peninsula (Ragni 2006); and 3) an insular population that has been confined to Sicily at least since the LGM (Pierpaoli *et al.* 2003). A remnant wildcat population in the western Alps-Ligurian Apennines might have gone extinct in the first decades of 1980's (Ragni *et al.* 2012). These three populations live in areas included in the Alpine, Continental and Mediterranean biogeographical regions, which are characterized by different habitat types and climates (Schönswetter *et al.* 2005)

The existence of fragmented populations and putative multiple LGM refugia suggests that extant European wildcat populations in Italy have distinct phylogeographical origins and are genetically differentiated (Ragni *et al.* 1994). This study tests this hypothesis by genotyping 235 presumed European wildcats, which were collected from their entire distribution range (excluding the vanished western Alpine wildcats), using a panel of 35 domestic cat-derived microsatellites. These samples showed the typical European wildcat coat colour pattern (Ragni & Possenti 1996), but because of the possible presence of cryptic hybrids, which might be not morphologically recognized; Randi *et al.* 2001; Lecis *et al.* 2006, and before accurate genetic analyses, such as admixture analysis, any wildcat sample was considered as a "putative wildcat". In order to support the identification of hybrids among the studied samples, 17 known European wildcat *x* domestic cat hybrids were also analyzed. Clustering and admixture analyses suggested that some putative wildcats were hybrids, and these cats were subsequently removed from the data set. Subsequent analyses of the identified putative wildcat hybrids will be considered in a forthcoming study on wildcat x domestic cat hybridization across Europe. To compare the extent of genetic diversity among European wildcat populations with divergence among subspecies, domestic cats (*F. s. catus*), and 17 Sardinian wildcats (*F. s. libyca*) were also evaluated. *F. s. libyca* is a subspecies that

originated by anthropochorus introductions of pre-domesticated cats of African or Near Eastern origin into Sardinia and in other Mediterranean islands about 9.000-8.000 years ago (Vigne *et al.* 2012; Ragni 2006). The results provide a better understanding of the historical factors that shaped the European wildcat population structure in Italy, and lead to the identification of significant conservation units.

MATERIALS AND METHODS

Sample collection and DNA extraction

DNA was isolated from 346 tissue, blood and skin samples obtained from 77 free-living or house domestic cats, 235 putative European wildcats, 17 African wildcats from Sardinia (Sardinian wildcats) and 17 previously described *silvestris x catus* hybrids, collected from 2003 to 2010 (Table 1). Seven of these hybrids were obtained from controlled crosses (Ragni 1993); 10 wild-living hybrids, including a family of five full-sibs extracted from the uterus of a road-killed apparently pure *F. s. silvestris* female, that were genetically identified in other studies (Pierpaoli *et al*. 2003; Lecis *et al*. 2006) and reanalyzed here. The European wildcats were opportunistically collected from found-dead or trapped animals in the eastern Italian Alps (n = 78), central (n = 132) and southern (n = 11) Italian peninsula and on the island of Sicily (n = 14; Fig. 1).

Table 1. Subspecies, sampling location and sample size of genotyped cats (*Felis silvestris*) used in this study (see Fig.1 and text for details). * One sample ID 1009 is a museum specimen collected in Maremma.

Subspecies	Sampling locations	Years	*n*	Collectors
Domestic cats *F. s. catus*	Eastern Alps	2003 - 2009	7	L. Lapini
	Central Ap and Maremma	2003 -2009	64	E. Randi, A. Sforzi, A. De Faveri, B. Ragni, A. Giuliani
	Southern Apennines	2009 - 2010	3	E. Mallia
	Sardinia	2007 - 2010	3	R. Oliveira, M. Delogu
European wildcats *F. s. silvestris*	Eastern Alps	2003 - 2009	78	L. Lapini, B. Ragni, A. De Faveri
	Central Ap and Maremma	2003 - 2009	132	A. Sforzi, B. Ragni, A. Giuliani, A. Di Croce, L. Gentile
	Southern Apennines	2003 - 2010	11	B. Ragni, E. Mallia
	Sicily	2003 - 2009	14	S. Anile, B. Ragni
Sardinian wildcats *F. s. libyca*	Sardinia	2003	17	B. Ragni
Captive *silvestris x catus* hybrids	Captivity	2003	7	B. Ragni
Wild-living *silvestris x catus* hybrids	Eastern Alps	2006	5	L. Lapini
	Central Ap and Maremma	2003	4	A. Sforzi, A. Giuliani, B. Ragni
	Sicily	2003	1	W. Trocchi

Climate and habitat types in the Alps and in peninsular Apennines are different, with prevailing coniferous forests and snowy late autumn-winter-spring seasons in the first and broad-leaved forests, shorter snowy winter, in the second (Bransford 2009). The islands of Sicily and Sardinia, have typical Mediterranean climate with mild winter, short raining spring and fall, long, dried, warm summer, and habitat dominated by evergreen sclerophyll wood, chaparral and matorral features. The cores of the European wild cat distribution areas in the Alps and Apennines are at least 550 km distant. Ancient deforestation (beginning before the Roman times; Williams 2006), has continued and extended intensively for about three millennia, and recent heavy urbanization in the lowlands and in the entire Po Valley, are the major anthropogenic barriers which might have caused the fragmentation of extant European wildcat populations (Ragni 2006). Wildcats were phenotypically identified by collectors according to diagnostic coat-colour traits (Ragni & Possenti 1996) and/or biometric indices (skull size and intestinal length; Schauenberg 1969, 1970, 1977), independent of any genetic information. Free-living domestic cats (n = 39) were randomly sampled within the range of wildcats. House cats (n = 38) were sampled from local catteries following veterinary rules (Abrams-Ogg ACG *et al*. 2010).

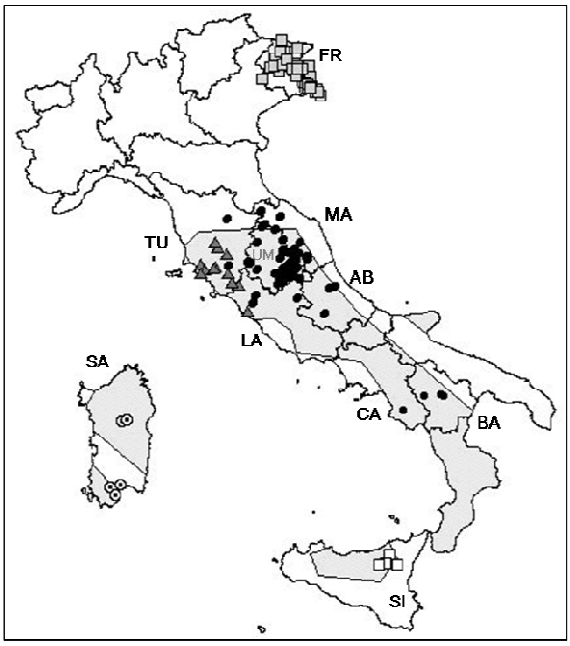


Figure 1. Sampling locations of European and Sardinian wildcats (*Felis silvestris*) used in this study. The grey areas indicate the approximate wildcat distribution ranges in Italy. Each symbol represents a population. Acronyms indicate the sampled regions: Friuli Venezia Giulia (FR) in the eastern Alps; Tuscany (TU), Lazio (LA), Marche (MA), Umbria (UM), Abruzzo (AB) in the central peninsula; Campania (CA); Basilicata (BA) in the southern peninsula; Sicily (SI); Sardinia (SA). The question mark indicates the probably extinct wildcat population in the western Alps – Ligurian Apennines.

Samples were stored in 5 volumes of 95% ethanol (tissues and skins) or Tris/SDS buffer (blood; Longmire *et al*. 1997), and kept at - 20°C. Total DNA was automatically extracted using a MULTIPROBE II^{EX} *Robotic Liquid Handling System* (Perkin Elmer) and the QIAGEN DNeasy tissue and blood extraction kits (Qiagen Inc, Hilden, Germany).

Selection of molecular markers

Thirty-five domestic cat-derived dinucleotide microsatellites (33 autosomal and two X-linked: FCA240 and FCA651; Menotti-Raymond *et al*. 2003), were chosen because of their wide chromosomal distribution and high heterozygosity in diverse domestic cat populations (Lipinski *et al*. 2008). Subsets of the STRs have been evaluated in other wildcat population genetic studies (Pierpaoli *et al.* 2003, Oliveira *et al.* 2008a,b, Eckert *et al.* 2009, O'Brien *et al.* 2009). These markers were amplified in eight multiplexed sets (Supplementary Table S1) using the Qiagen Multiplex PCR Kit with forward tailed-primers fluorescently labelled with 6-FAM, NED, PET, or VIC dyes (Applied Biosystems, Foster City, CA, USA). PCRs were performed in a total volume of 8 µl containing: 2.0 µl of DNA (20-40 ng), 1.0 µl of mix of primers and tails (10mM), 0.2 µl of 10 mg/ml bovine serum albumin, 4.0 µl of Qiagen master mix and 0.8 µl of RNase-free water. Amplifications were performed in a GeneAmp® *PCR* System 9700 (Applied Biosystems) using the touchdown cycling profile: 95 ºC for 15 min, 30 cycles of 30 s at 94°C, 90 s at 62°C (decreasing 0.16°C per cycle to 57°C), 60 s at 72°C, 8 cycles for the annealing of tails of 30 s at 94°C, 90 s at 53°C, 60 s at 72°C, final extension of 10 min at 72°C, followed by a 4°C hold. The genotype of one museum sample (ID 1009; see Table 1) collected in Maremma was consistently obtained after four replicated PCRs, following a multi-tube approach designed for low-quality DNA samples (Taberlet *et al.* 1996). PCR products were analyzed in an ABI 3130 XL (Applied Biosystems) automated sequencer and allele sizes were determined with GeneMapper 4.0 (Applied Biosystems). The power of markers to identify each unique genotype was evaluated calculating the probability-of-identity values (PID and PIDsibs; Mills *et al*. 2000; Waits *et al*. 2001) in GENALEX 6.41 (Peakall & Smouse 2006). Individual genotypes were matched to exclude replicates. About 10% of randomly selected samples were independently replicated twice to assess the rate of allelic dropout (ADO) and false alleles (FA). Presence of null alleles was assessed with MICROCHECKER (Van Oosterhout *et al*. 2004) with an adjusted *P* value corresponding to □ = 0.05 after Bonferroni correction (Rice 1989).

Analysis of genetic variation

Population genetic analyses were performed within and among domestic, Sardinian and European wildcat subspecies, and in the European wildcat sub-populations in Italy, excluding all the admixed genotypes (putative hybrids; see below), which would have affected the estimates of allele frequencies in wildcats. GENALEX was used to estimate allele frequencies, the number of private (N_P), average observed (N_A) and expected (N_E) alleles per locus, and the observed (H_O) and expected unbiased (H_E) heterozygosity for each locus and populations (Peakall & Smouse 2006). Allelic richness (N_{AR}) was computed with FSTAT 2.9.3.2 (Goudet *et al.* 2002). Deviations of loci from Hardy-Weinberg (HWE) and linkage (LE) equilibrium in all populations, and the significance of the Weir & Cockerham's (1984) F_{IS} estimator were evaluated using the Guo and Thompson's (1992) Monte Carlo Markov Chains (MCMC) in GENEPOP 4.0 (Rousset 2008). Sequential Bonferroni corrections for multiple comparisons were used to adjust the significance levels. Single and multilocus *F*- (Weir & Cockerham's 1984) and *R*-statistics (Slatkin 1995) were estimated using GENEPOP and FSTAT, respectively. The partition of genetic diversity within and between groups was obtained by analysis of molecular variance (AMOVA) on Euclidean pairwise genetic distances, using analogues of Wright's *F*-statistics as implemented in GENALEX.

Population structure, assignment testing and identification of hybrid cats

The genetic divergence and phylogeographic structure of the sampled wildcat populations were derived from the results of Bayesian cluster analyses implemented in STRUCTURE 2.3.1 (Pritchard *et al.* 2000; Falush *et al*. 2003). STRUCTURE was used to infer population structure and simultaneously assign the multilocus genotypes to their population (cluster) of origin. Each run of STRUCTURE was replicated five times, with 10^4 burn-in followed by 10^5 simulations, with or without the "admixture" model, with the correlated ("*F*") or the independent ("*I*") allele frequencies models. The optimal number of clusters (*K*) was identified using ΔK and ΔF_{ST} statistics (Evanno *et al*. 2005) in CORRSIEVE 1.6.1 (Campana *et al*. 2011). For each selected *K* value, we assessed: 1) the average proportion of membership (*Q*i) of the sampled populations to the inferred clusters; 2) the individual proportion of membership (q_i) to one or more than one (in case of admixed genotypes) of the inferred clusters; and 3) the 90% credibility intervals (CI) of the q_i values. We used STRUCTURE to perform the following analyses: 1) identification of hybrids in the European wildcats (using sample set *A* = European wildcats, domestic cats and known hybrids; *n* = 332), and in the Sardinian wildcats (using sample set *B* = Sardinian wildcats and domestic cats; *n* = 94); the admixed

genotypes were identified at threshold q_i = 0.90 (based on admixture analyses of observed and simulated cat datasets; Randi 2008, Oliveira *et al.* 2008a) and subsequently removed from the data set; 2) genetic differentiation among the three cat subspecies (using sample set *C* = domestic cats, European and Sardinian wildcats, hybrids excluded; *n* = 295); and 3) inference on population sub-structuring in the European wildcats sampled in peninsular Italy (using sample set *D* = European wildcats only, hybrids excluded; *n* = 202). Details on the STRUCTURE options used in these analyses are indicated in Table 2.

Table 2. Description of the cat sample sets used in this study. The known admixed cats include the captive hybrids (*n* = 7) and the previously identified (*n* = 10) hybrids.

Sample set	Subspecies	Populations	*n*	Analyses
Set A	Domestic cats	Italy	332	Structure analyses
	European wildcats	Italy		*K* = 1 - 10
	Known hybrids	Italy		Admixture, *F* and *I* models
				Option *usepopinfo* not active
Set B	Domestic cats	Italy	94	Structure analyses
	Sardinian wildcats	Sardinia		*K* = 1 - 10
				Admixture, *F* and *I* models
				Option *usepopinfo* not active
Set C	Domestic cats	Italy	295	Structure analyses
	European wildcats	Italy		*K* = 1 - 10
	Sardinian wildcats	Sardinia		No-admixture, *F* and *I* models
				Genetic variability within and among pop
Set D	European wildcats	Eastern Alps	202	Structure analyses
		Maremma		*K* = 1 - 10
		Central and Southern Apennines		No-admixture, *F* and *I* models
		Sicily		Locprior model (Hubisz *et al*. 2009).
				Spatial analyses
Set E	European wildcats	Maremma	105	Spatial analyses
		Central Apennines		

STRUCTURE models are based on the assumption of Hardy-Weinberg and linkage equilibrium (HWLE) in the inferred cluster, which might be violated in empirical data sets (Pritchard *et al.* 2000). Therefore, independently, on explicit population models, patterns of genetic differentiation among cat subspecies (set *C*) and European wildcat populations (set *D*; excluding all the hybrids) were analyzed by a Discriminant Analysis of Principal Components (DAPC) in ADEGENET (Jombart 2008). European wildcat genotypes (set *D*) were also depicted in an unrooted neighbor-joining tree using inter-individual genetic distances (Nei *et al*. 1983) computed with POPULATION 1.2.32 (Langella 2010).

Spatial analyses

The integration of both geographic and genetic information into spatial and landscape genetic models might improve the identification of weakly differentiated populations and yield accurate spatial locations of clusters and genetic barriers (Storfer *et al.* 2007). The spatial clustering of European wildcats in the Italian peninsula, excluding all the hybrids and seven cats without geographical coordinates, $n = 105$; set *E*, was determined by Spatial Principal Component Analysis (sPCA, Jombart *et al.* 2008). Connection networks among neighbouring samples (individuals or populations) were defined through the inverse Euclidean distances algorithm (Jombart 2008). The spatial structure was described by spatial auto-correlations based on Moran's *I* (Moran 1948; Cliff & Ord 1981) and tested by non-parametric randomized regressions of allelic frequencies to global (U+) or local (U-) Moran's Eigenvector Maps (MEMs). For each MEM, a mean coefficient of determination R^2 was generated and the highest values were summed to obtain the test statistic (Jombart 2008). The results were graphically displayed as positive and negative eigenvalues, respectively, for global and local population structure. The spatial distribution of European wildcat clusters was determined in a Poisson-Voronoi tessellation of the sampling areas, using GENELAND 3.3.0 (Guillot *et al.* 2005b). Each run was replicated five times with 100 thinning followed by 10^5 MCMC iterations and with both "*I*" and "*F*" models. GENELAND was run first to estimate the optimal number of sub-populations (with *K* from 1 to 10). Then the spatial structure was obtained by five replicated runs, with the previous parameter values and optimal $K = 5$ (see Results and Supplementary Figure S1E). The level of uncertainty of spatial coordinates was set to 1.4 km, based on estimates of the average wildcat home ranges in the Apennines, Maremma and Sicily (*ca.* 6 km^2; Anile *et al.* 2012; Bizzarri *et al.* 2010, Sforzi *et al.* 2010).

Estimates of population divergence time

Rough assessments of divergence times (in generations) between the European wildcat populations sampled in the eastern Italian Alps, peninsular Italy and Sicily were obtained by simulations. Using EASYPOP 2.0.1 (Balloux 2001), three populations of size $N = 250$, 500 or 1000 were constructed (assuming sex ratio = 1:1 (Tryjanowski *et al.* 2002), 33 loci with free recombination and same mutation rate $\mu = 0.0001$ (Hille *et al.* 2000), single step mutation model (Sainudiin *et al.* 2004), maximum allele number $A = 20$, initial genetic variability set to randomly assign alleles: *var in = max*), which have diverged without gene flow for 50, 500 or 5000 generations. Each run was replicated 10 times. The simulated average number of alleles per locus,

average heterozygosity and F_{ST} were compared to the observed values, and the most likely combinations of population size and divergence times that might have produced the observed genetic diversity were identified.

RESULTS

Genetic variability and identification of cat genotypes and subspecies

All 35 microsatellites were polymorphic, showing from eight (FCA453) to 22 (FCA045) alleles per locus, with the exception of FCA88 and FCA023, which were monomorphic in the Sicilian samples, and FCA035, which was monomorphic in the Sicilian and Sardinian samples (Supplementary Table S1). The independent replication of 10% of the samples provided no evidence for genotyping errors (ADO and FA were equal to zero). Likewise, none of the 35 loci showed significant presence of null alleles. No identical genotypes were observed, and genotype pairs mismatched at a minimum of two loci. Genetic variation statistics are presented in Table 3. The low values of PID suggested that cats in the study were not highly related: PID = 1.5×10^{-34}, PIDsibs = 1.2×10^{-13} in European wildcats; PID = 1.5×10^{-34}, PIDsibs = 7.6×10^{-14} in Sardinian wildcats; and PID = 4.4×10^{-40}, PIDsibs = 6.2×10^{-15} in domestic cats. The PID estimates excluded the 17 known hybrids and 37 newly identified admixed cats identified by admixture analyses (see below).

Allele richness, estimated by rarefaction for a sample size $n = 16$ (the number of successfully genotyped Sardinian wildcats), was similar in the European (5.7 ± 0.3) and Sardinian wildcats (5.8 ± 0.3), and slightly larger in the domestic cats (6.9 ± 0.3). Average values of heterozygosity were similar in the Sardinian ($H_O = 0.66$; $H_E = 0.70$) and domestic cats ($H_O = 0.65$; $H_E = 0.75$), and slightly lower in European wildcats ($H_O = 0.55$; $H_E = 0.69$). There were significant deficit of heterozygotes and departures from HWLE among European wildcats and domestic cats, suggesting population admixture or local sub-structuring. The Sardinian wildcats were in genetic equilibrium (Table 3).

Table 3. Summary of genetic variability at 35 microsatellite loci (33 autosomal and two X-linked STR) in two cat sample sets split into three subspecies and six populations. Sample size = n; N_A = mean number of alleles per locus; N_{AR} = allelic richness estimated by rarefaction on sample sizes n = 16 (genotypes in Sardinia; set C) and n = 11 (genotypes in Sicily; set D); H_O, H_E = observed and expected heterozygosity; F_{IS} = Weir & Cockerham (1984)'s fixation index computed excluding the two X-linked loci (*** significant departures from HWE at $p < 0.001$ corresponding to $p < 0.00028$ after Bonferroni correction for 33 independent comparisons computed excluding the two X-linked loci); HWE = number of loci out of Hardy-Weinberg equilibrium over the total; LE = number of pairwise loci comparisons out of linkage equilibrium over the total; PID = cumulative probability-of-identity; PIDsibs = cumulative Hardy-Weinberg-expected PID among full sib dyads; standard errors in parentheses.

Samples	Subspecies	Populations	n	N_A	N_{AR}*	H_O	H_E	F_{IS}	HWE	LE	PID	PIDsibs
Set C	Domestic cat	Italy	77	11.3	6.9	0.65 (0.02)	0.75 (0.02)	0.11***	1/33	21/595	4.4×10^{-40}	6.2×10^{-15}
	Sardinian wildcat	Sardinia	16	6.4	5.8	0.66 (0.03)	0.70 (0.02)	0.05	0/33	1/595	1.5×10^{-34}	7.6×10^{-14}
	European wildcat	Italy	202	10.2	5.7	0.55 (0.03)	0.69 (0.03)	0.18***	20/33	105/595	1.5×10^{-34}	1.2×10^{-13}
Set D	European wildcat	Eastern Alps	74	7.6	2.5	0.57 (0.03)	0.64 (0.03)	0.09***	2/33	19/595	3.3×10^{-30}	1.9×10^{-12}
		Apennines	117	8.8	2.6	0.54 (0.03)	0.65 (0.03)	0.18***	11/33	10/595	6.4×10^{-32}	6.6×10^{-13}
		Sicily	11	3.6	2.1	0.46 (0.03)	0.50 (0.03)	0.13	0/33	0/595	2.4×10^{-20}	1.9×10^{-9}
Set D	European wildcat	Easter Alps	74	7.6	2.5	0.57 (0.03)	0.64 (0.03)	0.09***	2/33	8/595	3.3×10^{-30}	1.9×10^{-12}
		Maremma	23	6.0	2.4	0.51 (0.03)	0.62 (0.03)	0.18***	2/33	0/595	1.2×10^{-28}	4.5×10^{-12}
		Central &Southern Apennines	94	8.3	2.5	0.55 (0.03)	0.65 (0.03)	0.16***	9/33	2/595	2.4×10^{-31}	9.7×10^{-13}
		Sicily	11	3.6	2.1	0.46 (0.03)	0.50 (0.03)	0.13	0/33	0/595	2.4×10^{-20}	1.9×10^{-9}

The R_{ST} distances among the three subspecies were 2.6 times higher than the corresponding F_{ST} distances (R_{ST} = 0.385; F_{ST} = 0.147, on average), but R_{ST} distances were only 1.7 times higher among the three main European wildcat geographic populations sampled in the Alps, Italian peninsula and Sicily (R_{ST} = 0.192; F_{ST} = 0.109, on average; Table 4). The genetic divergence between domestic cats and Sardinian (African) wildcats (Φ_{ST} = 0.164) was about 30% lower than between domestic cats and European wildcats (Φ_{ST} = 0.222 - 0.265), and between Sardinian and European wildcats (Φ_{ST} = 0.265; AMOVA; Table 4B). The genetic divergence between the European wildcats in the eastern Alps and in Sicily (Φ_{ST} = 0.299) was higher than that between the domestic cats and the two wildcat subspecies (Table 5).

Table 4. (A) Analysis of molecular variance (AMOVA) for cat subspecies and European wildcat populations computed in GENALEX, using Φ_{ST}. The F_{ST} and R_{ST} values were obtained over all loci with GENEPOP and FSTAT, respectively.

Cat subspecies	Populations	Source of variation	Variance	% var	PhiST	F_{ST}	R_{ST}
Domestic cat	Italy (n = 76)	Among groups	8.603	22.64%	0.226 (P < 0.001)	0.147	0.385
Sardinia wildcat	Sardinia (n = 16)	Within groups	29.402	77.36%			
European wildcat	Italy (n = 202)						
	Eastern Alps (n = 74)	Among groups	5.663	17.76%	0.178 (P < 0.001)	0.109	0.192
	Apennines (n = 117) Sicily (n = 11)	Within groups	26.209	82.23%			

Table 5. Genetic divergence among cat subspecies and European wildcat geographic populations. Lower triangular matrix: pairwise estimates of Φ_{ST} (Weir and Cockerham 1984); upper triangular matrix: pairwise estimates of R_{ST} (Michalakis and Excoffier 1996). All values are estimated at 35 microsatellite loci and are highly significant (P < 0.001)

Cat subspecies	Populations	Domestic cat	Sardinia wildcat	European wildcat
Domestic cat	Italy (n = 76)	-	0.170	0.408
Sardinia wildcat	Sardinia (n = 16)	0.164	-	0.522
European wildcat	Italy (n = 202)	0.222	0.265	-
European wildcat	**Populations**	**Eastern Alps**	**Peninsula**	**Sicily**
	Eastern Alps (n = 74)	-	0.195	0.309
	Peninsula (n = 117)	0.165	-	0.097
	Sicily (n = 11)	0.299	0.173	-
European wildcat	**Populations**	**Eastern Alps**	**Maremma**	**Central-Southern Apennines**	**Sicily**
	Eastern Alps (n = 74)	-	0.207	0.197	0.309

Population assignment and admixture analyses

The highest values of ΔK and ΔF_{ST} were obtained in STRUCTURE with K = 2 in both the European (set *A*) and Sardinian (set *B*) wildcats (Supplementary Table S2 and Supplementary

Figure S1). Domestic cats (77 individuals with $Q_I = 0.980$ and individual q_d ranging from 0.881 to 0.996) and European wildcats (202 individuals with $Q_{II} = 0.982$ and $q_W = 0.904$ - 0.997) were assigned to two distinct clusters either using the "*I*" or "*F*" allele frequency model. The Sardinian wildcats (16 individuals with $Q_{II} = 0.990$ and $q_W = 0.910 – 0.998$) and the domestic cats (average $Q_I = 0.993$ and $q_d = 0.950 – 0.999$) were also assigned to distinct clusters. At threshold $q_i = 0.80$, all the 17 known hybrids were confirmed as admixed. Moreover, 36 new admixed European wildcats (with individual q_W ranging from 0.127 to 0.888), and one new admixed Sardinian wildcat (with individual $q_S = 0.680$) were detected in sample set *A* and *B*, respectively. All these putatively admixed samples would have complicated the detection of population structure, and were, therefore, removed from the dataset for subsequent analyses.

Genetic and spatial clustering of European wildcat populations

STRUCTURE analyses on sample set *D* (admixed genotypes excluded), with the "no-admixture" and "*I*" or "*F*" models, no prior information, showed that at $K = 2$ the European wildcats sampled in the eastern Alps clustered separately from the ones sampled in the Italian peninsula and in Sicily; at $K = 3$ also the European wildcats from Sicily clustered separately; and finally the European wildcats split into four distinct subpopulations at $K = 4$, with the identification of a sub-population with geographical distribution restricted to the plains and lower hills of Tuscany and Lazio Maremma. This pattern of population sub-structuring was confirmed using Hubisz *et al*. (2009) sampling location model (Fig. 2A; Supplementary Table S3). GENELAND clustering (with the "*F*" model) splits the European wildcats in set *D* into five clusters ($K = 5$; see Supplementary Figure S1E) including wildcats from: 1) eastern Alps; 2) Mediterranean areas of Tuscany and Lazio Maremma; 3) central peninsular Apennines; 4) southern peninsular Apennines; and 5) Sicily (Supplementary Figure S2). GENELAND analyses performed with only the Italian peninsular wildcats confirmed the existence of two clusters roughly separated by the Apennines ridge, namely grouping: 1) the European wildcat sampled from Tuscany and Lazio Maremma, on the western side of the ridge; and 2) the European wildcats from the Apennines (Emilia-Romagna, Umbria, Marche and Abruzzo regions; Fig. 2B).

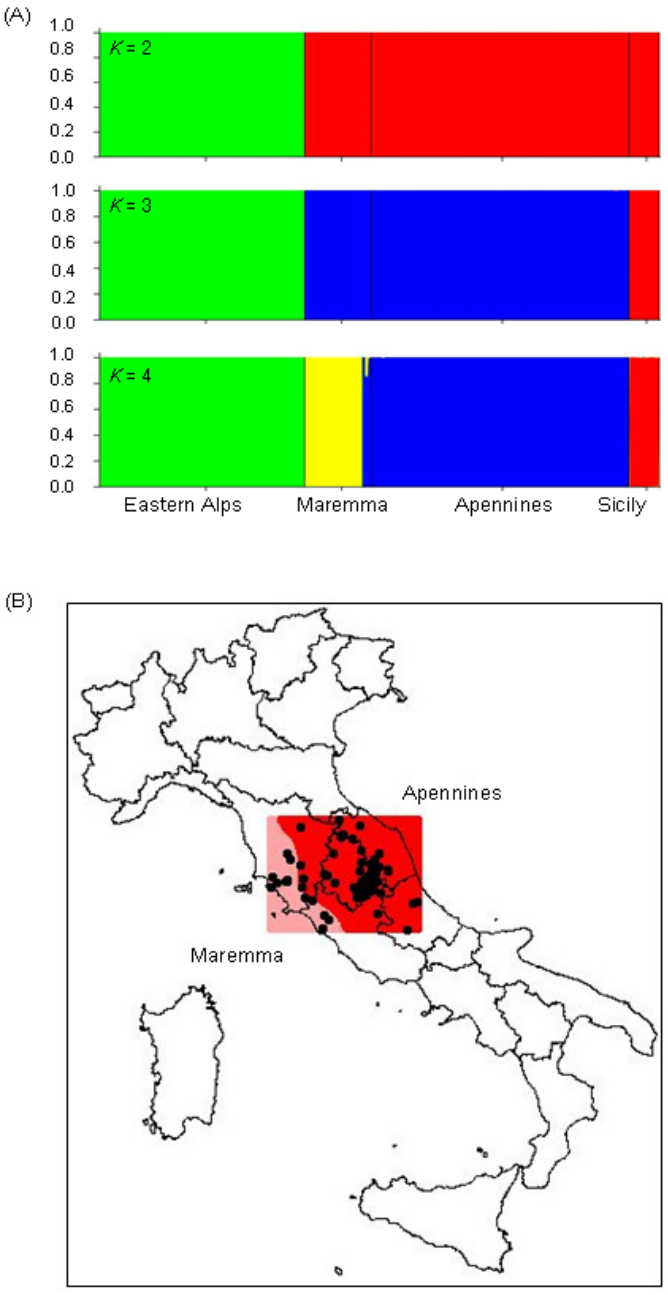


Figure 2. (A) Population clustering obtained in STRUCTURE (with the 'sampling location prior model' and assuming *K* = 2, 3 or 4 genetic clusters) of European wildcats sampled in the eastern Alps, Maremma (areas in Tuscany and Lazio Maremma in the western Italian peninsula; Apennines (areas in Marche, Umbria, Abruzzo, Campania and Basilicata regions), and in Sicily. Each cat genotype is represented by a vertical bar split in *K* coloured sections, according to its relative assignment to the *K* genetic clusters. (B) Maps of posterior probability of European wildcats sampled in central Italy (Maremma and central Apennines) and assigned to two spatial clusters identified by GENELAND.

The DAPC scatter-plots confirmed a sharp distinction among the three cat subspecies (Fig. 3A) and among the four European wildcat sub-populations in Italy (Fig. 3B). The NJ clustering based on Nei's inter-individual genetic distance showed also congruent results: wildcats sampled in the Alps, peninsula and Sicily belong to three distinct clades (respectively labelled 1, 2 and 3 in Fig. 4). However, wildcats sampled in the eastern (Apennines) and western (Lazio and Tuscany Maremma) sides of the central peninsula are only partially split into distinct subclades (*Fsi* E, *Fsi* W). Moreover, some wildcat sampled in the southern Apennines (*Fsi* S) are closely linked to the Sicilian clade, suggesting shared ancestries.

The sPCA analysis of the four European wildcat sub-populations showed a significant correlation between genetic and geographic distances (p-value = 0.047; Mantel test with 999 permutations), which revealed a global structure mainly explained by the first global principal component λ_1. The spatial genetic pattern is visualized in the interpolated gradient map of individual scores (Fig. 5A) and by the individual scores (Fig. 5B). A deep separation occurred along an east-west direction between the European wildcats sampled in the Lazio and Tuscany Maremma and those distributed in the eastern side of the central peninsula (Apennines).

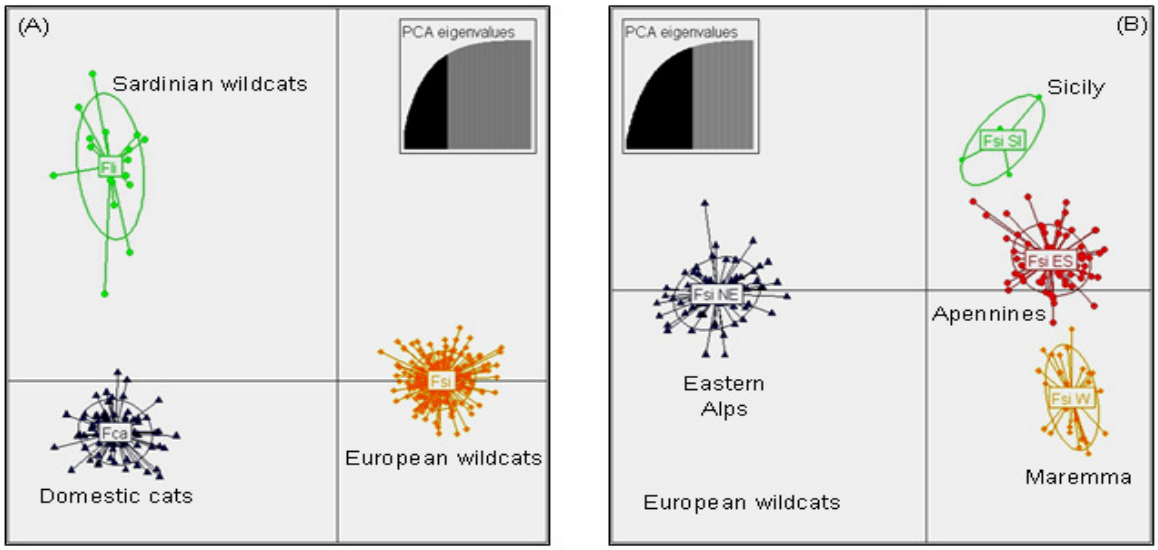


Figure 3. Scatterplot of a Discriminant Analysis of Principal Component (DAPC) obtained with ADEGENET showing genetic distinctions among: (A) three cat subspecies (set *C*, including domestic cats, Sardinian wildcats and European wildcats); (B) all European wildcat subpopulations in Italy (set *D*), including samples from the eastern Alps, Maremma (areas in Tuscany and Lazio Maremma in the western Italian peninsula), Apennines (areas in Marche, Umbria, Abruzzo, Campania and Basilicata regions), and in Sicily. The barplots in the inserts show the proportion of genetic diversity described by each Principal Component (PCA eigenvalues). In all plots the first PC describes 86.81% and 85.61% of the genetic diversity respectively among cat subspecies and populations.

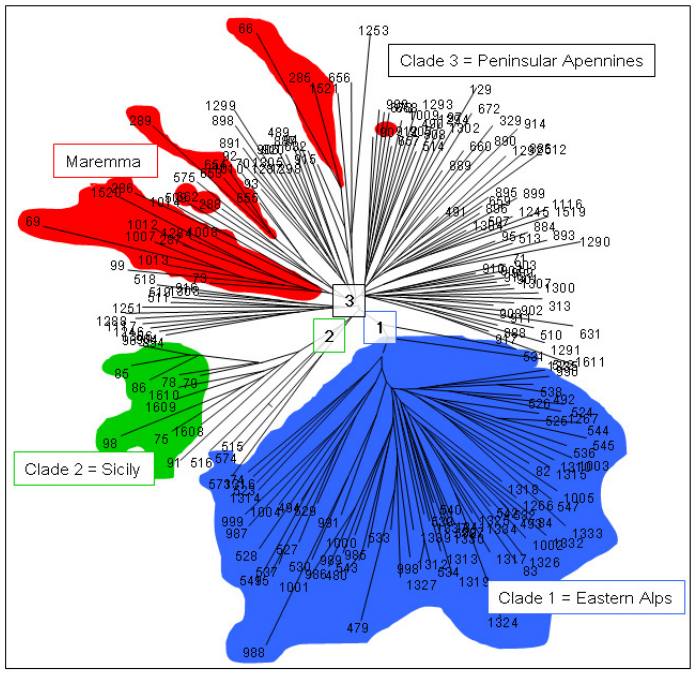


Figure 4. Neighbor-joining tree clustering the pairwise Nei's *et al*. (1983) genetic distances among individual multilocus genotypes of the European wildcats sampled in Italy. The European wildcat sub-populations and the main clades are indicated.

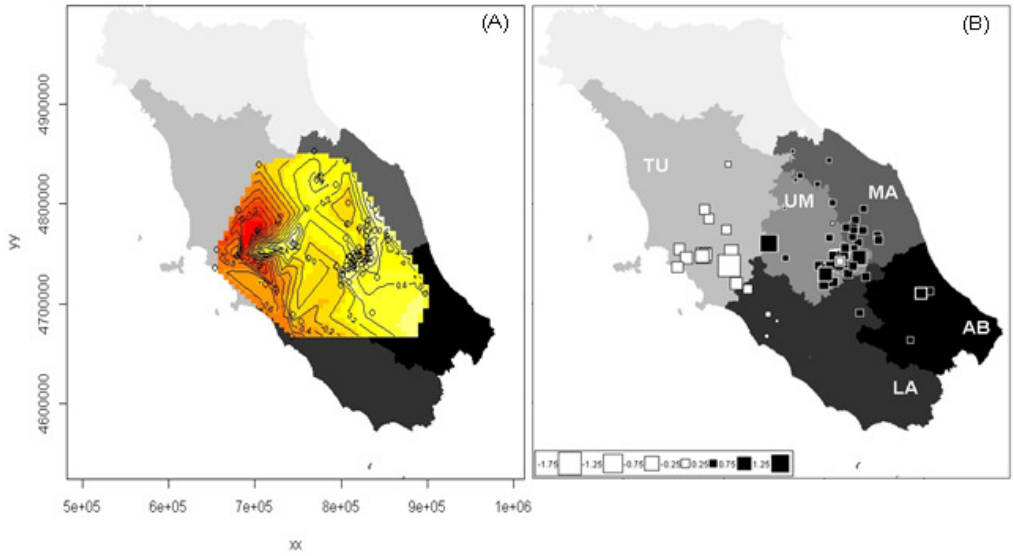


Figure 5. Spatial Principal Component Analysis (sPCA, obtained with ADEGENET) of European wildcats sampled in the central Apennines. (A) Interpolation of the individual genotype scores. The contour lines quantify the degree of genetic differentiation among individuals; circles represent the individual genotypes. (B) Assignment of individual genotypes to their population of origin. Black and white squares represent individual genotype scores on the first principal component (the only significant PC, represented by λ_1 and explaining a significant proportion of spatial structuring). Large white squares indicate individuals with high negative scores; large black squares indicate individuals with highly positive local scores; square dimension is proportional to the degree of differentiations (high for large squares, low for small squares). Letters indicate the sampled regions.

Estimates of population divergence time

Results from EASYPOP simulated populations indicated that a combination of large population size (N = 500 or 1000 breeding individuals) and long divergence times (5000 generations) produced the values of genetic parameters that are most closely correspond to the observed (Fig. 6). Simulated values with N = 1000 and population divergence protracted for 5000 generations (A = 11; H_O = 0.63; F_{ST} = 0.21), compared well with the corresponding observed values (average allele number per locus in the Alpine and peninsular wildcats A = 9; average heterozygosity H_O = 0.57; average divergence between the Alpine and peninsular wildcats F_{ST} = 0.165). Simulations showed also that at smaller population sizes, the effects of genetic drift were too strong as compared to the observed values (F_{ST} values were close to 0.59 and 0.39, allele number were 5 and 8 with n = 250 and 500 respectively after 5000 generations). Moreover 50 or 500 generations of independent evolution were not sufficient to produce the observed values of A, H_O and F_{ST} (for instance, with n = 1000, after 50 or 500 generation the simulated values of F_{ST} were 0.003 and 0.030, respectively, much smaller than the observed F_{ST} = 0.09). Assuming that generation time in the European wildcat is 2 – 3 years, the estimate divergence time between populations in the eastern Alps and in the Italian peninsula should correspond to 10000 – 15000 years, correlating with the end of the LGM.

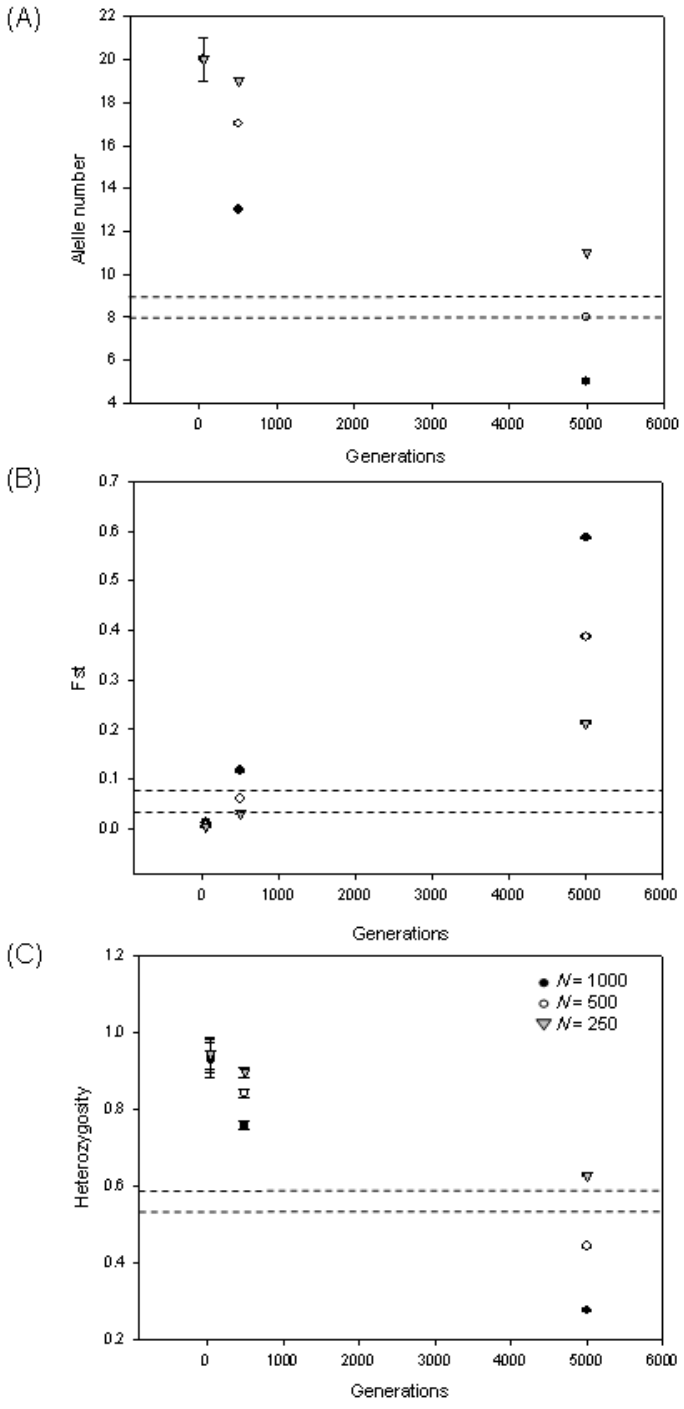


Figure 6. Plot of the average number of alleles per locus (A), F_{ST} values (B) and observed heterozygosity (C), computed simulating two populations of different size (N = 250, 500 and 1000), that were allowed to evolve independently at 33 unlinked autosomal microsatellite loci for 50, 500 and 5000 generations. Horizontal grey blocks identify the average number of alleles per locus (A = 8.77± 0.29), the average heterozygosity (H_O = 0. 57 ± 0.03) observed in European wildcats sampled in the Alps and peninsular Apennines, and the observed F_{ST} (F_{ST} = 0.09) value between them. Dots indicate the average values (+/- their standard errors) obtained from 10 replicate simulations of each parameter settings in EASYPOP.

DISCUSSION

In this study, using an extensive sample size genotyped by a higher number of microsatellite loci than in previous works (Pierpaoli *et al*. 2003; Lecis *et al*. 2006), we describe the main patterns of genetic subdivision of the European wildcat populations distributed in Italy. The cat samples from Italy show values of genetic diversity comparable to other wildcat or domestic cat populations in Europe (Pierpaoli *et al*. 2003, Oliveira *et al*. 2008a,b, O'Brien *et al*. 2009, Eckert *et al*. 2009). Despite risks of genetic erosion in the Sardinian wildcats due to founder effects during the historical introduction process, and in the European wildcats due to long-term population decline and fragmentation, these populations are not genetically depleted. All populations showed allelic richness and heterozygosity comparable to random bred cat populations from the region. The known wildcat x domestic cat hybrids were easily distinguished in the population assignments (q_i = 0.80), providing valid controls and a threshold limit for hybridization detection in the remaining wildcats.

Within Italy, this new dataset shows a subdivision of European wildcats in at least four genetically distinct subpopulations, as it is concordantly supported by model-based, multivariate or distance-based clustering procedures. The European wildcats in the eastern Alps, in peninsular Italy and in Sicily are sharply differentiated. Further, the wildcats in the central Italian peninsula are split into two distinct groups; the first one distributed in the Apennine mountain-hills; the other one in the Maremma hill-plain and coastal areas. This eastern-western subdivision of European wildcats in the central peninsula is clearly described by landscape genetic methods. The assignment of individual genotypes to these two subpopulations is always unambiguous, with the exception of one sample (ID 1009), a museum specimen collected in Maremma, that was partially assigned to the western peninsular subpopulation by both STRUCTURE and sPCA. However, a labelling error can not be excluded for this sample, which was typed from DNA extracted from a museum skin.

This pattern of population structure is concordant with the Pleistocene biogeographical framework of the Italian peninsula, and is congruent with the distribution of fossil cats before, within and after the LGM (Sommer *et al*. 2006). The main population subdivision: eastern Alps, peninsular Italy and Sicily, fits well within a scenario of LGM isolation of European wildcat populations in Mediterranean refuges in southern Italy, on the island of Sicily, and in Cisalpine refuges around the borders of the south-eastern Alps. The subdivision in the central Italian peninsula might be the consequence of more recent expansions of historically or ecologically distinct European wildcat sub-populations associated to distinct habitat types. In particular, populations in the western sector of the range might have experienced periods of isolation and local

adaptation to a peculiar Mediterranean-type habitat known as Maremma. The central Apennines and the Maremma regions, although parts of the same latitudinal range, represent two distinct bioclimatic and ecological regions (Piovesan *et al*. 2005). The Apennines are characterized by temperate-fresh summer-autumns and cold-fresh snowing-raining winter and springs. Deciduous-broad leaved forests and pastures are prevalent in the mountain habitats used by European wildcats. On the contrary, climate in Maremma (the western-coastal part of Tuscany and Latium regions) is influenced by the Tyrrhenian Sea and presents dry and warm spring and summer, temperate and relatively raining autumn and winter, with Mediterranean vegetation composed by sclerophyll and evergreen forest, maquis and garriga (Sforzi & Ragni 1997).

The observed genetic diversity within and among the European wildcat populations in the Alps and in the Italian peninsula might have been jointly generated by large population size (N = 500 or 1000 breeding individuals) and long divergence times (5000 generations). Thus, both observed and simulated genetic parameters suggest that extant European wildcat populations in Italy did not undergo deep historical declines of their effective population sizes, and that genetic divergence among populations can not be explained by recent fragmentation, but by extended periods of isolation without gene flow (in the order of 5000 generations). A scenario of ancient isolation in LGM Alpine and Mediterranean refuges is further supported by the observed genetic divergence between European wildcats in the eastern Alps and in Sicily, which was larger than that between the domestic cats and the two wildcat subspecies. Both the R_{ST} and the F_{ST} distances between the domestic cats, the European wildcats and the African wildcats from Sardinian support previous studies that domestic cats are more closely associated with North African subspecies of wildcats than the European subspecies (Driscoll *et al*. 2007; Lipinski *et al*. 2008). The phylogeography of the European wildcat might be further refined by expanding the samples to include other populations, and by expanding the markers by the use of genome-wide SNPs scans. For instance, the European wildcat population in the eastern Alps might be in contact with neighbouring populations in Slovenia and Croatia. An integrated data set, including Italian plus Slovenian and Croatian samples will help to better delineate the phylogeography of European wildcats, and to assess if wildcats in the eastern Alps have been isolated in the hypothesized LGM refuge, or originated by post-glacial expansions of south Balkan source populations.

The integration of STR and SNP data sets, or the analysis of genome-wide SNP scans will also accomplish the need to develop more realistic phylogeographic events (Nussberger *et al*. 2013). The poorly known mutation mechanisms, and the rapid molecular evolutionary rates of STR loci, with the consequent risk of homoplasy, might complicate the reconstruction of the dynamics of populations that have been genetically isolated for thousand of generations. In contrast, the simpler

mutation mechanism of SNPs, and the possibility to select nucleotide substitutions in different regions of the cat genome, might help to reconstruct ancient evolutionary events (Miller *et al*. 2012). In this study, we did not use mtDNA sequences because of: 1) the extensive presence of nuclear-mitochondrial copies (*numts*) in the cat genome (Antunes *et al*. 2007); 2) the uncertain distinction among domestic and wildcats mtDNA haplotypes (but see: Driscoll *et al*. 2011). In these conditions we feel that any mtDNA phylogeographic reconstruction of populations which might have hybridized in the recent past, or that are still hybridizing, is risky, because mtDNA haplotypes of undetected domestic origin might confuse the evolutionary reconstructions of the European wildcat populations.

Conclusions and conservation perspectives

Results in this study add novel details to the reconstructions of the European wildcat population structure in Italy, contributing to better identify significant conservation units that are relevant for wildcat conservation strategies. The discovery of distinct refugial populations dictates the need of conservation plans focusing on the priority to guarantee long term survival of both population networks. The western versus eastern population subdivision in central Italy might be related to peculiar processes of local adaptations to different habitat types, which need to be better understood. The ongoing transition from conservation genetics to conservation genomics will help to answer both theoretical and practical wildcat conservation issues. Whole- or wide-genome screening might identify mutations showing sharp frequency changes among populations, indicative of functional divergence and adaptation to variable ecological conditions and/or domestication. The discovery of selected loci will clarify the evolutionary dynamics of local adaptations in wildcats in the widest contest of comparative ecological genomics (Martin *et al*. 2003). Wildcat populations are threatened by hybridization with free-ranging domestic cats. The discovery of novel diagnostic molecular markers will also help to identify hybrid individuals and areas of genetic introgression which greater precision and efficiency (Nussberger *et al*. 2013).

ACKNOWLEDGEMENTS

We thank everybody who collaborated for providing the cat samples. We are also grateful to all anonymous veterinarians and biologists that assisted in samples collection, and to E. Velli for the cover image. Funding for this study was provided by ISPRA, the Italian Ministry of Environment, the University of California - Davis, Center for Companion Animal Health. Rita Oliveira was

supported by Fundação para a Ciência e a Tecnologia (FCT) through a PhD grant SFRH/BD/24361/2005.

2.1 REFERENCES

Abrams-Ogg ACG, Schneider A (2010) Schalm's veterinary hematology. In: *Principles of Canine and Feline Blood Collection, Processing, and Storage*. (eds. Douglas J, Weiss K, Wardrop J), **94**, pp.155–188

Anile S, Amico C, Ragni B (2012) Population Density Estimation of the European Wildcat (*Felis silvestris silvestris*) in Sicily Using Camera Trapping. *Wildlife Biology in Practice,* **1**, 1-12

Antunes A, Pontius J, Ramos MJ, O'Brien S, Johnson E (2007) Mitochondrial introgressions into the nuclear genome of the domestic cat. *Journal of Heredity*, **98(5)**, 414-420

Balloux F (2001) EASYPOP (version 1.7): A computer program for population genetics simulations. *Journal of Heredity*, **92(3)**, 301-302

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London*, **B265**, 1219-1226

Bizzarri L, Lacrimini M, Ragni B (2010) Live capture and handling of the European wildcat in central Italy. *Hystrix Italian Journal of Mammalogy*, **21(1)**, 73–82

Bransford T (2009) Ecoregions of Italy. In: *Encyclopedia of Earth.* (eds. Hogan CM, Cleveland CJ), Washington DC

Campana MG, Hunt HV, Jones H, White J (2011) CORRSIEVE: software for summarizing and evaluating Structure output. *Molecular Ecology Resources*, **11**, 349–352

Cliff AD, Ord JK (1981) Spatial processes: models and applications, Taylor& Francis

Davison J, Ho SYW, Bray SC, Korsten M, Tammeleht E, Hindrikson M, Østbye K, Østbye E, Lauritzen SE, Austin J, Cooper A, Saarma U (2011) Late-Quaternary biogeographic scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quaternary Science Reviews*, **30**, 418-430

Deffontaine V, Libois R, Kotlik P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux J (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology*, **14**, 1727–1739

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ, Macdonald DW (2007) The Near Eastern origin of cat domestication. *Science*, **317**, 519-523

Driscoll CA, Nowell K (2010) Felis silvestris. In: *IUCN Red List of Threatened Species. International Union for Conservation of Nature* (http://www.iucnredlist.org/)

Driscoll CA, Yamaguchi N, O'Brien S, Macdonald DW (2011) A suite of genetic markers useful in assessing wildcat (*Felis silvestris* ssp.) - domestic cat (*Felis silvestris catus*) admixture. *Journal of Heredity*, **102(1)**, 87-90

Eckert I, Suchentrunk F, Markov G, Hartl GB (2009) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mammalian Biology*, **75(2010)**, 160-174

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 611-2620

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics,* **164**, 1567-1587

Fløjgaard C, Normand S, Skov F, Svenning, JC (2009) Ice age distributions of European small mammals: insights from species distribution modelling. *Journal of Biogeography*, **36**, 1152-1163

Fløjgaard C, Normand S, Skov F, Svenning JC (2011) Deconstructing the mammal species richness pattern in Europe towards an under standing of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. *Global Ecology and Biogeography*, **20**, 218–230

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of Southern European Refugia. Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds. Weiss S, Ferrand N), pp.155–188, Springer, Amsterdam

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology*, **11**, 1103-1114

Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–71

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*, **48,** 203-208

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913

Hille A, Pelz O, Trinzen M, Schlegel M, Peters G (2000) Using microsatellite markers for genetic individualization of European wildcats (*Felis silvestris*) and domestic cats. *Bonner Zoologische Beitrage*, **49**, 165–176

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9(5)**, 1322-1332

Johnson WE, Godoy JA, Palomares F, Delibes M, Fernandes M, Revilla E, O'Brien SJ (2004) Phylogenetic and Phylogeographic Analysis of Iberian Lynx Populations. *Journal of Heredity*, **95(1)**, 19-28

Jombart T (2008) ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101(1)**, 92-10

Langella O (2010) POPULATIONS 1.2.31.Website last modified on October 20, 2010 (accessed on November 1 2011), http://bioinformatics.org/~tryphon/populations/

Lapini L (1989) Il gatto selvatico nella regione Friuli-Venezia Giulia. In: *Fauna* (*1)*, pp. 64-67, Udine

Lapini L (2006) Attuale distribuzione del gatto selvatico *Felis silvestris silvestris* Schreber, 1775 nell'Italia nord-orientale (Mammalia: Felidae). In: *Atti del Museo Civico di Storia Naturale di Venezia* (*57)*, pp. 221–234

Lecis R, Pierpaoli M, Biro` S, Szemethy L, Ragni B, Vercillo F, Randi E (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119-131

Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, Pedersen NC, Lyons LA (2008) The ascent of cat breeds: Genetic evaluations of breeds and worldwide random-bred populations. *Genomics*, **91**, 12-21

Longmire JL, Maltbie M, Baker RJ (1997) Use of 'lysis buffer' in DNA isolation and its implication for museum collections. In: *Occasional Papers of Museum of Texas Tech University* (*163)*, pp. 1-3

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latałowa M, Litt T, Paule L, Roure JM, Tantau I, Van der Knaap WO, Petit RJ, De Beaulieu JL (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221

Maliouchenko O, Palme AE, Buonamici A, Vendramin GG, Lascoux M (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. *Journal of Biogeography*, **34**, 1601–1610

Martin E, Feder ME, Thomas Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Genetics*, **100**, 101-102

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin.*Journal of Biogeography*, **36**, 1333-1345

Menotti-Raymond M, David VA, Agarwala R, Schäffer AA, Stephens R, O'Brien SJ, Murphy WJ (2003) Radiation hybrid mapping of 304 novel microsatellites in the domestic cat genome. *Cytogenetic and Genome Research*, **102**, 272-276

Menotti-Raymond M, David VA, Chen ZQ, Menotti KA, Sun S, Schäffer AA, Agarwala R, Tomlin JF, O'Brien SJ, Murphy WJ (2003) Second-Generation Integrated Genetic Linkage/Radiation Hybrid Maps of the Domestic Cat (*Felis catus*). *Journal of Heredity*, **94**, 95-106

Michaux JR, Libois R, Paradis E, Filippucci MG (2004) Phylogeographic history of the yellow-necked fieldmouse (*Apodemus flavicollis*) in Europe and in the Near and Middle East. *Molecular Phylogenetics and Evolution*, **32**, 788–798

Michaux JR, Libois R, Filippucci MG (2005) So close and so different: comparative phylogeography of two small mammal species, the Yellownecked fieldmouse (Apodemus flavicollis) and the Woodmouse (*Apodemus sylvaticus*), in the Western Palearctic region. *Heredity*, **94**, 52–63

Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Lim Kim H, Burhans RC, Drautz DI, Wittekindt NE, Tomsho LP, Ibarra-Laclette E, Herrera-Estrella L, Peacock E, Farley S, Sage GK, Rode K, Obbard M, Montiel R, Bachmann L, Ingólfsson O, Aars J, Mailund T, Wiig Ø, Talbot SL, Lindqvist C (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1210506109

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. *Ecological Applications*, **10**, 283-294

Moran PAP (1948) The interpretation of statistical maps. *Journal of the Royal Statistical Society*, **10**, 243-251

Nei M, Tajima F (1983) Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics*, 105, 207-217

Nussberger B, Greminger MP, Grossen P, Keller LF, Wandeler P (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources*, doi: 10.1111/1755-0998.12075

O'Brien J, Devillard S, Say L, Vanthomme H, Léger F, Ruette 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 Conservation*, **18**, 2351-2360

Oliveira R, Godinho R, Randi E, Alves PC (2008a) Hybridizations versus conservation: are domestic cats threatening the genetic integrity of wildcats (*Felis silvestris silvestris*) in Iberian Peninsula? *Philosophical Transaction of the Royal Society B*, **363**, 2953-2961

Oliveira R, Godinho R, Randi E, Ferrand N, Alves, PC (2008b). Molecular analysis of hybridisation between wild and domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**,1-11

Peakall R, Smouse PE (2006) GENALEX v.6.1: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6,** 288-295

Petit RJ, Aguinagalde I, Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R Fineschi S Grivet D, Lascoux M, Mohanty A, Muller-Starck G, Demesure-Musch B, Palmé A, Pedro Marti J, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science,* **300**, 1563-1565

Pierpaoli M, Birò ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585-2598

Pilot M, Branicki W, Jędrzejewski W, Goszczyński J, Jędrzejewska B, Dykyy I, Shkvyrya M, Tsingarska M (2010) Phylogeographic history of grey wolves in Europe. *BMC Evolutionary Biology*, **10**, 104

Piovesan G, Biondi F, Bernabei M, Di Filippo A, Schirone B (2005) Spatial and altitudinal bioclimatic zones of the Italian peninsula identified from a beech (*Fagus sylvatica* L.) tree-ring network. *Acta Oecologica*, **27**, 197-210

Pritchard JK, Stephens M, Donnelly PJ (2000) Inference multilocus genotype data. *Genetics*, **155**, 945-959

Ragni B (1993) Status and conservation of the wildcat in Italy. In: *Seminar on the biology and conservation of the wildcat (Felis silvestris*) (ed. Council of Europe Enviromental Encounters), pp. 40-41, Council of Europe Press, Strasbourg

Ragni B, Possenti M, Sforzi A, Zavalloni D, Ciani F (1994) The wildcat in the central-northen Italian Peninsula: a biogeographical dilemma. *Biogeographia*, **17**, 401-414

Ragni B, Possenti M (1996) Variability of coat-colour and markings system in *Felis silvestris*. *Italian Journal of Zoology*, **63**, 285-292

Ragni B (2006) Il Gatto selvatico. In: *Salvati dall'Arca. WWF Italia* (ed. Perdisa A), pp. 35-56, Bologna

Ragni B, Bizzarri L, Tedaldi G, Salvidio S, Vercillo F, Borgna V, Gavagnin P, Randi E, De Curtis O (2012) A long-term study on a small carnivore: the case of *Felis s. silvestris* peninsular range in Italy. *Hystrix* (*in press*)

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

Randi E (2007) Phylogeography of South European Mammals. In: *Phylogeography of Southern European Refugia* (eds. Weiss S, Ferrand N), pp. 101-126, Kluwer Academic Publishers, Amsterdam

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology,* **17**, 285-293

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**,103-106

Sainudiin R, Durrett RT, Aquadro CF, Nielsen R (2004) Microsatellite Mutation Models. *Genetics*, **168(1)**, 383-395

Schauenberg P (1969) L'identification du Chat forestier d'Europe *Felis s. silvestris* Schreber, 1777 par une méthode ostéometrique. *Revue Suisse de Zoologie*, **76**(**2**), 433-441

Schauenberg P (1970) Note sur l'indice cranien du Chat domestique fèral. *Revue Suisse de Zoologie*, **78**(**1**), 209-215

Schauenberg P (1977) Longueur de l'intestin du Chat forestier (*Felis silvestris*). *Mammalia*, **41**, 357-360

Schönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547–3555

Sforzi A, Ragni B (1997) Atlante dei Mammiferi della Provincia di Grosseto. In: *Atti del Museo di Storia Naturale della Maremma Suppl. 16*, pp. 1-190

Sforzi A, Bizzarri L, Ragni B, Paoloni D (2010) Reconstitution of an European wildcat viable population in the Maremma Regional Park: a twenty years experience. In: *Biologia e*

Conservazione dei Felidi in Italia. Parco Nazionale Foreste Casentinesi (eds. Randi E, Ragni B, Bizzarri L, Agostini N, Tedaldi G), pp. 79-82, Forlì

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462

Sommer RS, Nadachowsky A (2006) Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Review*, **36**, 251-265

Sommer RS, Benecke N (2006) Late Pleistocene and Holocene development of the felid fauna (Felidae) of Europe: a review. *Journal of Zoology*, **269**, 7–19

Stewart JR, Lister AM (2001) Cryptic northen refugia and the origins of modern biota. *Trends in Ecology and Evolution*, **16**, 608-613

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the 'landscape' in landscape genetics. *Heredity*, **98**, 128-142

Szövényi P, Arroyo K, Guggisberg A, Conti E (2009) Effects of Pleistocene glaciations on the genetic structure of *Saxifraga florulenta* (Saxifragaceae), a rare endemic of the Maritime Alps. Taxon, **58 (2)**, 532–543

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **26**, 3189–3194

Tribsch A, Schönswetter P (2003) Patterns of endemism and comparative phylogeography confirm palaeo- environmental evidence for Pleistocene refugia in the Eastern Alps. T*axon*, **52**, 477–497

Tryjanowski P, Antczak M, Hromada M, Kuczynski L, Skoracki M (2002) Winter feeding ecology of male and female European wildcats *Felis silvestris* in Slovakia. *Zeitschrift Fur Jagdwissenschaft*, **48**, 49–54

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538

Vigne JD, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J (2012) First wave of cultivators spread to Cyprus at least 10,600 y ago. *Proceedings of the National Academy of Sciences*, **109(22)**, 8445-8449

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249-256

Weir BS, Cockerham CC (1984) Estimating *F*-statistic for the analysis of population structure. *Evolution*, **38**, 1358-1370

Weiss S, Ferrand N (2006) Phylogeography of Southern European Refugia. In: *Evolutionary perspectives on the origins and conservation of European biodiversity,* pp. 377, Springer

Williams M (2006) Deforesting the Earth: From Pre-History to Global Crisis. An Abridgement. Pp. 1-62, University of Chicago Press, Chicago

Willis KJ, Van Andel TH (2004) Trees or no trees? The environments of central and eastern Europe during the last glaciation. *Quaternary Science Reviews*, **23**, 2369–2387

DATA ACCESSIBILITY

Sample locations, IMa2 input files and microsatellite data: DRYAD entry XXX

2.2 SUPPORTING INFORMATION

Supplementary Table S1. Description of 35 microsatellite loci used to genotype the cat samples (*Felis silvestris*) analyzed in this study. Locus identifications (ID) and chromosome assignments are from Menotti-Raymond *et al.* (2003); the asterisk indicates imperfect dinucleotide microsatellites showing some intermediate alleles; primer tails were labelled to fit the design of eight multiplex sets. The allelic range (in base-pairs) and the observed number of alleles at each locus (N_A) are reported. F_{IS} = Weir & Cockerham (1984)'s fixation index for each locus in cat subspecies (data set E) and European wildcat populations (data set F) were computed excluding the admixed genotypes (see Results); *** significant departures from Hardy-Weinberg equilibrium at $p < 0.001$ ($p < 0.00028$ after Bonferroni-correction for 35 independent comparisons). F_{IS} values at X-linked loci FCA240 and FCA651 were computed only in females (N_f = number of females in samples). M = monomorphic loci in wildcats sampled in Sardinia and Sicily.

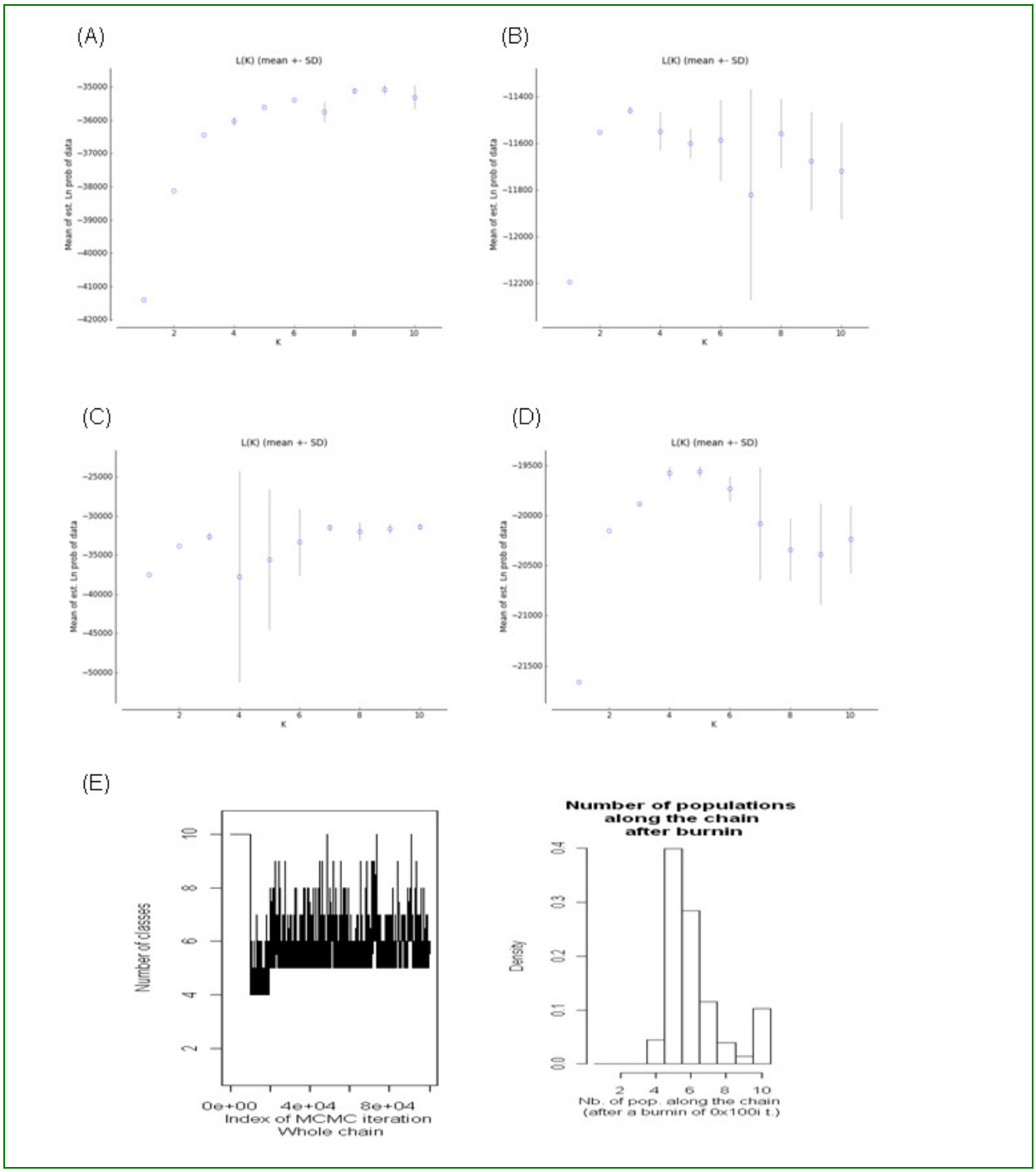
Multiplex	Locus ID	Tail	Chr	Allelic range	N_A	F_{IS} in cat subspecies (data set E)			F_{IS} in European wildcats (data set F)		
						F. s. catus (N_f = 35)	*F. s. silvestris* (N_f = 35)	*F. s. libyca* (N_f = 7)	Alps (N_f = 25)	Peninsular Italy (N_f = 37)	Sicily (N_f = 2)
M19	FCA058	NED	E2	216-249	12	0.07	0.19***	0.01	0.25	0.05	0.19
	FCA077	VIC	C2	156-174	12	0.16	0.18***	0.04	0.06	0.19	0.13
	FCA088	FAM	B3	118-140	12	0.17	0.25	0.25	0.04	0.38***	M
	FCA126	NED	B1	127-170	14	0.11	0.04	0.02	- 0.04	0.06	-0.29
	FCA453	FAM	A1	199-220	8	0.00	0.14	0.05	0.13	0.13	-0.03
M32	FCA023	FAM	B1	142-171	14	0.11	0.20***	- 0.14	0.09	0.15***	M
	FCA045*	NED	D4	157-175	22	0.14	0.13	0.17	0.10	0.15	0.01
	FCA080B	FAM	A3	225-253	14	0.07	0.26***	0.01	0.06	0.32***	0.28
	FCA094	NED	F2	230-259	20	0.10	0.26***	0.50	0.11	0.25***	0.35
	FCA097	VIC	B1	146-173	13	0.22	0.28***	- 0.22	0.11	0.26***	0.97
M34	FCA005	NED	E1	152-168	12	0.11	0.17***	- 0.17	0.06***	0.23***	0.11
	FCA090	FAM	A1	103-129	15	0.09	0.24***	0.03	0.04	0.18***	0.30
	FCA224	FAM	A3	165-188	13	0.29	0.09	0.07	0.14	0.03	0.08
	FCA262	VIC	D2	172-208	17	0.13	0.35***	- 0.01	0.14	0.33***	0.21
M50	FCA008	NED	A1	131-163	15	0.08	0.13***	0.01	0.06	0.08	- 0.16
	FCA043	VIC	C2	134-149	14	-0.05	0.06	0.05	0.07	0.04	-0.27
	FCA096	NED	E2	199-245	21	0.05	0.02	- 0.23	- 0.06	- 0.02	0.52
	FCA293	VIC	C1	196-210	11	0.03	0.05	0.02	0.05	0.04	- 0.09
M51	FCA026	FAM	D3	145-178	13	0.17	0.15***	- 0.26	0.03	0.19	0.19
	FCA035	VIC	D2	154-175	12	0.34	0.60***	M	0.42***	0.76***	M
	FCA132*	NED	D3	154-176	20	0.05	0.09	-0.03	0.03	0.09	- 0.55
	FAC223	PET	F1	210-252	20	0.09	0.40***	0.19	0.28	0.13	0.08
M52	FCA105	FAM	A2	189-220	17	0.06	0.19***	0.06	0.12	0.14	- 0.22
	FCA123	VIC	A1	153-175	11	0.03	0.15***	- 0.20	0.08	0.14	0.02
	FCA211	NED	B1	120-135	12	0.15	0.13***	0.21	0.13	0.14	- 0.23
	FCA305	NED	B2	205-215	14	0.14	0.48***	0.23	0.27	0.54***	0.66
	FCA698	FAM	D1	222-281	16	0.06	0.12	0.11	- 0.02	0.12	0.31
PP1	FCA075	NED	E2	126-159	14	0.06	0.18***	0.20	0.13	0.16***	0.14
	FCA220	FAM	F2	224-238	11	0.22	0.17***	0.49	0.18	0.12	0.20
	FCA229	NED	A1	162-188	14	0.16***	0.23***	- 0.17	0.17	0.16	0.73
	FCA441	FAM	D3	159-186	14	0.21	0.07	0.25	0.02	0.06	0.21
PP2	FCA149	FAM	B1	134-154	10	0.17	0.10	- 0.03	- 0.04	0.14	0.35
	FCA240	FAM	X	157-192	8	0.24	0.29	1.00	0.23	0.36	M
	FCA310	VIC	C2	137-159	15	0.08	0.07	0.17	- 0.05	0.24	- 0.03
	FCA651	NED	X	147-170	9	0.12	0.52***	- 0.25	0.32	0.37	- 0.33

Supplementary Table S2. Identification of the number of *K* clusters in STRUCTURE analyses of cat samples (see: Table 2). Optimal *K* values (in bold) were identified by the maximum increase (ΔK) of the mean *Ln* posterior probability (Mean lnPD) and of the mean F_{ST} values (ΔF_{ST}) between subsequent analyses. NA = not analysed.

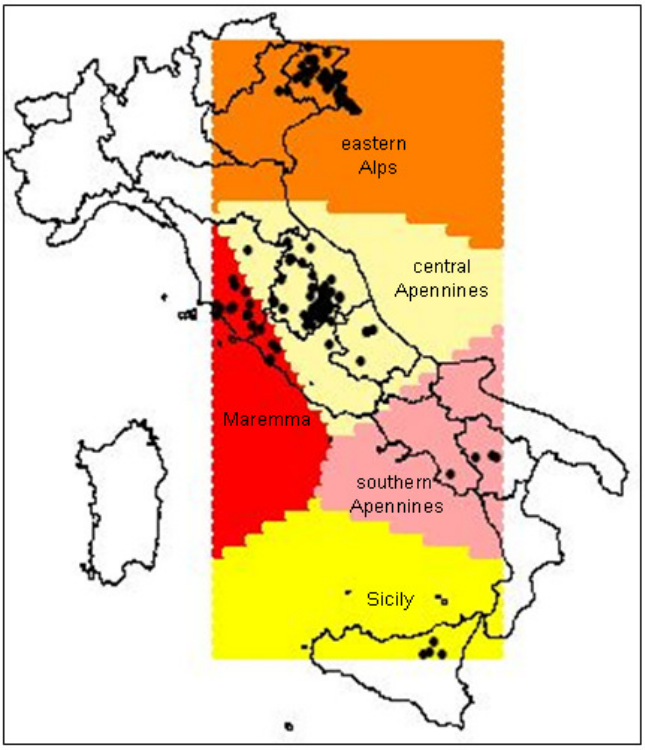
Sample set	Subspecies	Pop	*K*	Mean lnP(D)	ΔK	Mean F_{st}	ΔF_{st}
Set *A*	Domestic cats	Italy	1	-41397.56	NA	0.03005	NA
	European wildcats		**2**	**-38112.4**	**2353.18370**	**0.09881**	**0.75555**
	Known hybrids		3	-36449.06	492.37620	0.13384	0.06787
			4	-36037.46	0.16370	0.17297	0.34892
			5	-35605.14	64.88000	0.17843	0.09777
			6	-35401.36	74.58350	0.17553	0.14723
			7	-35753.88	3.43510	0.18502	0.03392
			8	-35125.60	7.29930	0.19716	0.06592
			9	-35080.54	2.10900	0.20380	0.08766
			10	-35311.25	NA	0.20343	NA
Set *B*	Domestic cats	Italy	1	-12192.74	NA	0.00086	NA
	Sardinian wildcats	Sardinia	**2**	**-11553.48**	**307.63741**	**0.09292**	**1.37619**
	Known hybrids		3	-11458.66	11.83049	0.08630	0.00375
			4	-11548.84	0.47607	0.07943	0.19427
			5	-11601.14	1.06514	0.08521	0.25381
			6	-11587.02	1.43833	0.10706	0.11505
			7	-11819.72	1.09732	0.11897	0.10155
			8	-11558.67	2.58416	0.14113	0.16967
			9	-11676.10	0.36632	0.14325	0.08171
			10	-11717.56	NA	0.15419	NA
Set *C*	Domestic cats	Italy	1	-37473.21	NA	0.01032	NA
	Sardinian wildcats	Italy	**2**	**-33787.65**	**5941.82804**	**0.08830**	**0.95296**
	European wildcats	Sardinia	3	-32608.96	13.51385	0.12182	0.29846
			4	-37754.4	0.54493	0.13884	0.21030
			5	-35569.98	0.00643	0.14369	0.00543
			6	-33327.75	0.09295	0.14888	0.09309
			7	-31478.84	6.38171	0.14714	0.16721
			8	-31950.78	0.72720	0.15791	0.24357
			9	-31618.94	0.18090	0.18674	0.38356
			10	-31397.36	NA	0.17651	NA
Set *D*	European wildcats	Eastern Alps	1	-21658.76	NA	0.02328	NA
		Maremma	**2**	**-20149.68**	**531.02230**	**0.08295**	**2.44673**
		Central and Southern Apennines	3	-19881.08	3.83910	0.08404	0.72811
		Sicily	4	-19571.1	5.18930	0.12372	0.77059
			5	-19558.06	3.64643	0.10360	0.31372
			6	-19731.16	1.42560	0.10766	0.01169
			7	-20079.87	0.16233	0.11254	0.04776
			8	-20337.54	0.69832	0.12061	0.07156
			9	-20383.0	0.38016	0.13344	0.18079
			10	-20237.82	NA	0.129558	NA

Supplementary Table S3. Average proportion membership (Q_i) of wildcat populations obtained by STRUCTURE with K = 2 – 4 and the "localities as prior model", using: the three cat subspecies (sample set *C*); European and Sardinian wildcats (domestic cats excluded); European wildcats (Sardinian wildcats excluded), sampled in eastern Alps, Maremma, central and southern Apennines and Sicily (sample set *D*, see Fig. 2A)

Sample set	**Populations**	***K* = 2**		***K* = 3**						
C		Q_{I}	Q_{II}	Q_{I}	Q_{II}	Q_{III}				
Domestis cats	Italy	1.000	0.000	1.000	0.000	0.000				
European wildcats	Italy	0.000	1.000	0.000	0.366	0.634				
Sardinian wildcats	Sardinia	1.000	0.000	1.000	0.000	0.000				
C		Q_{I}	Q_{II}	Q_{I}	Q_{II}	Q_{III}				
European wildcats	Italy	0.366	0.634	0.000	0.366	0.634				
Sardinian wildcats	Sardinia	1.000	0.000	1.000	0.000	0.000				
							***K* = 4**			
D		Q_{I}	Q_{II}	Q_{I}	Q_{II}	Q_{III}	Q_{I}	Q_{II}	Q_{III}	Q_{IV}
European wildcats	Eastern Alps	1.000	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000
	Maremma	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.887	0.112
	Central and Southern Apennines	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.012	0.987
	Sicily	0.000	1.000	0.000	1.000	0.000	0.000	0.999	0.000	0.001



Supplementary Figure S1. Plot of Log probability [L(*K*)] as a function of *K* averaged over five independent runs of STRUCTURE. The Y-error bars are standard deviation and *K* is the assumed number of genetic clusters. Each plot represents a sample set: (A) sample set *A* (European wildcats, domestic cats and known hybrids); (B) sample set *B* (Sardinian wildcats, domestic cats and known hybrids); (C) sample set *C* (domestic cats, European and Sardinian wildcats); (D) sample set *D* (European wildcats only). (E) Inference of the number of genetic clusters in the study area: posterior distribution of the number of populations estimated using GENELAND.



Supplementary Figure S2. Maps of posterior probability of European wildcats identified by GENELAND. Samples are split into five clusters (K = 5): 1) eastern Alps; 2) Maremma; 3) central peninsular Apennines; 4) southern peninsular Apennines; and 5) Sicily.

2.3 LIST OF TABLE

Table 1. Subspecies, sampling location and sample size of genotyped cats (*Felis silvestris*) used in this study (see Fig.1 and text for details).

Table 2. Description of the cat sample sets used in this study. The known admixed cats include the captive hybrids ($n = 7$) and the previously identified ($n = 10$) hybrids.

Table 3. Summary of genetic variability at 35 microsatellite loci (33 autosomal and two X-linked STR) in two cat sample sets split into three subspecies and six populations. Sample size = n; N_A = mean number of alleles per locus; N_{AR} = allelic richness estimated by rarefaction on sample sizes $n = 16$ (genotypes in Sardinia; set C) and $n = 11$ (genotypes in Sicily; set D); H_O, H_E = observed and expected heterozygosity; F_{IS} = Weir & Cockerham (1984)'s fixation index computed excluding the two X-linked loci (*** significant departures from HWE at $p < 0.001$ corresponding to $p < 0.00028$ after Bonferroni correction for 33 independent comparisons computed excluding the two X-linked loci); HWE = number of loci out of Hardy-Weinberg equilibrium over the total; LE = number of pairwise loci comparisons out of linkage equilibrium over the total; PID = cumulative probability-of-identity; PIDsibs = cumulative Hardy-Weinberg-expected PID among full sib dyads; standard errors in parentheses.

Table 4. Analysis of molecular variance (AMOVA) for cat subspecies and European wildcat populations computed in GENALEX, using Φ_{ST}. The F_{ST} and R_{ST} values were obtained over all loci with GENEPOP and FSTAT, respectively.

Table 5. Genetic divergence among cat subspecies and European wildcat geographic populations. Lower triangular matrix: pairwise estimates of Φ_{ST} (Weir and Cockerham 1984); upper triangular matrix: pairwise estimates of R_{ST} (Michalakis and Excoffier 1996). All values are estimated at 35 microsatellite loci and are highly significant ($P < 0.001$).

Supplementary Table S1. Description of 35 microsatellite loci used to genotype the cat samples (*Felis silvestris*) analyzed in this study. Locus identifications (ID) and chromosome assignments are from Menotti-Raymond *et al.* (2003); the asterisk indicates imperfect dinucleotide microsatellites showing some intermediate alleles; primer tails were labelled to fit the design of eight multiplex sets. The allelic range (in base-pairs) and the observed number of alleles at each locus (N_A) are

reported. F_{IS} = Weir & Cockerham (1984)'s fixation index for each locus in cat subspecies (data set E) and European wildcat populations (data set F) were computed excluding the admixed genotypes (see Results); *** significant departures from Hardy-Weinberg equilibrium at $p < 0.001$ ($p < 0.00028$ after Bonferroni-correction for 35 independent comparisons). F_{IS} values at X-linked loci FCA240 and FCA651 were computed only in females (N_f = number of females in samples). M = monomorphic loci in wildcats sampled in Sardinia and Sicily

2.4 LIST OF FIGURES

Figure 1. Sampling locations of European and Sardinian wildcats (*Felis silvestris*) used in this study. The grey areas indicate the approximate wildcat distribution ranges in Italy. Each symbol represents a population. Acronyms indicate the sampled regions: Friuli Venezia Giulia (FR) in the eastern Alps; Tuscany (TU), Lazio (LA), Marche (MA), Umbria (UM), Abruzzo (AB) in the central peninsula; Campania (CA); Basilicata (BA) in the southern peninsula; Sicily (SI); Sardinia (SA). The question mark indicates the probably extinct wildcat population in the western Alps – Ligurian Apennines.

Figure 2. (A) Population clustering obtained in STRUCTURE (with the 'sampling location prior model' and assuming K = 2, 3 or 4 genetic clusters) of European wildcats sampled in the eastern Alps, Maremma (areas in Tuscany and Lazio Maremma in the western Italian peninsula; Apennines (areas in Marche, Umbria, Abruzzo, Campania and Basilicata regions), and in Sicily. Each cat genotype is represented by a vertical bar split in K coloured sections, according to its relative assignment to the K genetic clusters. (B) Maps of posterior probability of European wildcats sampled in central Italy (Maremma and central Apennines) and assigned to two spatial clusters identified by GENELAND.

Figure 3. Scatterplot of a Discriminant Analysis of Principal Component (DAPC) obtained with ADEGENET showing genetic distinctions among: (A) three cat subspecies (set *C*, including domestic cats, Sardinian wildcats and European wildcats); (B) all European wildcat subpopulations in Italy (set *D*), including samples from the eastern Alps, Maremma (areas in Tuscany and Lazio Maremma in the western Italian peninsula), Apennines (areas in Marche, Umbria, Abruzzo, Campania and Basilicata regions), and in Sicily. The barplots in the inserts show the proportion of genetic diversity described by each Principal Component (PCA eigenvalues). In all plots the first PC

describes 86.81% and 85.61% of the genetic diversity respectively among cat subspecies and populations.

Figure 4. Neighbor-joining tree clustering the pairwise Nei's *et al*. (1983) genetic distances among individual multilocus genotypes of the European wildcats sampled in Italy. The European wildcat sub-populations and the main clades are indicated.

Figure 5. Spatial Principal Component Analysis (sPCA, obtained with ADEGENET) of European wildcats sampled in the central Apennines. (A) Interpolation of the individual genotype scores. The contour lines quantify the degree of genetic differentiation among individuals; circles represent the individual genotypes. (B) Assignment of individual genotypes to their population of origin. Black and white squares represent individual genotype scores on the first principal component (the only significant PC, represented by λ_1 and explaining a significant proportion of spatial structuring). Large white squares indicate individuals with high negative scores; large black squares indicate individuals with highly positive local scores; square dimension is proportional to the degree of differentiations (high for large squares, low for small squares). Letters indicate the sampled regions.

Figure 6. Plot of the average number of alleles per locus (A), F_{ST} values (B) and observed heterozygosity (C), computed simulating two populations of different size (N = 250, 500 and 1000), that were allowed to evolve independently at 33 unlinked autosomal microsatellite loci for 50, 500 and 5000 generations. Horizontal grey blocks identify the average number of alleles per locus (A = 8.77± 0.29), the average heterozygosity (H_O = 0. 57 ± 0.03) observed in European wildcats sampled in the Alps and peninsular Apennines, and the observed F_{ST} (F_{ST} = 0.09) value between them. Dots indicate the average values (+/- their standard errors) obtained from 10 replicate simulations of each parameter settings in EASYPOP.

Supplementary Figure S1. Plot of Log probability [L(K)] as a function of K averaged over five independent runs of STRUCTURE. The Y-error bars are standard deviation and K is the assumed number of genetic clusters. Each plot represents a sample set: (A) sample set *A* (European wildcats, domestic cats and known hybrids); (B) sample set *B* (Sardinian wildcats, domestic cats and known hybrids); (C) sample set *C* (domestic cats, European and Sardinian wildcats); (D) sample set *D* (European wildcats only). (E) Inference of the number of genetic clusters in the study area: posterior distribution of the number of populations estimated using GENELAND.

Supplementary Figure S2. Maps of posterior probability of European wildcats identified by GENELAND. Samples are split into five clusters (K = 5): 1) eastern Alps; 2) Maremma; 3) central peninsular Apennines; 4) southern peninsular Apennines; and 5) Sicily.

Chapter 3

European wildcat population structure

Population genetic structure of wildcats (*Felis silvestris silvestris*) in Europe

F. Mattucci[1*], R. Oliveira[2,3], L. A. Lyons[4], P. C. Alves[2,3,5], E. Randi[1]

[1] Laboratorio di genetica, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Via Cà Fornacetta 9, 40064 Ozzano dell'Emilia (BO), Italy

[2] Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), InBio - Laboratório Associado, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[3] Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal

[4] Comparative Medicine Department of Veterinary Medicine & Surgery College of Veterinary Medicine, University of Missouri – Columbia, 1600 E. Rollins Street, MO 65211, USA

[5] Wildlife Biology Program, Department of Ecosystem and Conservation Sciences, University of Montana, 32 Campus Drive, Missoula, Montana 59812, USA

[6] Aalborg University, Department 18/Section of Environmental Engineering, Sohngårdsholmsvej 57, 9000 Aalborg, Denmark

ABSTRACT

Extant populations of wildcats in Europe are highly fragmented, the likely consequence of local extirpations due to over-hunting, habitat loss and fragmentation. The genetic consequences of fragmentation, isolation and population decline (together with wild *x* domestic cat hybridization) may have relevant impacts of the evolutionary perspectives and conservation of wildcats in Europe. Therefore, sound conservation strategies should relay on the evaluation of genetic diversity within and among populations. This study was planned to assay genetic variation at 38 unlinked microsatellites (37 autosomal and one X-linked) in 1114 random bred domestic cats (n = 295), putative European (n = 670) and African (from Sardinia; n = 26) wildcats, sampled from most of their distribution ranges in 15 European countries. Through descriptive statistics, multivariate and Bayesian analyses we: a) evaluated the extent of genetic diversity within and between wild and domestic cat populations; b) estimated the extent of population structuring and fragmentation; c) identified genetic signatures of past and recent bottlenecks. Results confirmed that wild and domestic cats (plus African wildcats) belong to two well-differentiated genetic clusters (average Φ_{ST} = 0.158, R_{ST} = 0.387, P > 0.001; AMOVA), across most of the sampled locations, with the exception of the highly admixed wild x domestic cat populations in Hungary and Scotland. Genetic

evidences suggested, nevertheless, that cryptic hybrids are present also in other not deply introgressed populations. The African wildcats are related to domestic cats (average Φ_{ST} = 0.046, R_{ST} = 0.121, P > 0.001; AMOVA). The European wildcat populations were strongly structured, beign split into 10 independent clusters (average Φ_{ST} = 0.127, R_{ST} = 0.126, P > 0.001; AMOVA) geographically and genetically grouped into 5 macro populations corresponding to: the Italian peninsular & Sicily; Balkans & north-eastern Italy; Germany eastern; central Europe; and Iberian Peninsula. Population structuring may result from different historical, demographic and ecological detrminants. In particular: the central European population might be the consequence of extra-Mediterranean Würm ice age refuge areas in Europe (nortern Alps, Carpathians, and the Bulgarian mountain systems), while the divergence among and within the southern European populations well explain the Pleistocene biogeographical framework of Europe, with three southern refugia localized in the Balkans, Italian Peninsula and Iberic Peninsula. It is, therefore, clear that Europe does not constitute a unique biogeographical unit for wildcats and at least five major evolutionary significant units (ESU's) and ten minor subpopulations should be recognized in order to support the subspecies survival in the near future.

Keywords: *Felis silvestris*, European wildcat, African wildcat, domestic cat, microsatellites, Bayesian cluster analysis, population structure, fragmentation, conservation genetics

INTRODUCTION

The wildcat *Felis silvestris* is a polytypic species comprising six ecologically, geographically and genetically differentiated subspecies that inhabit the entire Old World (see Driscoll *et al*. 2007 for details). In Europe, three of them coexist: the European wildcat (*Felis silvestris silvestris*, Schreber 1777), from Portugal to Romania; the African wildcat (*Felis silvestris libyca*, Forster 1780), in the Mediterranean islands of Sardinia, Corsica and Crete (Randi and Ragni 1991; Driscoll *et al*. 2007); and the domestic descendant of *libyca* cats, the domestic cat (*Felis silvestris catus*) that has been spread throughout the entire continent. Archaeological remains suggest that the European subspecies probably appeared in the continent around 450,000-200,000 ya (Kitchener 1991; Sommer & Benecke 2006), descending from the Martelli's cat (*Felis lunensis*, Martelli 1906), which was found in Europe during the early Pleistocene (Kitchener 1991; Nowell & Jackson 1996). The presence of African wildcats in Mediterranean islands is a consequence of human translocations at very early stages of domestication, probably less than 11,000 ya by Neolithic navigators (Vigne *et al*. 2012).

Molecular studies and archeological remains suggest that cats' domestication likely began when humans started to build the first civilizations over the fertile crescent (Driscoll *et al*. 2007; Lipinski *et al.* 2008), and the earliest evidence of a cat-human close relationship was found in Cyprus deposits from 10,600 ya (Vigne *et al.* 2012). Succeeding domestication, cats promptly colonized the entire world and became common in Europe, spread through the major land and sea trade routes of Romans, Etruscans and Greeks (Clutton-Brock 1999). The sudden diffusion of free-ranging domestic cats in sympatry with wildcats created conditions for crossbreeding and introgression of domestic alleles into wildcats' genomes, compromising the evolutionary resilience of European wildcats (Beaumont *et al*. 2001; Pierpaoli *et al.* 2003; Lecis *et al.* 2006). Causes and outcomes of hybridization are complex, and the extent of introgression varied from limited in central and southern Europe (e.g. Germany, Italian and Iberian peninsulas) to widespread in Scotland and Hungary, suggesting that generalizations must be avoided (Beaumont *et al*. 2001; Randi *et al.* 2001; Pierpaoli *et al.* 2003; Lecis *et al.* 2006; Oliveira *et al.* 2008a, b; Eckert *et al.* 2010; Hertwig *et al.* 2009; O'Brien *et al.* 2009).

Current patterns of European wildcats distribution and genetic variability are most probably a reflection of both natural and anthropogenic events. In one hand, the range shifts suffered during the climatic oscillations of the Pleistocene contributed to shape wildcat's demographic history and genetic diversity (Kitchener and Rees 2009). However, a comprehensive phylogeography of wildcats in Europe is still missing, and the structure of population subdivision is not well known

(Mattucci *et al*. 2013). On the other hand, strong demographic declines in the 18th and 19th centuries due to deforestation, intensive human persecution (Stahl and Artois 1994), and local decline of major prey (e.g. Lozano *et al.* 2007; Monterroso *et al.* 2009) have been documented for most of the species range (Nowell and Jackson 1996). Such drastic habitat and demographic changes have resulted in high geographic fragmentation and reduced gene -flow, which is exposing small isolated populations to the risks of inbreeding depression, lowered fitness and loss of evolutionary potential and adaptation to environmental changes. Therefore, reconstructing the patterns of population structuring of wildcats in Europe, and estimating their within and between population genetic diversity, is needed to understand the genetic consequences of fragmentation, to identify eventual evolutionary or conservation units and forecast their conservation perspectives.

To delineate sound conservation strategies for European wildcat populations is, therefore, mandatory to improve the molecular tools for the understanding of their global genetic structure and variation of European populations, while reliably identifying admixture events. One of the most far-reaching conclusions that can be drawn from the published studies is that the number and choice of molecular markers is crucial for the resolution and robustness of population structure inferences, and they are still under the optimum for the precise identification of admixed cats (e.g. Oliveira *et al*. 2008; Nussberger *et al*. 2013). In this study, through multivariate analyses and bayesian clustering, we report the first range wide study on the structure of European wildcat populations with more than 1000 cats genotyped at 38 unlinked microsatellites. With such broad-scale analysis we intend to: a) estimate the extent of genetic diversity within and between wild and domestic cat populations; b) describe the main patterns of population structuring; c) infer the extent of their genetic isolation; and d) identify genetic signatures of past and recent bottlenecks.

MATERIAL AND METHODS

Sampling and laboratory procedures

A total of 1114 biological samples (tissue, blood, buccal swabs, hair and skin samples), belonging to 670 putative European wildcats, 27 African wildcats, 295 domestic cats, and 119 previously described admixed cats from Hungary (n = 98) and Scotland (n = 21) were collected from 15 localities of Europe in twelve years (from 1998 to 2010; Table 1). Aiming to help the identification of admixed wildcats in the European and African samples, we added 17 previously described *silvestris x catus* hybrids from Italy (Pierpaoli *et al*. 2003; Mattucci *et al*. 2013). Seven of

these hybrids were obtained from controlled crosses (Ragni 1993); the remaining 10 wild-living hybrids were genetically identified in other studies (Pierpaoli *et al*. 2003; Lecis *et al*. 2006) and reanalyzed here. The European wildcat sample was opportunistically collected from found-dead or trapped animals, covering the majority of species range in Europe, from Portugal to Romania (see Table 1 and Fig 1 for details). All putative wildcats were previously morphologically identified by collectors according to wildcat phenotype, life history, cranial and intestinal indexes, stomach content and/or biometric indices (Schauenberg 1969, 1977; French *et al.* 1988; Ragni and Possenti 1996).

The domestic cat sample included: i) free-ranging cats that had no regular feeding or housing connection to humans, ii) random bred cats having some connection to humans, and iii) cats owned and cared for by humans but not belonging to any specific breed. African wildcats were sampled in Sardinia (Italy), Corsica (France) and North Africa.

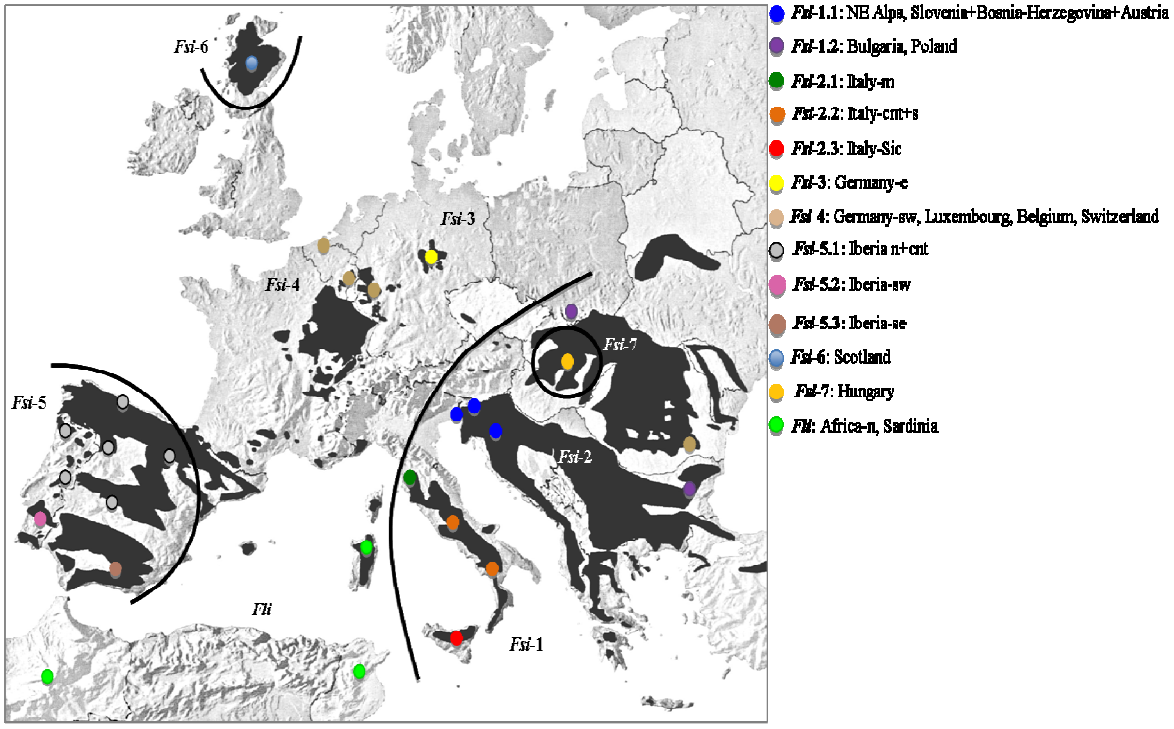


Figure 1. Approximate sampling locations of wildcats surveyed across Europe and North Africa. Colours represent the highest partition of samples into genetic clusters and black lines divide the major differentiated groups (macroareas) in the European wildcat, as identified by multivariate and bayesian analyses. Dark areas in the map correspond to the approximate current distribution of *Felis silvestris* in Europe (adapted from Grabe and Worel 2001).

Table 1. Sampling size and location of all genotyped cats. Hybrid cats that have not been identified through phenotypic evaluation and that have been initially included in the parental populations are reported between brackets ("admixed genotypes"). Acronym = symbol used, in this study, to identify the different populations

Subspecies	**Sampling locations**	**Acronym**	**N**
Domestic cats	Italy	*Fca*	71
Felis silvestris catus	Poland	*Fca*	17
n = 295	Slovenia	*Fca*	7
(admixed genotypes, *n* = 1)	Germany	*Fca*	31
	Netherlands	*Fca*	1
	Switzerland	*Fca*	2
	Portugal	*Fca*	83
	Spain	*Fca*	80(1)
	Sardinia & Corsica	*Fca*	3
African wildcats	Africa: northern	*Fli*	9
Felis silvestris libyca	Sardinia & Corsica	*Fli*	18(1)
n = 27			
(admixed genotypes, *n* = 1)			
Known hybrids (or possible hybrids)	Captivity (Italy)		7
Felis silvestris x catus	Italy: Northern-Eastern Alps		6
n = 136	Italy: Central Apennines and Maremma		3
	Italy: Sicily		1
	Scotland		21
	Hungary		98
Putative European wildcats	Italy: Northern-Eastern Alps	Ne Alps	77(2)
Felis silvestris silvestris	Slovenia + Bosnia & Herzegovina + Austria	Slo + B&H + A	39(2)
n = 670	Bulgaria	Bul	15(2)
(admixed genotypes, *n* = 60)	Poland	Pol	17(1)
	Italy: Maremma	Italy-m	32(8)
	Italy: Central + Southern Apennines	Italy-cnt + s	113(15)
	Italy: Sicily	Italy-Sic	12(2)
	Germany: Eastern	Germany-e	48
	Germany: Southern-Western	Germany-sw	194(17)
	Belgium: Wallonia	Bel	16
	Luxembourg	Lx	11(1)
	Romania	Rom	2(1)
	Switzerland	Swi	3
	Portugal: Northern + Central	Iberia-n + cnt	13(2)
	Spain: Northern + Central	Iberia-n + cnt	29(2)
	Portugal: Southern-Western	Iberia-sw	31(4)
	Spain: Southern-Eastern	Iberia-se	18(1)

Samples were stored in 5 volumes of 95% ethanol (tissues, skins and hairs) or Tris/SDS buffer (blood, buccal swabs; Longmire *et al*. 1997), and kept at -20°C. Total DNA was isolated using standard phenol-chloroform (Sambrook and Russell 2006), high-salt methods (Sambrook *et al.* 1989) or the QIAGEN DNeasy tissue and blood extraction kits (Qiagen Inc, Hilden, Germany), depending on the available quality and quantity of each sample. Two negative controls were included for each extraction run.

Thirty-eight domestic cat-derived dinucleotide microsatellites (37 autosomal and one X-linked: Fca240; Menotti-Raymond *et al*. 2003), were chosen according to their high heterozygosity, high polymorphism information content (PIC) and wide chromosomal distribution (Lipinski *et al*. 2008). All loci were amplified in eight PCR multiplex reactions using the Qiagen Multiplex PCR Kit following the manufacturer's protocol (Supplementary Table S1). All forward primers sequences were modified to include an additional universal tail fluorescently labelled with 6-FAM, NED, PET, or VIC dyes (Applied Biosystems, Foster City, CA, USA). A touch-down thermocycling protocol was performed as following: 95ºc for 15 min; 6 cycles of denaturation at 94ºc for 30s, primer annealing with temperatures between 62-57ºc, decreasing 1ºc every cycle for 60s, and sequence extension at 72ºc for 60s; followed by 25 cycles (35 for hair and museum samples) of the previous cycling protocol but with a permanent annealing temperature of 57ºc; eight additional cycles at 53ºc for labelled tails' incorporation and a final extension step at 60ºc for 30 min. All amplifications were performed in a total volume of 10 µl using an Applied Biosystems thermal cycler (GeneAmp® PCR System 9700) or a Bio-Rad thermal cyclers (Mycycler and Icycler). Hair and museum skin samples were amplified in four replicates, following a multi-tube approach designed for low-quality DNA samples (Taberlet *et al.* 1996) and using dedicated rooms. PCR products were, afterwards, analyzed in an ABI 3130 XL DNA Analyzer (Applied Biosystems Inc.) and allele sizes were determined with GENEMAPPER 4.1 (Applied Biosystems Inc.) by comparison with size standard fragments of GeneScan-500 LIZ (Applied Biosystems Inc.). All genotyping steps included the negative controls for extraction and PCR. Additionally, a reference positive control was always included to infer PCR success and to calibrate independent runs.

The power of markers to identify each unique genotype was evaluated calculating the probability-of-identity values (PID and PIDsibs; Mills *et al*. 2000; Waits *et al*. 2001) in GENALEX 6.41 (Peakall and Smouse 2006). Individual genotypes were matched to exclude replicates. About 10% of randomly selected samples were independently replicated twice to assess the rate of allelic dropout (ADO) and false alleles (FA). Presence of null alleles was assessed with MICROCHECKER (Van Oosterhout *et al*. 2004) with an adjusted *P* value corresponding to $\Delta = 0.05$ after Bonferroni

correction (Rice 1989). Additional laboratory details are available upon request.

Analyses of genetic diversity and differentiation

Microsatellites diversity was estimated separately for domestic (*Fca*), African (*Fli*) and European cat (*Fsi*) subspecies, excluding all admixed genotypes (known and putative hybrids) detected in the hybridization analyses and all cats from the introgressed populations of Scotland and Hungary (see below). We also evaluated genetic diversity within each cluster that split the European wildcats in different subpopulations (see below).

ARLEQUIN 3.5.1.2 (Schneider *et al.* 2000; Excoffier and Lischer 2010) was used to estimate allele frequencies, mean number of alleles per locus (N_A), observed (H_O) and expected heterozygosity (H_E). ARLEQUIN was used to test for deviations from Hardy–Weinberg Equilibrium (HWE), with a Markov Chain length of 10^5 and 3,000 dememorization steps, and pairwise linkage disequilibrium (LD), with 100 initial conditions followed by 16,000 permutations, for all locus-subpopulation combinations, based on the exact test of Guo and Thompson (1992). *P*-values were adjusted for multiple tests using a sequential Bonferonni correction (Rice 1989). Allelic richness (N_{AR}) and private alleles richness (N_{PAR}) were computed for each population following a rarefaction method that compensates for uneven sample sizes, as implemented in the software HP-RARE (Kalinowski 2005). Genetic differentiations among the predefined taxonomic groups and between pairs of subpopulations were estimated with pairwise F_{ST} (Weir and Cockerham's 1984) and R_{ST} (Slatkin 1995) in GENEPOP 4.1 (Rousset 2008) and FSTAT 2.9.3.2 (Goudet *et al.* 2002), respectively. An analysis of molecular variance (AMOVA) on Euclidean pairwise genetic distances was computed using analogues of Wright's *F*-statistics, as implemented in GENALEX (Peakall and Smouse 2006). AMOVA was conducted among and within groups, and the significance of these parameters was estimated by 10,000 permutations of the distance matrix.

Timing and magnitude of possible genetic bottlenecks were characterized using two methodological approaches: i) the first is most effective at detecting very recent bottlenecks of low magnitude and postulate: a) reductions in effective population size cause a temporary excess of heterozygote genotypes relative to the number of alleles in the population ("heterzygote excess test", Cornuet and Luikart 1997), and b) alleles with intermediate frequency will be most abundant because of rare allele loss during a bottleneck event ("mode-shift test", Luikart *et al.* 1998); and ii) the second follows Garza and Williamson's (2001) "*m*-ratio test", which is best suited for detecting more severe, older bottlenecks (up to 100 generations ago), where *m* is the ratio of the number of alleles (*k*) over the range in fragment sizes (*r*), which is predicted to decline in a bottleneck because

the number of alleles should decrease faster than the range in fragment sizes. For the first method, the software BOTTLENECK 1.2.02 (Cornuet and Luikart 1997) was used assuming a "two-phase mutational model" (TPM; Luikart *et al.* 1998) with 90% one-step mutations. The "two-tailed Wilcoxon signed rank test" was used to determine significance of the observed deviations (Cornuet and Luikart 1997). For the second, *m*-ratio was assessed using the software M_P_VAL for each locus and averaging the value of *m* over loci (Garza and Williamson 2001). The significance of *m* was determined by comparing to a critical value (*M*c), calculated from hypothetical populations in mutation-drift equilibrium using the program CRITICAL_M with 10,000 simulation replicates (Garza and Williamson 2001). We used a "two-phase mutation model" with 10% multi-step mutations, an average size of non one-step mutations of 3.5, and theta of 5 and 10 to evaluate the sensitivity of the method to this variable.

Population structure, individuals' assignment and admixture analyses

Bayesian analyses of population structure were implemented in STRUCTURE 2.3.3 (Pritchard *et al*. 2000; Falush *et al*. 2007; Hubisz *et al*. 2009) to simultaneously: i) infer the level of genetic structure in the data (number of *K* clusters) and estimate the average proportion of membership (Q_i) of all sampled populations to each inferred clusters; and ii) assign each multilocus genotype to one or more *K* clusters according to their posterior probability membership assignments (q_i coefficients) and their 90% credibility intervals (CI). Each run of the program was replicated five times, with 10^4 burn-in followed by 10^5 MCMC iterations, using a q_i threshold of 0.80 (based on admixture analyses of observed and simulated cat datasets; Randi 2008, Oliveira *et al.* 2008a) to assign each cat to a single cluster. The optimal number of clusters (*K*) was identified using ΔK and ΔF_{ST} statistics (Evanno *et al*. 2005) as implemented in CORRSIEVE 1.6.1 (Campana *et al*. 2011), see Supplementary Table S2. All computations focusing on hybrids detection and European wildcat populations' structure were performed combining the "admixture" and "*F*" models and both with or without any prior non-genetic information.

We started our computations by preliminarily identifying all genotypes with possible hybrid ancestry. STRUCTURE was run with two different datasets to assign individuals into two possible populations ($K = 2$): putative European wild cats (*Fsi*) versus domestic cats (*Fca*), and putative African wildcats (*Fli*) versus domestic cats. The analyses were replicated within each of the main European wildcat geographic-genetic clusters (resulting from STRUCTURE's clustering, see Results) to overcome a possible bias due to within-subspecies population structuring. Cats' ancestry was

computed using $K = 2$ with prior population information ("*usepopinfo* activated") for the domestic and wildcats that were genetically pre-identified in the first runs of STRUCTURE. All putatively admixed cats ($n = 79$) and the introgressed populations of Scotland and Hungary ($n = 119$) were subsequently excluded from the data set. The same computations were used to examine evidences of admixture within African wildcats.

Then we used a hierarchical population structure approach, that we used STRUCTURE to determine the divergence: 1) among the three subspecies (*Fca*, $n = 294$; *Fsi*, $n = 610$ and *Fli*, $n = 26$); 2) among the main European wildcat populations ($n = 610$); and 3) within each of the main European wildcat populations (varying the number of possible clusters (K) from 1 to 10).. Patterns of genetic differentiation among cat subspecies and European wildcat populations (excluding all hybrids) were also explored by Discriminant Analysis of Principal Components (DAPC) and Principal Component Analysis (PCA) implemented in the ADEGENET package (Jombart 2008).

Estimation of past demographic changes and divergence times in European wildcat populations

We used the $(\delta\mu)^2$ genetic distance (Goldstein *et al.* 1995*b*) and the equation $(\delta\mu)^2 = 2\mu T$ (μ = mutation rate; T = generations; Goldstein and Pollock 1997) to infer divergence times of among the European wildcat populations. Supposing the presence of mutation-drift equilibrium and constant effective population size, we applied two estimates of mutation rate for microsatellite loci in humans (5.6×10^{-4} and 2.05×10^{-3}), which were previously used by Driscoll *et al.* (2002). Current effective population size (N_E) was estimated with the software LD N_E (Waples and Do 2008), based on the extent of linkage disequilibrium generated by genetic drift in declining *populations* (Hill 1981; Waples 2006). We used the random mating model option and jacknife-adjusted 95% confidence intervals. The critical allele frequency value was set following the Waples *et al.* (2009) suggestions for balancing the precision-bias tradeoff of the LD method: for sample size (S) larger than 100 we chose a $P_{crit} = 0.01$ (alleles with frequency < 0.01 were excluded), for $S > 25$ a $P_{crit} = 0.02$, and for $S \leq 25$ a P_{crit} value ranging between $1/(2S)$ and $1/S$. The Nunney and Elam's (1994) equation and an N_e/N value adjusted for domestic cat population (Kaeuffer *et al.* 2004) of 0.38 led us to estimate total wildcat population size (N).

We used also an Approximate Bayesian Computation (ABC) approach (Beaumont 2008) implemented in the POPABC program (Lopes *et al.* 2009) to infer the divergence times and historical values of N_E in the main European wildcat population clusters, as they were defined by concordant geographical distributions and Bayesian clustering Divergence time and effective population size were estimated assuming a simple Isolation-Migration model (Nielsen and Wakeley

2001) in which populations have diverged from a single ancestral population in the past, and using a rejection step (Pritchard *et al.* 1999) followed by a regression algorithm (Beaumont *et al.* 2002).

Summary statistics (including heterozygosity, variance in allele length and number of alleles) were simulated 100,000 times using the stepwise mutation model with the varying mutation rate measured as the number of mutations per generation per locus. The mutation rates for each of the 37 loci were drawn from a normal distribution with mean of 0.0001, standard deviations set to 0 and standard deviation of mean to 0.0005. Prior population parameters were uniform distributions bound between minimum and maximum values (Table 9) without gene flow. A total of 1000 simulations resulting from a tolerance index of 0.01, were used to estimate the parameters. Coalescence simulations and rejection steps were computed with POPABC. The regression step was performed in R using scripts developed by M. Beaumont (http://code.google.com/p/popabc/model_choice.r) and modified to fit our analyses.

RESULTS

Genetic diversity at 37 autosomal unlinked microsatellites

All loci were polymorphic in the total sample of European wild (*Fsi*, n = 610), African wild (*Fli*, n = 26) and domestic (*Fca*, n = 294) cats (excluding the two hybridizing populations and the admixed cats in non-hybridizing populations; see below), showing between 4 (locus Fca035, for *Fli*) and 32 alleles (locus Fca628, for *Fca*) per locus. Across all loci and populations values of observed and expected heterozygosities ranged between H_O = 0.042 (locus Fca305, for *Fsi*) to 0.885 (locus Fca023 for *Fli*), and H_E = 0.064 (locus Fca305, for *Fsi*) to 0.911 (locus Fca628, for *Fli*), respectively. Global genetic diversity showed marked differences between the three subspecies (see Supplementary Table S1 for details on variability per locus). The average number of alleles per locus, estimated for 52 genes (correction accounting for *Fli* low sample size of 26 individual) varied between N_{AR} = 8.07 (*Fsi*), 8.93 (*Fca*) and 9.84 (*Fli*), and N_{PAR} = 1.39 (*Fsi*), 1.19 (*Fca*) and 1.99 (*Fli*), respectively. The mean values of observed and expected heterozygosity were similar between domestic (H_O = 0.664 ±0.103; H_E = 0.778 ±0.098) and African wildcats (H_O = 0.710±0.132; H_E = 0.811±0.100), but lower for European wildcats (H_O = 0.582±0.158; H_E = 0.741±0.180). European wildcats and domestic cats both showed average H_O values significantly lower than expected, with F_{IS} values significantly higher than zero (0.206 and 0.147, respectively; p<0.001). Although few

loci proved to significantly deviate from HW Equilibrium even when each wildcat genetic cluster (Table 2) was analyzed separately, none of them showed significant deviations in all subpopulations. The same pattern was observed for LE estimates. African wildcats were globally in HWE and LE, with only one locus (Fca649) showing significant deviations from equilibrium (Table 2, Supplementary Table S1).

Table 2. Variability at 37 autosomal microsatellites for the three cat subspecies analysed (*Fca, Fsi* and *Fli*) and for the European wildcat subpopulations detected in clustering analyses. All putative hybrids and admixed populations were excluded. N = sample size; N_A = mean number of alleles per locus; N_{AR} and N_{PAR} = allelic and private allelic richness; H_O, H_E = observed and expected heterozygosity; F_{IS} = inbreeding coefficient (* significant departures from HWE at $p <$ 0.001, Bonferroni corrected); HWE and LE = number of tests out of equilibrium at 37 loci and 666 pairwise comparisons, respectively. [A] populations' acronyms are used as in Table 1; [B] codes correspond to the symbols used to discriminate the different genetic clusters (see below) that subdivide European subpopulations; [C] N_{AR} and N_{PAR} were obtained for 52 and 22 genes when comparing subspecies and macro groups, respectively.

Spp or population[a]		Code[b]	N	N_A	N_{AR}[c]	N_{PAR}[c]	H_O	H_E	F_{IS}	HWE	LE
All	*F. s. catus*	*Fca*	294	14.950 (±4.770)	8.93	1.19	0.664 (±0.103)	0.778 (±0.098)	0.147*	23	8
	F. s. silvestris	*Fsi*	610	14.108 (±3.211)	8.07	1.39	0.582 (±0.158)	0.741 (±0.180)	0.206*	37	207
	F. s. libyca	*Fli*	26	9.838 (±2.820)	9.84	1.99	0.710 (±0.100)	0.811 (±0.100)	0.134	1	0
Macro group 1	*F. s. silvestris*	*Fsi*-1	140	9.811 (±2.283)	5.87	0.49	0.621 (±0.171)	0.700 (±0.176)	0.109	3	2
	Ne Alps, Slo+B&H+A	*Fsi*-1.1	111	8.595 (±2.166)	5.27	0.10	0.625 (±0.177)	0.687 (±0.174)	0.086	2	1
	Bul, Pol	*Fsi*-1.2	29	7.167 (±2.396)	5.66	0.17	0.601 (±0.186)	0.728 (±0.149)	0.148	1	0
Macro group 2	*F. s. silvestris*	*Fsi*-2	133	9.622 (±2.265)	5.95	0.46	0.562 (±0.184)	0.697 (±0.176)	0.170*	15	2
	Italy-m	*Fsi*-2.1	24	6.324 (±1.717)	5.33	0.11	0.543 (±0.203)	0.676 (±0.168)	0.182*	1	0
	Italy-cnt+s	*Fsi*-2.2	98	8.784 (±2.225)	5.48	0.15	0.578 (±0.191)	0.691 (±0182)	0.157*	10	0
	Italy-Sic	*Fsi*-2.3	11	3.794 (±1.274)	3.57	0.12	0.500 (±0.229)	0.564 (±184)	0.122	0	0
Macro group 3	Germany-e	*Fsi*-3	49	5.917 (±2.298)	4.21	0.13	0.534 (±0174)	0.637 (±0.173)	0.100	5	3
Macro group 4	Germany-sw, Lx, Bel, Swi, Rom	*Fsi*-4	206	9.919 (±3.022)	5.64	0.13	0.590 (±0.180)	0.707 (±0.186)	0.281*	16	38
Macro group 5	*F. s. silvestris*	*Fsi*-5	82	10.622 (±2.890)	6.20	0.51	0.559 (±0.160)	0.751 (±0.157)	0.213*	18	14
	Iberia-n+cnt	*Fsi*-5.1	38	8.139 (±2.072)	6.21	0.24	0.614 (±0.173)	0.757 (±0.122)	0.178*	7	3
	Iberia-sw	*Fsi*-5.2	17	5.306 (±1.582)	4.61	0.10	0.568 (±0.212)	0.637 (±0.144)	0.096	0	0
	Iberia-se	*Fsi*-5.3	27	6.472 (±2.535)	5.29	0.18	0.549 (±0.182)	0.697 (±0.175)	0.203*	3	0

The hierarchical AMOVA was performed by subdividing the non-hybridizing set of samples into the three taxonomic groups (European wildcats, African wildcats and domestic cats) and by grouping European wildcat samples into subsequent partitions as detailed in Table 3. Although most of the variation was found within groups (between 84% and 92%), results reflect high genetic divergence. Genetic variability was significantly partitioned among taxonomic groups ($\phi_{ST} = 0.158$; $F_{ST} = 0.115$; $R_{ST} = 0.387$) and among all genetically identified locations ($\phi_{ST} = 0.127$; $F_{ST} = 0.078$; $R_{ST} = 0.126$), indicating that wildcats are subdivided into distinct gene pools in Europe. Over all loci and populations, a substantial proportion of genetic variation was attributed to mutation (as measured by R_{ST}) especially when comparing the three cat subspecies ($R_{ST} = 0.387$; Table 3).

Table 3. Hierarchical analysis of molecular variance (AMOVA) computed in ARLEQUIN using ϕ_{ST} and values of F_{ST} and R_{ST} estimated under different perspectives of samples grouping. * all ϕ_{ST} values were highly significant at $p<0.001$

Grouping	Populations	Source of variation	Variance	%Var	Φ_{ST} *	F_{ST}	R_{ST}
By cat subspecies	*Fca*/ *Fsi*/ *Fli*	Among groups	6.887	16%	0.158	0.115	0.387
		Within groups	36.591	84%			
European wildcats by macro groups	*Fsi*-1/ *Fsi*-2 *Fsi*-3/ *Fsi*-4 *Fsi*-5	Among groups	4.314	12%	0.116	0.072	0.124
		Within groups	33.008	88%			
European wildcats by subpopulations	*Fsi*-1.1/ *Fsi*-1.2/ *Fsi*-2.1/ *Fsi*-2.2/ *Fsi*-2.3 *Fsi*-3/ *Fsi*-4 *Fsi*-5.1/ *Fsi*-5.2/ *Fsi*-5.3	Among groups	4.728	13%	0.127	0.078	0.126
		Within groups	32.462	87%			
European wildcats within *Fsi*-1	*Fsi*-1.2/ *Fsi*-1.2	Among groups	4.077	12%	0.119	0.038	0.090
		Within groups	30.151	88%			
European wildcats within *Fsi*-2	*Fsi*-2.1/ *Fsi*-2.2/ *Fsi*-2.3	Among groups	3.280	9%	0.095	0.066	0.078
		Within groups	31.174	91%			
European wildcats within *Fsi*-5	*Fsi*-5.1/ *Fsi*-5.2/ *Fsi*-5.3	Among groups	5.647	15%	0.109	0,076	0,078
		Within groups	32,512	85%			

Pairwise F_{ST} and R_{ST} estimations of genetic differentiation between pairs of populations revealed a substantial partition of the European wildcat population into subpopulations (Table 4), with most of the comparisons resulting in divergence estimates above 0.05. R_{ST} values between wildcat subpopulations were in average 2-3 times the divergence obtained with F_{ST}, reflecting the importance of allele size differences in splitting the wildcat group. Results reflect also the higher

genetic proximity between African wildcats and domestic cats, for which values of F_{ST} (0.046) and R_{ST} (0.107) were globally lower than values between African and European wildcats (F_{ST} = 0.117-0.212; R_{ST} = 0.316-0.552; Table 4).

Table 4. Genetic divergence parameters (Φ_{ST}, below diagonal, and R_{ST}, above diagonal) for pairwise comparison between all the non-hybridizing European wildcat subpopulations and domestic cats sampled in our study. The last two columns report the genetic differentiation between each wildcat subpopulations and all domestic cats.

	Fca	*Fsi*-1.1	*Fsi*-1.2	*Fsi*-2.1	*Fsi*-2.2	*Fsi*-2.3	*Fsi*-3	*Fsi*-4	*Fsi*-5.1	*Fsi*-5.2	*Fsi*-5.3	*Fli*
Fca		0.395	0.354	0.316	0.369	0.310	0.362	0.459	0.311	0.309	0.283	0.129
Fsi-1.1	0.209		0.053	0.097	0.088	0.232	0.150	0.123	0.057	0.141	0.141	0.591
Fsi-1.2	0.202	0.119		0.041	0.060	0.181	0.093	0.085	0.011	0.067	0.124	0.463
Fsi-2.1	0.193	0.155	0.158		0.015	0.104	0.130	0.107	0.026	0.101	0.081	0.419
Fsi-2.2	0.188	0.122	0.144	0.066		0.114	0.130	0.115	0.044	0.111	0.108	0.525
Fsi-2.3	0.262	0.232	0.211	0.173	0.119		0.249	0.171	0.230	0.298	0.218	0.382
Fsi-3	0.225	0.168	0.139	0.217	0.172	0.247		0.162	0.085	0.111	0.119	0.512
Fsi-4	0.192	0.113	0.104	0.165	0.143	0.216	0.137		0.118	0.209	0.177	0.620
Fsi-5.1	0.160	0.101	0.083	0.112	0.092	0.171	0.130	0.101		0.031	0.050	0.476
Fsi-5.2	0.240	0.215	0.237	0.252	0.195	0.326	0.266	0.236	0.150		0.070	0.426
Fsi-5.3	0.205	0.153	0.158	0.175	0.116	0.187	0.169	0.163	0.064	0.149		0.448
Fli	0.071	0.240	0.211	0.217	0.206	0.274	0.270	0.218	0.159	0.236	0.216	

Comparisons between observed and expected heterozygosities in BOTTLENECK provided no evidences of recent genetic bottlenecks among European wildcats, with loci fitting mutation-drift equilibrium under the "TPM model" (Table 5). The *m*-ratio test showed instead that the population from eastern Germany (*Fsi*-3) had an average *m*-value significantly lower than the critical values estimated for $\theta = 5$ (0.764) and $\theta = 10$ (0.746).

Table 5. Analyses of bottleneck signatures for each of the 10 wildcat subpopulations according to M-RATIO (Garza and Williamson, 2001) and BOTTLENECK (Cornuet and Luikart, 1997) probability tests. Following author's suggestion, populations with less than 15 individuals were not analysed for recent bottlenecks (n/a).

Wildcat		M-RATIO			BOTTLENECK
Population	N	M	Critical m ($\theta = 5$)	Critical m ($\theta = 10$)	$P<0.05$
Fsi-1.1	111	0.826	0.780	0.772	1.000
Fsi-1.2	29	0.795	0.748	0.720	0.957
Fsi-2.1	24	0.745	0.742	0.710	0.995
Fsi-2.2	98	0.814	0.779	0.769	1.000
Fsi-2.3	11	0.765	0.707	0.652	n/a
Fsi-3	49	0.711	0.764	0.746	0.928
Fsi-4	206	0.853	0.786	0.785	1.000
Fsi-5.1	38	0.875	0.758	0.734	0.987
Fsi-5.2	17	0.753	0.729	0.687	0.996
Fsi-5.3	27	0.852	0.747	0.718	0.631

Population structure: subspecies and European populations' clustering

Bayesian clustering analyses performed in STRUCTURE to assign individuals into two populations clearly suggested the presence of two well-differentiated genetic clusters that sharply split putative European wildcats and domestic cats. All domestic cats were assigned to cluster I (hereafter referred as *Fca*) with an average proportion of membership Q_{fca} = 0.967, while European wildcats sampled across the entire continent were assigned to cluster II (hereafter referred as *Fsi*) with variable Q_{fsi} values: Portugal = 0.827; Spain = 0.866; Scotland = 0.465; Belgium, Luxembourg and Switzerland = 0.962; Germany = 0.954; Italy = 0.897; Slovenia = 0.965; Hungary = 0.460; and Bulgaria and Poland = 0.925 (Supplementary Table S3 and Supplementary Figure S1).

The genetic composition of cats in Hungary and Scotland, especially when compared to the other European populations, confirmed their admixed structure. Individual assignment values were frequently intermediate between the wild and domestic clusters, with as much as 47.62% (n = 10) and 37.76% (n = 37) of the samples showing qi values between 0.15 and 0.85 in Scotland (n = 21) and Hungary (n = 98), respectively. The plot of individual scores in the first two principal variables of a Principal Component Analysis (Fig 2) also indicated the admixed structure of Hungarian and Scottish cats, with a high proportion of the samples plotting in an intermediate position between the European wild and domestic cat clusters. As a result of their highly admixed composition, prior non-genetic identification of wildcats in Hungary and Scotland is strongly compromised.

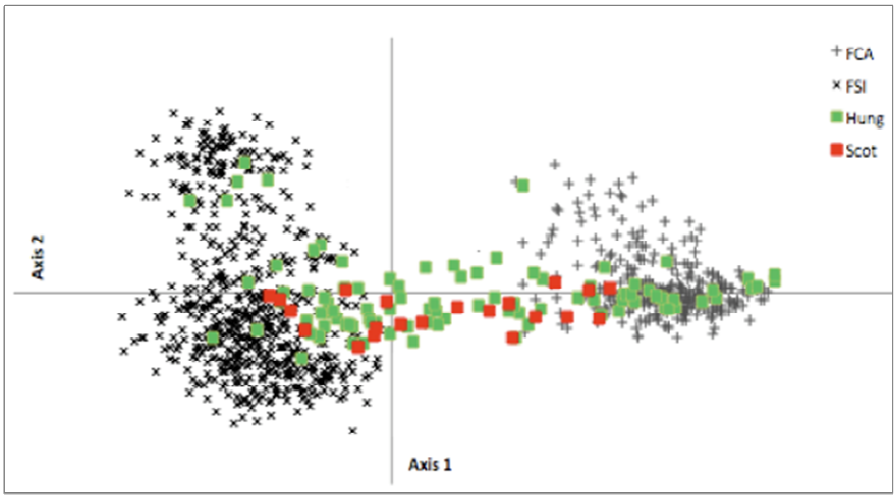


Figure 2. Principal Component Analysis displaying multivariate clustering of Hungarian (Hun) and Scottish (Sco) cats relatively to all sampled European wild (*Fsi*) and domestic cats (*Fca*).

Among non-hybridizing populations, a total of 79 individuals have shown evidences of possible hybrid ancestry. Moreover, all 7 captive bred hybrids were confirmed as admixed cats (see below the detailed analysis of admixture patterns for these individuals). When comparing putative African wildcats (*Fli*) and domestic cats, distinction between the two biological groups was sharp for $K = 2$, with domestic cats assigning with average $q_{fca} = 0.755$ and African wildcats clustering with $q_{fli} = 0.812$. However, no admixture inferences have been made for the *libyca* subspecies in this study, since no straightforward threshold value for individuals' assignment could be accurately inferred (data not shown). The issue of cat hybridization in Europe will be addressed in another paper (Mattucci *et al*. in prep.)

After excluding all putatively admixed cats from the dataset, sub-structuring of non-hybridizing European wildcats' populations was better achieved by the partition of the data into five main geographic-genetic clusters (Table 6 and Supplementary Figure S1): *Fsi*-1 (Balkan area); *Fsi*-2 (Italian peninsula); *Fsi*-3 (Germany eastern); *Fsi*-4 (Belgium, Luxembourg + Romania, Switzerland and Germany southern-western) and *Fsi*-5 (Iberian peninsula). By exploring further subdivision within the main groups (*Fsi*-1, *Fsi*-2 and *Fsi*-5), additional substructuring could be markedly detected by splitting the 3 main populations in 8 differentiated clusters that segregate macroarea *Fsi*-1 into two smaller groups: *Fsi*-1.1 (northern-eastern Alps, Slovenia + Bosnia & Herzegovina + Austria) and *Fsi*-1.2 (Bulgaria and Polonia); the macroarea *Fsi*-2 into three subpopulations: *Fsi*-2.1 (Maremma region), *Fsi*-2.2 (central-southern Apennine) and *Fsi*-2.3 (Sicily); and the macroarea *Fsi*-5 into three subclusters: *Fsi*-5.1 (northern-central Iberia), *Fsi*-5.2 (southern-western Iberia) and *Fsi*-5.3 (southern-eastern Iberia). Population structure inferences are summarized in Table 6, Figure 1 and Supplementary Figure S1.

Table 6. Summary of the Bayesian analysis performed in STRUCTURE for the best *K* value obtained either a) analysing all populations together, either b) analysing each of the wildcat macro groups defined across Europe.

		Inferred clusters									
		1	2	3	4	5	1	2	3	4	5
	Dataset	A) among all wildcat populations					B) within macro groups 1, 2 and 5				
Fsi-1	Italy: Northern-Eastern Alps				0.999					1.000	
	Slovenia + Bosnia&Herzegovina + Austria				0.999					1.000	
	Bulgaria	0.237	0.006		0.687	0.010				0.001	0.999
	Polonia	0.176	0.021		0.715					0.001	0.999
Fsi-2	Italy: Maremma					0.999		0.916	0.084		
	Italy: Central + Southern Apennines	0.010	0.016		0.010	0.963	0.023	0.020	0.936	0.020	
	Italy: Sicily	0.009				0.999	1.000				
Fsi-3	Germany: Eastern			1.000			0.980	0.020			
Fsi-4	Belgium	0.900	0.042		0.040			1.000			
	Luxembourg + Romania	0.932	0.068		0.002			1.000			
	Switzerland	0.894	0.080		0.024			1.000			
	Germany: Southern-Western	0.993			0.005			1.000			
Fsi-5	Iberia: Northern + Central		0.921		0.074			1.000			
	Iberia: Southern-Western		0.970				0.985	0.006	0.009		
	Iberia: Southern-Eastern		1.000					0.004	0.996		

Patterns of genetic variation graphically summarized by DAPC scatter-plots (Fig 3) sharply distinguished *catus*, *silvestris* and *libyca* subspecies, reflected the closer genetic similarity between African wild and domestic cats and exposed a broader partition in the space for the African subspecies (Fig 3a). Moreover, multivariate clustering confirmed the divergence among European wildcat sub-populations (Figure 3b, c and d). The first principal component (PC) describes most of the genetic diversity among cat subspecies (Fig 3a) and the five main geographic-genetic European wildcat groups (Fig 3b). The first two PCs proved to be important in diversifying subpopulations within the Italian and Balkanic sample (Fig 3c) and the Iberian wildcat samples (Fig 3d).

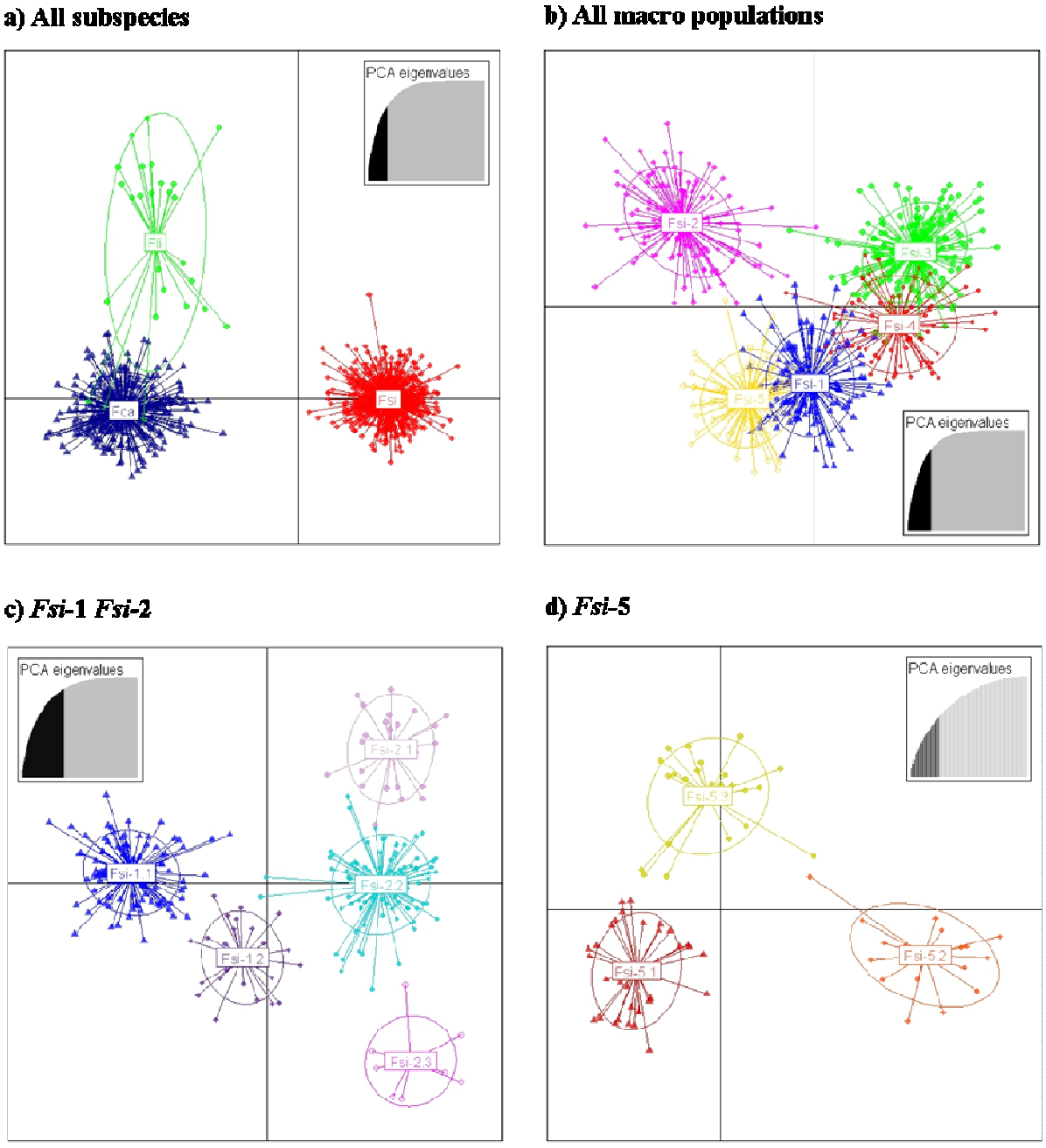


Figure 3. Plot of discriminant analysis of principal components (DAPC) obtained with ADEGENET, displaying genetic diversity among: a) European *felis silvestris* subspecies; b) populations in Southern-Central and Eastern Europe; c) Southern-Western populations (Iberian peninsula) and d) Central European wildcats. Individuals (dots) and populations (coloured ellipses) are positioned on the space maximizing separation between groups.

Past demographic changes in European wildcat populations

The LD-based estimate ranges from N_E = 37 (95%CI = 69-155; *Fsi*-3) to 143 (95%CI = 308-960; *Fsi*-4) in the five main European wildcat clusters, and from N_E = 10 (95%CI = 12-40; *Fsi*-2.3) to 100 (95%CI = 229-399; *Fsi*-1.1) in the subpopulations (Table 7). The harmonic mean of N_E computed on wildcat population sampled in Bulgaria and Polonia (*Fsi*-1.2) revealed some imprecisions (95% CI = infinity) probably due to small sample size and fixation of two loci in this population, reducing the power of the test to detect linkage disequilibrium (Waples 1989).

Table 7. Estimating demographic parameters computed for European wildcat macro groups and subpopulations. Harmonic means of estimated effective sizes (N_E), *S* sample size, *N* estimated using the Using Nunney & Elam's (1994) equation and a correction factor for domestic cat of 0.38 (Kaeuffe *et al.* 2004), 2.5% and 97.5% estimate quartiles (using the jacknife method), and mean numbers of independent allelic comparisons for the selected allele exclusion criteria (P_{crit}). P_{crit} was chosen following the Waples *et al.* (2009) suggestions: for sample size (*S*) larger than 100 P_{crit} = 0.01, for $S > 25$ P_{crit} = 0.02, and for $S \leq 25$ P_{crit} = 0.05.

Grouping	Pop	*S*	Harmonic mean N_E	*N*	95% jacknife CIs lower	upper	Mean Indip comp
European wildcats by macro groups	*Fsi*-1	140	119	314	230	348	31532
	Fsi-2	133	116	306	210	299	33324
	Fsi-3	49	37	97	69	155	11590
	Fsi-4	206	143	375	308	960	33029
	Fsi-5	82	64	167	66	83	29448
European wildcats by subpopulations	*Fsi*-1.1	111	100	264	229	399	25232
	Fsi-1.2	29	na	na	na	na	na
	Fsi-2.1	24	20	53	47	119	11028
	Fsi-2.2	98	88	232	504	5233	22571
	Fsi-2.3	11	10	26	12	40	3522
	Fsi-5.1	38	28	73	93	286	25561
	Fsi-5.2	17	16	41	16	27	7502
	Fsi-5.3	27	22	57	81	328	16623

The estimated posterior distributions for all demographic parameters used to infer pattern of demographic divergence of European wildcat populations are shown in Figure 4 and summarized in Table 8. Effective population size estimate for the central Europe population (Ger-sw, Lx, Bel, Swi, Rom; N_E2) resulted to be four time higher than the Eastern Germany one (Ger-e; N_E1) whose N_E value (~ 3,000) is remarkably similar to their ancestor effective population size (N_EA = 2,997). The divergence time between the two wildcat macro populations of central Europe (T_{EV} = 21,991) differs from the splitting time estimate using the microsatellite genetic distance $(\delta\mu)^2$ that ranged

from 10,441 ya to 7,740 ya (respectively for 5.6×10^{-4} and 2.05×10^{-3} microsatellite mutation rates).

The southern European group disclosed comparable effective population size values ranging from 9,952 (*Fsi*-1) to 11,302 (*Fsi*-5) and both the ancestor effective population sizes with an order of magnitude lowest. Empirical splitting times, ranging from 27,621ya to 7,545 ya (between *Fsi*-1 and *Fsi*-2) and from 29,143 ya to 7,961 ya (between *Fsi*-1 & *Fsi*-2 and *Fsi*-5), fit well with our estimate of divergence time (respectively $T_{EV}1$ = 25,269 $T_{EV}2$ = 34,531; see Table 8).

Within the Italian Peninsular group the effective population sizes detected were ranging from 12,055 of the central-southern Apennines population to 2,378 of the Sicilian one. The ancestor effective population sizes ranging from 1,047 (N_EA2) to 3,181 (N_EA1) resulted to be comparable with the N_E estimate for the Sicilian wildcat population. Divergence times between Maremma and central southern Appenines populations ($T_{EV}1$ = 16,762 ya) and between Apennines populations and Sicilian one ($T_{EV}2$,=,23,693 ya) are comparable with the estimated timing of these events, respectively 10,596-7,851 ya for the first split and 25,281-6,906 ya for the second split.

Within the Iberian Peninsular group both the ancestor effective population sizes are comparable with the N_E estimate for the southern-western Iberian population, while the effective population sizes detected for the southern-eastern and central-northern ones are ranging from 6,673 to 10,497. Divergence times fit weel with the estimated timing of these events, respectively 10,815-6,813 ya for the first split (Ip-sw / Ip-se divergence time) and 25,100-6,856 ya for the second split (Ip-sw & Ip-se / Ip-n+cnt divergence time)

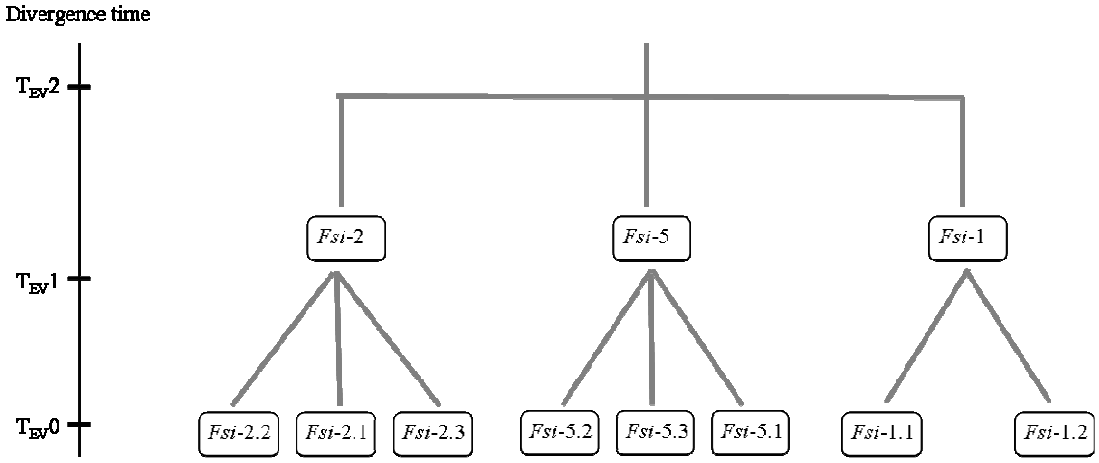


Figure 4.Plot of the assumed Isolation-Migration model of the historical evolution of European wildcats populations in Iberian, Italian and Balkans Peninsula. Divergence time scale is reported on the left.

Table 8. Summary of Prior distribution, mode and quantiles of Posterior distribution of the estimated demographic parameters (Symbol) for the European wildcat data.

Grouping	Populations	Sym	Description	Priors	Posterior distribution		
					Mode	0.0025 quantile	0.0975 quantile
Central European group	Fsi-3/Fsi-4						
		T_{EV}	Divergence time	Uniform (1000, 40000)	21990.76	10675.42	34571.42
		N_E1	Ger-e effective population size	Uniform (100, 10000)	3285.55	1400.164	5239.479
		N_E2	Ger-sw, Lx, Bel, Swi, Rom effective population size	Uniform (100, 20000)	11634.81	9636.317	13478.04
		N_EA	Ancestor effective population size	Uniform (100, 5000)	2996.607	740.7955	4922.23
Southern European group	Fsi-1/Fsi-2/Fsi-5						
		$T_{EV}2$	Fsi-2 / Fsi-1 divergence time	Uniform (1000, 40000)	25269.35	15207.74	35197.03
		$T_{EV}2$*	Fsi-2 & Fsi-1 / Fsi-5 divergence time	Uniform (1000, 40000)	34531.08	19708.13	52711.28
		N_E1	Fsi-2 effective population size	Uniform (5000, 10000)	10777.09	9126.842	12375.24
		N_E2	Fsi-1 effective population size	Uniform (5000, 10000)	9951.804	8279.544	11669.46
		N_E3	Fsi-5 effective population size	Uniform (5000, 10000)	11302.42	9529.421	12894.97
		N_EA1	Fsi-2 / Fsi-1 ancestor effective population size	Uniform (700, 1200)	1040.655	725.1477	1197.822
		N_EA2	Fsi-2 & Fsi-1 / Fsi-5 ancestor effective pop size	Uniform (3000, 7000)	2960.893	1153.137	5158.066
Italian Peninsular macro group	Fsi-2.1/Fsi-2.2/Fsi-2.3						
		$T_{EV}1$	It-cnt+s / It-m divergence time	Uniform (600, 40000)	16761.81	5097.528	29319.9
		$T_{EV}1$*	It-cnt+s & It-m / It-sic divergence time	Uniform (600, 40000)	23692.62	5329.621	43392.12
		N_E1	It-cnt+s effective population size	Uniform (100, 10000)	12054.76	9984.142	14170.59
		N_E2	It-m effective population size	Uniform (100, 10000)	7496.346	5351.176	9714.136
		N_E3	It-sic effective population size	Uniform (100, 10000)	2377.892	933.3744	4105.908
		N_EA1	It-cnt+s / It-m ancestor effective population size	Uniform (200, 5000)	3181.082	811.3184	5234.487
		N_EA2	It-cnt+s & It-m / It-sic ancestor effective pop size	Uniform (200, 5000)	1046.953	0	3055.264
Iberian Peninsular macro group	Fsi-5.1/Fsi-5.2/Fsi-5.3						
		$T_{EV}1$	Ip-sw / Ip-se divergence time	Uniform (600, 40000)	9446.901	0	21592.22
		$T_{EV}1$*	Ip-sw & Ip-se / Ip-n+cnt divergence time	Uniform (600, 40000)	15437.59	0	32525.88
		N_E1	Ip-sw effective population size	Uniform (100, 10000)	3976.574	1770.789	6330.391
		N_E2	Ip-se effective population size	Uniform (100, 10000)	6673.195	4319.57	9001.568
		N_E3	Ip-n+cnt effective population size	Uniform (100, 10000)	10496.69	8469.849	12519.43
		N_EA1	Ip-sw / Ip-se ancestor effective population size	Uniform (200, 5000)	4066.92	1778.024	5998.142
		N_EA2	Ip-sw & Ip-se / Ip-n+cnt ancestor effective pop size	Uniform (200, 5000)	3050.459	842.2992	5304.112

DISCUSSION

In the last decade, a number of molecular studies have focused in the study of genetic diversity and hybridization patterns among European wildcats. Most recent research has taken advantage of the revolutionary advances in molecular and statistical technics faced by conservation genetics, and some populations across Europe have now been investigated. Examples of detailed analyses can be found for Iberian peninsula (Oliveira *et al*. 2008a,b), France (O'Brien *et al*. 2009), Italy (Randi *et al*. 2001), Germany (Hertwig *et al*. 2009; Eckert *et al*. 2010), Hungary (Lecis *et al*. 2006) and Scotland (Beaumont *et al.* 2001), where the analysis of mitochondrial variation and/or microsatellites diversity suggested varying degrees of hybridization between wild and domestic cats. Until now, only Pierpaoli and colleagues have tried to draw the complete picture for the species across its European range by analyzing a total of 12 polymorphic microsatellites in 336 cats from 9 different countries (Pierpaoli *et al.* 2003). However, the low number of putative wildcat samples in some of the analyzed regions (eg. Iberian peninsula $n = 13$, Hungary $n = 17$) may have limited the deep analysis of population structure and demographic variation. Furthermore, the number of loci used to analyze admixture patterns ($n = 12$) remained far from the advised numbers according to convincing simulations by Vähä and Primmer (2006), since as much as 48 loci with average $F_{ST} = 0.21$ might be needed to distinguish F_1, F_2, backcrosses and parental individuals. While not exactly achieving these numbers, we used in this study 37 autosomal microsatellite loci (with average F_{ST} varying between 0.12 and 0.20 in European subpopulations) and one locus in the X chromosome (Fca240) to: a) describe genetic variation at 1114 cats and b) evaluate population structure within European wildcats.

Genetic diversity and genetic consequences of populations' declines

Patterns of genetic diversity among European wild, African wild and domestic cats have shown that the three subspecies represent highly variable taxa, with the *libyca* cats displaying the higher values of allelic richness, private alleles and levels of heterozygosity (Table 2). Subsequent higher diversity was found for domestic cats, with the exception of private alleles richness that was higher for European wildcats. The high genetic variability found for African wildcats might be a results of one or both of the following intrinsic characteristics of the evolutionary history of the subspecies: a) in one hand *libyca* cats are known to occur in an extremely wide distributional range (Driscoll *et al.* 2007) and to display a very broad habitat tolerance (Driscoll and Nowell 2010),

which might have protected the species against strong past population declines and promote gene flow among populations; and b) past episodes of crossbreeding between domestic cats and their wild ancestors might have occurred in multiple occasions and for a long time during the process of domestication (Driscoll *et al.* 2007) and might have continuously support the maintenance of high genetic diversity. These same reasons might also in part explain the levels of variability found within European domestic cats, for which patterns of high diversity may, in fact, reflect a diversity of evolutionary origins and possible admixture. Results from AMOVA and pairwise-F_{ST} and R_{ST} estimates further confirm the close genetic proximity of *libyca* and *catus* subspecies, for which a clearly significant divergence was hard to obtain based on our set of loci. Instead, significant differentiation was detected between domestic cats and ten divergent subpopulations of European wildcats, showing that, despite of long co-existence, two entities are well differentiated in Europe and introgressive hybridization has not been the major factor shaping both wild and domestic cats' gene pools (but see Hungarian and Scottish populations). Since our sampling is composed of substantially different numbers of individuals per subpopulation (e.g. 11 samples in Sicily (*Fsi*-2.3) against 206 wildcats in Belgium, Luxembourg + Romania, Switzerland and Germany Southern-Western (*Fsi*-4)), and deviations from HWE and LE for two of the detected subpopulations (North and Centre of Iberian peninsula (*Fsi*-5.1) might indicate that further substructure could be present at least in these regions, any comparison of subpopulations' "purity" could be misleading. Globally, our results confirm previous findings reported by Pierpaoli *et al*. 2003 and by studies within some European countries (Portugal and Spain, Oliveira *et al.* 2008a,b; France, O'Brien *et al*. 2009; Italy, Randi *et al.* 2001; and Germany, Hertwig *et al.* 2009 and Eckert *et al*. 2010), increasing today's confidence in the existence of true non highly hybridizing populations in Europe.

Since the amount of genetic variation is high within populations, but the populations are small and isolated, demographic declines and hybridization seem to pose a higher threat to the populations' persistence than strong inbreeding depression and low genetic variation. Carnivores such as the European wildcat are, in fact, considered to be particularly sensitive to population decline and local extinction (Gittleman *et al.* 2001), especially due to their low population densities, high generation times, sensibility to ecosystems modifications, and current exposition to important survival threats resulting from anthropogenic changes (Schipper *et al*. 2008). Among all the studied subpopulations, genetic signatures of population bottleneck were identified only for the Eastern population of Germany (*Fsi*-3), confirming predictions by Pierpaoli *et al*. 2003 and Eckert *et al.* 2010 and contrasting the absence of such event advocated by Hertwig *et al.* 2009. Although no significant signs of recent heterozygosity excess could be detected in our study (as assessed by the procedure of Cornuet and Luikart (1997)), the 49 samples genotyped from this region displayed

generally lower values of genetic diversity than other subpopulations (lower number of alleles, allelic richness, private alleles richness and heterozygosities; Table 2), and the ratio between the number of alleles and the range in allele size (*m*-ratio) suggested a significant reduction in effective population size. These results suggest that this population might have suffered past demographic declines that left detectable footprints on its genetic diversity. Although large population declines are known to have occurred across the entire species' range (Driscoll and Nowell 2010), no evidences could be found for any of the other sampled subpopulations.

LD-based estimates suggest that effective population sizes of both European wildcat macro groups and subpopulations appeared to be biased by the number of included animals per population (see Results), so that they should be regarded with caution. All N_E estimates, in fact, show a strong correlation with sample size estimates. For example, the N_E value of 116 individuals for *Fsi*-2 (n = 133) decreased to only 56.1 (95%CI = 105-152; data not shown) when sample was reduced twofold (n = 83), similarly to the others populations. The program LDNE have been shown to be reliable with use of 10–20 microsatellite loci and samples of at least 25–50 individuals, if the effective population size is smaller (i.e., less than approximately 500 individuals; Waples and Do 2010) than the sample size used to estimate it (England et *al.* 2006; Waples 2006). However, it should be noted that for a precise early detection of population decline adding more individuals appears more beneficial than adding more loci, especially for the LD method (Antao *et al.* 2010). For example, in case of overlapping generations N_E is more likely to be an estimate of the effective number of breeders producing the sample, rather than the effective size for a generation (Waples 2006). Thus, the minimum of 50 effective breeders that has been suggested as needed to prevent inbreeding depression in the short term (Franklin 1980) was only reached by the *Fsi*-1, *Fsi*-2, *Fsi*-4 and *Fsi*-5 macro groups and by the *Fsi*-1.1 and *Fsi*-2.2 subpopulations, probably as a result of their wide spatial distribution. Since Linkage disequilibrium models assume a closed and unstructured population (Waples 2006), sampling from two-three sub-populations violates this assumption and may increase disequilibrium above that caused by factors related to effective population size, such as drift. Thus, the N_E estimate for the *Fsi*-1, *Fsi*-2, *Fsi*-4 and *Fsi*-5 macro groups may be underestimated by admixture among sub-populations whitin them (Waples and England 2011). All the other populations that exhibited extremely small effective size (included the eastern Germany population *Fsi*-3) resulted to be very poorly sampled ($S < 50$), making impossible to state about their conservation status. The target effective population sizes of 500-700 recommended for securing long-term viability (Franklin 1980; Lande 1995; Reed *et al*. 2003) is clearly several times larger than those observed.

The loss of population fitness associated with a loss of genetic diversity is expected and well

known from many studies of rare species (Frankham and Ballou 2003). Furthermore, a loss of genetic diversity may not be a problem over the short term but may reduce the ability of the population to evolve following future changes in environmental conditions or directly in the target species, as may occur under the scenario of introgressive hybridization. Therefore, the wildcat populations from eastern Germany should be regarded as a conservation priority.

Range-wide population structure of European wildcats

Habitat fragmentation may disrupt original patterns of gene flow and lead to drift-induced differentiation among local population units. Top predators such as the wildcat may be particularly susceptible to this effect, given their low population densities, leading to small effective sizes in local fragments. Wildcats have a high dispersal rate and often disperse over long distances, suggesting that they may counteract this process and that there was probably little differentiation within European populations in the past. It is, however, evident that a significant decrease in gene flow and a recent increase in population fragmentation occurred. Our results indicate a clearly detectable genetic diversification among wildcat populations, with the entire sampling of non-hybridizing wildcats being partitioned at least in 10 well-defined genetic clusters. The first partition of the data in five macro groups (Table 6, Figure 1 and Supplementary Figure S1) reflects the Bayesian clustering reported by Pierpaoli *et al.* (2003), that subdivided the European population in southern and central Europe and separated the Eastern German population from all the other wildcat populations. However, our analysis clearly defined higher fragmentation of the European sampling, with the sharp separation of subpopulations within those clusters. Southern Europe could now be well differentiated in two main macro groups (Iberian peninsula: *Fsi*-5) and Italy (*Fsi*-2, Mattucci *et al.* 2013), Slovenia and Bosnia&Herzegovina (*Fsi*-1) with their respective further subdivisions (see Table 6 and Figures 1 and 2 for details). These results clearly demonstrate that larger sampling and better molecular definition was able to reveal additional information on populations' structure. It is, therefore, clear that Europe does not constitute a unique biogeographical unit for wildcats and at least five major evolutionary significant units (ESU's) and ten minor subpopulations should be recognized. It is however clear that some areas of the species distribution have maintained certain levels of reproductive contact, especially in Eastern Europe (Fig 1). Studying which mechanisms control gene flow among European wildcat populations would be particularly important for a better understanding of the results found in this study and, in general, for a better knowledge of the processes shaping wildcat's evolutionary history. While one may predict that a considerable

proportion of today's fragmentation might result from habitat degradation and direct persecution (among others), many other geographical, historical and ecological factors may contribute to explain genetic differentiation among local wildcat populations. For example, ecological factors (climate, habitat types and diet composition) proved to strongly influence the amount of gene flow among European grey wolf populations, rather than topographic barriers or historical populations' fragmentation (Pilot *et al*. 2006). This example of a carnivore species that is also widely mobile, wide distributed and suffers from similar conservation threats (e.g. Hybridization with domestic relatives) highlights the importance of further studies aimed at understanding the direct mechanism that links population ecology and population genetic structure in wildcats (Pilot *et al.* 2006).

For a better and more comprehensive evaluation of the reported genetic partition we further recommend the European-wide evaluation of phylogeographic patterns based on mtDNA diversity and nuclear sequencing in the near future. Although using mtDNA for studying hybridizing taxa might be prone to errors due to its uniparental inheritance, the fact that wildcats from the entire Europe are today genotyped for 38 microsatellites provides great confidence to select the "purest" wildcat samples for subsequent mtdna analyses.

The pattern of demographic parameters estimated for the southern European populations (*Fsi*-1, *Fsi*-2 and *Fsi*-5) is concordant with the Pleistocene biogeographical framework of Europe. The main population subdivision: Balkans, Italian Peninsula and Iberic Peninsula, fits well within a scenario of LGM isolation of European wildcat populations in Mediterranean refuges in southern Europe (Table 9). The subdivision in the central Europe between the Germany eastern population and the Belgium, Luxembourg + Romania, Switzerland + Germany southern-western populations, minght be the consequence of extra-Mediterranean Würm ice age refuge areas in Europe. In particular, most of these refugia are assumed to be geographically small and situated in climatically buffered pockets in the landscape of the nortern Alps, Carpathians, and the Bulgarian mountain systems (Stewart *et al.* 2001; Schmitt and Varga 2012). The populations of these scattered extra-Mediterranean refugial pockets could expand and hybridise among each other during the milder interstadial phases of the Würm and also between the LGM and the younger Dryas period (Stewart *et al.* 2001). Thus, these ice age rear edges became the leading edges of the postglacial northwards range expansions, strongly impacting the genetic constitution of Central and North Europe in many plant and animal species.

The particular importance of a Carpathian refugium as extra-Mediterranean retreat has repeatedly been suggested in several groups of vertebrates, e.g. in the moor frog (*Rana arvalis*), the agile lizard (*Lacerta agilis*), the bank vole (*Clethrionomys glareolus*), the common and the field vole (*Microtus arvalis, M. agrestis*), the wild boar (*Sus scrofa*), the roe deer (*Capreolus capreolus*)

and the red deer (*Cervus elaphus*), and even in the brown bear (*Ursus arctos*), e.g. (Kotlík *et al.* 2006; Sommer *et al.* 2006; Jaarola *et al.* 2002; Surget-Groba *et al.* 2006).

The observed genetic diversity within the Iberian and Italian macro groups might have been jointly generated by and long divergence times (ranging from 25,269 to 34,531 ya, see Table 8). Thus, both observed and simulated genetic parameters suggest that extant genetic divergence among European wildcat populations of Iberia and Italy cannot be explained by recent fragmentation, but by extended periods of isolation without gene flow (in the order of 5,000-10,000 generations).

Future perspectives

Wildcats are fully protected across most of its range in Europe, are listed on the eu habitats and species directive (ANNEX IV), protected by the bern convention (APPENDIX II), included on cites APPENDIX II and classified as threatened at the National Level in many European states. However, there have been no recent large-scale surveys of the species that provide a global picture of its European status (Driscoll and Nowell 2010). Because population fragmentation and introgressive hybridization have been pointed out as major threats to the survival of wildcat's natural populations, we have focused our research in European-wide analyses of both menaces and provided new insights into the knowledge of species. Nevertheless, many questions persist and arise especially for what regards the mechanisms and dynamics of introgressive hybridization. For example, a detailed analysis of the geographical location of hybrids relatively to a precise map of distribution for each wildcat subpopulation would help understanding if crossbreeding between wild and domestic cats has been restricted to peripheral areas of wildcat range as previously advocated by Oliveira *et al.* (2008a,b) in Iberia and Randi *et al.* (2001) in Italy, or if admixture events might also take place in the core of the species distribution. Furthermore, identifying the direction of hybridization through the additional genotyping of uniparentally inherited loci would be essential for the proper understanding of hybridization dynamics in natural populations. For example, while observations of wildcats' spatial activities suggested that most of admixture events are probably occurring between male wildcats and domestic females (Birò *et al.* 2004), observations in Ardennes Mountains in France indicate that hybridization might not involve domestic females, but probably the low densities of wild males promote the crossbreeding between wild females and domestic males (Germain *et al.* 2008).

Among the dataset used in this study, unbalanced efforts and sampling strategies have been

applied for each subpopulation, with geographical areas being represented by samples collected during long-lasting ecological studies of the species (e.g. Hungary or Italy) and others for which cat samples were obtained only opportunistically (e.g. North and Centre of Iberia). The unbalanced number of domestic and wildcat sampled in *Fsi*-1 or *Fsi*-2 and *Fsi*-5 is most representative of this heterogeneity. Moreover, in areas other than Hungary and Scotland, it is possible that samples displaying domestic phenotypes would not be intentionally collected, but they could still be backcrosses with domestic cats and not pure domestic sympatric individuals (a problem that as also been underlined for the study of wolf x dog hybridization by Godinho *et al.* 2011). We, therefore, recommend future range-wide works to be based in an European agreement for homogenizing sampling strategies, with the collection of all putatively wild, domestic and hybrid phenotypic samples.

Finally, wildcat hybridization studies should soon take advantage of the increasing knowledge of the domestic cat genome and of the resulting availability of new informative markers other than microsatellites. For example, the analysis of multiple SNPs (Single Nucleotide Polymorphisms) and the selection of the most informative ones for the differentiation between wild and domestic cats might soon help to improve our understanding of hybridization at a genome-wide level. In this context, the analysis of genetic variants that are responsible for the vast morphological, physiological and behavioural diversity occurring among domesticates might be particularly interesting. For example, since basal morphology in European wildcats is unchanging, specific mutations determining variable patterns in domestic cats are very rare or absent in natural populations, and alternative variants or random polymorphism are expected. At the same time, explicit genetic variants might benefit the way of living in nature and may have remained fixed in the wild populations (e.g. Camouflage patterns crucial for hiding and hunting behaviour), while variable genetic expressions might be maintained in domestic individuals due to relaxation of selective pressures. Analysing levels of genetic diversity in mutations known to have been under different types of selection during domestication and/or breeds' improvement may, thus, revolutionize wildcat hybridization studies.

ACKNOWLEDGEMENTS

We thank Herrman M, Suchentrunk F, Liberek M, Natural Museum of Scotland, Romy P, Szolt B , Sforzi A, Ragni B, Lapini L, De Faveri A, Hupe K, Eckert I, Potocnik H, Moes M, Cagnolati G, Vercillo F, Godoy J, Malsaña M, Mejias M, Fernández JM, Robles JL, Penafiel GD, Duperón EB, Moleón M, Monterroso P, Álvares F, Rodrigues J, BTVS/ICNB: Portuguese National Tissue

Bank/National Institute for Nature and Biodiversity Conservation, Lyberakis P and their collaborators for providing putative wildcat samples. We are also grateful to all anonymous veterinarians and biologists that assisted in samples collection. Rita Oliveira was supported by fundação para a ciência e a tecnologia (fct) through the phd grant sfrh/bd/24361/2005 and the research project ptdc/cvt/71683/2006.

3.1 REFERENCES

Antao T, Pérez-Figueroa A, Luikart G (2010) Early detection of population declines: high power of genetic monitoring using effective population size estimators. *Evolutionary Applications*, **4**, 144-154

Beaumont MA (2008) Joint determination of topology, divergence time, and immigration in population trees. Simulation, Genetics and Human Prehistory, McDonald Institute Monographs (Eds Matsumura, Forster SP and Renfrew C), Pp. 134-1541. McDonald Institute Monographs, UK

Beaumont M, Barratt EM, Gottelli D, Kitchener AC, Daniels MJ, Pritchard JK, Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319-336

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025-2035

Biró Z, Szemethy L, Heltai M (2004) Home range sizes of wildcats (*Felis silvestris*) and feral domestic cats (*Felis silvestris catus*) in a hilly region of Hungary. *Mammalian Biology*, **69**, 302-310

Cornuet JM and Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014

Campana MG, Hunt HV, Jones H, White J (2011) Corrsieve: software for summarizing and evaluating structure output. *Molecular Ecology Resources*, **11**, 349-352

Clutton-brock J (1999) A natural history of domesticated mammals, Cambridge University Press

Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, and O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome Research*, **12**, 414-423

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ, MacDonald DW (2007) The near eastern origin of cat domestication. *Science*, **317**, 519-523

Driscoll C and Nowell K (2010) *Felis silvestris*. In: IUCN 2012. IUCN red list of threatened species. Version 2012.2

Eckert I, Suchentrunk F, Markov G, Hartl GB (2010) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mammalian Biology - zeitschrift fur saugetierkunde*, **75(2)**, 160-174

England PR, Cornuet JM, Berthier P, Tallmon DA, Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias using small samples. *Conservation Genetics*, **7**, 303-308

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 611-2620

Excoffier L and Lischer HEL (2010) Arlequin suite Ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular ecology notes*, **7**, 574-578

Frankham R and Ballou JD (2003) Introduction to conservation genetics. University of Cambridge, Cambridge, UK

Franklin IR (1980) Evolutionary change in small populations. In *Conservation biology: an evolutionary-ecological perspective:* 135-150. Soulé ME & Wilcox BA (Eds). Sunderland, MA: Sinauer

French D, Corbett LK, Easterbee N (1988) Morphological discriminants of Scottish wildcats (*Felis silvestris*), domestic cats (*F. catus*) and their hybrids. *Journal of Zoology, London*, **161**, 75-123

Garza JC and Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305-318

Germain E, Benhamou S, Poulle ML (2008) Spatio-temporal sharing between the European wildcat, the domestic cat and their hybrids. *Journal of Zoology*, **276**, 195-203

Gittleman JL, Funk SM, MacDonald DW, Wayne RK (2001) Carnivore conservation (Conservation Biology, vol. 5). Lavoisier, Cambridge

Godinho R, Llaneza L, Blanco L, Lopes S, Alvares F, García E, Palacios V, Cortés Y, Talegón J, Ferrand N (2011) Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian peninsula. *Molecular Ecology*, **20**, 5154-5166

Goldstein DB and Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, **88**, 335-342

Goldstein DB, Unares AR, Cavalli-Sforza LL, Feldman MW (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences USA*, **92**, 6723-6727

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology*, **11**, 1103-1114

Grabe H and Worel G (2001) Die Wildkatze. Zurück auf leisen Pfoten, Pp. 14-27. Buch and Kunstverlag Oberpfalz, Amberg

Guo SW and Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361-372

Hertwig ST, Jungnickel A, Stepanow S, Böhle UR, Fischer MS (2009) Regionally high rates of hybridization and introgression in German wildcat populations (Felis silvestris, Carnivora, Felidae). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 283-297

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, **38**, 209-216

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9(5)**, 1322-1332

Jaarola M, Searle JB (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11(12)**, 2613-2621

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24(11)**, 1403-1405

Kaeuffer, R., Pontier, D., DeVillard, S., and Perrin, N. (2004). Effective size of two feral domestic cat populations (Felis catus L.): effect of the mating system. *Molecular Ecology*, **13**, 483-490

Kitchener AC (1991) The natural history of the wild cats. Cornell University Press, New Yrk. Pp.280

Kitchener AC and Rees EE (2009) Modelling the dynamic biogeography of the wildcat: implications for taxonomy and conservation. *Journal of Zoology*, **279**, 144-155

Kotlík P, Deffontaine V, Mascheretti S, Zima J, Michaux JR, Searle JB (2006) A northern glacial refugium for bank voles (*Clethrionomys glareolus*). *Proceedings of the national academy of sciences of the USA*, **103**, 14 860-14 864

Hertwig ST, Schweizer M, Stepanow S, Jungnickel A, Böhle UR, Fischer MS (2009) Regionally high rates of hybridization and introgression in German wildcat populations (*Felis silvestris*, Carnivora, Felidae). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 283-297

Kalinowski ST (2005) HP-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187-189

Lande R (1995) Mutation and conservation. *Conservation Biology*, **9**, 782-791

Lecis R, Pierpaoli M, Birò ZS, Szemethy L, Ragni B, Vercillo F, Randi E (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119-131

Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, et al. (2008) The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics*, **91**, 12-21

Longmire JL, Maltbie M, Baker RJ (1997) Use of 'lysis buffer' in DNA isolation and its implication for museum collections. In: *Occasional Papers of Museum of Texas Tech* University (163), pp. 1-3

Lopes J, Balding D, Beaumont MA (2009). PopABC: a program to infer historical demographic parameters. *Bioinformatics*, **25**, 2747-2749

Lozano J, Virgós E, Cabezas–Díaz S, Mangas JG (2007) Increase of large game species in Mediterranean areas: is the European wildcat (*Felis silvestris*) facing a new threat? *Biological Conservation*, **138**, 321-329

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89(3)**, 238-247

Mattucci F, Oliveira R, Bizzarri L, Vercillo F, Anile S, Ragni B, Lapini L, Sforzi A, Alves PC, Lyons LA, Randi E (2013) Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecology and Evolution,* **3(8)**, 2443-2458

Menotti-Raymond M, et al. (2003) Second-generation integrated genetic linkage/radiation hybrid maps of the domestic cat (*Felis catus*). *Journal of Heredity*, **94**, 95-106

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA (2000) Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. *Ecological Applications*, **10**, 283-294

Monterroso P, Brito JC, Ferreras P, Alves PC (2009) Spatial ecology of the European wildcat in a Mediterranean ecosystem: dealing with small radio–tracking datasets in species conservation. *Journal of Zoology*, **279**, 27-35

Nielsen R and Wakeley J (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885-896

Nowell K and Jackson P (1996) European wildcat, *Felis silvestris*, silvestris group Schreber, 1775. In wild cats: status survey and conservation action plan: 110-113. Nowell K and Jackson P (Eds). Gland, Switzerland: IUCN

Nunney L and Elam DR (1994) Estimating the effective size of conserved populations. *Conservation Biology*, **8**, 175-184

Nussberger B, Greminger MP, Grossen P, Keller LF, Wandeler P (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources,* **13**, 447-460

O'Brien SJ, Devillard S, Say L, Vanthomme H, Léger F, Ruette 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(9)**, 2351-2360

Oliveira R, Godinho R, Randi E, Ferrand N, Alves PC (2008a) Molecular analysis of hybridisation between wild and domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**, 1-11

Oliveira R, Godinho R, Randi E, Alves PC (2008b) Hybridization vs conservation: are domestic cats threatening the genetic integrity of European wildcat (*Felis silvestris silvestris*) populations in Iberian peninsula? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **363**, 2953-2961

Peakall R and Smouse PE (2006) Genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295

Pierpaoli M, Birò ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585-2598

Pilot M, Jedrzejeski W, Branicki W, Sidorovich VE, Jedrejewska B, Stachura K, Kunk SM (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, **15**, 4533-4553

Pritchard CE, Fornerod M, Kasper LH, van Deursen JM (1999) RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains. *The Journal of Cell Biology*, **145**, 237-254

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology,* **17**, 285-293

Randi E and Ragni B (1991) Genetic variability and biochemical systematics of domestic and wildcat populations (*Felis silvestris*, Felidae). *Journal of Mammology*, **72**, 79-88

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

Ragni B (1993) Status and conservation of the wildcat in Italy. In: *Seminar on the biology and conservation of the wildcat (Felis silvestris*) (ed. Council of Europe Enviromental Encounters), pp. 40-41, Council of Europe Press, Strasbourg

Ragni B and Possenti M (1996) Variability of coat-colour and markings system in *Felis silvestris*. *Italian Journal of Zoology*, **63**, 285-292

Reed DH, O'Grady JJ, Brook BW, Ballou JD, Frankham R (2003) Estimates of minimum viable population sizes for vertebrates and factors influencing those estimates. *Biological Conservation,* **113**, 23-34

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**,103-106

Sambrook E, Fritsch F, Maniatis T (1989) Molecular cloning. Cold Spring Harbour Press, Cold Spring Harbour, New York

Sambrook J and Russell DW (2006) Purification of nucleic acids by extraction with Phenol-Chloroform. Cold Spring Harbor protocols. Doi:10.1101/pdb.prot4455

Schauenberg P (1969) L'identification du chat forestier d'Europe, *Felis s. silvestris* Schreber 1777, par une méthode ostéométrique. *Revue Suisse de Zoologie*, **76**, 433-441

Schauenberg P (1977) La stature du chat forestiere *Felis silvestris* Schreber, 1777, et la variabilité morphologique de l'espèce. *Revue Suisse de Zologie*, **84**, 3-337

Schipper J, Hoffmann M, Duckworth JW, Conroy J (2008) The 2008 IUCN red listings of the world's small carnivores. *Small Carnivore Conservation*, **39**, 29-34

Schmitt T and Varga Z (2012) Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology*, **9**, 22-33

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis, genetics and biometry laboratory, Department of Anthropology, University of Geneva, Geneva

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462

Sommer M, Kaczorek D, Kuzyakov Y, Breuer J (2006) Silicon pools and fluxes in soils and landscapes – a review (169, 310 pp, 2005), Journal of Plant Nutrition and Soil Science, **169**, 582-582

Sommer RS and Benecke N (2006) Late Pleistocene and Holocene development of the Felid fauna (felidae) of Europe: a review. *Journal of Zoology*, **269**, 7-19

Stahl P and Artois M (1994) Status and conservation of the wildcat (*Felis silvestris*) in Europe and around the Mediterranean rim. Convention on the conservation of European wildlife and natural habitats standing committee. *Council of Europe Nature and Environment series*, **69**, 1-76

Stewart JR, Lister AM (2001) Cryptic northen refugia and the origins of modern biota. *Trends in Ecology and Evolution*, **16**, 608-613

Surget-Groba Y, Heulin B, Guillaume CP, Puky M, Semenov D, Orlova V, Kupriyanova L, Ghira I, Smajda B (2006). Multiple origins of viviparity, or reversal from viviparity to oviparity? The European Common Lizard (*Zootoca vivipara*, Lacertidae) and the evolution of parity. *Biological Journal of Linnean Society*, **87**, 1-11

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **26**, 3189–3194

Vähä JPK and Primmer CR (2006) Detecting hybridization between individuals of closely related populations – a simulation study to assess the efficiency of model-based Bayesian methods to detect hybrid individuals. *Molecular Ecology*, **15**, 63-72

Van Oosterhout C, Hutchinson WF, Wills DPM, Schipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538

Vigne JD, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J (2012) First wave of cultivators spread to Cyprus at least 10,600 y ago. *Proceedings of the national academy of sciences of the USA*, **109(22)**, 8445-8449

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249-256

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167-184

Waples RS and Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753-756

Waples RS and Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244-262

Waples RS and England PR (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, **189**, 633-644

Waples EP, Antes AL, Murphy ST, Connelly S, Mumford MD (2009) A Meta-Analytic Investigation of Business Ethics Instruction. *Journal of Business Ethics*, **87(1)**, 133-151

Weir BS and Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370

3.2 SUPPORTING INFORMATION

Supplementary Table S1. Description of the 38 microsatellites used to genotype all cat (*Felis silvestris*) samples excluding admixed genotypes (see results). Locus identifications (ID) and chromosome assignments (Chr) are from Menotti-Raymond *et al*. (2003). Primer tails were labelled to fit the design of eight multiplexes. The number of repetitions (r), the allelic range (in base pairs), the observed number of alleles (N_A), the allelic richness (N_{AR}) based on 52 genes and the expected values of heterozygosity (H_E) are reported at each locus for domestic cats (*Fca*, n = 294), European wildcats (*Fsi*, n = 610) and African wildcats (*Fli*, n = 26).* dinucleotides showing intermediate alleles

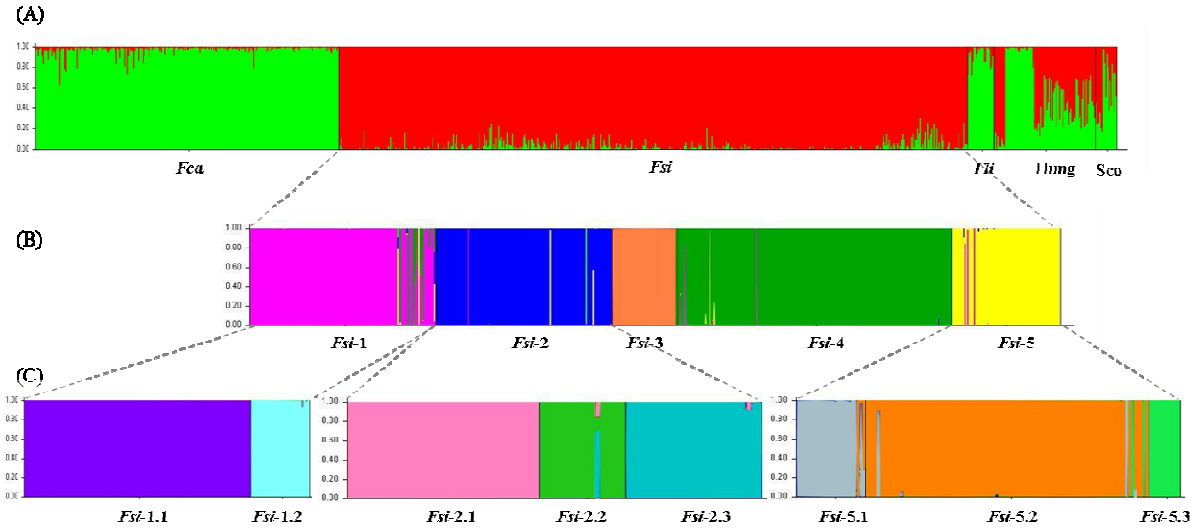
						Fca			*Fsi*			*Fli*		
Multiplex	ID	Label	r	Chr	Allele range	N_A	N_{AR}	H_E	N_A	N_{AR}	H_E	N_A	N_{AR}	H_E
M19	Fca058	Ned	2	E2	118-234	11	9.285	0.711	11	6.546	0.659	10	10	0.829
	Fca077	Vic	2	C2	132-162	10	7.587	0.747	11	8.184	0.797	7	7	0.752
	Fca088	Fam	2	B3	99-125	13	10.927	0.845	11	5.532	0.651	12	12	0.854
	Fca126	Ned	2	B1	103-157	14	10.326	0.817	13	6.475	0.755	10	10	0.852
	Fca453	Fam	4	A1	174-206	9	7.102	0.658	10	5.067	0.657	6	6	0.719
M32	Fca023	Fam	2	B1	124-154	14	12.956	0.784	14	6.333	0.752	11	11	0.799
	Fca045	Ned	2*	D4	135-165	21	17.798	0.916	22	11.947	0.854	14	14	0.865
	Fca080	Fam	2	A3	210-250	18	11.517	0.751	16	8.572	0.821	11	11	0.879
	Fca094	Ned	2	F2	207-249	16	13.41	0.883	16	9.042	0.754	14	14	0.890
	Fca097	Vic	2	B1	124-154	12	9.880	0.841	15	10.543	0.861	12	12	0.817
M34	Fca005	Ned	2	E1	124-152	13	9.342	0.762	11	8.063	0.799	7	7	0.808
	Fca035	Vic	2	D2	116-168	15	8.953	0.533	18	11.414	0.897	4	4	0.322
	Fca090	Fam	2	A1	82-118	16	13.717	0.808	14	8.097	0.792	12	12	0.853
	Fca262	Vic	2	D2	163-195	15	12.153	0.839	16	10.229	0.855	12	12	0.828
M50	Fca008	Ned	2	A1	112-148	16	12.474	0.863	18	8.692	0.807	9	9	0.820
	Fca043	Vic	2	C2	106-146	13	9.689	0.714	13	7.297	0.656	9	9	0.833
	Fca096	Ned	2	E2	96-235	21	14.354	0.517	18	11.079	0.884	8	8	0.715
	Fca293	Vic	2	C1	175-255	10	8.116	0.747	11	6.955	0.760	9	9	0.842
	Fca649	Fam	2	C1	114-152	19	12.989	0.841	18	7.546	0.690	12	12	0.882
M51	Fca026	Fam	2	D3	126-162	16	11.758	0.836	17	9.833	0.842	10	10	0.799
	Fca132	Ned	2	D3	127-161	24	17.436	0.878	15	9.690	0.845	11	11	0.858
	Fca391	Ned	4	B3	219-267	13	9.870	0.708	12	9.913	0.887	6	6	0.741
	Fca628	Vic	2	E3	77-155	32	23.204	0.910	22	10.281	0.834	16	16	0.911
M52	Fca105	Fam	2	A2	169-209	18	12.863	0.841	13	9.256	0.863	15	15	0.897
	Fca123	Vic	2*	A1	125-155	13	9.663	0.837	11	7.182	0.802	8	8	0.771
	Fca211	Ned	2	B1	98-120	9	7.537	0.644	12	7.565	0.744	8	8	0.845
	Fca305	Ned	2	B2	174-224	12	9.038	0.681	14	2.587	0.064	8	8	0.772
	Fca698	Fam	2	D1	210-272	17	15.311	0.893	16	8.576	0.775	14	14	0.884
Pp1	Fca069	Fam	2	B4	86-124	14	11.375	0.833	16	8.569	0.780	8	8	0.738
	Fca075	Ned	2	E2	104-142	16	10.470	0.830	13	9.852	0.876	10	10	0.873
	Fca220	Fam	2	F2	202-224	11	7.987	0.594	12	8.472	0.813	7	7	0.821
	Fca229	Ned	2	A1	142-176	14	11.933	0.774	16	9.370	0.723	8	8	0.776
	Fca441	Fam	4	D3	127-173	16	10.711	0.744	15	8.747	0.754	8	8	0.78
Pp2	Fca149	Fam	2	B1	116-134	9	7.234	0.768	9	5.597	0.697	6	6	0.775
	Fca223	Pet	2	F1	196-240	22	17.196	0.862	13	6.001	0.610	11	11	0.909
	Fca310	Vic	2	C2	106-140	13	10.400	0.777	11	2.612	0.068	13	13	0.909
	Fca678	Fam	2	A1	196-236	8	6.855	0.810	9	6.987	0.739	8	8	0.784
	Fca240	Fam	2	X	142-186	11	6.719	0.780	11	3.786	0.195	6	6	0.685

Supplementary Table S2. Identification of clusters number (*K*) in STRUCTURE analyses of cat samples. Optimal *K* values (in bold) were identified by the maximum increase (Δ*K*) of the mean *Ln* posterior probability (Mean lnPD) and of the mean F_{ST} values (ΔF_{ST}) between subsequent analyses. NA = not analysed.

Grouping	Populations	*K* value	Mean F_{st}	ΔF_{st}	Mean lnP(D)	ΔK
by cat subspecies	Fca	1	0.03976	NA	-143857.3	NA
	Fsi	2	**0.08632**	**0.72**	**-131734.82**	**3865.5**
	Fli	3	0.11034	0.34	-128943.34	29.405
	Hun	4	0.11865	0.16	-126957.86	23.784
	Sco	5	0.12042	0.28	-125702.44	1.7222
		6	0.13291	0.03	-124736.88	1.472
		7	0.14721	0.18	-123298.98	14.672
		8	0.14849	0.25	-122551.7	0.3875
		9	0.16585	0.18	-121915.84	0.0437
		10	0.17045	NA	-121294.06	NA
European wildcats by macro groups no loc prior	Fsi-1	1	NA	NA	-71972.86	NA
	Fsi-2	2	0.0018	0.84	-69252.32	9.3148
	Fsi-3	3	0.002	2.74	-67492.4	2.2765
	Fsi-4	4	0.03412	0.62	-65987.16	139.28
	Fsi-5	5	**0.03806**	**0.02**	**-64795.94**	**334.79**
		6	0.02575	0.12	-64049.88	171.11
		7	0.02145	0.08	-63924.94	1.1256
		8	0.01995	0.04	-63357.62	5.391
		9	0.0241	0.06	-63114.18	1.6318
		10	0.02955	NA	-62682.74	NA
European wildcats within Fsi-1	Fsi-1.1	1	0.011	NA	-15127.460	NA
	Fsi-1.2	**2**	**0.041**	**1.906**	**-14788.900**	**85.356**
		3	0.045	3.087	-14772.100	0.029
		4	0.092	0.089	-14744.500	0.066
		5	0.130	0.252	-14735.020	0.299
		6	0.129	0.243	-14811.360	0.130
		7	0.159	0.211	-14854.940	2.287
		8	0.154	0.246	-15483.020	1.452
		9	0.184	0.441	-15064.800	1.135
		10	0.148	NA	-15209.120	NA
European wildcats within Fsi-2	Fsi-2.1	1	0.002	NA	-14611.520	NA
	Fsi-2.2	2	0.038	3.205	-14251.840	0.485
	Fsi-2.3	**3**	**0.113**	**1.037**	**-13882.760**	**233.418**
		4	0.092	0.470	-13820.960	8.244
		5	0.114	0.281	-13826.780	2.295
		6	0.108	0.005	-13988.740	0.190
		7	0.103	0.159	-14100.000	0.821
		8	0.110	0.012	-14091.180	0.712
		9	0.117	0.121	-14282.480	1.949
		10	0.115	NA	-14973.500	NA
European wildcats within Fsi-5	Fsi-5.1	1	0.001	NA	-9279.820	NA
	Fsi-5.2	**2**	**0.094**	**1.146**	**-8834.600**	**115.244**
	Fsi-5.3	3	0.088	0.436	-8590.480	38.048
		4	0.109	0.084	-8433.140	3.121
		5	0.135	0.088	-8317.080	1.614
		6	0.153	0.173	-8248.380	1.248
		7	0.152	0.115	-8240.960	4.408
		8	0.162	0.090	-8703.820	0.527
		9	0.163	0.016	-8565.580	0.297
		10	0.161	NA	-8550.480	NA

Supplementary Table S3. Average proportion membership (Q_i) of wildcat populations obtained by STRUCTURE with K = 2, the admixture and F model, using the three cat subspecies and detailed all the sampling European wildcat populations.

Populations		***K* = 2**	
		Q_I	Q_{II}
Domestic cats		0.039	0.961
European wildcats	Portugal	0.844	0.156
	Spain	0.846	0.154
	Scotland	0.467	0.533
	Belgium + Luxembourg	0.940	0.060
	Switzerland	0.949	0.051
	Germany	0.956	0.044
	Italy	0.910	0.090
	Slovenia	0.965	0.035
	Hungary	0.461	0.539
	Bulgary	0.900	0.100
	Romania	0.906	0.094
Lybica wildcats		0.089	0.911



Supplementary Figure S1. (A) Bayesian clustering analyses performed in STRUCTURE to distinguish cat subspecies and admixed populations of Hungary and Scotland. (B) Population clstering of European wildcats populations into 5 main geographic-genetic groups, and into (C) 8 sub structured groups, run respectively without and with the 'sampling location prior model' in STRUCTURE. Each cat genotype is represented by a vertical bar split in *K* coloured sections, according to its relative assignment to the *K* genetic clusters.

3.3 LIST OF TABLE

Table 1. Sampling size and location of all genotyped cats. Hybrid cats that have not been identified through phenotypic evaluation and that have been initially included in the parental populations are reported between brackets ("admixed genotypes"). Acronym = symbol used, in this study, to identify the different populations

Table 2**.** Variability at 37 autosomal microsatellites for the three cat subspecies analysed (*Fca, Fsi* and *Fli*) and for the European wildcat subpopulations detected in clustering analyses. All putative hybrids and admixed populations were excluded. N = sample size; N_A = mean number of alleles per locus; N_{AR} and N_{PAR} = allelic and private allelic richness; H_O, H_E = observed and expected heterozygosity; F_{IS} = inbreeding coefficient (* significant departures from HWE at $p < 0.001$, Bonferroni corrected); HWE and LE = number of tests out of equilibrium at 37 loci and 666 pairwise comparisons, respectively. [A] populations' acronyms are used as in Table 1; [B] codes correspond to the symbols used to discriminate the different genetic clusters (see below) that subdivide European subpopulations; [C] N_{AR} and N_{PAR} were obtained for 52 and 22 genes when comparing subspecies and macro groups, respectively.

Table 3. Hierarchical analysis of molecular variance (AMOVA) computed in ARLEQUIN using ϕ_{ST} and values of F_{ST} and R_{ST} estimated under different perspectives of samples grouping. * all ϕ_{ST} values were highly significant at $p<0.001$

Table 4. Genetic divergence parameters (Φ_{ST}, below diagonal, and R_{ST}, above diagonal) for pairwise comparison between all the non-hybridizing European wildcat subpopulations and domestic cats sampled in our study. The last two columns report the genetic differentiation between each wildcat subpopulations and all domestic cats.

Table 5. Analyses of bottleneck signatures for each of the 10 wildcat subpopulations according to M-RATIO (Garza and Williamson, 2001) and BOTTLENECK (Cornuet and Luikart, 1997) probability tests. Following author's suggestion, populations with less than 15 individuals were not analysed for recent bottlenecks (n/a).

Table 6. Summary of the Bayesian analysis performed in STRUCTURE for the best *K* value obtained either a) analysing all populations together, either b) analysing each of the wildcat macro groups defined across Europe.

Table 7**.** Estimating demographic parameters computed for European wildcat macro groups and subpopulations. Harmonic means of estimated effective sizes (N_E), *S* sample size, *N* estimated using the Using Nunney & Elam's (1994) equation and a correction factor for domestic cat of 0.38 (Kaeuffe *et al.* 2004), 2.5% and 97.5% estimate quartiles (using the jacknife method), and mean numbers of independent allelic comparisons for the selected allele exclusion criteria (P_{crit}). P_{crit} was chosen following the Waples *et al.* (2009) suggestions: for sample size (*S*) larger than 100 P_{crit} = 0.01, for $S > 25$ $P_{crit} = 0.02$, and for $S \leq 25$ $P_{crit} = 0.05$.

Table 8. Summary of Prior distribution, mode and quantiles of Posterior distribution of the estimated demographic parameters (Symbol) for the European wildcat data.

Supplementary Table S1. Description of the 38 microsatellites used to genotype all cat (*Felis silvestris*) samples excluding admixed genotypes (see results). Locus identifications (ID) and chromosome assignments (Chr) are from Menotti-Raymond *et al.* (2003). Primer tails were labelled to fit the design of eight multiplexes. The number of repetitions (r), the allelic range (in base pairs), the observed number of alleles (N_A), the allelic richness (N_{AR}) based on 52 genes and the expected values of heterozygosity (H_E) are reported at each locus for domestic cats (*Fca*, *n* = 294), European wildcats (*Fsi*, *n* = 610) and African wildcats (*Fli*, *n* = 26).* dinucleotides showing intermediate alleles

Supplementary Table S2. Identification of the number of *K* clusters in STRUCTURE analyses of cat samples. Optimal *K* values (in bold) were identified by the maximum increase (ΔK) of the mean *Ln* posterior probability (Mean lnPD) and of the mean F_{ST} values (ΔF_{ST}) between subsequent analyses. NA = not analysed.

Supplementary Table S3. Average proportion membership (Q_i) of wildcat populations obtained by STRUCTURE with *K* = 2, the admixture and F model, using the three cat subspecies and detailed all the sampling European wildcat populations.

3.4 LIST OF FIGURES

Figure 1. Approximate sampling locations of wildcats surveyed across Europe and North Africa. Colours represent the highest partition of samples into genetic clusters and black lines divide the major differentiated groups (macroareas) in the European wildcat, as identified by multivariate and bayesian analyses. Dark areas in the map correspond to the approximate current distribution of *Felis silvestris* in Europe (adapted from Grabe and Worel 2001).

Figure 2. Principal Component Analysis displaying multivariate clustering of Hungarian (Hun) and Scottish (Sco) cats relatively to all sampled European wild (*Fsi*) and domestic cats (*Fca*).

Figure 3. Plot of discriminant analysis of principal components (DAPC) obtained with ADEGENET, displaying genetic diversity among: a) European *felis silvestris* subspecies; b) populations in Southern-Central and Eastern Europe; c) Southern-Western populations (Iberian peninsula) and d) Central European wildcats. Individuals (dots) and populations (coloured ellipses) are positioned on the space maximizing separation between groups.

Figure 4. Plot of the assumed Isolation-Migration model of the historical evolution of European wildcats populations in Iberian, Italian and Balkans Peninsula. Divergence time scale is reported on the left.

Supplementary Figure S1. (A) Bayesian clustering analyses performed in STRUCTURE to distinguish cat subspecies and admixed populations of Hungary and Scotland. (B) Population clstering of European wildcats populations into 5 main geographic-genetic groups, and into (C) 8 sub structured groups, run respectively without and with the 'sampling location prior model' in STRUCTURE. Each cat genotype is represented by a vertical bar split in *K* coloured sections, according to its relative assignment to the *K* genetic clusters.

Chapter 4

Improving the molecular toolbox

Combining use of most informative autosomal SNPs with uniparental markers (mtDNA and Y-chromosome) for the assessment of hybridization in European wildcats (*Felis silvestris*).

Federica Mattucci[1†], Edoardo Velli [1†], Leslie A Lyons[4], Paulo Célio Alves[2,3,5], Rita Oliveira[2,3] and Ettore Randi[1,6]

[1] Laboratorio di genetica, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Via Cà Fornacetta 9, 40064 Ozzano dell'Emilia (BO), Italy

[2] Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), InBio - Laboratório Associado, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[3] Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal

[4] Comparative Medicine Department of Veterinary Medicine & Surgery College of Veterinary Medicine, University of Missouri – Columbia, 1600 E. Rollins Street, MO 65211, USA

[5] Wildlife Biology Program, Department of Ecosystem and Conservation Sciences, University of Montana, 32 Campus Drive, Missoula, Montana 59812, USA

[6] Aalborg University, Department 18/Section of Environmental Engineering, Sohngårdsholmsvej 57, 9000 Aalborg, Denmark

[†] these authors contributed equally to this work

ABSTRACT

Rates of hybridization and introgression are increasing dramatically worldwide as a consequence of translocations of organisms and habitat modifications by humans. Because of a recent divergence and entirely overlapping ranges, introgressive hybridization between free-ranging domestic cats and European wildcats (*Felis silvestris silvestris*), might locally threaten the survival and conservation of indigenous wildcats populations. Identifying pure wildcats and investigating the ancestry of admixed individuals is thus crucial for supporting appropriate conservation and managing programs of European wildcat. However, introgression is difficult to detect and the available morphologic and genetic markers for both Felis subspecies proved to be not sufficient to

reliably detect hybrids beyond first generation. In this study, we present a single nucleotide polymorphism (SNP) based approach combined with mitochondrial and Y-linked chromosome polymorphisms that allows the identification of introgressed individuals. First, we analyzed the genetic variation of 151 highly informative SNPs (F_{ST} > 0.8) on 187 European cat samples, including 45 village domestic cats, 100 putative European wildcats and 42 previously known or putative wild *x* domestic hybrids. The power of all the loci to accurately identify admixture events and discriminate the different hybrid categories was evaluated. Results from Bayesian model-based computations of simulated and real genotypes show that the 151 SNPs provide successful estimates of admixture, with 100% hybrid individuals (up to second generation backcrosses) being correctly identified in STRUCTURE analyses and 100% using the NEWHYBRIDS' algorithm. None of the unclassified cats were wrongly allocated to another hybrid class. The chromosome Y-linked markers further analyzed, proved to be useful for identifying wild and domestic cat males based on distinctive polymorphism. Furthermore, both mitochondrial sequenced regions (ND5 and part of the control region), clearly separated the subspecies in two well distinct haplogroups, with the exception of a few domestic shared haplotypes, suggesting the occurrence of ancient introgressive events into wildcats populations or the unreliability of the mitochondrial markers used. The integration of both uniparental and nuclear markers assignments, provided a complete insight of introgression level in wildcats populations analyzed. Finally, a total of 11 hybrids were identified: 8 were detected by both nuclear and uniparental markers, and 3 were identified because of the presence of nuclear genotype wild and shared domestic mitochondrial haplotype. This approach may be useful to further reconstruct both the historical and recent evolution of wildcat populations and, hopefully, to develop sound conservation guidelines for its legal protection in Europe.

Keywords: *Felis silvestris*, European wildcat, domestic cat, hybridization, introgression, single nucleotide polymorphisms, mitochondrial DNA, Y Chromosome, conservation genetics

INTRODUCTION

Anthropogenically driven changes of the spatial distribution of species are increasing the incidence of hybridization events (Reusch and Wood 2007), critically threatening the native fauna (Wayne and Brown 2001; Randi 2008). Especially in endangered taxa, hybridization (and introgression) occurring between wild species and their domestic counterparts may disintegrate the genetic integrity of the wild conspecific (Rhymer and Simberloff 1996; Allendorf et al. 2001). Genetic introgression of domestic alleles into native gene pools may, in fact, introduce genes favored under artificial selection that are maladaptive in the natural environment, disrupting locally adaptation or increasing genetic homogenization (Rhymer and Simberloff 1996). Both of these processes can reduce the fitness of wild species raising risks of genetic extinction, loss of local adaptations or outbreeding depression (Lynch & O'Hely 2001; Rhymer and Simberloff 1996; Allendorf et al. 2001; McGinnity *et al*. 2003; Hutchings and Fraser 2008).

Interbreeding between domesticated and wild counterparts has been observed in terrestrial carnivores, ungulates, fowl, anurans and many fish species (Rhymer and Simberloff 1996; Williams *et al*. 2002). But one of the most remarkable example to understand the consequences of anthropogenic hybridization on natural populations is the current situation of the European wildcat (Felis silvestris silvestris). The human-mediated pan-global dispersal of the domestic cat, together with the past demographic decline of European wildcats' population and the fragmentation of suitable habitat (McOrist and Kitchener 1994) have, in fact, increased the risk of anthropogenic hybridization and have promoted the extinction of some natural populations during the last century. In addition, the fertility of the hybrid offspring (Pierpaoli et al. 2003), could have facilitated the continuing dilution of the wild genotype over progressive generations.

Since evidence of extinction of localized wildcat populations has already been detected in central Europe (Suminski 1962), the prevention of hybridization has been identified as the greatest priority for the persistence of the subspecies (Driscoll and Nowell 2010). Hence, accurate detection of hybrid individuals and quantification of introgression rate in potentially threatened populations are the main challenges for assessing wildcats conservation's status and, subsequently, developing appropriate conservation measures.

Over the last decade, the hybridization pattern with feral domestic cats had been assessed through molecular approaches. In particular, the genotyping of several highly polymorphic molecular markers, specifically microsatellites (short tandem repeats - STR), and partial mitochondrial DNA sequences, combined with new Bayesian statistical tools have radically improved knowledge of the genetics of European wildcats (e.g. Beaumont *et al.* 2001; Randi *et al.*

2001; Pierpaoli *et al.* 2003; Kitchener *et al.* 2005; Lecis *et al.* 2006; Oliveira *et al.* 2008a,b; Eckert *et al.* 2010; O'Brien *et al.* 2009; Hertwig *et al.* 2009). The empirical evidence available so far suggests that hybridization occurred sporadically in some European countries (i.e. Spain, France, Germany and Italy), but extensively in others, as in Scotland or in Hungary (Beaumont *et al*. 2001; Pierpaoli *et al*. 2003), where, most likely, the presence of forests patches and traditional agriculture with human settlements facilitates the meeting between free-ranging domestic cats and wildcats.

The recent domestication (Vigne *et al.* 2012) and the protracted coexistence of domestic and wild cats raised fear that widespread interbreeding would have led to genetic pollution and rendered uncertain any identification of 'pure' wildcat populations. Until recently, the type and number of markers showed a limited power of hybrid detection after the first few generations of backcrossing (Oliveira *et al.* 2008a,b; Hertwig *et al*. 2009; Say *et al.* 2012). Thereby, a set of more powerful markers is required to assess the level of introgression in natural wildcat populations (Randi 2008; Vaha and Primmer 2006).

Over the last decades, next-generation sequencing (NGS) technologies generated a large quantity of nucleotide sequence data, promising to improve vastly the ability to study hybridization and introgression by using both molecular phylogenetic approaches and population genetic studies. For backcross detection, single nucleotide polymorphism (SNP) markers appear promising because of: i) reduced propensity for homoplasy due to lower mutation rates; ii) higher density and more uniform distribution in genomes; iii) suitability for successful high-throughput genotyping and straightforward comparability and transportability across laboratories and detection protocols, and iii) highly successful application in fragmented DNA samples, e.g. non-invasive and historical DNA (see Brumfield *et al*. 2003; Morin *et al*. 2004; Garvin *et al*. 2010 for reviews). Nussberger *et al.* (2013), recently identified a promising set of 48 nuclear SNPs for detecting European wildcats, domestic cats and their admixed progeny. However, the reference samples used in this work were limited to Switzerland, and the SNPs power for hybrid detection has probably been overestimated by computing simulations using individuals too much differentiated than average.

We developed a 2-step protocol for assessing 'pure' reference cats, collected over 13 different localities of European wildcat home range, and for estimating introgression in conspecific wildcats.

First, we used the Illumina Infinium iSelect 63K Cat DNA Array to amplify 250 cat samples, including three *Felis silvestris* subspecies (*silvestris, lybica* and *catus*) and a number of known and putative admixed individuals previously identified with high polymorphic microsatellites panel set (Pierpaoli *et al.* 2003; Mattucci *et al.* in prep). Since this work was largely motivated by the need to routinely identify admixture events in conservation studies of European

wildcat populations, we extracted the minimum number of highly informative SNPs able to efficiently detect current levels of hybridization between wild (*Felis s. silvestris*) and domestic (*Felis s. catus*) cats, regardless of their geographical origin. Finally, we tested the hybrid discrimination power of 151 most informative SNPs on 187 cat samples, excluding from this work all individuals belonging to the *Felis s. lybica* subspecies, and to the highly introgressed populations of Hungary and Scotland (Beaumont *et al.* 2001).

Second, we analyzed the genetic variation occurred in the Y chromosome and in both the control region and the NADH dehydrogenase subunit 5 of the mtDNA. The integrated molecular panel set obtained combining the 151 most informative SNPs with the uniparental markers, was then used to accurately identify 'pure' reference cat and second generation backcrosses and to investigate ancient introgressive events occurred in wildcat populations, in order to promote and prioritize conservation efforts for the subspecies survival in the near future.

MATERIALS AND METHODS

Genotyping cats and screening for the most informative SNPs

We used the Illumina Infinium iSelect 63K Cat DNA Array to genotype 100 putative European wildcats (*Felis s. silvestris*), 45 village domestic cats (*Felis s. catus*), and 41 previously known or putative wild *x* domestic hybrids (sampling information and locations are shown in Table 1 and Fig 1) at 62,897 autosomal SNPs (referred to as the 63K panel) and 2735 X-chromosome SNPs (X-SNPs).

All samples included in this work were previously analyzed with 35 or 38 unlinked microsatellites loci for both their admixture and genetic diversity patterns (Mattucci *et al.* 2013; Mattucci *et al.* in prep).

European wildcats were opportunistically collected from found-dead or trapped animals, across 10 European geographic localities (Table 1). Sampling was performed by randomly selecting a few available individuals from each location and by taking in consideration the natural distribution of European wildcat and the genetic fragmentation of its populations across Europe (Mattucci *et al*. in prep). All putative wildcats were identified both morphologically by collectors, according to phenotypic traits (Schauenberg 1969, 1977; French *et al.* 1988; Ragni and Possenti 1996), and genetically with microsatellites loci (Mattucci *et al.* 2013; Mattucci *et al.* in prep).

Ten known hybrids were added aiming to help the identification of admixed wildcats in the European samples. Four of these hybrids were obtained from controlled crosses (Ragni 1993); six wild-living hybrids, including a family of five full-sibs extracted from the uterus of a road-killed apparently pure *F. s. silvestris* female, were genetically identified in other studies and reanalyzed here (Pierpaoli *et al*. 2003; Lecis *et al*. 2006). 31 putative hybrids were further included in the analysis aiming to verify the introgression level detected in previous studies (Mattucci *et al.* 2013; Mattucci *et al.* in prep).

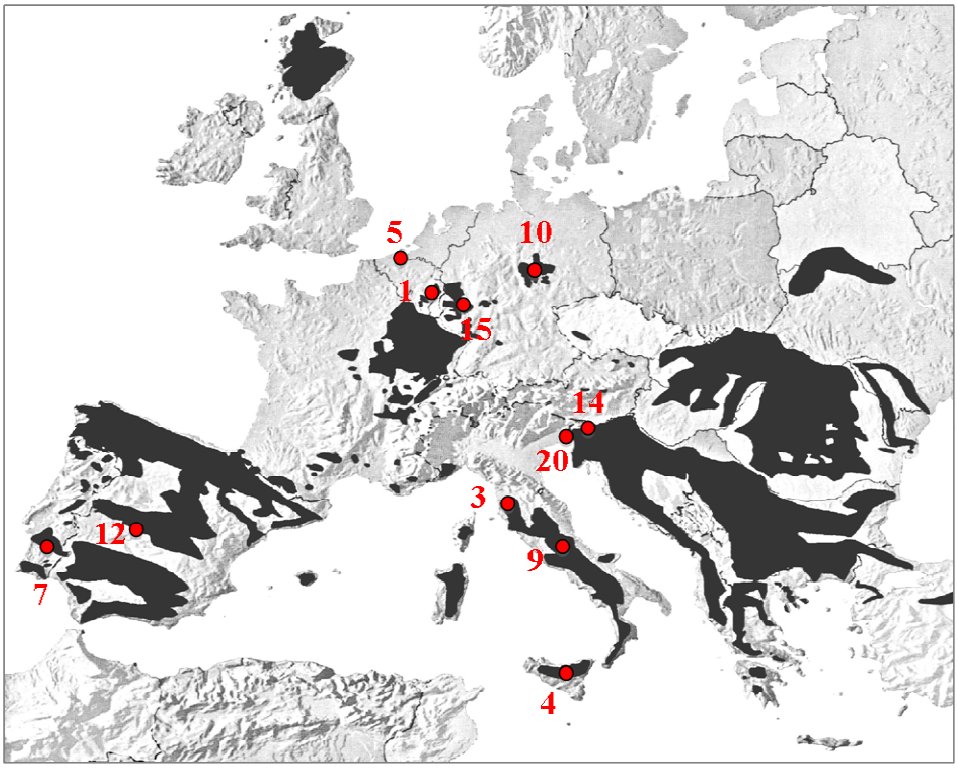


Figure 1. Sampling locations of putative European wildcats. Shaded areas correspond to the approximate current distribution of Felis silvestris in Europe (adapted from Grabe and Worel 2001).

Table 1. Sampling size and location of all genotyped cats. Hybrid cats that have not been identified through phenotypic evaluation and that have been initially included in the parental populations are reported between brackets ("admixed genotypes").

Subspecies	**Sampling locations**	**N**
Domestic cats	Italy	18
Felis silvestris catus	Poland	4
n = 45	Portugal	10
	Spain	12
	Greece	1
known hybrids (or putative hybrids)	Captivity (Italy)	4
Felis silvestris x catus	Italy: Northern-eastern Alps	6
n = 42	Italy: Central Apennines and Maremma	14
	Italy: Southern Apennines	6
	Luxembourg	2
	Portugal	6
	Spain	1
	Bosnia & Herzegovina	1
	Germany: Southern-western	2
Putative European wildcats	Italy: Northern-Eastern Alps	20
Felis silvestris silvestris	Slovenia	14
n = 100	Italy: Maremma	3
	Italy: Central	9
	Italy: Sicily	4
	Germany: Central	10
	Germany: Southern-western	15
	Belgium: Wallonia	5
	Luxembourg	1
	Portugal	7
	Spain	12

A total of 37 cats, belonging to the wild, domestic and admixed populations, were removed from the initial dataset, because of a missing rate per individual (MIND) < 0.2. Hence, we obtained a reduced dataset of 150 individuals that we used for all the subsequently elaborations.

The initial 63K panel set was pruned using PLINK (Purcell *et al.* 2007) for loci that were invariant, showing individual missing rates per SNPs (GENO) < 0.2, or minor allele frequency MAF < 0.05. We further pruned the panel for loci in linkage disequilibrium (LD), filtering SNPs with $r^2 < 0.5$ within 50 SNP sliding windows, shifted and recalculated every 5 SNPs, see Table for

details. Finally we obtained a panel of 26,361 SNPs (referred to as the 26K panel) from which we preliminarily excluded the X-SNPs because of the unbalanced (*n* males = 114; *n* females = 21; sex not revealed *n* = 15) sex-ratio occurred in the reduce dataset (*n* = 150) analyzed.

The package HIERFSTAT (Goudet 2005a) was used to estimate the F-statistics and variance components among cat subspecies, and identify the most informative SNPs, among the pruned 26K panel. Based on the highest F_{ST} value (F_{ST} >0.80) among European wild and domestic cat populations, a number of 151 SNPs were selected as the most informative loci, and, subsequently, were analyzed for their hybridization diagnostic value by computing their informativeness for assignment index (*In*) with INFOCALC (Rosenberg *et al.* 2003, 2005). *In* is a mutual information based statistics. From a likelihood perspective, *In* gives the expected logarithm of the likelihood ratio that an allele is assigned to one of the populations compared with a hypothetical 'average' population whose allele frequencies equal the mean allele frequency across sub-populations. Correlations between the F_{ST} and *In* ranks were further computed using the Spearman Rank correlation coefficient (Kendall 1970).

The power of the 151 markers to identify each unique genotype was evaluated calculating the probability-of-identity (PID and PIDsibs; Mills *et al.* 2000; Waits *et al.* 2001) in GENALEX 6.41 (Peakall and Smouse 2006).

Statistical analyses of the SNP markers

Summary statistics were used to describe levels of genetic variability and differentiation on the wild and domestic subspecies, excluding all admixed genotypes (known and putative hybrids) detected in the hybridization analyses (see below). The individuals excluded were, afterwards, included to the dataset for hybridization analyses.

We computed values of observed (H_O) and unbiased expected (H_E; Nei 1978) heterozygosity for all locus-population combinations, Hardy-Weinberg and linkage disequilibrium tests (HWLE; using the Markov chain exact test with a chain length of 100,000 and 3,000 dememorization steps), AMOVA and F-statistics (testing the null hypothesis of no differentiation by permuting genotypes between populations with 10,000 replicates at $P < 0.001$) in ARLEQUIN 3.5.1.2 (Excoffier and Lischer 2010). Allelic richness (N_{AR}) was estimated in FSTAT 2.9.3.2 (Goudet *et al.* 2002).

Admixture analyses and assignment of the individual genotypes

We used STRUCTURE 2.3.3 (Pritchard *et al*. 2000; Falush *et al.* 2007; Hubisz *et al.* 2009) to identify the most likely population clusters in the sample and perform admixture analyses. Each run of STRUCTURE was replicated five times, with 10^4 burn-in followed by 10^5 simulations, combining the "admixture" model with correlated and independent (respectively "*F* and *I*") allele frequency models, both with or without prior non-genetic information (option *usepopinfo* = 1 or = 0, respectively). The optimal number of clusters (*K*) was identified by the ΔK and ΔF_{ST} statistics (Evanno *et al*. 2005) in CORRSIEVE 1.6.1 (Campana *et al*. 2011). For each selected *K* value, we assessed: 1) the average proportion of membership (Q_i) of the sampled populations to the inferred clusters; 2) the individual proportion of membership (q_i) to one or more than one (in case of admixed genotypes) of the inferred clusters; and 3) the 90% credibility intervals (CI) of the q_i values.

STRUCTURE was run using the pruned dataset (n = 150) of putative European wild cats (*Fsi*), domestic cats (*Fca*) and putative or known admixed cats and the 151 most informative SNPs.
Hybrid class of the admixed genotypes were identified using NEWHYBRIDS (Anderson and Thompson 2002), under the default computational parameters of the program.

NEWHYBRIDS estimates the posterior probability of belonging to each of six genotypic classes corresponding to hybrid categories (H_i): parental subspecies (domestic or wild cats), F1, F2, and the first backcrosses. Jeffreys priors were chosen to down-weight the influence of an allele that might be rare in one species and absent in the other. Ten independent runs were performed to test for stability.

The power of the 151 SNPs to detect different hybrid classes was assessed by the analysis of the assignment accuracy obtained for simulated genotypes. Forty multi-locus genotypes of each parental (wildcat x wildcat; domestic cat x domestic cat), F1 (wildcat x domestic cat), F2 (F1 x F1) and backcrosses categories were generated with the software HYBRIDLAB 1.0 (Nielsen *et al.* 2006) and, afterwards, analysed using STRUCTURE and NEWHYBRIDS under the same setting of the admixture analysis described above. Q_i threshold values for all analyses where established by the minimum value for which all parental cats could be correctly assigned. Observed genotypes that displayed admixed genetic assignments or for which molecular assignments opposed their prior morphological identifications in the hybridization analyses of STRUCTURE and NEWHYBRIDS were also analysed together with the simulated genotypes.

Mitochondrial sequencing and chromosome Y characterization

We sequenced 877 bp (including the primers) of the mtDNA NADH dehydrogenase subunit 5 (ND5; nucleotides 13131 - 14007 mapped on the mitochondrial genome of the domestic cat; NCBI Reference Sequence NC001700), which, according to Driscoll *et al.* 2011, contains 7 diagnostic SNPs discriminating European wildcats (*Felis silvestris silvestris*) and domestic cats (*Felis silvestris catus*). This sequence was amplified using PCR primers F2B (5'-TGCCGCCCTACAAGCAAT-3') and R3B (5'-TAAGAGACGTTTAATGGAGTTGAT-3') (Driscoll *et al.* 2011). In addition, we sequenced 719 bp of the mtDNA control-region (sites 16236 - 16955) using primers CHF3 (5'-CTC CCT AAG ACT TCA AGG AAG-3'; Freeman *et al.* 2001) and CHR3 (5'-CCT GAA GTA AGA ACC AGA TG-3'; Tiedemann *et al.* 1996). Each 10 µL PCR reaction contained 2 µL of DNA (*c*. 50 ng), 0.8 µL of 10X Taq Buffer advanced with self-adjusting Mg^{2+} (Eppendorf, Milano, Italy), 0.80 µL of 0.2% bovine serum albumin (Sigma-Aldrich, St Luis, USA), 0.36 µL of 2.5 mM dNTPs (Eppendorf, Milano, Italy), 0.15 µL of each 10 mM primer solution (Bionordika, Stockholm, Sweden), 0.04 µL of 5U/µL HotStart Taq polymerase (Eppendorf, Milano, Italy) and 5.70 µL of purified water (Eppendorf, Milano, Italy). PCRs were performed in a Veriti Thermal Cycler (Life Technologies, Carlsbad, USA) with the following thermal profile: 94°C for 15 min for initial denaturation and Taq activation, followed by 50 cycles of 30 s at 94°C, 60 s at 55°C and 60 s at 72°C. The PCR cycling was followed by a final extension for 10 min at 72°C. PCR products were stored at 4°C and then purified by exonuclease digestions (1 µL of EXO-SAP per samples, incubated at 37° C for 30 min, then at 80° C for 15 min). The purified amplicons were Sanger-sequenced. Each 10 µL reaction contained 1 µL of amplified DNA, 1 µL of BigDye 1.1 (Life Technologies, Carlsbad, USA), 0.2 µL of either the forward or reverse primer and 7.8 µL of purified water. Sequencing was performed in a Veriti Thermal Cycler with 25 cycles of 10 s at 96° C, 5 s at 55° C, 4 min at 60° C and a storage at 4° C. Sequences were cleaned from unincorporated label nucleotide by precipitation adding to each PCR product a 12 µL mix composed by 2 µL of NaOAc 3M and 10 µL of purified water. Then were added 50 µL of 100% EtOH. The mix was then centrifuged at 12500 rpm for 10 min. The supernate was extracted and eliminated manually using a tranfer pipette. The precipitate was washed with 70 µL of 70% EtOH and centrifuged at 12500 rpm for 7 min. The supernate was eliminated and the precipitate was left to dry in the dark. The purified product was combined with 10 µL of Hi-DI formamide (Life Technologies, Carlsbad, USA) and denatured for 3 min at 95°C. Products were separated on an ABI 3130 DNA Analyzer.

We sequenced 376 bp of Y-linked SRY gene (Nussberger *et al.* 2013), using the primers SRYF (5'–GGCCTGTGTGTCGTTTAACA –3') and SRYR (5'- GTTTTTCCACAGGAGGGATG -3') using the same reagents recipes and thermal conditions used for the sequencing of ND5 region. Furthermore we amplified a biallelic microsatellite locus SMCY-7 STR on Chr Y. Both these markers present a polymorphism that seems to be fixed with different alleles in the two subspecies under study (Pecon-Slattery *et al.* 2004; King *et al.* 2007; Luo *et al.* 2007; Nussberger *et al.* 2013). We considered as reliable only samples that yielded a genotype for both the two markers. For the STR amplification each 10 µL reaction contained 1,5 µL of DNA (> 50 ng), 0.8 µL of 10X Taq Buffer advanced with self-adjusting Mg^{2+} (Eppendorf, Milano, Italy), 0.80 µL of bovine serum albumin 0.2% (Sigma-Aldrich, St Luis, USA), 0.36 of dNTPs 2.5 mM (Eppendorf, Milano, Italy), 0.2 µL of each primer 10 mM (Bionordika, Stockholm, Sweden), 0.04 of Taq polymerase 5U/µL (Eppendorf, Milano, Italy) and 5.80 µL of purified water (Eppendorf, Milano, Italy). Each reaction was amplified under the following PCR condition using a Veriti Thermal Cycler (Life Technologies, Carlsbad, USA): 94°C for 2 min for initial denaturation and taq activation, followed by 10 cycles of 40 s at 94°C, 30 s at 60°C and 30 s at 72°C. Every cycle temperature decreased by 0,5° C. The PCR cycling was followed by 25 cylces of 40 s at 94°C, 30 s at 55°C and 30 s at 72°C with a final extension for 10 min at 72°C. PCR products were stored at 4°C. The amplified product was combined with 10 µL of Hi-DI formamide (Life Technologies, Carlsbad, USA) and denatured for 3 min at 95°C. Products were separated on an ABI 3130 DNA Analyzer. Table x provides a complete list of the markers and the primers used for amplification.

Sequences were aligned using SEQSCAPE software 2.5 (Life Technologies, Carlsbad, USA). The sequence of full mtDNA genome of the domestic cat (NCBI Reference Sequence: NC_001700), trimmed at the above-mentioned positions, was used as reference sequence. The seven known variants were tagged on the reference sequence and all sequences were analyzed to find all the other variants using a conservative criterium. The exported sequences of the two regions were trimmed using BIOEDIT 7.1.11 (Hall 1999) respectively into equal sequences of 671 bp (positions 13243 – 13913) and 600 bp (16302 – 16901) for independent sequence analyses to maintain full-length, double-stranded, high-quality sequence data across all samples. STR fragment were analyzed and corrected using GENEMAPPER 4.1 (Life Technologies, Carlsbad, USA). Haplotypes, genetic diversity and basic statistics were extrapolated using DNASP 5.10.01. Software NETWORK 4.6 (Fluxus Technology Ltd) was used to construct the phylogenetic network and verify the separation between the two population (wild and domestic). The analysis of molecular variance (AMOVA) performed in ARLEQUIN was used to estimate the degree of variation within each population and the degree of differentiation between them.

RESULTS

SNPs screening and variability

A total of 33,833 SNPs was excluded from the initial 63K SNPs panel after PLINK pruning (see Table 2) for missing rate (i.e. no genotyping (n = 33) and GENO filter (n = 649)), allele frequencies (i.e. fixed (n = 30) and MAF filter (n = 14,513), INDEL and Linkage Disequilibrium (7 and 18,601 respectively), improving the SNP genotype call rate to 97% in the remaining analyzed cat samples (n = 150).

Table 2. SNPs and genotypes preliminary pruning computed in PLINK.

n SNPs to remove	Filter type	Description
7	INDEL	Insertion and Deletion
30	fixed	Fixed with the same mutation among the 3 Subsps analyzed
33	no genotyping	No genotyping results
649	GENO	Missing rate per SNP < 0.2
14513	MAF	Minor Allele Frequencies ≥ 0.05
18601	LD	Linkage Disequilibrium based SNP pruining < 0.5
n individuals to remove		
37	MIND	Missing rate per person < 0.2

Descriptive statistics are presented in Table 3, excluding all the admixed cats detected through the admixture analysis (see below). All SNPs were polymorphic among the wildcats (n = 89), implying a MAF > 5%, however, 7 SNPs (about 5%) were monomorphic amongst domestic cats (n = 44) and 52 SNPs (about 34%) showed a MAF < 5%. None significant deviations from HWE, following Bonferroni correction (P < 0.000065), was detected in anyone SNP loci. All SNPs were significantly portioned between wildcat and domestic cats (average F_{CT} = 0.864; P <0.001), with single-locus F_{CT} pairwise values ranging between 0.797 (ChrUn13.13324872) and 0.982 (ChrC1.124364347; AMOVA P < 0.001). The mean value of H_E was 0.101 (± 0.005), ranging from 0.000 (at 7 loci monomorphic in domestic cats; i.e ChrC1.216357902, ChrD1.132172994, ChrUn.34671542, ChrB3.26272231, ChrUn1.10713841, ChrC1.120057704, ChrUn5.2974540) to 0.325 (ChrC1.63091997) in the domestic group; while a mean value of H_E = 0.127 (± 0.004), ranging between 0.011 (ChrC1.59406628) and 0.236 (ChrUn26.10046275), was disclosed in the wildcat group. Average value of allelic richness proved to be significantly high both in domestic (A_R = 1.930) and wildcat group (A_R = 1.998; P < 0.001).

The average informativeness for assignment score (I_N) was 0.459, ranging from 0.002 (ChrB1.193634290; ChrB2.117030105) to 0.641 (ChrC1.105529441), see Table 3. The low Spearman's Rank Correlation coefficient (r = -0.11) revealed the absent of any correlation among the informativeness scores estimated using both the F_{ST} and *In* method (see Table 3). Any correlation was further displayed among the distribution of H_E values and both the informativeness estimates (data not shown), despite markers with lowest values of He in both groups should displayed the highest values of differentiation as they represented high frequencies of the two possible alternate variants.

Table 3. Genetic description of 151 most informative SNP loci used in this study: Allelic richness (A_R); expected heterozygosity (H_E); Inbreeding coefficient (F_{IS}); pairwise estimate of genetic differentiation between European wild and domestic cats (AMOVA pw-F_{CT}); informativeness for assignment index (INFOCALC I_N) computed for each SNP locus.

SNP	Chr	Position Mb	A_R		H_E		F_{IS}		pw F_{CT}	I_N
			Fca	Fsi	Fca	Fsi	Fca	Fsi		
chrC1.124364347	C1	124364347	1,814	1,972	0.023	0.014	mono	mono	0.982	0.544
chrD2.28093358	D2	28093358	1,967	2,000	0.045	0.028	-0.012	-0.007	0.964	0.544
chrC1.120057704	C1	120057704	1,000	2,000	0.000	0.061	mono	0.386	0.959	0.436
chrD1.156500121	D1	156500121	2,000	1,992	0.082	0.034	-0.030	-0.012	0.949	0.484
chrUn.34671542	UN	34671542	1,000	2,000	0.000	0.078	mono	0.311	0.947	0.459
chrA1.154897381	A1	154897381	1,795	2,000	0.022	0.066	mono	0.315	0.946	0.457
chrC1.243444286	C1	243444286	1,795	2,000	0.022	0.087	mono	0.481	0.931	0.407
chrA1.67830943	A1	67830943	2,000	2,000	0.069	0.064	-0.025	0.385	0.929	0.587
chrD3.66450925	D3	66450925	1,995	2,000	0.067	0.066	-0.024	-0.030	0.928	0.455
chrB4.142126995	B4	142126995	2,000	1,992	0.133	0.034	-0.065	-0.012	0.926	0.465
chrUn1.10713841	UN1	10713841	1,000	2,000	0.000	0.107	mono	-0.055	0.925	0.440
chrB1.162220276	B1	162220276	1,993	2,000	0.066	0.072	-0.024	-0.033	0.924	0.439
chrE2.44868665	E2	44868665	1,974	2,000	0.046	0.086	-0.012	0.220	0.922	0.393
chrD2.101514624	D2	101514624	1,967	2,000	0.045	0.088	-0.012	0.219	0.921	0.543
chrC1.59406628	C1	59406628	2,000	1,787	0.187	0.011	-0.105	mono	0.920	0.510
chrB2.117030105	B2	117030105	1,999	2,000	0.089	0.064	0.485	0.385	0.920	0.002
chrUn31.884121	UN31	884121	1,999	2,000	0.087	0.067	-0.036	-0.030	0.920	0.414
chrC2.71556573	C2	71556573	1,999	2,000	0.091	0.065	-0.038	-0.029	0.919	0.423
chrA3.123183917	A3	123183917	1,999	2,000	0.087	0.068	0.485	-0.031	0.918	0.400
chrF1.24859656	F1	24859656	1,999	2,000	0.091	0.071	-0.038	-0.032	0.914	0.446
chrUn38.11625325	UN38	11625325	1,993	2,000	0.066	0.087	-0.024	-0.042	0.913	0.411
chrA1.166331725	A1	166331725	1,814	2,000	0.023	0.112	mono	-0.057	0.913	0.416
chrA1.31688322	A1	31688322	1,993	2,000	0.066	0.090	-0.024	-0.044	0.911	0.544
chrD4.36674221	D4	36674221	2,000	2,000	0.110	0.066	0.373	-0.030	0.911	0.477
chrUn5.2974540	UN5	2974540	1,000	2,000	0.000	0.127	mono	0.111	0.910	0.409
chrC1.236293313	C1	236293313	1,996	2,000	0.069	0.092	-0.025	0.479	0.908	0.418
chrA3.40321217	A3	40321217	1,960	2,000	0.044	0.107	-0.012	0.157	0.907	0.402
chrD4.41337045	D4	41337045	1,967	2,000	0.045	0.110	-0.012	0.156	0.905	0.426
chrF2.34671501	F2	34671501	1,995	2,000	0.067	0.100	-0.024	0.185	0.902	0.398
chrUn.59973692	UN	59973692	1,854	2,000	0.024	0.128	mono	0.128	0.900	0.582
chrB1.90775428	B1	90775428	1,999	2,000	0.089	0.096	0.485	0.186	0.896	0.488
chrD4.115991773	D4	115991773	1,999	2,000	0.089	0.098	-0.037	-0.049	0.895	0.460

chrA1.186398902	A1	186398902	2,000	1,991	0.201	0.033	0.333	-0.012	0.894	0.426
chrD1.132172994	D1	132172994	1,000	2,000	0.000	0.151	mono	0.072	0.893	0.548
chrB3.1681903	B3	1681903	1,960	2,000	0.044	0.126	-0.012	0.290	0.893	0.481
chrF1.58410096	F1	58410096	1,999	2,000	0.089	0.101	-0.037	0.651	0.892	0.514
chrE1.127826462	E1	127826462	2,000	2,000	0.130	0.078	-0.063	-0.037	0.892	0.392
chrB3.74631411	B3	74631411	2,000	2,000	0.110	0.092	-0.050	0.218	0.890	0.476
chrE2.69893420	E2	69893420	1,995	2,000	0.067	0.117	-0.024	0.366	0.890	0.443
chrC1.92123224	C1	92123224	1,974	2,000	0.046	0.128	-0.012	-0.068	0.889	0.440
chrD4.32906349	D4	32906349	1,996	2,000	0.069	0.116	-0.025	0.327	0.889	0.526
chrC1.216357902	C1	216357902	1,000	2,000	0.000	0.156	mono	0.204	0.889	0.569
chrA3.5987880	A3	5987880	2,000	2,000	0.107	0.097	-0.049	0.186	0.887	0.529
chrA2.194725092	A2	194725092	1,795	2,000	0.022	0.146	mono	-0.081	0.886	0.427
chrB4.119147	B4	119147	1,960	2,000	0.044	0.135	-0.012	-0.073	0.885	0.428
chrC1.26196706	C1	26196706	2,000	2,000	0.165	0.066	0.186	-0.030	0.885	0.380
chrB4.123591019	B4	123591019	1,960	2,000	0.044	0.142	-0.012	0.643	0.883	0.397
chrA3.2785813	A3	2785813	2,000	2,000	0.110	0.104	-0.050	0.417	0.881	0.467
chrB2.146764150	B2	146764150	1,999	2,000	0.089	0.116	-0.037	0.327	0.880	0.514
chrE3.69030840	E3	69030840	2,000	2,000	0.110	0.106	0.373	0.417	0.879	0.451
chrC2.170223552	C2	170223552	1,999	2,000	0.089	0.118	-0.037	0.132	0.879	0.450
chrUn.9699383	UN	9699383	2,000	2,000	0.110	0.107	-0.050	-0.054	0.878	0.582
chrE2.20288683	E2	20288683	2,000	2,000	0.130	0.096	0.649	-0.048	0.878	0.472
chrA1.169461878	A1	169461878	2,000	2,000	0.165	0.076	-0.089	-0.035	0.877	0.469
chrUn12.17303165	UN12	17303165	1,999	2,000	0.087	0.126	-0.036	-0.067	0.874	0.426
chrB3.114518607	B3	114518607	1,999	2,000	0.087	0.127	-0.036	0.646	0.873	0.445
chrUn12.8261513	UN12	8261513	1,999	2,000	0.089	0.126	-0.037	-0.067	0.873	0.433
chrD3.124519058	D3	124519058	2,000	2,000	0.127	0.106	-0.062	-0.054	0.871	0.508
chrA1.257652083	A1	257652083	2,000	2,000	0.107	0.118	-0.049	0.519	0.870	0.390
chrUn.77050150	UN	77050150	1,999	2,000	0.087	0.133	-0.036	0.125	0.869	0.526
chrD4.110867181	D4	110867181	2,000	2,000	0.112	0.117	-0.051	-0.061	0.869	0.481
chrC1.33827163	C1	33827163	1,967	2,000	0.045	0.158	-0.012	0.224	0.868	0.484
chrF1.10508463	F1	10508463	2,000	2,000	0.150	0.097	-0.077	0.419	0.867	0.495
chrD3.121660315	D3	121660315	2,000	2,000	0.133	0.107	0.293	-0.055	0.867	0.500
chrA2.74350148	A2	74350148	1,993	2,000	0.066	0.146	-0.024	-0.081	0.867	0.445
chrF1.1700092	F1	1700092	2,000	2,000	0.172	0.086	-0.093	0.481	0.866	0.518
chrD2.98140067	D2	98140067	1,960	2,000	0.044	0.165	-0.012	0.377	0.865	0.420
chrB1.95582849	B1	95582849	1,795	2,000	0.022	0.175	mono	0.159	0.864	0.465
chrA3.17994107	A3	17994107	1,999	2,000	0.087	0.138	-0.036	-0.075	0.864	0.413
chrB2.94207958	B2	94207958	1,999	2,000	0.087	0.140	-0.036	-0.076	0.863	0.526
chrA1.271452674	A1	271452674	1,999	2,000	0.087	0.140	-0.036	0.257	0.863	0.447
chrC1.50317920	C1	50317920	1,960	2,000	0.044	0.165	-0.012	0.043	0.862	0.503
chrA2.143540215	A2	143540215	2,000	2,000	0.165	0.093	0.186	-0.045	0.861	0.391
chrE2.61902026	E2	61902026	1,993	2,000	0.066	0.156	-0.024	0.536	0.860	0.500
chrA2.123045290	A2	123045290	2,000	2,000	0.146	0.107	-0.075	-0.055	0.860	0.442
chrC1.105529441	C1	105529441	2,000	2,000	0.165	0.096	-0.089	0.186	0.860	0.641
chrA1.245760324	A1	245760324	1,993	2,000	0.066	0.156	-0.024	-0.088	0.859	0.466
chrB1.178805063	B1	178805063	1,993	2,000	0.066	0.156	0.661	0.204	0.859	0.397
chrB3.70368504	B3	70368504	2,000	2,000	0.112	0.130	-0.051	0.110	0.858	0.452
chrA2.171627840	A2	171627840	1,999	2,000	0.087	0.146	0.485	0.075	0.857	0.521
chrUn15.2682639	UN15	2682639	1,999	2,000	0.087	0.146	-0.036	0.384	0.857	0.462
chrA2.132974752	A2	132974752	1,967	2,000	0.045	0.172	-0.012	-0.099	0.856	0.461
chrE1.132017960	E1	132017960	1,999	2,000	0.087	0.148	-0.036	0.384	0.856	0.466
chrA3.120458264	A3	120458264	2,000	2,000	0.107	0.137	-0.049	0.258	0.855	0.436
chrA1.210108596	A1	210108596	2,000	2,000	0.206	0.077	0.106	0.261	0.855	0.466
chrF2.42999512	F2	42999512	1,814	2,000	0.023	0.185	mono	0.014	0.855	0.526
chrUn.45668217	UN	45668217	1,999	2,000	0.087	0.150	-0.036	0.383	0.855	0.472

chrB3.150053764	B3	150053764	2,000	2,000	0.127	0.126	-0.062	-0.067	0.854	0.390
chrB3.51419880	B3	51419880	2,000	2,000	0.127	0.127	-0.062	0.111	0.853	0.486
chrA2.222548225	A2	222548225	2,000	2,000	0.201	0.080	0.109	0.260	0.853	0.483
chrB3.127289249	B3	127289249	1,993	2,000	0.066	0.164	-0.024	0.181	0.853	0.507
chrD3.78037429	D3	78037429	1,967	2,000	0.045	0.178	-0.012	-0.103	0.853	0.460
chrB1.118680910	B1	118680910	1,995	2,000	0.067	0.164	-0.024	0.044	0.852	0.539
chrB1.100367105	B1	100367105	1,795	2,000	0.022	0.191	mono	0.122	0.851	0.392
chrUn13.12209356	UN13	12209356	2,000	2,000	0.127	0.133	-0.062	-0.071	0.849	0.437
chrB3.39630826	B3	39630826	1,795	2,000	0.022	0.200	mono	0.005	0.847	0.454
chrB3.86869224	B3	86869224	1,999	2,000	0.087	0.159	-0.036	0.056	0.847	0.466
chrB4.68540749	B4	68540749	1,960	2,000	0.044	0.184	-0.012	0.016	0.847	0.458
chrA1.180113591	A1	180113591	2,000	2,000	0.206	0.090	0.332	0.219	0.844	0.432
chrUn.41472022	UN	41472022	2,000	2,000	0.165	0.115	0.186	-0.059	0.843	0.533
chrA2.62766160	A2	62766160	1,999	2,000	0.087	0.165	-0.036	0.181	0.842	0.479
chrC2.132017434	C2	132017434	2,000	2,000	0.165	0.118	0.186	-0.062	0.841	0.421
chrC2.262161	C2	262161	2,000	2,000	0.184	0.107	-0.103	0.369	0.841	0.524
chrA2.205830088	A2	205830088	2,000	2,000	0.107	0.154	0.374	0.059	0.840	0.476
chrF1.75274841	F1	75274841	1,999	2,000	0.089	0.165	-0.037	0.318	0.840	0.491
chrD4.55661288	D4	55661288	1,795	2,000	0.022	0.206	mono	0.326	0.840	0.463
chrC1.201204339	C1	201204339	2,000	2,000	0.110	0.156	-0.050	0.058	0.838	0.519
chrB1.71995132	B1	71995132	2,000	2,000	0.268	0.056	0.162	-0.024	0.837	0.499
chrD1.91944678	D1	91944678	1,993	2,000	0.066	0.185	-0.024	0.014	0.835	0.526
chrA3.101023230	A3	101023230	2,000	2,000	0.228	0.088	0.071	-0.042	0.834	0.524
chrB4.80902801	B4	80902801	1,960	2,000	0.044	0.201	-0.012	0.234	0.834	0.415
chrB4.109683320	B4	109683320	2,000	2,000	0.169	0.126	-0.091	-0.067	0.833	0.520
chrA2.66162743	A2	66162743	1,795	2,000	0.022	0.214	mono	0.192	0.833	0.365
chrB1.193634290	B1	193634290	1,999	2,000	0.089	0.175	-0.037	0.289	0.833	0.002
chrUn12.6652777	UN12	6652777	2,000	2,000	0.187	0.116	0.143	-0.060	0.832	0.473
chrC1.52927181	C1	52927181	2,000	2,000	0.172	0.126	-0.093	0.290	0.832	0.360
chrA1.50468195	A1	50468195	1,996	2,000	0.110	0.164	0.373	0.044	0.831	0.446
chrA1.275002460	A1	275002460	2,000	1,998	0.298	0.044	0.095	-0.017	0.831	0.470
chrB3.26272231	B3	26272231	1,000	2,000	0.000	0.229	mono	0.354	0.830	0.444
chrA3.164199478	A3	164199478	1,999	2,000	0.089	0.178	-0.037	0.547	0.830	0.457
chrB2.8692400	B2	8692400	2,000	2,000	0.236	0.087	0.047	-0.042	0.829	0.003
chrUn15.2234009	UN15	2234009	2,000	2,000	0.150	0.145	-0.077	0.230	0.827	0.461
chrD3.10653740	D3	10653740	2,000	2,000	0.130	0.158	0.294	0.349	0.826	0.474
chrUn7.2511376	UN7	2511376	2,000	2,000	0.127	0.161	0.295	0.201	0.825	0.566
chrA3.97197194	A3	97197194	2,000	2,000	0.146	0.150	-0.075	0.073	0.824	0.452
chrUn5.4430690	UN5	4430690	1,795	2,000	0.022	0.225	mono	-0.038	0.824	0.475
chrUn8.1386135	UN8	1386135	2,000	2,000	0.107	0.176	0.374	0.288	0.822	0.419
chrA2.179540281	A2	179540281	2,000	2,000	0.223	0.106	0.074	-0.054	0.821	0.439
chrC2.105768020	C2	105768020	2,000	2,000	0.236	0.097	0.239	-0.048	0.820	0.481
chrUn13.14266848	UN13	14266848	1,993	2,000	0.066	0.203	0.661	0.215	0.820	0.607
chrB4.30520254	B4	30520254	2,000	2,000	0.223	0.107	0.074	0.369	0.819	0.471
chrB1.158491011	B1	158491011	2,000	2,000	0.219	0.110	-0.132	-0.056	0.819	0.472
chrA1.122465543	A1	122465543	1,967	2,000	0.045	0.217	-0.012	-0.032	0.818	0.547
chrC2.34543191	C2	34543191	2,000	2,000	0.146	0.158	-0.075	-0.089	0.818	0.389
chrD4.78705483	D4	78705483	1,999	2,000	0.087	0.195	-0.036	0.355	0.817	0.497
chrC1.15780754	C1	15780754	2,000	2,000	0.184	0.135	0.391	0.093	0.817	0.489
chrB4.95831693	B4	95831693	2,000	2,000	0.165	0.148	-0.089	0.229	0.816	0.502
chrB1.81168124	B1	81168124	1,795	2,000	0.022	0.233	mono	0.427	0.815	0.500
chrE2.3933982	E2	3933982	2,000	2,000	0.201	0.126	-0.117	0.112	0.815	0.545
chrA1.194996231	A1	194996231	1,993	2,000	0.066	0.210	-0.024	0.087	0.814	0.472
chrE2.65383456	E2	65383456	2,000	2,000	0.187	0.137	-0.105	0.092	0.813	0.474
chrB2.112444547	B2	112444547	2,000	2,000	0.127	0.176	0.295	0.288	0.812	0.483

chrD1.4140039	D1	4140039	1,999	2,000	0.087	0.201	-0.036	0.328	0.811	0.582
chrB4.43191231	B4	43191231	1,993	2,000	0.066	0.217	-0.024	0.279	0.808	0.006
chrC1.46354234	C1	46354234	2,000	2,000	0.219	0.126	0.076	0.468	0.805	0.412
chrC1.95227611	C1	95227611	1,999	2,000	0.087	0.208	-0.036	0.196	0.805	0.623
chrUn26.10046275	UN26	10046275	1,960	2,000	0.044	0.236	-0.012	0.041	0.803	0.526
chrC1.63091997	C1	63091997	2,000	2,000	0.325	0.055	0.173	0.388	0.802	0.417
chrA2.116555292	A2	116555292	2,000	2,000	0.283	0.086	0.127	0.220	0.801	0.478
chrUn30.18116	UN30	18116	2,000	2,000	0.127	0.191	0.295	0.357	0.799	0.423
chrUn13.13324872	UN13	13324872	1,999	2,000	0.087	0.217	-0.036	0.176	0.797	0.477

Individuals' assignment and admixture analyses

The admixture analysis computed on simulated genotypes (generated by HYBRIDLAB), running the "admixture" with both the allele frequencies models, and using $K = 2$ in STRUCTURE, was able to efficiently recognise 100% of the parental individual at a threshold of qi = 0.80 (see Supplementary Table S1) and correctly identify all the F1- F2 hybrids and the backcross genotypes of I-II generation. However, all the backcross genotypes of III- IV generation, showed a qi > 0.80 to one single cluster and could not be distinguished from parental individuals. All hybrids detected by simulations revealed a very wide 90% CI, ranging between 0,003 and 0.996. Accordingly, we choose the threshold of qi > 0.80 for the following Bayesian analyses, performing with and without prior information for parental samples and reporting results computed without prior non-genetic information and allele frequencies model ("*F*"), since they yielded globally identical results (data not shown).

At the best assignment cluster $K = 2$ (showing the highest values of ΔK and ΔF_{ST}, see Supplementary Figure S1) 44 domestic cats ($Q_I = 0.989$ and individual q_d ranging from 0.933 to 1.000) and 89 European wildcats ($Q_{II} = 0.965$ and $q_W = 0.918$ - 1.000) were clearly assigned to their expected cluster with high NEWHYBRIDS' membership probabilities ($qi > 0.99$, see Table 4 and Fig 2). 10 previously identified admixed European wildcats (Pierpaoli *et al.* 2003) were confirmed as admixed, showing qi from 0.470 to 0.655 (90% CI = 0.420 - 0.702) to the wildcat cluster (see Table 4). Moreover, they were mostly assigned to their known hybrid category: 6 individuals from the southern Apennines ($n = 1$) and the northern-eastern Alps ($n = 5$) as F1, 4 captive-breed individuals as BxFSI (ID 57), BxFSI/F2 (ID 60, ID 62) and F2 (ID 63), see Table 4 and Figure 3. 6 of the 19 putative admixed European wildcats (Mattucci *et al.* 2013; Oliveira *et al.* in prep) analyzed in this study showed individual q_W ranging from 0.525 to 0.751: two from Italy (central Apennines and Maremma), one from Luxembourg, one from Bosnia & Herzegovina and two from Portugal. Moreover, they were clearly classified as BxFSI (ID 269, ID 1323, ID 586, ID 1056) and F2 (ID 700, ID 712) with high posterior probabilities ($qi > 0.99$). However, the remain

13 putative admixed European wildcats partially showed q_i values to the domestic cluster ranging from 0.928 to 0.978 (CI 0.898 - 0.993) for three individuals from Germany (southern-western, *n* = 1), Portugal (*n* = 1) and Spain (*n* = 1); and mainly were assigned to the wildcat cluster with an individual q_i value ranging between 0.800 - 0.976 (CI 0.759 - 0.991) for ten individual from Italy (central Apennines and Maremma, *n* = 4 and north-eastern Alps, *n* = 1; southern Apennines, *n* = 4) and Germany (southern-western, *n* = 1), see Table 4. NEWHYBRIDS' clustering proved to be highly efficient to correctly allocated 10 of the 13 previously misclassified putative admixed cats to their parental category (*qi* > 0.99); with the only exception of ID 1518 (from Italian central Apennines and Maremma) and ID 1629 (from the southern-western part of Germany) classified as BxFSI (*qi* > 0.99) and ID 1248 (From Italian southern Apennines) partially classified as Pure wildcat (FSI) and BxFSI (0.695 < *qi* < 0.979), see Fig 3.

Table 4. Individual membership proportions (q_i) and inferred ancestry of 16 putatively admixed cats according to the Bayesian analyses performed in STRUCTURE and NEWHYBRIDS. In STRUCTURE, individuals were assigned into two cluster corresponding to the domestic (Q_d) and wild (Q_w) groups excluding prior population information. In NEWHYBRIDS, individuals were assigned into six different genotypes classes: pure domestic and wild cat, F1, F2, Bx I (backcross with domestic cat) and Bx II (backcross with wildcat). STRUCTURE 90% credibility intervals (CI) are shown in brackets.

Dataset	ID	Populations	STRUCTURE		NEWHYBRIDS					
			Q_d	Q_w	D	W	F1	F2	Bx I	Bx II
Domestic cats			0.957	0.043	1.000					
Felis s. catus n = 44			(0.933-0.977)	(0.023-0.066)						
European wildcats			0.015	0.985		1.000				
Felis s, silvestris n = 89			(0.035-0.083)	(0.916-0.964)						
Known admixed cats	57	Captivity (Italy)	0.345	0.655						1.000
Felis silvestris x catus			(0.298-0.394)	(0.606-0.702)						
n = 10	60	Captivity (Italy)	0.381	0.619				0.341		0.659
			(0.334-0.430)	(0.570-0.666)						
	62	Captivity (Italy)	0.373	0.627				0.220		0.780
			(0.325-0.423)	(0.577-0.675)						
	63	Captivity (Italy)	0.417	0.583				0.989		0.011
			(0.368-0.467)	(0.533-0.632)						
	228	Italy: Southern Apennines	0.501	0.499			1.000			
			(0.447-0.554)	(0.446-0.553)						
	992	Italy: Northern-eastern Alps	0.496	0.504			1.000			
			(0.445-0.547)	(0.453-0.555)						
	993	Italy: Northern-eastern Alps	0.529	0.471			1.000			
			(0.479-0.579)	(0.421-0.521)						
	994	Italy: Northern-eastern Alps	0.515	0.485			1.000			
			(0.464-0.566)	(0.434-0.536)						

	995	Italy: Northern-eastern Alps	0.530 (0.479-0.580)	0.470 (0.420-0.521)			1.000		
	996	Italy: Northern-eastern Alps	0.507 (0.457-0.558)	0.493 (0.442-0.543)			1.000		
Putative admixed cats *Felis silvestris x catus* *n* = 19	67*	Italy: Central Apennines and Maremma	0.144 (0.111-0.181)	0.856 (0.819-0.889)		0.994			0.006
	123*	Italy: Central Apennines and Maremma	0.110 (0.080-0.143)	0.890 (0.857-0.920)		1.000			
	627*	Italy: Central Apennines and Maremma	0.126 (0.092-0.164)	0.874 (0.836-0.908)		1.000			
	677*	Italy: Central Apennines and Maremma	0.147 (0.112-0.184)	0.853 (0.816-0.888)		1.000			
	1269	Italy: Central Apennines and Maremma	0.249 (0.206-0.293)	0.751 0.707-0.794)					1.000
	1323	Italy: North-eastern Alps	0.258 (0.215-0.302)	0.742 (0.698-0.785)					1.000
	1518*	Italy: Central Apennines and Maremma	0.180 (0.141-0.221)	0.820 (0.779-0.859)		0.021			0.979
	77*	Italy: Southern Apennines	0.024 (0.009-0.043)	0.976 (0.957-0.991)		1.000			
	1248*	Italy: Southern Apennines	0.164 (0.127-0.203)	0.836 (0.797-0.873)		0.695			0.305
	1250*	Italy: Southern Apennines	0.129 (0.095-0.166)	0.871 (0.834-0.905)		1.000			
	1252*	Italy: Southern Apennines	0.116 (0.085-0.150)	0.884 (0.850-0.915)		1.000			
	586	Luxembourg	0.293 (0.247-0.341)	0.707 (0.659-0.753)					1.000
	1056	Bosnia & Herzegovina	0.267 (0.223-0.313)	0.733 (0.687-0.777)					1.000
	421*	Germany: Southern-western	0.978 (0.959-0.993)	0.022 (0.007-0.041)	1.000				
	629*	Germany: Southern-western	0.200 (0.161-0.241)	0.800 (0.759-0.839)					1.000
	700	Portugal	0.475 (0.425-0.524)	0.525 (0.476-0.575)				1.000	
	706*	Portugal	0.971 (0.951-0.987)	0.029 (0.013-0.049)	1.000				
	712	Portugal	0.460 (0.411-0.510)	0.540 (0.490-0.589)				1.000	
	717*	Spain	0.928 (0.898-0.955)	0.072 (0.045-0.102)	1.000				

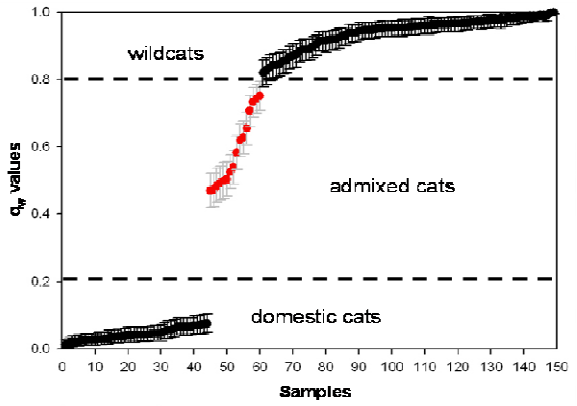


Figure 2. Plot of individual *q*w values (to a wildcat cluster) and their 90% credibility intervals CI. The *q*i thresholds = 0.20 - 0.80 (interrupted lines) define the admixed genotypes. Pure wild and domestic cats are in black, admixed individuals are in red .

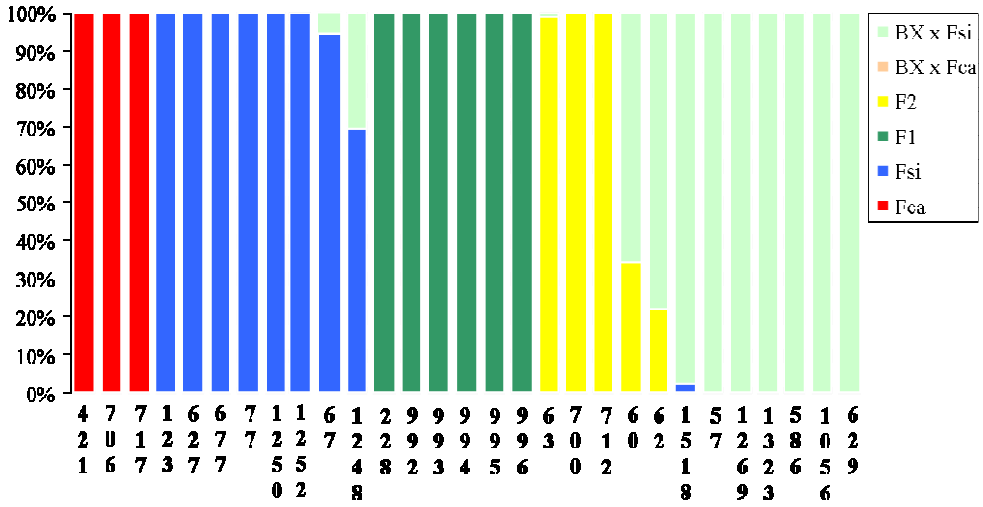


Figure 3. NEWHYBRIDS' assignment of 10 known hybrids and 19 'putative' admixed individuals to the different hybrid categories. Each individual is represented by a single vertical bar colored according to the proportion of their genome descending from each of the inferred clusters or hybrid class.

mtDNA sequences

We obtained reliable mtDNA ND5 sequences in 134/150 samples (89%). The total haplotype diversity was h = 0.87 ± 0.013 (Table 5). Eighteen different haplotypes were detected. Seven unique haplotypes were identified. This last result could be derived from the small sample size combined with its high geographic representativeness.

Table 5. Genetic variability based on ND5 region of mtDNA observed within the main genetic haplogroups (Haplogroup D, Haplogroup W)

	N	Number of haplotypes	Nucleotide diversity π (±SD percentage)	Haplotype diversity (h ± SD)
Total	134	18	0.885 ± 0.047	0.870 ± 0.013
Haplogroup D	81	10	0.283 ± 0.180	0.815 ± 0.021
Haplogroup W	53	8	0.182 ± 0.130	0.595 ± 0.060

A median joining network provided to describe the relationships among haplotypes (Fig 4). Two major groups were clearly detected (here called W and D), separated by the seven known mutations. Haplogroup W counted 8 haplotypes (H6, H7, H8, H10, H11, H14, H16, H18), for a total 53 individuals (39%). Haplotypes H7 and H6 are the most frequents and represent respectively 58,5% and 26,4% of this haplogroup (Fig 5). The intra-group gene diversity was estimated to be 0.595 ± 0.060. Haplogroup D counted 10 different haplotype (H1, H2, H3, H4, H5, H9, H12, H13, H15, H17) for a total of 82 individuals. Of these 66 (80%) was represented by 4 haplotype (H1,H2,H5,H9). The intra-group gene diversity was estimated to be 0.815 ± 0.021. AMOVA analysis (Table 6) showed that 84.55% of total variation was explained between the two haplogroups (F_{ST} = 0.845, p < 0.001) while 15.45% within them.

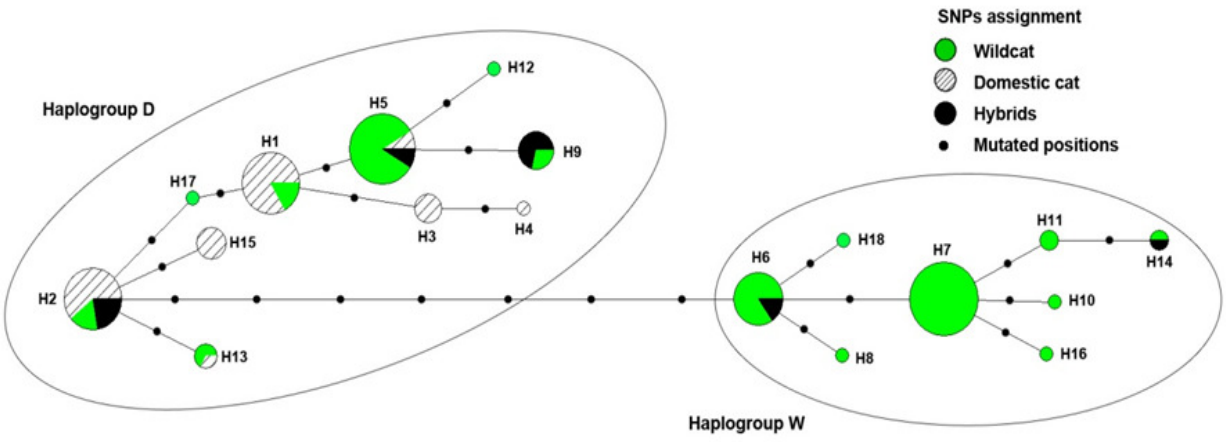


Figure 4. NETWORK diagram of ND5 cat haplotypes. Each haplotype is named with a code and is represented by a circle whose size is scaled according to its abundance in the cat sample analyzed. Different colors refer to highly informative nuclear SNPs assignments of genotypes (see Table 4). Mutated positions are indicated on braches.

Table 6. Analyses of molecular variance based on ND5 region of mtDNA data from the main haplogroups (Haplogroup D. Haplogroup W)

Source of variation	d.f	Variance components	Percentage of variance	P
Among haplogroups	1	4.472	84.55	<0.001
Within haplogroups	132	0.817	15.45	<0.001
Fixation index Fst:	0.845			

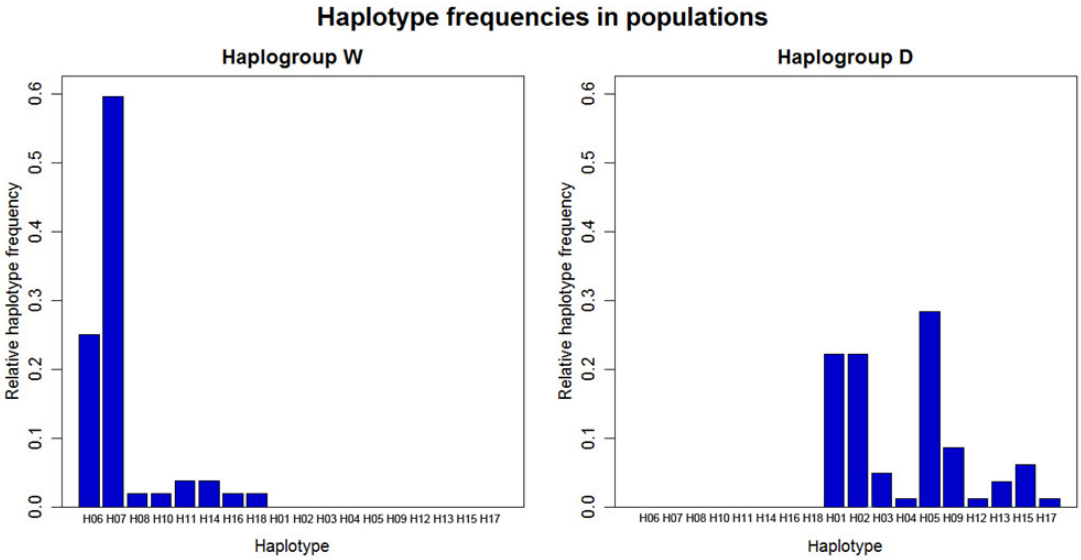


Figure 5. Frequency of ND5 cat haplotypes for both the haplogroups W and D.

Crossing this data with the ones from SNPs analyses it came up that of the 53 individuals of haplogroup W 50 (94%) are classified as pure wildcat and 3 (6%) as wild-domestic hybrids according to the STRUCTURE results (see SNPs section). No domestic cat clustered within the haplogroup W. Among the 82 samples of haplogroup D grouped all the 42 (52%) samples identified as domestic cat by the SNPs analyses, together with 11 (13%) wild-domestic hybrids. Furthermore 32 (35%) individuals classified as pure wildcat by SNPs analyses clustered within haplogroup D. Of these 19 were found within the same haplogroup H5. This particular haplotype represented by 23 samples contained only 2 domestic and 2 hybrid individuals.

In order to verify these results we performed the same network analysis on the Control Region (Fig 6). Only 126 samples (84%) yielded reliable sequences. The analyses found 50 haplotypes with a total haplotype diversity was 0,96 ± 0,007. In the Figure 6 are displayed the three major haplogroups (1,2 and 3) among which is explained the 75.2% of total variance (F_{ST} = 0.75, p

< 0,01). In haplogroup 1 clustered the haplotypes (20) represented by the same wildcat and hybrids individuals that we found in haplogroup W. The haplogroup 2 appeared as a complex cluster of haplotypes (17) represented by wild, domestic and hybrid individuals, with a high value of within-group gene diversity (0,92 ± 0,014) . In particular the wild individuals of this haplogroup are the same that presented the domestic mutations on ND5 region and clustered in haplogroup D. From haplogroup 2 derived haplogroup 3 represented by 18 individuals (13 haplotypes). Of these, 15 samples resulted belonging to domestic cats and 3 to wildcats.

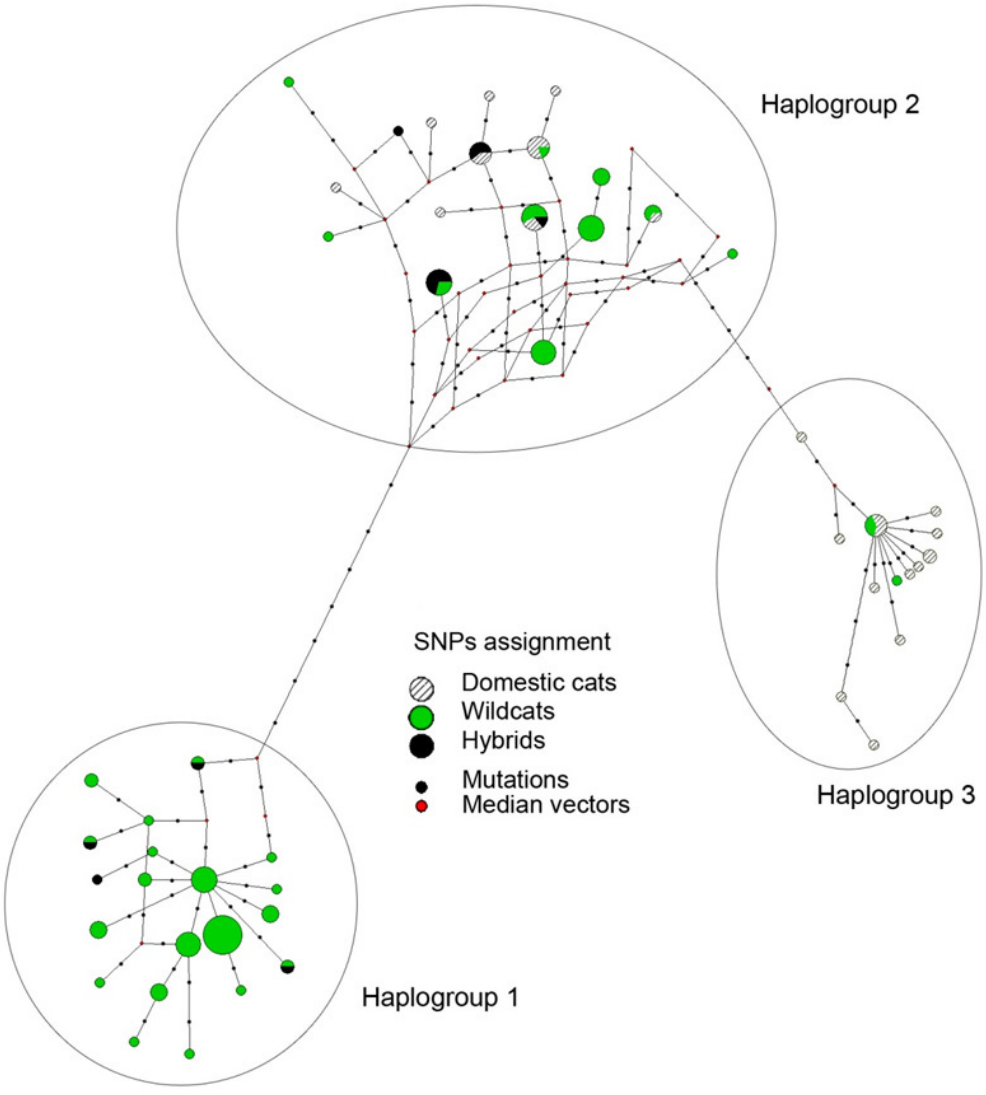


Figure 6. NETWORK diagram of Control Region (CR) cat haplotypes. Each is represented by a circle whose size is scaled according to its abundance in the cat sample analyzed. Different colors refer to highly informative nuclear SNPs assignments of genotypes (see Table 4). Mutated positions (black nodes) and theoretical intermediary haplotypes predicted by network analysis (red nodes) are indicated on braches.

Table 7 show the degrees of differentiation between haplotype 1 and haplotypes 2 and 3 are quite greater compared to the one between haplotypes 2 and 3.

Table 7. Population pairwise Fst matrix based on Control Region of mtDNA data from the main haplogroups (Haplogroup 1=1. Haplogroup 2=2. Haplogroup 3=3)

	1	2	3
1	0.000		
2	0.711	0.000	
3	0.900	0.598	0.000

Y chromosome genotyping

Concerning the Chr Y markers, 89 samples of male individuals yielded reliable genotypes for both the markers. The polymorphic site was found at absolute position 1956 of the SRY gene with two alleles (A and G). The microsatellite presented two allele respectively of 271 and 273 bp. In all amplified samples the two markers showed a coherent genotype to one another. So allele A of SRY always came with the allele 271 (henceforth "haplotype WY") of STR and consequently allele G came with allele 273 (henceforth "haplotype DY". Haplotype WY was found in 52 samples (58%) while haplotype DY in 37 (42%). Crossing these results with the SNPs assignation (Table 8) we found that all 42 samples assigned as pure wildcat except 1 showed the haplotype WY; of 29 samples assigned as domestic cat 27 were characterized by haplotype DY and 2 by haplotype WY; the 18 hybrid individuals presented both the haplotype equally.

Table 8. Chromosome Y haplotypes based on SMCY-7 STR and SRY gene SNP

		Haplotype DY	Haplotype WY
SNPs assignment	FCA	27	2
	FSI	1	41
	HYB	9	9
	Tot	37	52

DISCUSSION

In this study, using a sample of European cats (domestic and wild) collected by taking in consideration the home range distribution of the wildcat subspecies and the genetic-geographic

substructure of wildcats populations in Europe (Mattucci *et al.* in prep), we describe a multi-locus protocol for admixture detection, with the main goal to improve the reliability of hybrids identification and deeply investigate ancient introgressive events. Thus, 151 most informative SNPs (with average F_{ST} varying between 0.797 and 0.982 in European subpopulations) were selected from the Illumina Infinium iSelect 63K Cat DNA Array, and combined with two pairs of uniparental markers, showing distinctive diagnostic mutations for wild and domestic cats. In particular, the ND5 region and part of the control region (Driscoll *et al*. 2007; Freeman *et al.* 2001; Tiedemann *et al.* 1996) for maternal lineage and Y-linked markers (Pecon-Slattery *et al.* 2004; King *et al.* 2007; Luo *et al.* 2007; Nussberger *et al.* 2013) for paternal lineage, helped to provide a complete overview of hybridization in European wildcats populations.

To assess the diagnostic power of markers, it appeared crucial choose the right reference, avoiding any hybrids or introgressed individuals, and ensuring the representativeness of the genetic diversity of the parental populations (Mattucci *et al.* in prep). Although this last factor might influence the correct allocation of backcrossed individuals through Bayesian algorithms (Falush *et al*. 2003), markers power has been assessed on a reduced sample set (n = 187), compared to recent STRs study across Europe (n = 1114; Mattucci *et al.* in prep), that however provided a representation of the 5 macro populations of wildcat detected with 38 STRs in Mattucci *et al.* (in prep).

SNPs were ascertained from a 63K cat array, which has been designed on ~9.55 million SNPs from the three combined genome sequencing efforts and subsequently has been tested on 12 different breeds, 10 wildcats, 10 western and 10 eastern random bred cats, in addiction to five trios, the Abyssinian (Cinnamon) and the 6 cats from the Hill's SNP discovery project. Thus, the selected SNPs would be expected to be generally diverse amongst random bred cats compared to wildcats. But genetic diversity, including Ar and H_E, proved to be significantly high both in domestic and wildcat group, showing similar mean values, most likely as a result of the heterogeneous sampling planned by taking into account the genetic-geographic repartition of wildcats in Europe. Generally, genetic variability is expected to be lower in domesticated forms relatively to their wild counterparts, due to the domestication bottleneck caused by the low number of founder individuals and the restricted gene flow imposed by human constrains (Doebley *et al.* 2006). However, the progenitor of the domestic cat is considered the *Felis s. lybica* subspecies (Driscoll *et al*. 2007; Lipinski *et al*. 2008), thus a direct comparison between the proposed wild progenitor and domesticate cannot be evaluated in this study.

The selected panel of 151 most informative SNPs proved to be efficient in identifying 100% of parental genotypes and first-second generation hybrids with a posterior probability of >0.80, but none of third-fourth generation hybrids. Increasing the posterior probability to >0.95, according to Nussberger *et al.* (2013), whose informative SNPs panel identified over 86% of all hybrids, we correctly identified 100% of all simulated individuals (including both parental and different generation hybrids) but only 90% of empiric domestic cats and 66% of empiric wildcats. Thus, we chose an assignment threshold of 0.80 less stringent but more discriminant than previous studies (Nussberger *et al.* 2013; Oliveira *et al.* in prep), allowing the correct identification of all sampled parental cats and all firs-second generation introgressed cats.

To verify the accuracy of the assignment tests, we further included ten known hybrids (Pierpaoli *et al.* 2003) and 31 putative admixed individuals (reduced to 19 after preliminary pruning of dataset), previously identified with a panel of 35-38 microsatellites loci, that showed membership probabilities lower than the established assignment threshold and/or wide 90% Credibility Intervals (Mattucci *et al.* 2013; Mattucci *et al.* in prep). By comparing Bayesian admixture outcomes for known and putative hybrids, using highly informative SNPs and STRs (see Supplementary Table S2), it has been possible evaluate the performance of both markers in hybrid's detection. 10 known hybrids were clearly identified as admixed by both markers with *qi* to the wildcat cluster ranging from 0.470 - 0.655 (90% CI = 0.420 - 0.702) for SNPs to 0.321 - 0.546 (90% CI = 0.185 - 0.697) for STRs (see Supplementary Table S2). Although the evidence of comparable *qi* assignment, the 90% CI proved to be wider for STRs than for SNPs.

On the contrary, differences in markers performance were shown in putative admixed identification. The admixture ancestry of the 19 putative hybrids, has been, in fact, confirmed by both markers, in 6 individuals, namely: 2 individuals from Italy (respectively ID 1269 from central and Maremma Apennines; ID 1323 from north-eastern Alps), 2 from Portugal (ID 700; ID 712), one from Bosnia & Herzegovina (ID 1056), and one individual from Luxembourg (ID 586), whose previously assignment with STRs (Mattucci *et al.* in prep) showed the lower value of Credibility Intervals to the wildcats cluster of 0.694 (namely under the assignment threshold). However, the range of their membership probabilities to the wildcat cluster and related CI resulted to be extremely narrow with SNPs when compared to STRs values (respectively 0.525 - 0.742 with 90% CI = 0.476 - 0.794, and 0.357 - 0.846 with 90% CI = 0.187 - 0.997). 3 further individuals might be considered as admixed because, even if their assignment probabilities with SNPs are ranging from 0.800 to 0.836, the lower value of Credibility Intervals resulted to be under the assignment threshold (ranging from 0.759 to 0.797). On the contrary, 10 individuals previously misclassified as putative hybrids based on STRs assignment, using the panel of most informative SNPs were

identified as 'pure wildcat' from central-southern Apennines and Maremma (ID 67, ID 627, ID 677, ID 77, ID 123, ID 1250, ID 1252), and as 'pure domestic cats' from southern-western Germany (ID 421), Portugal (ID 706), and Spain (ID 717), see Supplementary Table S2.

The uncertain assignment of admixture individuals revealed with 38 STRs (Mattucci *et al.* in prep) has been overcame with SNPs, that proved to be more reliable than STRs, providing narrow IC ranges and individual membership probabilities clearly partitioned between wild and domestic clusters, even if the reference cats and the putative hybrids belonging to different genetic-geographic populations (Mattucci *et al.* in prep). Most likely, the high allelic richness in combination with homoplasy might have reduced the microsatellites diagnostic power for hybrid recognition, since there are more possibilities of allele sharing between two hybridizing taxa. Thus, microsatellites appeared to be more suited to recognize genetic population structure (Guichoux *et al.* 2011) than identify introgression events. Recent studies have already demonstrated the SNPs potentiality to equal or even outperform microsatellites not only for individual ancestry (Lao *et al.* 2008), but also for population assignment (e.g. Seddon *et al.* 2005; Narum *et al.* 2008; Smith and Seeb 2008; Coates *et al.* 2009) and pedigree studies (Santure *et al.* 2010, Hauser *et al.* 2011), and proved to have large allele frequency differences among populations (Freamo *et al.* 2011).

Moreover, contrary to Nussberger *et al.* (2013) recent outcomes, that limit the power of their SNPs panel on a regional level (specially beyond the Swiss borders) while encouraging their use in a wider sample, our set of 151 highly informative SNPs proved to be surely efficient and applicable with all genetic populations of cats disclosed in Europe (Mattucci *et al.* in prep), since different genetic-geographic reference cats has already been tested for ancestry detection analyses.

The repeatedly crossbreeding occurred in the past with both parental groups and admixed individuals (Bewick 1807; Driscoll *et al.* 2009; Driscoll *et al.* 2007; Suminski 1962) and the fertility of wild x domestic hybrids, might have diluted the proportion of domestic alleles through the generations into the wildcat gene pool, leaving traces only in non-recombining mtDNA or Y chromosome regions. Thus, the uniparental markers might allow a further detection of hybridization by crossing maternal and paternal diagnostic polymorphism between domestic and wild cats with SNPs Bayesian assignments (see Supplementary Table S3).

The chromosome Y-linked markers proved to be useful for identifying wild and domestic cat males based on distinctive polymorphism. Both mitochondrial sequenced regions (ND5 and part of the control region), clearly separated the subspecies in two well distinct haplogroups, with the exception of a few haplotypes characteristic of domestic cats that have been found in some wild individuals. The presence of these sharing haplotypes opened the way to two different hypothesizes. In one hand it might be possible that, as a consequence of introgressive events occurred in the past,

most likely between a domestic female and a wild male, a domestic signature entered in the wild mitochondrial ND5 region through these common haplotypes (namely H1, H2, H5, H7, H9, H12, H13). Over time, the wild individuals carrying the domestic haplotypes might have accumulated mutations on the mitochondrial control region until now, generating the haplogroups 2. On the other hand, it might be possible that none admixture events occurred but probably these haplotypes are sharing between the two subspecies, and thus the markers selected in this work are not reliable and diagnostic to distinguish wild from domestic cats and to investigate on introgressive episodes.

In order to verify and clarify these hypothesizes, it should be recommendable compute a complete phylogenic analysis, by combining both mitochondrial regions analyzed, extending the sample over the entire home range distribution of the subspecies in Europe and mapping evolution and distribution of each haplotype.

Both uniparental markers confirmed the admixture assignment, previously detected with nuclear SNPs and STRs, of all know hybrids. Four of them showed, in fact, a complementary wild and domestic uniparental assignment, while the remains 6, reported the domestic patrilineal mutation and the sharing domestic haplotypes, most likely resulting by ancient introgressive event. Among the 9 putative hybrids detected with high informative SNPs (see above), 6 individuals, whose admixture ancestry has been assessed with both nuclear markers (SNPs and STRs), evidenced wild haplotypes in both mitochondrial and chromosome Y, with the exception of samples lacking uniparental informations (for Y ID 1323 and ID 1056 and for both markers ID 700), and ID 1056 from Bosnia & Herzegovina that showed the shared domestic mitochondrial haplotype. The remains 3 individuals, recognized as 'putative hybrids' with SNPs because of their wide CI range, presented wild haplotype for both markers (for ID 1518) and complementary haplotypes (for ID 1248), while the individual ID 629 showed only the shared domestic mitochondrial haplotype.

For what concern the individuals re-classified as 'pure parental' with SNPs (n = 10), we found congruencies between SNPs and mitochondrial identification only in 5 cats but completely lacking the paternal information, one individual missed both uniparental assignment (ID 717), while 3 showed incongruencies between SNPs and mitochondrial identification (namely nuclear genotype wild and shared domestic mitochondrial haplotype), see Supplementary Table S3.

The integration of both uniparental and nuclear markers assignments, provided a complete insight of introgression level in wildcats populations analyzed (see Supplementary Table S3). Finally, a total of 11 hybrids were identified among the 19 'putative' hybrids of which: 8 were detected by both nuclear (STRs and SNPs) and uniparental (mtDNA and Y) markers, and 3 were identified because of the presence of nuclear genotype wild and shared domestic mitochondrial haplotype (ID 123, ID 1250, ID 1252). Because of the long history of domestic/ wildcat

introgression (Bewick 1807; Driscoll *et al.* 2009; Driscoll *et al.* 2007; Suminski 1962), individuals showing only STR and SNP data have been considered as incomplete for admixture analysis. Thus, the individual ID 700 from Portugal, clearly identified with both SNPs and STRs markers as admixed, has not been considered in the final hybrids count (n = 11), being lacking of uniparental informations, and should be further analyzed improving the markers performance.

Hence, the uniparental markers analyzed proved to be extremely efficient in subspecies identification, providing matrilineal and patrilineal signature of potentially distant admixture events, once crossed with nuclear (SNPs or STRs) assignment informations. To ensure a finely detection of cryptic hybrids in wildcats populations and assess the introgression degree of highly compromised populations of Scotland and Hungary (Beaumont *et al.* 2001), it should be needed extend the admixture analyses, by using this combined and useful multi locus panel, and increase the representativeness of cats genetic diversity in Europe, by implementing samples from each macro and sub populations detected by Mattucci *et al.* (in prep).

The new throughput technologies under development for domestic cats will soon allow the evaluation of the entire genome of *Felis silvestris* species, supporting the identification of more diagnostic loci and potentially indicating areas of the genome involved with domestication. Other phenotypic polymorphisms, such as melanism at the *Agouti* locus (*ASIP*) (Eizirik *et al.* 2003), would likely be an important diagnostic for domestic cat introgression into wildcats might be equally highly informative. Thus, combined repertoires of highly informative autosomal SNPs, X and Y-linked markers and mtDNA variants promise to involve a deep investigation of cat domestication effects on wildcats populations.

ACKNOWLEDGEMENTS

We thank Herrman M, Suchentrunk F, Liberek M, Natural Museum of Scotland, Romy P, Szolt B , Sforzi A, Ragni B, Lapini L, De Faveri A, Hupe K, Eckert I, Potocnik H, Moes M, Cagnolati G, Vercillo F, Godoy J, Malsaña M, Mejias M, Fernández JM, Robles JL, Penafiel GD, Duperón EB, Moleón M, Monterroso P, Álvares F, Rodrigues J, BTVS/ICNB: Portuguese National Tissue Bank/National Institute for Nature and Biodiversity Conservation, Lyberakis P and their collaborators for providing putative wildcat samples. We are also grateful to all anonymous veterinarians and biologists that assisted in samples collection.

4.1 REFERENCES

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613-622

Anderson EC and Thompson EA (2002) A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics*, **160(3)**, 1217-1229

Beaumont M, Barratt EM, Gottelli D, Kitchener AC, Daniels MJ, Pritchard JK, Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319-336

Bewick TA (1807) General history of the quadrupeds. Newcastle upon Tyne (UK): Bewick and Hodgeson

Brumfield R, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249-256

Campana MG, Hunt HV, Jones H, White J (2011) Corrsieve: software for summarizing and evaluating structure output. *Molecular Ecology Resources*, **11**, 349-352

Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell*, **127**,1309-1321

Driscoll C and Nowell K (2010) *Felis silvestris*. In: IUCN 2012. IUCN red list of threatened species. Version 2012.2

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ, MacDonald DW (2007) The near eastern origin of cat domestication. *Science*, **317**, 519-523

Driscoll CA, Macdonald DW, O'Brien SJ (2009) From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, **106(1)**, 9971-9978

Driscoll C, Yamaguchi N, O'Brien SJ, Macdonald DW (2011) A Suite of Genetic Markers Useful in Assessing Wildcat (*Felis silvestris ssp.*) - Domestic Cat (*Felis silvestris catus*) Admixture. *Journal of Heredity*, **102**, 87-90

Eckert I, Suchentrunk F, Markov G, Hartl GB (2010) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and

mitochondrial DNA sequences. *Mammalian Biology - zeitschrift fur saugetierkunde*, **75(2)**, 160-174

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ (2003) Molecular genetics and evolution of melanism in the cat family. *Current Biology*, **13**, 448-53

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 611-2620

Excoffier L and Lischer HEL (2010) Arlequin suite Ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular ecology notes*, **7**, 574-578

Freeman G, Shepperd S, Robinson I, Ehrich K, Richards S (2001) Continuity of Care: Report of a scoping exercise for the NCCSDO, London: National Coordinating Centre for Service Delivery and Organisation

French D, Corbett LK, Easterbee N (1988) Morphological discriminants of Scottish wildcats (*Felis silvestris*), domestic cats (*F. catus*) and their hybrids. *Journal of Zoology, London*, **161**, 75-123

Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. *Molecular Ecology Resources*, **10**, 915-934

Goudet J (2002) Fstat 2.9.3.2. URL: http://www2.unil.ch/popgen/softwares/fstat.htm

Goudet J (2005a) HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes,* **5(1)**, 184-186

Grabe H and Worel G (2001) Die Wildkatze. Zurück auf leisen Pfoten, Pp. 14-27. Buch and Kunstverlag Oberpfalz, Amberg

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C et al. (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources*, **11**, 591–611

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98

Hertwig ST, Schweizer M, Stepanow S, Jungnickel A, Böhle UR, Fischer MS (2009) Regionally high rates of hybridization and introgression in German wildcat populations (*Felis silvestris*, Carnivora, Felidae). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 283-297

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9(5)**, 1322-1332

Hutchings JA and Fraser DJ (2008) The nature of fisheries and farming induced evolution. *Molecular Ecology,* **17**, 294-313

Kendall MG (1970) *Rank correlation methods* (4th ed.). London: Griffin

King TE, Parkin EJ, Swinfield G, Cruciani F, Scozzari R, Rosa A, Lim SK, Xue Y, TylerSmith C, J obling MA (2007) Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy. *European Journal of Human Genetics*, **15**, 288-293

Kitchener AC, Yamaguchi N, Ward JM, Macdonald DW (2005) A diagnosis for the Scottish wildcat (*Felis silvestris*): a tool for conservation action for a critically-endangered felid. *Animal Conservation*, **8**, 223-237

Lao O, Lu TT, Nothnagel M, Junge O, Freitag-Wolf S, Caliebe A, Balascakova M et al. (2008) Correlation between Genetic and Geographic Structure in Europe. Current Biology, **18(16)**, 1241-1248

Lecis R, Pierpaoli M, Birò ZS, Szemethy L, Ragni B, Vercillo F, Randi E (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119-131

Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR et al. (2008) The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics*, **91**, 12-21

Luo SJ, Johnson WE, David VA, Menotti-Raymond M, Stanyon R, Cai QX, Beck T, Yuhki N, Pecon-Slattery J, Smith JLD, et al. (2007) Development of Y chromosome intraspecific polymorphic markers in the Felidae. *Journal of Heredity*, **98**, 400-413

Lynch M and O' Hely M (2001) Captive breeding and the genetic fitness of natural populations. *Conservation Genetics*, **2**, 363-378

Mattucci F, Oliveira R, Bizzarri L, Vercillo F, Anile S, Ragni B, Lapini L, Sforzi A, Alves PC, Lyons LA, Randi E (2013) Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecology and Evolution,* **3(8)**, 2443-2458

McGinnity P, Prodhl P, Ferguson A, Hynes R, Maoilidigh N, Baker N, Cotter D, O'Hea B, Cooke D, Rogan G, Taggart J, Cross T (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 2443-2450

McOrist S, Kitchener AC (1994) Current threats to the European wildcat, *Felis silvestris*, in Scotland. *Ambio*, **23**, 243-245

Mills L, Citta J, Lair K, Schwartz M, Tallmon D (2000) Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. *Ecological Applications*, **10**, 283-294

Morin PA, Luikart G, Wayne RK, SNP-Workshop Group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**, 208-216

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **76**, 379-390

Nielsen EE, Bach LA, Kotlicki P (2006) HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Resources*, **6(4)**, 971-973

Nussberger B, Greminger MP, Grossen P, Keller LF, Wandeler P (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources,* **13**, 447-460

O'Brien SJ, Devillard S, Say L, Vanthomme H, Léger F, Ruette 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(9)**, 2351-2360

Oliveira R, Godinho R, Randi E, Ferrand N, Alves PC (2008a) Molecular analysis of hybridisation between wild and domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**, 1-11

Oliveira R, Godinho R, Randi E, Alves PC (2008b) Hybridization vs conservation: are domestic cats threatening the genetic integrity of European wildcat (*Felis silvestris silvestris*) populations in Iberian peninsula? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **363**, 2953-2961

Peakall R and Smouse PE (2006) Genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295

Pecon-Slattery J, Wilkerson AJ, Murphy WJ, O'Brien SJ (2004) Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. *Molecular Biology and Evolution*, **21(12)**, 2299-2309

Pierpaoli M, Birò ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585-2598

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, **81(3)**, 559-75

Ragni B (1993) Status and conservation of the wildcat in Italy. In: *Seminar on the biology and conservation of the wildcat (Felis silvestris*) (ed. Council of Europe Enviromental Encounters), Pp. 40-41, Council of Europe Press, Strasbourg

Ragni B and Possenti M (1996) Variability of coat-colour and markings system in *Felis silvestris*. *Italian Journal of Zoology*, **63**, 285-292

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology,* **17**, 285-293

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

Reusch TBH and Wood TE (2007). Molecular ecology of global change. *Molecular Ecology*, **16**, 3973-3992

Rhymer JM and Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109

Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. *The American Journal of Human Genetics*, **73**, 1402-1422

Rosenberg NA, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, Feldman MW (2005) Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genetics,* **1**, 660-671

Say L, Devillard S, Léger F, Pontier D, Ruette S (2012) Distribution and spatial genetic structure of European wildcat in France. *Animal Conservation*, **15**, 18-27

Schauenberg P (1969) L'identification du chat forestier d'Europe, *Felis s. silvestris* Schreber 1777, par une méthode ostéométrique. *Revue Suisse de Zoologie*, **76**, 433-441

Schauenberg P (1977) La stature du chat forestiere *Felis silvestris* Schreber, 1777, et la variabilité morphologique de l'espèce. *Revue Suisse de Zologie*, **84**, 3-337

Suminski P (1962) Research in the native form of wildcat (*Felis silvestris* Schreber) on the background of its geographical distribution. *Folia Forestalia Polonica*, **8**, 1-81

Tiedemann R, Harder J, Gmeiner C, Haase E (1996) Mitochondrial DNA sequence patterns of harbour porpoises (*Phocoena phocoena*) from the North and the Baltic Seas. *Z. Säugetierkd*, **61**, 104-111

Vähä JPK and Primmer CR (2006) Detecting hybridization between individuals of closely related populations – a simulation study to assess the efficiency of model-based Bayesian methods to detect hybrid individuals. *Molecular Ecology*, **15**, 63-72

Vigne JD, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J (2012) First wave of cultivators spread to Cyprus at least 10,600 y ago. *Proceedings of the national academy of sciences of the USA*, **109(22)**, 8445-8449

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249-256

Wayne RK and Brown DM (2001). Hybridization and conservation of carnivores. In: Gittleman JL, Funk SM, Macdonald D, Wayne RK (Eds) Carnivore Conservation, Cambridge University Press: Cambridge, UK, Pp 145-162

Williams JL, Eggen A, Ferretti L, Farr C, Gautier G, Amati G, Ball G, Caramori T, Critcher R, Costa S, Hextall P, Hills D, Jeulin A, Kiguwa SL, Ross O, Smith AL, Saunier KL, Urquhart BGD, Waddington D (2002) A bovine whole genome radiation hybrid panel and outline map. *Mammalian Genome*, **13**, 469-474

4.2 SUPPORTING INFORMATION

Supplementary Table S1. Average proportions membership (Q_i) and inferred ancestry of sampled domestic and wild cats combining with 12 different hybrid classes of genotypes generated by HYBRIDLAB. All populations are running with "admixture" and "correlated allele frequencies" model in STRUCTURE. On the left are reported the average proportion membership of each populations to wild and domestic clusters (Q_w = wildcats cluster; Q_d = domestic cluster). On the right are numbered the genotypes assigned at six different threshold values to own genetic cluster: simulated domestic and wild cats (Fca and Fsi) are correctly assigned at a threshold of 0.95. first and second generation hybrids (F1 and F2) are assigned equally to both two clusters. I and II generation backcrosses with wild and domestic cats are assigned at a threshold of 0.75 while III and IV generation backcrosses with wild and domestic cats are chiefly assigned at a threshold of 0.85. Empiric wild and domestic cats are correctly identified at a threshold of 0.80 and 0.90 respectively.

Pop	Q_w	Q_d	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	**N**
Fca sim	0.022	0.978	30	30	30	30	30	30	30
Fsi sim	0.970	0.030	30	30	30	30	30	30	30
Fca ref	0.031	0.969	10	10	10	10	10	9	10
Fsi ref	0.938	0.062	53	53	53	49	40	35	53
F1	0.490	0.510							
F2	0.492	0.508							
BX1_Fca	0.256	0.744	29	7					30
BX1_Fsi	0.730	0.270	29	14					30
BX2_Fca	0.247	0.753	27	8					30
BX2_Fsi	0.730	0.270	30	20					30
BX3_Fca	0.135	0.865	30	30	30	25			30
BX3_Fsi	0.852	0.148	30	30	30	20			30
BX4_Fca	0.138	0.862	30	30	30	25	2		30
BX4_Fsi	0.851	0.149	30	30	30	19			30

Supplementary Table S2. Individual membership proportions (*qi*) of known and putatively admixed cats according to the Bayesian analyses performed in STRUCTURE with admixture and correlated allele frequencies model. excluding prior population information (POP = 0). and by using 151 most informative SNPs and 38 STRs (Mattucci *et al*. 2013; Mattucci *et al.* in prep). STRUCTURE *qi* values correspond to allocations with *K*=2 to the domestic (Q_d) and wild (Q_w) inferred clusters. with their 90% credibility intervals (CI). Cats resulting with both markers admixed are shaded in light grey.

Dataset	ID	Populations	155 SNPs		38 STRs	
			Q_d	Q_w	Q_d	Q_w
Known admixed cats *Felis silvestris x catus n* = 10	ID 57	Captivity (Italy)	0.345 (0.298-0.394)	0.655 (0.606-0.702)	0.454 (0.303-0.609)	0.546 (0.391-0.697)
	ID 60	Captivity (Italy)	0.381 (0.334-0.430)	0.619 (0.570-0.666)	0.479 (0.320-0.644)	0.521 (0.356-0.680)
	ID 62	Captivity (Italy)	0.373 (0.325-0.423)	0.627 (0.577-0.675)	0.485 (0.324-0.649)	0.515 (0.351-0.676)
	ID 63	Captivity (Italy)	0.417 (0.368-0.467)	0.583 (0.533-0.632)	0.490 (0.326-0.657)	0.510 (0.343-0.674)
	ID 228	Italy: Southern Apennines	0.501 (0.447-0.554)	0.499 (0.446-0.553)	0.679 (0.533-0.815)	0.321 (0.185-0.467)
	ID 992	Italy: Northern-eastern Alps	0.496 (0.445-0.547)	0.504 (0.453-0.555)	0.493 (0.342-0.647)	0.507 (0.353-0.658)
	ID 993	Italy: Northern-eastern Alps	0.529 (0.479-0.579)	0.471 (0.421-0.521)	0.575 (0.416-0.730)	0.425 (0.270-0.584)
	ID 994	Italy: Northern-eastern Alps	0.515 (0.464-0.566)	0.485 (0.434-0.536)	0.575 (0.416-0.729)	0.425 (0.271-0.584)
	ID 995	Italy: Northern-eastern Alps	0.530 (0.479-0.580)	0.470 (0.420-0.521)	0.546 (0.388-0.701)	0.454 (0.299-0.612)
	ID 996	Italy: Northern-eastern Alps	0.507 (0.457-0.558)	0.493 (0.442-0.543)	0.500 (0.345-0.656)	0.500 (0.344-0.655)
Putative admixed cats *Felis silvestris x catus n* = 19	ID 67	Italy: Central Apennines and Maremma	0.144 (0.111-0.181)	0.856 (0.819-0.889)	0.238 (0.111-0.380)	0.762 (0.620-0.889)
	ID 123	Italy: Central Apennines and Maremma	0.110 (0.080-0.143)	0.890 (0.857-0.920)	0.127 (0.000-0.281)	0.873 (0.719-1.000)
	ID 627	Italy: Central Apennines and Maremma	0.126 (0.092-0.164)	0.874 (0.836-0.908)	0.344 (0.171-0.522)	0.656 (0.478-0.829)
	ID 677	Italy: Central Apennines and Maremma	0.147 (0.112-0.184)	0.853 (0.816-0.888)	0.214 (0.078-0.370)	0.786 (0.630-0.922)
	ID 1269	Italy: Central Apennines and Maremma	0.249 (0.206-0.293)	0.751 0.707-0.794)	0.342 (0.200-0.495)	0.658 (0.505-0.800)
	ID 1323	Italy: North-eastern Alps	0.258 (0.215-0.302)	0.742 (0.698-0.785)	0.246 (0.123-0.385)	0.754 (0.615-0.877)
	ID 1518	Italy: Central Apennines and Maremma	0.180 (0.141-0.221)	0.820 (0.779-0.859)	0.533 (0.351-0.714)	0.467 (0.286-0.649)
	ID 77	Italy: Southern Apennines	0.024 (0.009-0.043)	0.976 (0.957-0.991)	0.156 (0.036-0.297)	0.844 (0.703-0.964)
	ID 1248	Italy: Southern Apennines	0.164 (0.127-0.203)	0.836 (0.797-0.873)	0.285 (0.144-0.436)	0.715 (0.564-0.856)
	ID 1250	Italy: Southern Apennines	0.129 (0.095-0.166)	0.871 (0.834-0.905)	0.255 (0.120-0.408)	0.745 (0.592-0.880)
	ID 1252	Italy: Southern Apennines	0.116 (0.085-0.150)	0.884 (0.850-0.915)	0.361 (0.217-0.517)	0.639 (0.483-0.783)

	ID 586	Luxembourg	0.293 (0.247-0.341)	0.707 (0.659-0.753)	0.154 (0.003-0.306)	0.846 (0.694-0.997)
	ID 1056	Bosnia & Herzegovina	0.267 (0.223-0.313)	0.733 (0.687-0.777)	0.399 (0.232-0.571)	0.601 (0.429-0.768)
	ID 421	Germany: Southern-western	0.978 (0.959-0.993)	0.022 (0.007-0.041)	0.829 (0.639-1.000)	0.171 (0.000-0.361)
	ID 629	Germany: Southern-western	0.200 (0.161-0.241)	0.800 (0.759-0.839)	0.417 (0.263-0.573)	0.583 (0.427-0.737)
	ID 700	Portugal	0.475 (0.425-0.524)	0.525 (0.476-0.575)	0.404 (0.243-0.570)	0.596 (0.430-0.757)
	ID 706	Portugal	0.971 (0.951-0.987)	0.029 (0.013-0.049)	0.593 (0.430-0.750)	0.407 (0.250-0.570)
	ID 712	Portugal	0.460 (0.411-0.510)	0.540 (0.490-0.589)	0.643 (0.467-0.813)	0.357 (0.187-0.533)
	ID 717	Spain	0.928 (0.898-0.955)	0.072 (0.045-0.102)	0.675 (0.515-0.822)	0.325 (0.178-0.485)

Supplementary Table S3. Power to detect wild x domestic cat hybrids combining individual Bayesian assignment performed in STRUCTURE by using 151 most informative SNP and haplotypes classification detected by using maternal (mtDNA) and paternal (Y Chromosome) diagnostic polymorphism between domestic (D) and wild (W) haplogroups. Cats resulting with both nuclear (SNPs) and uniparental markers admixed are shaded in light grey. *presence of domestic haplotype shared with wildcats. The three control region (CR) haplogroups were numbered. Individuals needing further analyses to overcome the current lacking of uniparental informations. despite the clear SNPs classification. are underlined.

Dataset	ID	Populations	STRUCTURE assignemnt with 155 SNPs		mtDNA		Y Chr
			Q_d	Q_w	ND5	CR	
Known admixed cats *Felis silvestris x catus n* = 10	ID 57	Captivity (Italy)	0.345 (0.298-0.394)	0.655 (0.606-0.702)	D*	2*	W
	ID 60	Captivity (Italy)	0.381 (0.334-0.430)	0.619 (0.570-0.666)	D*	2	W
	ID 62	Captivity (Italy)	0.373 (0.325-0.423)	0.627 (0.577-0.675)	D*	2	W
	ID 63	Captivity (Italy)	0.417 (0.368-0.467)	0.583 (0.533-0.632)	D*	2	D
	ID 228	Italy: Southern Apennines	0.501 (0.447-0.554)	0.499 (0.446-0.553)	W	1	D
	ID 992	Italy: Northern-eastern Alps	0.496 (0.445-0.547)	0.504 (0.453-0.555)	D*	2*	D
	ID 993	Italy: Northern-eastern Alps	0.529 (0.479-0.579)	0.471 (0.421-0.521)	D*	2*	D
	ID 994	Italy: Northern-eastern Alps	0.515 (0.464-0.566)	0.485 (0.434-0.536)	D*	2*	D
	ID 995	Italy: Northern-eastern Alps	0.530 (0.479-0.580)	0.470 (0.420-0.521)	D*	2*	D
	ID 996	Italy: Northern-eastern Alps	0.507 (0.457-0.558)	0.493 (0.442-0.543)	D*	2*	D
Putative admixed cats *Felis silvestris x catus n* = 19	ID 67	Italy: Central Apennines and Maremma	0.144 (0.111-0.181)	0.856 (0.819-0.889)	W	1	F
	ID 123	Italy: Central Apennines and Maremma	0.110 (0.080-0.143)	0.890 (0.857-0.920)	D*	3	D
	ID 627	Italy: Central Apennines and Maremma	0.126 (0.092-0.164)	0.874 (0.836-0.908)	W	1	F
	ID 677	Italy: Central Apennines and Maremma	0.147 (0.112-0.184)	0.853 (0.816-0.888)	W	1	F
	ID 1269	Italy: Central Apennines and Maremma	0.249 (0.206-0.293)	0.751 0.707-0.794)	W	1	W
	ID 1323	Italy: North-eastern Alps	0.258 (0.215-0.302)	0.742 (0.698-0.785)	W		
	ID 1518	Italy: Central Apennines and Maremma	0.180 (0.141-0.221)	0.820 (0.779-0.859)	W	1	W
	ID 77	Italy: Southern Apennines	0.024 (0.009-0.043)	0.976 (0.957-0.991)	W		
	ID 1248	Italy: Southern Apennines	0.164 (0.127-0.203)	0.836 (0.797-0.873)	D*	2*	W
	ID 1250	Italy: Southern Apennines	0.129 (0.095-0.166)	0.871 (0.834-0.905)	D*	2*	W
	ID 1252	Italy:	0.116	0.884	D*	2*	W

	Southern Apennines	(0.085-0.150)	(0.850-0.915)			
ID 586	Luxembourg	0.293	0.707	W	1	W
		(0.247-0.341)	(0.659-0.753)			
ID 1056	Bosnia & Herzegovina	0.267	0.733	D*		F
		(0.223-0.313)	(0.687-0.777)			
ID 421	Germany: Southern-western	0.978	0.022	D*		F
		(0.959-0.993)	(0.007-0.041)			
ID 629	Germany: Southern-western	0.200	0.800	D*		
		(0.161-0.241)	(0.759-0.839)			
ID 700	Portugal	0.475	0.525			
		(0.425-0.524)	(0.476-0.575)			
ID 706	Portugal	0.971	0.029	W		
		(0.951-0.987)	(0.013-0.049)			
ID 712	Portugal	0.460	0.540	W	1	W
		(0.411-0.510)	(0.490-0.589)			
ID 717	Spain	0.928	0.072			
		(0.898-0.955)	(0.045-0.102)			

4.3 LIST OF TABLE

Table 1. Sampling size and location of all genotyped cats. Hybrid cats that have not been identified through phenotypic evaluation and that have been initially included in the parental populations are reported between brackets ("admixed genotypes").

Table 2. SNPs and genotypes preliminary pruning computed in PLINK.

Table 3. Genetic description of 151 most informative SNP loci used in this study: Allelic richness (A_R); expected heterozygosity (H_E); Inbreeding coefficient (F_{IS}); pairwise estimate of genetic differentiation between European wild and domestic cats (AMOVA pw-F_{CT}); informativeness for assignment index (INFOCALC I_N) computed for each SNP locus.

Table 4. Individual membership proportions (q_i) and inferred ancestry of 16 putatively admixed cats according to the Bayesian analyses performed in STRUCTURE and NEWHYBRIDS. In STRUCTURE, individuals were assigned into two cluster corresponding to the domestic (Q_d) and wild (Q_w) groups excluding prior population information. In NEWHYBRIDS, individuals were assigned into six different genotypes classes: pure domestic and wild cat, F1, F2, Bx I (backcross with domestic cat) and Bx II (backcross with wildcat). STRUCTURE 90% credibility intervals (CI) are shown in brackets.

Table 5. Genetic variability based on ND5 region of mtDNA observed within the main genetic haplogroups (Haplogroup D, Haplogroup W)

Table 6. Analyses of molecular variance based on ND5 region of mtDNA data from the main haplogroups (Haplogroup D. Haplogroup W)

Table 7. Population pairwise Fst matrix based on Control Region of mtDNA data from the main haplogroups (Haplogroup 1=1. Haplogroup 2=2. Haplogroup 3=3)

Table 8. Chromosome Y haplotypes based on SMCY-7 STR and SRY gene SNP.

Supplementary Table S1. Average proportions membership (Q_i) and inferred ancestry of sampled domestic and wild cats combining with 12 different hybrid classes of genotypes generated by HYBRIDLAB. All populations are running with "admixture" and correlated allele frequencies model

in STRUCTURE. On the left are reported the average proportion membership of each populations to wild and domestic clusters (Q_w= wildcats cluster; Q_d = domestic cluster). On the right are numbered the genotypes assigned at six different threshold values to own genetic cluster: simulated domestic and wild cats (Fca and Fsi) are correctly assigned at a threshold of 0.95. first and second generation hybrids (F1 and F2) are assigned equally to both two clusters. I and II generation backcrosses with wild and domestic cats are assigned at a threshold of 0.75 while III and IV generation backcrosses with wild and domestic cats are chiefly assigned at a threshold of 0.85. Empiric wild and domestic cats are correctly identified at a threshold of 0.80 and 0.90 respectively.

Supplementary Table S2. Individual membership proportions (*qi*) of known and putatively admixed cats according to the Bayesian analyses performed in STRUCTURE with admixture and correlated allele frequencies model. excluding prior population information (POP = 0). and by using 151 most informative SNPs and 38 STRs (Mattucci *et al*. 2013; Mattucci *et al.* in prep). STRUCTURE *qi* values correspond to allocations with K=2 to the domestic (Q_d) and wild (Q_w) inferred clusters. with their 90% credibility intervals (CI). Cats resulting with both markers admixed are shaded in light grey.

Supplementary Table S3**.** Power to detect wild x domestic cat hybrids combining individual Bayesian assignment performed in STRUCTURE by using 151 most informative SNPs and haplotypes classification detected by using maternal (mtDNA) and paternal (Y Chromosome) diagnostic polymorphism between domestic (D) and wild (W) haplogroups. Cats resulting with both nuclear (SNPs) and uniparental markers admixed are shaded in light grey. *presence of domestic haplotype shared with wildcats. The three control region (CR) haplogroups were numbered. Individuals needing further analyses to overcome the current lacking of uniparental informations. despite the clear SNPs classification. are underlined.

4.4 LIST OF FIGURES

Figure 1. Sampling locations of putative European wildcats. Shaded areas correspond to the approximate current distribution of Felis silvestris in Europe (adapted from Grabe and Worel 2001).

Figure 2. Plot of individual *q*w values (to a wildcat cluster) and their 90% credibility intervals CI. The *q*i thresholds = 0.20 - 0.80 (interrupted lines) define the admixed genotypes. Pure wild and domestic cats are in black, admixed individuals are in red .

Figure 3. NEWHYBRIDS' assignment of 10 known hybrids and 19 'putative' admixed individuals to the different hybrid categories. Each individual is represented by a single vertical bar colored according to the proportion of their genome descending from each of the inferred clusters or hybrid class.

Figure 4. NETWORK diagram of cat ND5 haplotypes. Each haplotype is named with a code and is represented by a circle whose size is scaled according to its abundance in the cat sample analyzed. Different colors refer to highly informative nuclear SNPs assignments of genotypes (see Table 4). Mutated positions are indicated on braches.

Figure 5. Frequency of ND5 cat haplotypes for both the haplogroups W and D.

Figure 6. NETWORK diagram of Control Region (CR) cat haplotypes. Each is represented by a circle whose size is scaled according to its abundance in the cat sample analyzed. Different colors refer to highly informative nuclear SNPs assignments of genotypes (see Table 4). Mutated positions (black nodes) and theoretical intermediary haplotypes predicted by network analysis (red nodes) are indicated on braches.

Chapter 5

General Discussion

During the 18th and 19th centuries, the European wildcat (*Felis silvestris silvestris*) strongly declined as a consequence of historical loss of habitat, deforestation and human persecution occurring throughout most of central and western Europe. Concomitantly, the anthropogenic diffusion of domestic cats originated widespread free-ranging populations in agricultural ecosystems, raising concerns about the genetic integrity of wildcats (McOrist & Kitchener 1994). Genetic infiltration of domestic alleles into native gene pools may result in the detrimental breakup of locally adaptive autochonous wildcat gene complexes (Rhymer & Simberloff 1996; Allendorf *et al*. 2001; McGinnity *et al.* 2003). Moreover, the fertility of hybrid recombinant genotypes and their mating with the parental groups, in extreme situations, may lead to the disappearance of the 'pure' parental genomes (Mallet 2005), and the formation of hybrid swarm (Allendorf *et al*. 2001). Since admixed individuals are expected to be less fit than either parental type because they have never been tested by natural selection (Turelli *et al.* 2001), they may further compromise the long-term survival of the threatened species and promote further population decline (Rhymer & Simberloff 1996; Allendorf *et al*. 2001). Wildcat populations thus become less suited to their natural environments, more likely to occupy those habitats favored by domestics, and are so pushed along a slippery slope of greater admixture and compromised genetic integrity (Germain *et al.* 2008).
Hence, accurate detection of hybrid individuals and introgression rate, combining with a deep investigation on the genetic consequences of fragmentation and isolation, are the main challenges for assessing wildcats conservation's status and, subsequently, providing thoughtful knowledge to appropriately support conservation plans. Under this light, it has been developed this thesis.

5.1. Genetic structure of European wildcats in Italy

Current distributions and structure of natural biomes in Europe have been mainly affected by the Quaternary glaciations. Europe, in fact, during the Quaternary ice ages experienced massive and intense climate fluctuations culminating in the Last Glacial Maximum (LGM 18–20 kyr, Strandberg *et al.* 2011). Following climate warming during the Holocene, species that survived in refugia expanded northwards giving rise to the current phylogeographic structuring of populations (Hewitt 2004). In addition to the classical Southern Mediterranean refuge model identified in Iberia, Italy and the southern Balkans (Hewitt 2000; Weiss & Ferrand 2006), recent studies of diverse taxa have suggested the presence of cryptic Northern refugia, putatively located in sheltered areas scattered in central Europe north of the Pyrenees and the Alps (Stewart & Lister 2001; Deffontaine *et al*. 2005) and the possibility of multiple colonization waves of species migrating into continental Europe from eastern European and Asian refugia (Bilton *et al*. 1998; Randi 2007; Fløjgaard *et al*. 2011).

In particular, feline archaeological remains suggest that during the Last Glacial Maximum (LGM) both refuges in the southern Mediterranean (Spain and Italy), and in the western and eastern Alpine peripheral regions (Sommer & Benecke 2006), preserved habitat patches suitable to populations of forest-associated small-mammals and their predators, including the European wildcat (Sommer & Benecke 2006).

Hence, we used molecular tools to verify the phylogeographical origin and evolution of current Italian wildcats populations distributed in three separate geographic areas: in the eastern Italian Alps (Friuli Venezia Giulia and Veneto), which are presumably connected with neighbour populations in Slovenia and Croatia (Lapini 1989; 2006); in the central and southern Apennines ridge (Ragni 2006); and in Sicily.

We preliminarily analyzed the entire dataset of 346 cat samples aiming at assessing the purity of putative European wildcats and eventually the occurrence of admixture between domestic and wild cats; we also verified the power of the 35 microsatellites used for individuals and subspecies discrimination and hybrids detection.

Therefore, the analysis of genetic variation evidenced the power of microsatellite panel set to discriminate individual highly related in all the three subspecies, with probabilities of PID and PIDsibs with order of magnitude of 10^{-30} and 10^{-13}, respectively. Moreover, the three subspecies have been clearly identified and differentiated by using both Bayesian analysis and multivariate tools. Comparable values with other wildcat or domestic cat populations of Europe (Pierpaoli *et al.* 2003; Oliveira *et al*. 2008a,b; Eckert *et al*. 2009; O'Brien *et al.* 2009) were obtained from Italian samples for genetic diversity indices, suggesting that fragmentation and consequent demographic decline of the current population did not genetically depleted wildcats in Italy.

Bayesian admixture inferences were further computed because of the possible presence of cryptic hybrids, not morphologically recognizable (Randi *et al.* 2001; Lecis *et al.* 2006), inside the putative European wildcat population. Hence, choosing a threshold $qi = 0.90$ (based on admixture analyses of observed and simulated cat data sets; Randi 2008; Oliveira *et al.* 2008a), 17 known hybrids, analyzed as control, were confirmed as admixed, and 36 new admixed European wildcats and one new admixed Sardinian wildcat were further identified. This samples were removed from the dataset but deeply investigated in the *Paper III* with highly informative markers, such as SNPs and diagnostic mtDNA sequences.

We used landscape genetic tools, by combining Multivariate and Bayesian statistical tools with geographic and spatial patterns, aiming at understanding if and how landscape and environmental

features influenced the microevolutionary processes that generated current pattern of genetic structure and local adaptation of wildcat populations across space (Manel *et al*. 2003).

We further used simulations to provide an estimation of diverge time among populations detected with landscape genetic tools.

Within Italy, four genetically distinct subpopulations were disclosed: in the eastern Alps, in the central-southern Apennine mountain-hills, in the Maremma hill-plain and coastal areas, and in Sicily. Congruently with the Pleistocene biogeographical framework of the Italian peninsula, the three main populations in the eastern Alps, in the peninsular Apennines and in Sicily, most likely sharply differentiated during the LGM as a consequence of their geographic isolation in the southern Mediterranean refuges of Sicily and in Cisalpine refuges of the south-eastern Alps.

Evidences of this are provided by the divergence time estimates that suggest a period of isolation without gene flow of about 5000 generations (corresponding to 10,000-15,000 years), to explain the current genetic variation pattern of the Alpine and the Italian peninsula populations. In addition, both observed and simulated genetic parameters suggest that extant European wildcat populations in Italy did not experience deep historical declines of their effective population sizes and that genetic divergence among populations cannot be explained by recent fragmentation.

The subdivision in the central Italian peninsula, most likely, originated from a more recent adaptation of distinct European wildcat subpopulations to distinct habitat types. The central Apennines and the Maremma regions, although parts of the same latitudinal range, represent, in fact, two distinct bioclimatic and ecological regions (Piovesan *et al*. 2005), with peculiar vegetation, climate and temperature. In particular, populations in the western sector of the range might have experienced periods of isolation and local adaptation to a peculiar Mediterranean-type habitat known as Maremma.

It is widely accepted among conservation geneticists that understanding levels of genetic diversity and population genetic structure is a prerequisite for conservation programs for endangered species (e.g. Frankham *et al.* 2002). This study, for the first time, has provided evidence of genetic structure of Italian wildcats; however, the phylogeography of the European wildcat might be further refined by expanding samples and markers. For instance, the spatial connection across the Alpine ridge among eastern Italian, Slovenian and Croatian samples, should be considered by implementing the Alpine sampling, in order to better delineate the origin of wildcats in the eastern Alps. Therefore, it should be proved if they have been isolated in the hypothesized LGM refuge, or have been originated by post-glacial expansions of south Balkan source populations. Moreover, creating a multi locus panel set with highly informative SNPs selected from the entire cat genome, it might

help to reconstruct ancient evolutionary events (Miller *et al.* 2012), overcoming the limit of homoplasy that can occur using STRs loci. In addition, the discovery of selected loci might help to clarify the evolutionary dynamics of local adaptations of wildcats in the widest contest of comparative ecological genomics (Martin *et al*. 2003).

The current geographic and genetic pattern of Italian wildcats evidenced two widespread population networks distributed across the central-southern Apennines ridge and the Italian eastern Alpine-Balkanic area, and two populations confined in the Mediterranean Maremma area for adaptative reasons and in Sicily for geographic isolation from the mainland. In a conservation-oriented view, it became essential, especially for the latter populations ecologically and spatially limited, to assess admixture events between wild and domestic cats, clearly identifying origin and level of introgression in wildcat individuals and opportunely mapping the admixture area, in order to provide thoughtful knowledge of wildcat populations and eventually support the identification of significant conservation units (ESUs), namely population that require separate management because of their high conservation priority (Ryder 1986) or management units (MUs).

Conservation biologists, in fact, assign population distinctiveness by classifying populations as evolutionarily significant units (ESUs). Historically, this classification was developed as a supplemental way to describe evolutionarily distinct groups when taxonomy was either inadequate or too controversial to reflect these distinctions (Ryder 1986). Still now, there is no general agreement on the criteria to use for ESUs' definition. For instance, character-based methods have been proposed (Amato 1991; Dowling *et al.* 1992; Vogler & DeSalle 1994) to identify ESUs based on the presence of unique "diagnostic" characters that set apart groups of individuals or populations from others. However, in case of units that are demographically subdivided but have not experienced an evolutionary divergence time sufficient to allow the accumulation of diagnostic characters (Waples 1991; Moritz 1994), this criteria appear to be not applicable. Hence, further categories have recently been proposed to recognize demographically distinct populations that, despite the less genetic divergence, should be managed to ensure the viability of the larger ESU (Moritz 1994). In particular, the identification of management units (MUs) appears to be crucial to monitor, for example, the effects of human activity upon species abundance and fragmentation.

5.2. Genetic structure of European wildcat populations in Europe

Current patterns of European wildcats distribution and genetic variability are most probably a reflection of both natural and anthropogenic events. Drastic habitat and demographic changes, exposed small isolated populations to the risks of inbreeding depression, lowered fitness and loss of evolutionary potential and adaptation to environmental changes.

Only when anthropic pressure on wildcat populations and their habitat was reduced (e.g., Parent 1975; Easterbee *et al.* 1991) it has been possible a recovery of the subspecies. Nevertheless, it has been slow due to the isolation and fragmented distribution of many populations (Stahl & Artois 1991).

In order to understand the genetic consequences of fragmentation, to identify eventual evolutionary or conservation units and forecast their conservation perspectives, is needed to assess the population structure of wildcats in Europe, estimating the genetic diversity within and between populations. Previous studies evidenced how type and number of molecular markers is crucial for the resolution and robustness of population structure inferences, and for hybrids and introgression detection (e.g. Oliveira *et al*. 2008; Nussberger *et al*. 2013).

Hence, in this study we analyzed for the first time a wide sampling of 1114 cats, which covered the majority of species range in Europe (from Portugal to Romania), with a wide panels set of 38 unlinked microsatellites. Moreover, taking advantage of multivariate analyses and Bayesian clustering, we tried to describe the main patterns of population structuring, by inferring the extent of their genetic isolation and identifying genetic signatures of past and recent bottlenecks.

We preliminarily analyzed the entire dataset of 1114 cat samples aiming at assessing the purity of putative European wildcats and eventually the occurrence of admixture between domestic and wild cats; we also described patterns of genetic diversity among European wild, African wild (collected in Sardinia) and domestic cats.

Therefore, high level of genetic variability has been found for both African wildcats and domestic cats, opening the view to two different hypothesis: i) the wide distributional range (Driscoll *et al.* 2007) and habitat tolerance (Driscoll & Nowell 2010) might have promoted gene flow among populations, rescuing the subspecies against strong past population declines; or ii) past episodes of crossbreeding between domestic cats and their wild ancestors might have occurred in multiple occasions and for a long time during the process of domestication (Driscoll *et al.* 2007) and might have continuously supported the maintenance of high genetic diversity.

Bayesian admixture analyses led us to confirm the highly compromised genetic composition of cats in Hungary and Scotland, especially when compared to the other European populations, with individual assignment values frequently intermediate between the wild and domestic clusters. This might result from long lasting hybridization that seems not to have occurred in any of the other European regions (Beaumont *et al*. 2001; Pierpaoli *et al.* 2003; Lecis *et al.* 2006). Moreover, among non-hybridizing populations, a total of 79 individuals have shown evidences of possible hybrid ancestry and 7 captive bred hybrids were confirmed as admixed cats. This samples were removed from the dataset but deeply investigated in the *Paper III* with highly informative markers, such as SNPs and diagnostic mtDNA sequences.

Analyses of genetic diversity performing AMOVA and pairwise-F_{ST} and R_{ST} estimates confirmed the close genetic proximity of *lybica* and *catus* subspecies and evidenced a significant differentiation between domestic cats and all European wildcats subpopulations. Despite the heterogeneity of samplings among the wildcats' subpopulations and the possibility of a further substructure in regions affected by deviations from HWE and LE, globally it has been confirmed the confidence in the existence of true non highly hybridizing populations in Europe (in Europe, Pierpaoli *et al*. 2003; Portugal and Spain, Oliveira *et al.* 2008; France, O'Brien *et al*. 2009; Italy, Randi *et al.* 2001; and Germany, Hertwig *et al.* 2009 and Eckert *et al*. 2010).

Multivariate and Bayesian statistical tools were used to detected the genetic structure of non-hybridizing wildcats. Five macro groups were first identified, subdividing the European population in southern and central Europe and separated the Eastern German population from all the other others, as Bayesian clustering reported by Pierpaoli *et al.* (2003). In particular, Southern Europe includes two main macro groups: Iberian peninsula, and Italy (Mattucci *et al.* 2013), Slovenia and Bosnia & Herzegovina area. A more fine clustering was reconstructed within those macro groups, and a total of 10 geographic and genetic wildcats populations were finally identified. Oliveira *et al.* (2008) first identified a preliminary geographically separation between wildcats from Portugal and Spain, with the exception of some Portuguese and Spanish cats that were assigned to the other cluster without any apparent biological/ecological reason, such as translocation or geographical proximity of animals. However the low panel set used (12 unlinked microsatellites) was not reliable to perform fine substructure analysis across wildcat populations, and sharply differentiated the northern from the southern-eastern and southern-western Iberian wildcats.

Among all the studied subpopulations, genetic signatures of population bottleneck were identified only for the Eastern population of Germany, confirming predictions by Pierpaoli *et al*. 2003 and Eckert *et al.* 2010 and contrasting the absence of such event advocated by Hertwig *et al.* 2009. Although no significant signs of recent heterozygosity excess could be detected in our study (as assessed by the procedure of Cornuet & Luikart 1997), the 49 samples genotyped from this region displayed lower values of genetic diversity than other subpopulations (lower number of alleles, allelic richness, private alleles richness and heterozygosity); and the ratio between the number of alleles and the range in allele size (*m*-ratio) suggested a significant reduction in effective population size. In this area, the prevalence of just few scattered habitat spots populated by few individuals might have acted as a barrier to individuals' dispersal and, thus, limited gene flow and promoted allele diversity loss due to genetic drift (Pierpaoli *et al.* 2003; Eckert *et al.* 2010). Although large population declines are known to have occurred across the entire species' range (Driscoll & Nowell 2010), the high level of genetic diversity registered might indicate that the demographic decline did not produce noticeable signs of genetic depletion.

The LD-base population sizes estimates provided for each European wildcat populations, revealed the presence of the minimum effective breeders size (n = 50), needed to prevent inbreeding depression in the short term (Franklin 1980), only in the Italian, Balkan and Iberian macro populations and in the Italian north-eastern Alps and central-southern Apennines subpopulations, probably as a result of their wide spatial distribution. Since Linkage disequilibrium models assume a closed and unstructured population (Waples 2006), the N_E estimate for all the macro groups, with the exception of eastern Germany, may be underestimated by admixture among sub-populations within them (Waples & England 2011). The remain populations resulted to be very poorly sampled (S < 50), and the extremely small effective size (included the eastern Germany population) estimated make impossible to state about their conservation status. Nevertheless, the target effective population sizes of 500-700 recommended for securing long-term viability (Franklin 1980; Lande 1995; Reed *et al*. 2003) is clearly several times larger than those observed.

The importance in conservation of demographic inference has been first recognized by Lande (1998). Demographic factors (i.e. population sizes and the connectivity between populations) are useful to detect early warning signs of population subdivision, local extinction and loss of connectivity, providing a global perspective on all the events that structure species and communities and giving the possibility to make predictions about the future of species (Waits *et al.* 1998; Allendorf & Luikart 2007). We used Approximate Bayesian Computation (ABC) method to infer

the divergence times and historical values of N_E in the main European wildcat population clusters, as they were defined by concordant geographical distributions and Bayesian clustering.

The pattern of demographic parameters estimated for the southern European populations is concordant with the Pleistocene biogeographical framework of Europe. The three main population in Balkans, Italian and Iberian Peninsula, most likely sharply differentiated during the LGM as a consequence of their geographic isolation in the southern Mediterranean refuges.

While, the subdivision in the central Europe between the Germany eastern population and the Belgium, Luxembourg + Romania, Switzerland + Germany southern-western populations, might be the consequence of extra-Mediterranean Würm ice age refuge areas in Europe, especially in the northern Alps, Carpathians, and the Bulgarian mountain systems (Stewart *et al.* 2001; Schmitt & Varga 2012). These northern refugia would have been in areas of sheltered topography that provided suitable stable microclimates, facilitating the hybridization among populations of these scattered extra-Mediterranean refugial pockets during the milder interstadial phases of the Würm and also between the LGM and the younger Dryas period (Stewart *et al.* 2001). Thus, they became the leading edges of the postglacial northwards range expansions, strongly impacting the genetic constitution of Central and North Europe in many plant and animal species.

On the contrary, the extant genetic divergence among European wildcat populations of Iberia and Italy resulted to be by extended periods of isolation without gene flow (in the order of 5,000-10,000 generations), and not by more recent fragmentation.

The significant population structure found across the entire European distribution range suggests that current macro and subpopulations might represent evolutionary significant units (ESU's). The small sizes and isolation registered in the majority of these populations, might, in fact, worsen the effects that eventual demographic declines and hybridization can produce on populations' survival, especially due to their low population densities, high generation times, sensibility to ecosystems modifications and anthropogenic changes (Schipper *et al.* 2008). In particular, the wildcat populations from eastern Germany should be regarded as a conservation priority, since the geographical and genetic isolation concurred to deplete their genetic diversity and consequently to reduce their ability to adapt and survive under the scenario of introgressive hybridization.

However, some populations in the Eastern Europe have maintained certain levels of reproductive contact. Hence, further study aimed to understand the processes shaping wildcat's evolutionary history are recommended. In particular, it became relevant assess which geographical, historical or ecological factors mostly influenced the genetic differentiation among local wildcat. Recently, Pilot *et al.* (2006) found that the genetic differentiation among local populations of European grey wolf

was correlated with climate, habitat types, and wolf diet composition rather than topographic barriers or historical populations' fragmentation. Hence, it might be important also investigate the direct mechanism that links population ecology and population genetic structure in wildcats (Pilot *et al.* 2006)

In a conservation-oriented view, it became needed to assess the occurrence of admixture events between wild and domestic cats, clearly identifying origin and level of introgression in wildcat individuals and opportunely mapping the admixture area, in order to understand if crossbreeding between wild and domestic cats has been restricted to peripheral areas of wildcat range as previously advocated by Oliveira *et al.* (2008) in Iberia and Randi *et al*. (2001) in Italy, or if admixture events might also take place in the core of the species distribution. Under this light, it appears useful to increase the sampling of some populations of European wildcats in order to create heterogeneity among the studied areas and to develop sampling strategies not opportunistically aiming at collecting the majority of all putatively wild, domestic and hybrid phenotypic samples.

5.3. Improving the molecular tools in wildcat studies for assessing admixture

5.3.1. Detection of most distinctiveness SNPs between European wild and domestic cats

The intricate history of simpatry and introgression, and the remarkable resemblance between European wild and domestic cats have deeply complicated the frameworks to genetically discriminate parental groups and consequently hybrids individuals.

In the last decade, a number of molecular studies have focused in the study of genetic diversity and hybridization patterns among European wildcats. Most recent research has taken advantage of the revolutionary advances in molecular and statistical techniques faced by conservation genetics, and some populations across Europe have now been investigated. Examples of detailed analyses can be found for Iberian peninsula (Oliveira *et al*. 2008), France (O'Brien *et al*. 2009), Italy (Randi *et al*. 2001), Germany (Hertwig *et al*. 2009; Eckert *et al*. 2010), Hungary (Lecis *et al*. 2006) and Scotland (Beaumont *et al.* 2001), where the analysis of mitochondrial variation and/or microsatellites diversity suggested varying degrees of hybridization between wild and domestic cats.
Until now, only Pierpaoli and colleagues have tried to draw the complete picture for the species across its European range by analyzing a total of 336 cats from 9 different countries with 12 polymorphic microsatellites (Pierpaoli *et al.* 2003). However, the number of loci used to analyze

admixture patterns (n = 12) remained far from the advised numbers according to convincing simulations by Vähä & Primmer (2006), since as much as 48 loci with average F_{ST} = 0.21 might be needed to distinguish F_1, F_2, backcrosses and parental individuals.

Preliminary admixture analyses computed in *PaperI* and *PaperII*, evidenced the high methodological effort computed to triplicate (n = 35 for Italy and n = 38 for Europe) the number of loci previously used to analyze hybridization rates at both Italian (n = 12, Randi *et al*. 2001) and European levels (n = 12, Pierpaoli *et al.* 2003). However, Bayesian analysis on simulated genotypes revealed that only the parental, F1 and F2 individuals could be correctly identified by the STRUCTURE's algorithm using an assignment threshold of qi = 0.80 and 35 and 38 microsatellites respectively for Italy and Europe (referred to *PaperI* and *PaperII*, data not shown). In particular, in Italy 15% of I and II generation backcrosses and 83% of III and IV generation backcrosses resulted to be assigned to the parental cluster; while in Europe, the percent of individual misclassified is 14% of domestic backcrosses and 12% of wildcat backcrosses. Moreover, assignment values for NEWHYBRIDS revealed that in both Italy and Europe the ancestry of most of the past-generation hybrids remained unknown (referred to *PaperI* and *PaperII*, data not shown).
Hence, the uncertainty on achieving accurate inferences of admixture proportions disclosed in *PaperI* and *PaperII*, led us to look for more reliable markers, instead of further increase the number of unlinked microsatellites to approximately 50 loci or/and add a subset of linked STRs (Lecis *et al*. 2006), possibility that appeared too expensive and laborious.

Advances in genetic and genomic resources for domestic cats have provided a great amount of single nucleotide polymorphisms (SNPs) that are recognized as the most promising tools to apply among wildlife conservation-oriented studies (Brumfield *et al.* 2003; Morin *et al.* 2004; Seddon *et al.* 2005; Morin *et al.* 2009). Offering less variability per locus than STRs, SNPs provide: i) low propensity for homoplasy due to lower mutation rates; ii) higher density and uniformity distribution in genomes; iii) successful high-throughput genotyping, and iii) highly successful application in fragmented DNA samples, e.g. non-invasive and historical DNA (see Brumfield *et al*. 2003; Morin *et al*. 2004; Garvin *et al*. 2010 for reviews). Successfully contributing to one of the first SNPs genotyping arrays performed to analyze world wide domestic cat populations, we assessed the genetic variability of 62,897 autosomal SNPs (referred to as the 63K panel) in 100 putative European wildcats, 45 village domestic cats, and 42 previously known or putative wild *x* domestic hybrids, in *Paper III*.

Preliminary filters were used to prune the initial 63K panel set and to extract a reduced panel set containing the most informative SNPs for the admixture analyses. Analyses of genetic variation revealed that, although no diagnostic fixed differences were detected, 151 SNPs showed the highest genetic divergence (F_{ST}) between wild and domestic cats, with values ranging from 0.797 to 0.982. Most likely, such level of differentiation might be resulted from past episodes of positive, balancing and/or purifying selection occurred during cat's domestication that have created distinct signatures in the genome between wild and domestic cats. Since these highly informative SNPs, randomly distributed in all domestic cat chromosomes, resulted to be useful to detect such level of differentiation, the possibility to verify the role of disruptive selection during domestication process requires further analysis. In fact, while the random-bred domestic cats used in this study should represent most of the genetic variability that can be found among the domestic cat gene pool (since they are not constrained by breed standards), European wildcats do not readily represent the ancestral wild population from which domestic cats have been domesticated, and results from this work do not allow to make considerations about loci selected during the domestication process itself. Instead, it would be recommended to detailed investigate the role of these mutations in cat's domestication in specimens of *Felis silvestris lybica*, which is the ancestor of all domestic cats (Driscoll *et al.* 2007).

Bayesian analyses computed on simulated genotypes (generated by HYBRIDLAB), revealed that all parental, F1, F2, and I-II generation backcrossed genotypes could be correctly identified by the STRUCTURE's algorithm using the 151 most informative SNPs (*Paper III*), and a threshold of qi = 0.80. However, all III- IV generation backcrosses showed a qi > 0.80 to one single cluster and could not be distinguished from parental individuals. Moreover, none of the referred cats were significantly assigned to one of the other remaining hybrid categories, preventing any case of misclassification also with the NEWHYBRIDS' algorithm.

Even though we chose a threshold of qi > 0.80 for the Bayesian analyses performed in both STRUCTURE and NEWHYBRIDS, the individual membership assignments proved to be higher for parental genotypes (qi > 0.95 in both domestic and wild cats), and intermediated between the wild and domestic cluster in all the putative and know admixed genotypes. In both cases, the 90% of credibility intervals resulted to be very narrow, suggesting the lacking of uncertainty or misclassification of both admixed and parental genotypes assignment and identification.

The same pattern has been observed for NEWHYBRIDS's assignment, whose membership probabilities proved to be very stringent (qi > 0.99) and accurate to assign individual to the relative hybrid classes (table tot *PaperIII*).

Evidences of the admixture outperforming power of the 151 most informative SNPs have been proved by analyzing the same sample set with 38 STRs that were previously used to reconstruct biogeographical patterns in Italian and European wildcats populations (see *PaperI* and *PaperII*). By comparing the results obtaining with both marker set, several limit were disclosed in the STRs STRUCTURE's clustering analyses (see Supplementary Table S3 in *PaperIII*).

10 previously identified admixed European wildcats (known hybrids in Pierpaoli *et al.* 2003) were confirmed as admixed, and were mostly assigned to their known hybrid category, with the exception of two individuals that remained unclassified due to their high partition in two categories (namely, BxFSI and F2). While 9 putative admixed European wildcats (Mattucci *et al.* 2013; Oliveira *et al.* in prep) were clearly classified with high posterior probabilities. In both known (n = 10) and confirmed putative hybrids (n = 9) the range of their membership probabilities to the wildcat cluster and related CI resulted to be extremely wide with STRs when compared to SNPs values.

Moreover, 10 individuals previously misclassified as putative hybrids based on STRs assignment, using the panel of most informative SNPs were identified as 'pure wildcat' from central-southern Apennines and Maremma (ID 67, ID 627, ID 677, ID 77, ID 123, ID 1250, ID 1252), and as 'pure domestic cats' from southern-western Germany (ID 421), Portugal (ID 706), and Spain (ID 717), see Supplementary Table S3.

Finally, we identified a total of 19 hybrids individuals of which 10 represent the known hybrids previously identified with a panel of 12 STRs by Pierpaoli *et al.* (2003) and used as control, and 9 represent the putative hybrids preliminarily identified in *PaperI* and *PaperII* with a panel of 35 and 38 STRs and here confirmed as admixed by taking advantage of 151 most highly informative SNPs.

All 10 known hybrids were collected in the southern Appennines and north-eastern Alps in Italy, while the new 9 hybrids were collected in Portugal (n = 2), Bosnia & Herzegovina (n = 1), Luxembourg (n = 1), southern western Germany (n = 1), and in the Italian southern and Central Apennines and Maremma (n = 4).

Even though our small sampling is not representative of the wide genetic macro populations of wildcats extant in Europe (see *PaperII*), globally we disclosed with SNPs different levels of hybridizations in line with the crossbreeding estimates detected with 8-27 STRs microsatellites panels in previous studies by Oliveira *et al.* (2008) for Iberian Peninsula; by Randi *et al.* (2001) and Lecis *et al.* (2006) for Italy and North-eastern Alps; by Eckert *et al.* (2010) and Hertwing *et al.* (2009) for southern-western Germany; and finally by Hertwing *et al.* (2009) for eastern Germany (see Table below).

Sampling locations	N putative FSI	Current HYB	Previously HYB	References
Iberian Peninsula	21	2 (9.5%)	6.9%	Oliveira *et al*. 2008 a,b
Italy	27	4 (14.8%)	2% - 8%	Randi *et al.* 2001; Lecis *et al.* 2006
NE Alps + Balkans	35	1(3%)	2% - 8%	Randi *et al.* 2001; Lecis *et al.* 2006
SW Germany + central Europe	24	2(8.3%)	3%- 18%	Eckert *et al.* 2010; Hertwing *et al.* 2009
E Germany	10	0	4,2%	Hertwing et al. 2009

Although the level of hybridization detected in this work is higher than other European studies, it is important to bear in mind that quantity and quality of molecular loci, but also sampling strategies and statistical models vary greatly among studies. For example, the estimation of genetic cluster, the assignment of individuals and the proportion of cats assigned as admixed or parental are sensitive to the genetic diversity and structure of the population analyzed (Vaha & Primmer 2006; Randi 2008). In addition, the representation of parentals' variability is crucial to correctly identify unstable individuals' assignment (Falush *et al.* 2003), such as backcrossed individuals. Hence, to assess any comparison of hybridization levels in the main five macro populations or 10 subpopulations of European wildcats detected in *PaperII*, it would be needed extend the Bayesian admixture analysis of this work, by means of 151 most informative SNPs, for all the European wildcats genetically and geographically clustered in *PaperII*.

NEWHYBRIDS's assignment suggest the recent origin of the 9 hybrids detected. Two individuals from Portugal were, in fact, classified as F2, while the remains were identified as backcrosses with wildcat relates. The limit of the software to investigate individual ancestry older than first or second generation backcrosses, leave open the possibility that a recent crossbreeding might be occurred or, most likely, an ancient crossbreeding might have generated first generation hybrids that progressively mated with the wildcat parental groups or with other hybrids, leaving a trace of introgression inside the wild gene pool.

Thus, SNPs proved to be more reliable than STRs for admixture detection, providing narrow IC ranges and individual membership probabilities clearly partitioned between wild and domestic clusters, even if the reference cats and the putative hybrids belonging to different genetic-geographic populations (Mattucci *et al.* in prep).

Recently, Nussemberg *et al.* (2013), by taking advantage of high-throughput sequencing of reduced representation libraries, developed a diagnostic marker set containing 48 SNPs ($F_{ST} > 0.8$), which

allowed a reliable identification of first and second generation hybrids and 86.5% third generation hybrids. However he used only 42 cat samples collected from different breeds and regions of Switzerland. Since a good representation of parentals' variability influenced the correct allocations of backcrossed individuals (Falush *et al.* 2003), and results shown in *PaperII* evidenced the strong genetic diversity and structure of wildcat populations in Europe, we analyzed a sampling of 150 individuals collected from all the five macro populations detected in *PaperII* in order to ensure adequate representation of genetic diversity in both wild and domestic reference samples. Moreover, we used all the 44 domestic and 89 wild cats, preliminarily identified with STRUCTURE, to generate simulated genotypes needed for the SNPs power assignment, avoiding the bias occurred in Nussemberg *et al.* (2013) to overestimate the SNPs power for hybrid identification by enriching the samples with individuals more differentiated that average.

In fact, the choice of highly differentiated traits/loci from just a small panel of individuals and genomes has been referred as a possible reason to overlook differential population structure in non-diagnostic/less differentiating traits (Brumfield *et al.* 2003; Schlotterer 2004; Morin *et al.* 2009), and to distort our perception of the actual levels of introgressive hybridization in nature (Yuri *et al.* 2009).

The reliability on introgression assignments in natural populations of wildcats might be promising improved identifying SNPs displaying genomic signals of selection between wild and domestic cats. In this context, it might be interesting to analyze the genetic variants that generate the wide morphological, physiological and behavioural diversity occurring among domesticates. For example, since basal morphology in European wildcats is unchanging, specific mutations determining variable patterns in domestic cats are very rare or absent in natural populations, and alternative variants or random polymorphism are expected. At the same time, explicit genetic variants might benefit the way of living in nature and may have remained fixed in the wild populations (e.g. Camouflage patterns crucial for hiding and hunting behaviour), while variable genetic expressions might be maintained in domestic individuals due to relaxation of selective pressures. Analysing levels of genetic diversity in mutations known to have been under different types of selection during domestication and/or breeds' improvement may, thus, revolutionize wildcat hybridization studies.

5.3.2. The combining use of most informative autosomal SNPs with uniparental markers (mtDNA and Y-linked markers)

Over the past years, numerous previous studies in various distribution regions have recommended the use of several different genetic markers to better assess the introgressive degree of wildcats populations and provide a fine detection of historic phylogeography and demographic processes between or within the subspecies (French *et al.* 1988; Hubbard *et al.* 1992; Hille *et al.* 2000; Beaumont *et al.* 2001; Daniels & Corbett 2003; Pierpaoli *et al.* 2003; MacDonald *et al.* 2004; Kitchener *et al.* 2005; Oliveira *et al.* 2008a,b). Hertwing *et al.* (2009) in Germany and Randi *et al.* (2001) in Italy, first experimented the combined use of microsatellites and mitochondrial markers. By taking advantage of the highly mutation rate of STRs (Vähä & Primmer 2006) and the slow evolution rate of mtDNA (Gottelli *et al.* 1994; Ward *et al.* 1999; Adams *et al.* 2003), they tried to identified recent interbreeding and traces of historical splitting events. However, the presence of sharing haplotypes among wild and domestic cats and the low resolution power of STRs to identify backcrosses, suggested the needed to develop a suitable set of markers able to improve the resolution of past demographic and admixture events. Recently, Driscoll *et al.* (2011) tried to overcome the uncertainty of previously reported mitochondrial identifications, providing a subspecies informative mitochondrial DNA (mtDNA) markers. In particular, he disclosed the presence of 11 fixed nucleotide positions over 2,604 bp inside the ND5/ND6 genes that proved to be extremely efficient for distinguishing domestic and wild cats (Driscoll *et al.* 2007), for reconstructuring the phylogenies of *Felis silvestris* and the origin of domestication, and for assessing the mtDNA lineage purity of reference cats in Scotland (McEwing *et al.* 2009).
Great advances for the Felidae phylogeny and the valuable patrilineal perspective in species evolution in the cat family were further obtained by sequencing several genes and SINEs within the Y chromosome (Pecon-Slattery & O'Brien 1998; Pecon-Slattery *et al.* 2000a; Pecon-Slattery *et al.* 2000b, Pecon-Slattery *et al.* 2004). Recently, Luo *et al.* (2007) identified 4 intronic DNA markers (*SMCY3*, *SMCY7*, *UTY11*, and *DBY7*) and 1 microsatellite (SMCY-STR) to describe intraspecific male lineage polymorphism in Felidae species, and thus, to discriminate wildcat from domestic cats.

We developed the fist suitable integrating multi locus panel set, combining 151 most informative SNPs, with two mitochondrial markers (specifically the ND5 region and part of the control region (Driscoll *et al.* 2007; Freeman *et al.* 2001; Tiedemann *et al.* 1996) and two Y-linked markers (Pecon-Slattery *et al.* 2004; King *et al.* 2007; Luo *et al.* 2007; Nussberger *et al.* 2013), see *PaperIII*.

Mitochondrial sequencing and chromosome Y characterization were computed in order to first assess the power of distinctiveness between wild and domestic cat subspecies.

Mitochondrial results evidenced the presence of two main haplogroups (D and W), separated by seven known mutations (F_{ST} = 0.845). In particular, 10 haplotypes proved to be distributed mainly in domestic cats (haplogroup D), while 8 haplotypes resulted distributed only in EU wildcats (haplogroup W).

Crossing this data with the SNPs Bayesian assignments, we disclosed that the haplogroup W mainly included individuals previously classified as pure wildcat (94%) and as wild x domestic hybrids (6%). On the contrary, the haplogroup D revealed the presence of 52% of samples identified as domestic cat by the SNPs analyses, 13% as hybrid, and 35% as pure wildcats.

Further network analyses computed on the Control Region revealed the presence of three major haplogroups (1, 2 and 3; F_{ST} = 0.75) that confirm the genetic differentiation of wildcats and hybrids individuals found in haplogroup W (and now regrouped in haplogroup 1), from all the others. On the contrary, individuals previously found in haplogroup D were splitted into haplogroup 2 and the derivated haplogroup 3.

Concerning the Chr Y markers, two mutations at genes SRY (SNP) and two alleles at SMCY-7 (STR) were disclosed at a very different frequencies in domestic and wild cats. Crossing these results with the SNPs assignation, we found that all samples identified as pure wildcat except 1 showed the haplotype WY; and the majority of samples assigned as domestic cat except 2 were characterized by haplotype DY; all hybrid individuals presented both the haplotype equally.

Once we verified the diagnostic power of the uniparental markers to distinguish wild from domestic cats, we further investigated the ancestry of known and putative admixed individuals, by crossing SNPs STRUCTURE's assignments with mitochondrial and Y subspecies identifications. In particular, the chromosome Y markers disclosed distinctive polymorphism diagnostic for wild and domestic male cats identification. While both mitochondrial sequenced regions (ND5 and part of the control region), revealed the presence of domestic haplotypes sharing between the two subspecies, as a consequence of i) introgressive events occurred in the past, most likely between a domestic female and a wild male, or ii) a marker not reliable and diagnostic to distinguish wild from domestic cats and to investigate on introgressive episodes.

The integration of both uniparental and nuclear markers assignments provided a complete insight of introgression level in wildcats populations (see Supplementary Table S4). Finally, all 10 known hybrids were confirmed as admixed and a total of 11 hybrids were identified among the 19 'putative' hybrids. The 11 hybrids include both the admixed individuals detected with high

informative SNPs (see paragraph above), and three more cases showing some incongruences between SNPs and uniparental markers assignment. In particular, samples ID 123 from central Apennines and Maremma (Italy) and ID 1250 and ID 1252 from southern Apennines (Italy), disclosed a membership probability to the wildcat cluster ranging from 0.871 to 0.890 (with 90% CI ranging from 0.834 to 0.920), with a domestic mitochondrial assignment, suggesting the occurrence of an occasional gene flow between wild and domestic cats during their long history of coexistence in Europe (see also Driscoll *et al.* 2007).

Moreover, the individual ID 700 from Portugal, clearly identified with both SNPs and STRs markers as admixed, has not been considered in the final hybrids count (n = 11) being lacking of uniparental informations and should be further analyzed improving the markers performance.

Ideally, the identification of the first generation hybrids (in particular F1, F2) should be achieved by the presence of nuclear membership probabilities partitioning in both parental clusters and the complementary wild and domestic uniparental assignment. However, the repeatedly crossbreeding occurred in the past with both parental groups and admixed individuals (Bewick 1807; Driscoll *et al.* 2009; Driscoll *et al.* 2007; Suminski 1962) and the fertility of wild x domestic hybrids, might have diluted the proportion of domestic alleles through the generations into the wildcat gene pool, leaving traces only in non-recombining mtDNA or Y chromosome regions.

Under this light, it appears crucial the use of the multi-locus panel to detect all introgression traces in both mitochondrial and nuclear genome, provide a support to establish origin and level of introgression and furnish wide-integrated frames of evolutionary history of this endangered subspecies.

Our work provides a convincing demonstration of the enormous potential of the combined marker systems improving the identification of ancient introgressive events and selecting reference cat to use for fine hybridization analyses, and for improving high-throughput sequencing of wildcats, as suggested by Nussemberg *et al.* (2013).

Moreover, this multi-locus panels set might be extremely useful for a better and more comprehensive evaluation of the reported genetic partition disclosed in Italy and Europe (*PaperI* and *PaperII*), and for reconstructing both the matrilineal and patrilineal evolution of wildcats

5.4 Future perspectives and final remarks of the PhD work

Studies on evolutionary processes in natural populations have been greatly enabled by technological advances related to whole genome sequence data, from a variety of domesticated species (Allendorf

et al. 2010). Access to a large number of loci, often with annotated positions within the genome of the investigated species, may allow to overcome methodological limitations associated with the analysis of a small number of genetic markers. In particular, great advances may be achieved for i) reconstructing the admixture patterns among closely related species (vonHoldt *et al.* 2011, Miller *et al.* 2012), ii) detecting nuclear genotypes that may be associated with introgression and / or phenotypic selection that occurred during domestication, iii) identifying the genetic basis of parallel adaptations (Hohenlohe *et al*. 2010, Zulliger *et al.* 2013), and iv) investigating demographic effects of past climate change (Miller *et al.* 2012, Zhao *et al.* 2013).

One of the main application of the high-resolution marker set, that over the past ten years has become available also in cat subspecies, may be the discovery of informative genomic signals of selection between wild and domestic cats, aiming at enforcing thoughtful knowledge on European wildcats. When two populations are, in fact, subjected to different selective pressure, as occurred in wild and domestic cats, some level of natural and artificial selections are expected to cause divergence in different parts of their genome. By scanning several independent loci, it might be possible display atypical patterns of non-neutral diversity, and use these outliers loci to assess differential rates of introgression over the entire genome. A number of approaches might be used to detect this under-selection traits (Kohn *et al*. 2006).

For example, the candidate gene approach has founded great application in the identification of MHC genes (Hedrick 2001; Aguilar *et al.* 2004), which resulted to be associated with individual variation in parasite load (Madsen *et al*. 2004), local adaptations (Evans *et al.* 2010), maternal-fetal interactions (Hedrick *et al.* 1988), life-time reproductive success (Kalbe *et al*. 2009), and mate choice (Penn *et al.* 1999; see Ujvari & Belov 2011 for more references).

Recently, studies on cat MHC sequences (Yuhki *et al.* 2007, Morris 2009), provided a powerful and promising insight into feline hereditary and infectious disease models (such as Feline Leukemia Virus (Hoover & Mullins 1991), Feline Immunodeficiency Virus (Pederson *et al*. 1987), and Feline Infections Peritonitis Virus (Andrew 2000)) and humans viral diseases (namely, Human T Cell Leukemia and Acquired Immunodeficiency Syndrome (AIDS) (Winkler *et al*. 1989)).

On the other hand, selective sweep approaches lead to identify phenotypic traits that might have left a genetic signature in the genome (Luikart *et al.* 2003; Vasemagi & Primmer 2005), by using a wide genome scan and the low-variation surrounding the gene under selection (Pollinger *et al.* 2005). The great advantage of this method is the possibility to use molecular markers without prior knowledge about phenotypic traits under selection (Shikano *et al.* 2010). Therefore, it is especially suitable for

organisms in which breeding experiments are difficult to conduct and for identifying signatures of selection even on physiological traits. Over the past years, an increasing number of studies have conducted mapping selective sweeps in a variety of organisms (reviewed in Holderegger *et al.* 2008; Stinchcombe & Hoekstra 2008; Nosil *et al*. 2009), proving their feasibility also for finding genes under intense selection during domestication (Berry *et al.* 1991; Begun & Aquadro 1992; Kohn *et al.* 2000; Matsuoka *et al.* 2002; Schloetterer 2002; Luikart *et al.* 2003; Akey *et al.* 2004; Storz 2005; Wright *et al.* 2005).

In this context, the analysis of genetic variants that are responsible for the wide morphological, physiological and behavioural diversity occurring among domesticates might be particularly interesting. For example, since basal morphology in European wildcats is unchanging, specific mutations determining variable patterns in domestic cats are very rare or absent in natural populations, and alternative variants or random polymorphism are expected. At the same time, explicit genetic variants might benefit the way of living in nature and may have remained fixed in the wild populations (e.g. Camouflage patterns crucial for hiding and hunting behaviour), while variable genetic expressions might be maintained in domestic individuals due to relaxation of selective pressures. Analysing levels of genetic diversity in mutations known to have been under different types of selection during domestication and/or breeds' improvement may, thus, revolutionize wildcat hybridization studies.

Another method is the LD mapping approach that uses marker loci inherited in a non-random fashion (Przeworski 2002; Kim & Nielsen 2004; Palaisa *et al.* 2004) and that found application on the investigation of pig domestication and breeds history (Amaral *et al.* 2008). The first autosomal genetic linkage (GL) map generated in a domestic cat pedigree (Menotti-Raymond *et al.* 2009), have offered a valuable resource for the detection of implicated molecular variants and their biochemical and physiological impacts. Recent advances proved that this genomic tool may be useful also for disease-association studies in dogs (Goldstein *et al.* 2006) and cats (Alhaddad *et al.* 2013), as the LD within breeds largely reflects their population history and breeding strategies. Promising outcomes might be expected by analyzing LD decay in both wild and domestic cats populations to carefully explore their admixture inference levels, since the LD registered in wild populations of canids (Gray *et al.* 2009), mice (Laurie *et al*. 2007) and rabbits (Carneiro *et al.* 2011) proved to decade faster than the domestic relates one.

In search of molecular tools displaying genomic signals of selection between wild and domestic cats, it appears crucial the methods to apply and inferences to be made. Since both wild and domestic cats represent hierarchically substructured populations (as shown in *Paper I* and *Paper II*), that suffered significant demographic changes and shared a common recent genetic evolution, it's

recommended to primarily improve the representativeness of the entire population of both subspecies, allowing the possibility to identify the most informative under selection traits suitable also for distinguishing wild from domestic and identifying hybrids.

The second main application of the potentially high-resolution marker set may be, in fact, the discovery of molecular tools efficient in improving the detection of admixed individuals and hybridization degree. In this work, we proved that a multi locus approach, combining autosomal informative SNPs with uniparental markers, may help to decipher the domestication of the cat and the dynamics of wild and domestic cat populations around the world. However, many others aspects of the intricate complex of introgressive hybridization should be deeply investigated. For example, it would be important assess which genes are introgressed and how they influenced the ecological features (e.g. habitat selection, feeding ecology, individuals' home range) of populations, which is the sexually direction of domestic alleles during the introgression. Moreover, in a conservation oriented view, it would be crucial establish the fitness of admixed individuals, in comparison with the parentals' one.

Another advantageous application of using multiple loci released by the next generation sequencing (NGS) technologies may be the possibility to infer wildcat populations histories by reconstructing their phylogeography and phylogenetics evolution roots. However, until now, the use of NGS tools in this discipline has been limited by the low cost-effectiveness to sequence orthologous DNA fragments for many individuals, and the difficulty to estimate demographic parameters by using SNPs with insufficient variability for modeling gene coalescence (Kuhner 2009; Pinho & Hey 2010). However, recently, promising protocols have been proposed and developed to overcome previous drawbacks. For example, sequencing a subset of the genome, i.e., a reduced representation genomic library (Barbazuk *et al.* 2005) and combining individuals into a single run (Glenn 2011), might allow a financially and computationally feasible alternative (McCormack *et al.* 2011). Moreover, producing longer reads with the advent of third generation sequencing platforms (e.g., PacBio, Ion Torrent, Starlight) might facilitate gene tree analysis and allow the possibility to use SNPs for testing demographic hypotheses by involving gene flow (Durand *et al.* 2011; McCormack *et al.* 2011).

The work presented in this thesis has been developed by taking in account the main conservation threats of this endangered feline and the needed to overcome previous drawbacks in European wildcats' researches, in order to yield important insights for the supporting conservation guidelines

of the subspecies. The threatening introgression of domestic cat genes into European wildcat (*Felis silvestris silvestris*) populations, concomitantly with habitat loss and fragmentation, resulted in a strong reduction of wildcats' range in Europe that involved important demographic declines and high levels of isolation. Both events might have strongly jeopardized wildcat's fitness and evolutionary survival. Thus, we address the questions related with the artificial hybridization and populations' fragmentation by developing a multi locus analysis protocol and by increasing the sample over the wide subspecies' home range distribution in Europe. In particular, we focused on improving the identification of hybrids and introgression level, considering the deep genetic-geographic substructure of wildcat populations in Europe and the absent of a barrier to crossbreeding among hybrid and parental individuals.

Thus, the conservation of this endangered feline, may not refrain to evaluate the anthropogenic consequences of habitat destruction and urbanization on small and isolated fragmented populations, in addition to the role that hybrids have in maintaining hybridization levels. Although evolutionary mechanisms, such as the maintaining of genetic and morphological integrity of most wildcat populations in Europe, should have been happened to prevent the complete amalgamation with domestic relatives, pure wildcat populations without at least genetic traces of past introgression no longer seem to exist in most parts of the wildcats distribution area in Central Europe. Hence, conservation efforts and legal protection should focus on saving the local functional wildcats (see also Daniels & Corbett 2003), for instance, protecting the environmental conditions that favor pure wildcats, mapping the distribution of non-introgressed natural populations, and enforcing controls of the genetic status of game stocks reproduced in captivity and potentially usable for restocking (Randi 2008).

Several molecular studies on wildcats' admixture revealed the presence of diverse degrees of hybridization in Europe, suggesting that hybridization does not constitute a uniform threat throughout the entire range of the subspecies. Thus, until the situation is not entirely compromised with the formation of a hybrid swarm that lead to the disappearance of the 'pure' parental genomes (Mallet 2005), it should be taken in consideration the possibility to accept a proportion of admixture to preclude protection of wildcat individuals (Allendorf *et al.* 2001). However, setting and arbitrary threshold appears to be extremely hard because of i) the limited power of diagnostic markers often used to detect introgression levels, ii) the difficulty to distinguish between a small proportion of admixture (e.g. <5%) and natural polymorphisms, and iii) the erosion of the genetic integrity of the parental taxon by constantly lowering the definition of 'pure' (Allendorf *et al.* 2001). Moreover, in order to assess the potential value of a hybridized population, it should be considered how many pure populations of the taxon would remain, and how much the continued existence of hybridized

populations would affect the survival of pure ones. So that, the better solution would be define an acceptable introgression threshold based on the biology of the species or populations under study.

Although our work represents a notable effort on wildcat's genetic research, it can only provide a first insight to understand how European wildcat genome is being affected by ongoing hybridization and fragmentation. Thus, it is recommended to implement this study in order to point out crucial questions regarding i) which and how domestic genes introgress and threaten the subspecies survival, ii) which ecological and genetic variables may influence and be influenced by admixture, iii) if hybridization may help natural population to adapt to the highly humanized European environment by introducing new genetic variation that might improve the population fitness (Allendorf *et al.* 2001) instead of promoting outbreeding depression.

5.5 References

Adams KL, Cronn R, Percifield R, Wendel JF (2003) Genes duplicated by polyploidy show unequal contribution to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences USA*, **100**, 4649-4654

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences USA*, **101**, 3490-3494

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biology*, **2,** 1591-1599

Alhaddad H, Khan R, Grahn RA, Gandolfi B, Mullikin JC, et al. (2013) Extent of Linkage Disequilibrium in the Domestic Cat, *Felis silvestris catus,* and Its Breeds. *PLoS ONE*, **8(1)**, e53537

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613-622

Allendorf FW, Luikart G (2007) Conservation and the Genetics of Populations. Blackwell Publishing, Malden MA, USA

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11(10)**, 697-709

Amaral AJ, Megens H-J, Crooijmans RPMA, Heuven HCM, Groenen MAM (2008) Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, **179**, 569-579

Andrew SE (2000) Feline infectious peritonitis. *Veterinary Clinics of North America: Small Animal Practice*, **30(5)**, 987-1000

Barbazuk WB, Bedell JA, Rabinowicz PD (2005) Reduced representation sequencing: A success in maize and a promise for other plant genomes. *Bioessays*, **27**, 839-848

Beaumont M, Barratt EM, Gottelli D, Kitchener AC, Daniels MJ, Pritchard JK, Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat.*Molecular Ecology*, **10**, 319-336

Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in Drosophila melanogaster. *Nature*, **356,** 519-520

Berry AJ, Ajioja JW, Kreitman M (1991) Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. *Genetics*, **129**, 1111-1117

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London*, **B265**, 1219-1226

Brumfield R, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249-256

Carneiro M, Blanco-Aguiar JA, Villafuerte R, Ferrand N, Nachman MW (2010) Speciation in the European rabbit (*Oryctolagus cuniculus*): islands of differentiation on the X chromosome and autosomes. *Evolution*, **64**, 3443-3460

Cornuet JM and Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014

Daniels MJ, Corbett L (2003) Redefining introgressed protected mammals: when is a wildcat a wild cat and a dingo a wild dog? *Wildlife Research,* **30(3)**, 213-218

Deffontaine V, Libois R, Kotlik P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux J (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology*, **14**, 1727-1739

Dowling TE, DeMarais BD, Minckley WL, Douglas ME, Marsh PC (2002) Use of genetic characters in conservation biology. *Conservation Biology*, **6**, 7-8

Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, Harley EH, Delibes M, Pontier D, Kitchener AC, Yamaguchi N, O'Brien SJ (2007) The Near Eastern origin of cat domestication. *Science*, **317**, 519-523

Driscoll CA, Macdonald DW, O'Brien SJ (2009) From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, **106(1)**, 9971-9978

Driscoll CA, Nowell K (2010) Felis silvestris. In: *IUCN Red List of Threatened Species. International Union for Conservation of Nature* (http://www.iucnredlist.org/)

Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Molecular Biology and Evol*ution, **28**, 2239-2252

Easterbee N (1991) The Wildcat. In: The Handbook of British Mammals (eds Corbet GB, Harris S), Pp. 431-437. Blackwell Scientific Publications, Oxford

Eckert I, Suchentrunk F, Markov G, Hartl GB (2009) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mammalian Biology*, **75**, 160-174

Eckert I, Suchentrunk F, Markov G, Hartl GB (2010) Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mammalian Biology - Zeitschrift fur Saugetierkunde*, **75**, 160-174

Evans ML, Neff BD, Heath DD (2010) MHC genetic structure and divergence across populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Heredity*, **104**, 449-459

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics,* **164**, 1567-158

Fløjgaard C, Normand S, Skov F, Svenning JC (2011) Deconstructing the mammal species richness pattern in Europe towards an under standing of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. *Global Ecology and Biogeography*, **20**, 218–230

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK

Franklin IR (1980). Evolutionary change in small populations. In *Conservation biology: an evolutionary-ecological perspective:* 135–150. Soulé ME & Wilcox BA (Eds). Sunderland, MA: Sinauer.

Freeman G, Shepperd S, Robinson I, Ehrich K, Richards S (2001) Continuity of Care: Report of a scoping exercise for the NCCSDO, London: National Coordinating Centre for Service Delivery and Organisation

French DD, Corbett LK, Easterbee N (1988) Morphological discriminant functions of Scottish wildcats (*Felis silvestris*), domestic cats (*F. catus*) and their hybrids. *Journal of Zoology (London)*, **214**, 235-259

Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. *Molecular Ecology Resources*, **10**, 915-934

Germain E, Benhamou S, Poulle ML (2008) Spatio-temporal sharing between the European wildcat, the domestic cat and their hybrids. *Journal of zoology*, **276**, 1 95-203

Glenn TC (2011) Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, **11**, 759-769

Goldstein O, Zangerl B, Pearce-Kelling S, Sidjanin DJ, Kijas JW, Felix J, Acland GM, Aguirre GD (2006) Linkage disequilibrium mapping in domestic dog breeds narrows the progressive rod-cone degeneration interval and identifies ancestral disease-transmitting chromosome. *Genomics*, **88**, 541-550

Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA, Wayne R (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology*, **3**, 301-312

Gray MM, Granka JM, Bustamante CD, Sutter NB, Boyko AR, Zhu L, Ostrander E, Wayne R (2009) Linkage disequilibrium and demographic history of wild and domestic Canids. *Genetics*, **181**, 1493-1505

Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology and Evolution*, **16**, 629-636

Hedrick SM, Engel I, McElligot DL, Fink PJ, Hsu ML, Hansburg D, Matis A (1988) Selection of amino acid sequences in the p chain of T cell antigen receptor. *Science*, **239**, 1541-1544

Hertwig ST, Jungnickel A, Stepanow S, Böhle UR, Fischer MS (2009) Regionally high rates of hybridization and introgression in German wildcat populations (Felis silvestris, Carnivora, Felidae). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 283-297

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907-913

Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the* Royal Society B, **359**,183-195

Hille A, Pelz O, Trinzen M, Schlegel M, Peters G (2000) Using microsatellite markers for genetic individualization of European wildcats (*Felis silvestris*) and domestic cats. *Bonn. Zool. Beitr*, **49**, 165-176

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomic analysis of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6,** e1000862

Holderegger R, Herrmann D, Poncet B, Gugerli F, Thuiller W, Taberlet P, Gielly L, Rioux D, Brodbeck S, Aubert S, Manel S (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, **1**, 273-283

Hoover EA, Mullins JI (1991) Feline leukemia virus infection and diseases. *Journal of American Veterinary Medical Association*, **199**, 1287-1297

Hubbard AL, McOris S, Jones TW, Boid R, Scott R, Easterbee N (1992) Is survival of European wildcats Felis silvestris in Britain threatened by interbreeding with domestic cats? *Biological Conservation,* **61**, 203-208

Kalbe M, Eizaguirre C, Dankert I, Reusch TBH, Sommerfeld RD,Wegner KM, Milinski M (2009) lifetime reproductive success is maximized with optimal major histocompatibility complex diversity. *Proceedings of the Royal Society B*, **276**, 925

Kim Y, Nielsen R (2004) Linkage disequilibrium as a signature of selective sweeps. *Genetics*, **167**, 1513-1524

King TE, Parkin EJ, Swinfield G, Cruciani F, Scozzari R, Rosa A, Lim SK, Xue Y, TylerSmith C, J obling MA (2007) Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy. *European Journal of Human Genetics*, **15**, 288-293

Kitchener AC, Dugmore AJ (2000) Biogeographical change in the tiger, Panthera tigris. *Animal Conservation*, **3**, 113-24

Kitchener AC, Yamaguchi N, Ward JM, Macdonald DW (2005) A diagnosis for the Scottish wildcat (*Felis silvestris*): a tool for conservation action for a critically-endangered felid. *Animal Conservation*, **8**, 223-237

Kohn MH, Pelz HJ, Wayne RK (2000) Natural selection mapping of the warfarin-resistance gene. *Proceedings of the National Academy of Sciences* USA, **97(14)**, 7911-7915

Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *Trends in Ecology and Evolution*, **21**, 629-37

Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology and Evolution,* **24**, 86-93

Kurtén B (1965) The Carnivora of the Palestine Caves. *Acta Zoologica Fennica*, **107**, 1-74

Lande R (1995) Mutation and conservation. *Conservation Biology*, **9**, 782-791

Lande R (1998) Demographic stochasticity and Allele effect on a scale with isotropic noise. *Oikos*, **83**, 353-358

Langley PJW, Yalden DW (1977) The decline of the rarer carnivores in Great Britain during the nineteenth century. *Mammal Review*, **7**, 95-116

Lapini L (1989) Il gatto selvatico nella regione Friuli-Venezia Giulia. In: *Fauna* (*1)*, Pp. 64-67, Udine

Lapini L (2006) Attuale distribuzione del gatto selvatico *Felis silvestris silvestris* Schreber, 1775 nell'Italia nord-orientale (Mammalia: Felidae). In: *Atti del Museo Civico di Storia Naturale di Venezia* (*57)*, Pp. 221-234

Laurie CC, Nickerson DA, Anderson AD, Weir BS, Livingston RJ, et al. (2007) Linkage disequilibrium in wild mice. *PLoS Genetics*, **3**, e144

Lecis R, Pierpaoli M, Biro ZS, Szemethy L, Ragni B, Vercillo F, Randi E (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119-131

Lozano J, Virgós E, Malo AF, Huertas DL, Casanovas JG (2003) Importance of scrub-pastureland mosaics for wild-living cats occurrence in a Mediterranean area: implications for conservation of the wildcat (*Felis silvestris*). *Biodiversity and Conservation*, **12**, 921-935

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews in Genetics*, **4**, 981-94

Luo SJ, Johnson WE, David VA, Menotti-Raymond M, Stanyon R, Cai QX, Beck T, Yuhki N, Pecon-Slattery J, Smith JLD, et al. (2007) Development of Y chromosome intraspecific polymorphic markers in the Felidae. *Journal of Heredity*, **98**, 400-413

Madsen T, Ujvari B, Olsson M (2004) Novel genes continue to enhance population growth in adders (*Vipera berus*). Biological Conservation, **120**, 145-147

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, **20 (5)**, 229-237

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189-197

Martin E, Feder ME, Thomas Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Genetics*, **100**, 101-102

Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proceedings of the National Academy of Sciences, **99**, 6080-6084

McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of aphelocoma jays. *Evolution*, **65**, 184-202

McDonald DW, Daniels MJ, Driscoll C, Kitchener A, Yamaguchi N (2004) The Scottish Wildcat. Analyses for Conservation and an Action Plan. Wildlife Conservation Research Unit, Oxford.

McEwing R, Kitchener AC, Holleley C, Kilshaw K, O'Donoghue P (2012) An allelic discrimination SNP assay for distinguishing the mitochondrial lineages of European wildcats and domestic cats. *Conservation Genetics Resources*, **4**, 163-165

McGinnity P, Prodhl P, Ferguson A, Hynes R, Maoilidigh N, Baker N, Cotter D, O'Hea B, Cooke D, Rogan G, Taggart J, Cross T (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 2443-2450

McOrist S, Kitchener AC (1994) Current threats to the European wildcat, *Felis silvestris*, in Scotland. *Ambio*, **23**, 243-245

Menotti-Raymond M, David VA, Schäffer AA, Tomlin JF, Eizirik E, Phillip C, Wells D, Pontius JU, Hannah SS, O'Brien SJ (2009) An autosomal genetic linkage map of the domestic cat, Felis silvestris catus. *Genomics*, **93**, 305-313

Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Lim Kim H, Burhans RC, Drautz DI, Wittekindt NE, Tomsho LP, Ibarra-Laclette E, Herrera-Estrella L, Peacock E, Farley S, Sage GK, Rode K, Obbard M, Montiel R, Bachmann L, Ingólfsson O, Aars J, Mailund T, Wiig Ø, Talbot SL, Lindqvist C (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1210506109

Morin PA, Luikart G, Wayne RK, SNP-Workshop Group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**, 208-216

Morin PA, Martien KK, Taylor BL (2009) Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, **9**, 66-73

Moritz C (1994) Defining "evolutionarily significant units" for conservation. *Trends in Ecology and Evolution,* **9**, 373-375

Morris K (2009) The feline major histocompatibility complex. *University of Sydney Undergraduate Research Journal*, **1**, 75

Nosil P, Funk DJ, Ortiz-Barrientos D (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375-402

Nussberger B, Greminger MP, Grossen P, Keller LF, Wandeler P (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources,* **13**, 447-460

O'Brien J, Devillard S, Say L, Vanthomme H, Léger F, Ruette 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 Conservation*, **18**, 2351-2360

Oliveira R, Godinho R, Randi E, Ferrand N, Alves PC (2007) Molecular analysis of hybridisation between wild and domestic cats (*Felis silvestris*) in Portugal: implications for conservation. *Conservation Genetics*, **9**, 1-11

Oliveira R, Godinho R, Randi E, Alves PC (2008) Hybridizations versus conservation: are domestic cats threatening the genetic integrity of wildcats (*Felis silvestris silvestris*) in Iberian Peninsula? *Philosophical Transactions of the* Royal Society *B: Biological*, **363**, 2953-2961

Palaisa K, Morgante M, Tingey S, Rafalski A (2004) Long-range patterns of diversity and linkage disequilibrium surrounding the maize *Y1* gene are indicative of an asymmetric selective sweep. *Proceedings of the National Academy of Sciences USA*, **101**, 9885-9890

Parent GH (1975) La migration récente, a caractére invasionnel, du chat sauvage, *Felis silvestris silvestris* Schreber, en Lorraine belge. *Mammalia*, **39**, 251-288

Pecon Slattery J, O'Brien SJ (1998) Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. *Genetics*, **148**, 1245-1255

Pecon-Slattery J, Murphy WJ, O'Brien SJ (2000a) Patterns of diversity among SINE elements isolated from three Y-chromosome genes in carnivores. *Molecular Biology and Evolution*, **17**, 825-829

Pecon-Slattery J, Sanner-Wachter L, O'Brien SJ (2000b) Directed gene conversion occurs between mammalian X-Y homologs outside the pseudoautosomal region. *Proceedings of the National Academy of Sciences USA*, **97**, 5307-531

Pecon-Slattery J, Wilkerson AJ, Murphy WJ, O'Brien SJ (2004) Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. *Molecular Biology and Evolution*, **21(12)**, 2299-2309

Pedersen NC, Ho EW, Brown ML, Yamamoto JK (1987) Isolation of a T-lymphotropic virus from domestic cats with an immunodeficiency-like syndrome. *Science*, **235**, 790-793

Pierpaoli M, Biro ZS, Herrmann M, Hupe K, Fernandes M, Ragni B, Szemethy L, Randi E (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585-2598

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology and Systematics*, **41**, 215-230

Pilot M, Jędrzejewski W, Branicki W, Sidorovich VE, Jedrejewska B, Stachura K, Kunk SM (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, **15**, 4533-4553

Piovesan G, Biondi F, Bernabei M, Di Filippo A, Schirone B (2005) Spatial and altitudinal bioclimatic zones of the Italian peninsula identified from a beech (*Fagus sylvatica L.*) tree-ring network. *Acta Oecologica*, **27**, 197-210

Pollinger JP, Bustamante CD, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK (2005) Selective sweep mapping of genes with large phenotypic effects. Genome Research, **15**, 1809-1819

Przeworski M (2002) The signature of positive selection at randomly chosen loci. *Genetics*, **160**, 1179-1189

Ragni B (2006) Il Gatto selvatico. In: *Salvati dall'Arca. WWF Italia* (ed. Perdisa A), Pp. 35-56, Bologna

Randi E (2007) Phylogeography of South European Mammals. In: *Phylogeography of Southern European Refugia* (eds. Weiss S, Ferrand N), Pp. 101-126, Kluwer Academic Publishers, Amsterdam

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, **17**, 285-293

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

Reed DH, O'Grady JJ, Brook BW, Ballou JD, Frankham R (2003) Estimates of minimum viable population sizes for vertebrates and factors influencing those estimates. *Biological Conservation,* **113**, 23-34

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9-10

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109

Schipper ELF, Cigarán MP, Hedger MM (2008) *Adaptation to climate change: the new challenge for development in the developing world*. United Nations Development Programme (UNDP), New York, New York, USA

Schlötterer C (2002) A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics*, **160,** 753-763

Schlötterer C (2004) The evolution of molecular markers just a matter of fashion? *Nature Reviews Genetics***, 5**, 63-69

Schmitt T, Varga Z (2012) Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology*, **9**, 22-33

Seddon JM, Parker HG, Ostrander EA, Ellegren H (2005) SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology*, **14**, 503-511

Shikano T, Ramadevi TJ, Merila J (2010) Identification of Local- and Habitat-Dependent Selection: Scanning Functionally Important Genes in Nine-Spined ticklebacks (*Pungitius pungitius*). *Molecular Biology and Evolution*, **27**, 2775-2789

Sommer RS, Benecke N (2006) Late Pleistocene and Holocene development of the felid fauna (Felidae) of Europe: a review. *Journal of Zoology*, **269**, 7–19

Stahl P, Artois M (1991) Status and conservation of the wildcat (*Felis silvestris*) in Europe and around the Mediterranean rim. Convention on the Conservation of European Wildlife and Natural habitats Standing Committee. Council of Europe Nature and Environment Series, **69**, 1-76

Stahl P, Artois M (1994) Status and conservation of the wildcat (Felis silvestris) in Europe and around the Mediterranean rim. Nature and Environment Series, No. 69, Council of Europe Press, Strasbourg, Pp. 78

Stewart JR, Lister AM (2001) Cryptic northen refugia and the origins of modern biota. *Trends in Ecology and Evolution*, **16**, 608-613

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158-170

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14,** 671-688

Strandberg G, Brandefelt J, Kjellstrom E, Smith B (2011) High resolution regional simulation of last glacial maximum climate in Europe. *Tellus*, **63A**, 107-125

Suminski P (1962) Research in the native form of wildcat (*Felis silvestris* Schreber) on the background of its geographical distribution. *Folia Forestalia Polonica*, **8**, 1-81

Tiedemann R, Harder J, Gmeiner C, Haase E (1996) Mitochondrial DNA sequence patterns of harbour porpoises (*Phocoena phocoena*) from the North and the Baltic Seas. *Z. Säugetierkd*, **61**, 104-111

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends of Ecology Evolution*, **16**, 330-343

Ujvari B, Belov K (2011) Major histocompatibility complex (MHC) markers in conservation biology. *International Journal of Molecular Science*, **12**, 5168-5186

Waits LP, Paetkau D, Strobeck C (1998) The genetics of the bears of the world. In: Bear Conservation Act (ed. Servheen C), Pp. 25-32. IUCN, Gland

Waples RS (1991) Pacific salmon, Orcorhynchus spp. and the definition of "species" under the Endangered Species Act. *Marine Fisheries Review*, **53**, 11-22

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167-184

Waples RS, England PR (2011) Estimating contemporary effective population size based on linkage disequilibrium in the face of migration. *Genetics*, **189**, 633-644

Ward TJ, Bielawski JP, Davis SK, Templeton SW, Derr JN (1999) Identification of domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA markers and the parametric bootstrap. *Animal Conservation*, **2(1),** 51-57

Weiss S, Ferrand N (2006) Phylogeography of Southern European Refugia. In: *Evolutionary perspectives on the origins and conservation of European biodiversity,* Pp. 377, Springer

Winkler G, Maxwell SE, Ruemmler C, Stollar V (1989) Newly synthesized dengue-2 virus nonstructural protein NSI is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. *Virology*, **171**, 302-305

Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS (2005) The effects of artificial selection on the Maize genome. *Science*, **308,** 1310-1314

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63-72

Vasemägi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14(12)**, 3623-42

Vogler AP, DeSalle R (1994) Diagnosing units of conservation management. *Conservation Biology*, **6**, 170-178

vonHoldt B, Pollinger J, Earl D, Knowles J, Boyko A, et al. (2011) A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Research,* **21**, 1294-1305

Yuhki N, Beck T, Stephens RM, Nishigaki Y, Newmann K, O'Brien SJ (2003) Comparative genome organization of human, murine, and feline MHC class II region. *Genome Research,* **13,** 1169-1179

Zhao CY, Wang Y, Zhou XZ, et al. (2013) Changes in climatic factors and extreme climate events in Northeast China during 1961-2010. *Advances in Climate Change Research*, **4(2)**, doi: 10.3274/SP.J.1248.2013.092

Zulliger D, Schnyder E, Gugerli F (2013) Are adaptive loci transferable across genomes of related species? Outlier and environmental association analyses in Alpine Brassicaceae species. *Molecular Ecology,* **22**, 1626-1639

ACKNOWLEDGEMENTS

Ed ora veniamo a noi!!!

L'aver speso tutte le energie per elaborare, scrivere ed assemblare questo lavoro, ha di fatto penalizzato quella che, ahimè, resta la parte più letta, dopo ovviamente figure ed immagini a colori.

Ma siccome qui mi limiterò a scrivere poche righe essenziali, sarete costretti a leggervi la tesi..mi spiace!!!

Il primo ringraziamento doveroso va al boss, Ettore, che ha reso possibile questo mio viaggio supportandomi (e sopportandomi) scientificamente ed economicamente (tramite i vari contratti ISPRA) per tutta la durata del dottorato, permettendomi di crescere e di mettermi in discussione sin dall'inizio. L'essere riuscita a svolgere sei mesi di formazione presso UC Davis, ha inevitabilmente modificato la mia visione del 'fare ricerca', avvicinandomi ad un mondo che sebbene molto competitivo è estremamente professionale, dinamico e meritocratico!!! E per questo, oltre al boss, devo ringraziare tanto la Leslie, che mi ha accolto nel suo gruppo di ricerca coinvolgendomi in tutte le attività (lavorative e non), quanto il suo staff…in special modo Razib (il mio compagno di ufficio e di chiacchere surreali), Hasan (che ha cercato fino all'ultimo di tenermi a bada prima di mollare la spugna e capire che sono totalmente anarchica ☺) e Barbara (che mi ha inserita nella UC Davis Spanish community e mi ha fonito il supporto per soprvvivere in lab e al di fuori☺)!!!

Quella californiana, per quanto breve, è stata in assoluto l'esperienza più intensa della mia vita, dottorato compreso, per tutte le persone meravigliose che ho conosciuto, i posti visitati, le abitudini acquisite, per tutte le chiacchiere a bordo piscina, le cinnamom rolls che magicamente comparivano a colazione, i km 'biciclettando' nella perfetta cittadina di Davis, le attività di dog-sitter e gardening che puntualmente finivano in birre-chiacchiere e "ma che fine ha fatto il cane??", i BBQs organizzati ovunque e per ogni occasione dalla Spanish Community, per tutto il supporto che ho ricevuto, per essermi sentita parte di una comunità fatta di persone di tutte le etnie e nazionalità!!! Dovrei nominarli tutti ma sarebbero in troppi…alcuni di loro me li porto con me…ovunque!!!

E poi c'è Bologna..anzi Ozzangeles…ne son passate veramente tante di persone negli ultimi anni… alcune di queste hanno lasciato qualcosa di veramente speciale…un grazie va alla mia amica Alice (compagna di chiacchiere e bevute in terra belga e italiana) e alla grande Rita (con la quale ho avuto il piacere di lavorare e soprattutto di condividere tante esperienze)!!! Un ringraziamento va anche a Nadia, Marco, Francesca, Chiara, Alessandra, Valentina, Edoardo, Luca, Patrizia che con professionalità e spontaneità hanno contribuito e contribuiscono a rendere ogni giorno il 'moody and crazy lab' di Ozzangeles un 'normale posto di lavoro'!!! Per Edo c'è un ringraziamento

speciale anche per il lavoro svolto sul gatto negli ultimi due anni e per il supporto dato durante le elaborazioni e durante il talk a Belfast!!!

Manca qualcosa per concludere il quadro…un ringraziamento speciale va a Chiara e Francesca, compagne di sventura, chiacchiere, risate, discussioni e chi più ne ha più ne metta…perchè avete vissuto ogni singola giornata di questo mio percorso…tutti i miei deliri durante le elab, gli alti ed i bassi che inevitabilmente contrassegnano la vita di ognuno di noi, con assoluto rispetto avete assecondato le mie grandi e rivoluzionarie 'pensate' sperando che l'indomani mi fossero passate, avete sopportato il caldo e poi il freddo e poi l'oust e poi le mie fasi di pulizia maniacale seguite da quelle di caos più totale…avete condiviso con me il dramma delle blatte e delle scatapuzze!!!Ed un grazie anche e soprattutto per tutto il supporto logistico e psicologico di quest'ultimo anno!!!

Ed infine..un grazie speciale va alla mia grande famiglia e agli amici di sempre (sparsi ovunque tra l'Italia e l'Europa) per essere stati con me in ogni istante di questo lungo viaggio e per essere accorsi in ogni modo quando avevo bisogno!!!

Oky…ora che sono finiti i ringraziamenti..tornate alla prima pagina per leggere anche la tesi ☺ !!!