Unione europea
Fondo sociale europeo

REPUBBLICA ITALIANA

REGIONE AUTÒNOMA DE SARDIGNA
REGIONE AUTONOMA DELLA SARDEGNA

FSE 2007-2013
POR SARDEGNA
obiettivo competitività regionale e occupazione

A.D. MDLXII

UNIVERSITÀ DEGLI STUDI DI SASSARI

Dissertation for the Degree of Doctor of Philosophy in Environmental Biology
presented at Sassari University in 2013

XXVI cycle

RELATIONSHIPS BETWEEN ENVIRONMENTAL VARIABLES AND GENETIC STRUCTURE OF WILD UNGULATE POPULATIONS

PH.D. CANDIDATE: ***Dr. Daniela Biosa***

DIRECTOR OF THE SCHOOL: ***Prof. Marco Curini Galletti***

SUPERVISOR: ***Dr. Massimo Scandura***

CO-SUPERVISOR: ***Prof. Marco Apollonio***

La presente tesi è stata prodotta nell'ambito della scuola di dottorato in Scienze della Natura e delle sue Risorse dell'Università degli Studi di Sassari, a.a. 2010/2011–XXVI ciclo, con il supporto di una borsa di studio finanziata con le risorse del P.O.R. SARDEGNA F.S.E. 2007-2013 - Obiettivo competitività regionale e occupazione, Asse IV Capitale umano, Linea di Attività l.3.1.

INDEX

Summary

[ENGLISH] Understanding how environmental modifications, and especially those caused by humans, influence biological systems represents the first step toward the establishment of effective approaches for the conservation of global biodiversity. However, aspects concerning the genetic consequences of such changes would require the combination of knowledges on population genetics, eco-ethology, spatial and landscape ecology, and are still improving, thank to the increase of studied cases and to the development of new statistical approaches. The Landscape Genetics tries to integrate approaches of population genetics with landscape ecology with the aim to explain the observed genetic patterns.

Here I analyzed the genetic variability and structure of populations of the two most abundant ungulate species in Europe, wild boar (*Sus scrofa*) and roe deer (*Capreolus capreolus*). In each species, genetic variation was analyzed at two different geographical scales, highlighting how climate, land use and landscape features, differently and together with direct human action, can influence gene flow. Phylogeographic patterns in the European continent seem to be influenced mostly by past climatic events (Last Glacial Maximum), while local genetic differentiation results from local landscape characteristics and human management of populations.

Key words: Landscape Genetics, gene flow, spatial behaviour, genetic structure, environmental barriers

Riassunto

[ITALIANO] Comprendere il modo in cui le modificazioni ambientali operate dall'uomo influenzano i sistemi biologici rappresenta il primo passo verso l'individuazione di efficaci approcci per la conservazione della biodiversità a livello mondiale. Tuttavia, gli aspetti inerenti le conseguenze genetiche di tali cambiamenti prevedono l'integrazione di conoscenze di genetica delle popolazioni, di ecolo-etologia e di ecologia spaziale e del paesaggio, e sono ancora in fase di affinamento, grazie all'aumento di casi di studio descritti e allo sviluppo di metodologie statistiche per il trattamento dei dati. La Landscape Genetics cerca di integrare approcci di genetica di popolazione e di ecologia del paesaggio nella descrizione dei pattern genetici osservati.

Nel presente studio ho analizzato la variabilità e la struttura genetica di popolazioni delle due specie di ungulati più abbondanti in Europa, il cinghiale (*Sus scrofa*) e il capriolo (*Capreolus capreolus*). In ciascuna specie la diversità genetica è stata analizzata a due differenti scale geografiche, evidenziando come il clima, l'uso del suolo e le caratteristiche del paesaggio possono influenzare, in modo diverso ed unitamente all'azione diretta dell'uomo, la dispersione dei geni. I pattern filogeografici nel continente europeo sembrano risentire principalmente di eventi climatici passati (ultima glaciazione), mentre il differenziamento genetico locale è frutto di caratteristiche paesaggistiche locali e di pratiche gestionali.

Parole chiave: Landscape Genetics, flusso genico, comportamento spaziale, struttura genetica, barriere ambientali.

Introduction

Introduction

Genetic diversity represents one of the most important biodiversity components and is the base for the species evolution (Allendorf & Luikart, 2007). Indeed, a high genetic diversity gives species the ability to adapt to different environments and tolerate sudden changes in climate and/or environmental conditions. Thus, it is very important to preserve a certain degree of genetic variability, particularly in wildlife populations, that can allow species to be plastic to sudden external modifications (Allendorf & Luikart, 2007). Already in 1836, Darwin noticed the loss of vigor in deer populations inhabiting British natural parks due to a small population size and isolation, emphasizing the importance of maintaining an adequate genetic variation to facilitate the persistence of natural populations. Thus, current species genetic variability is the result of complex relationships between the genome and the environment, and is by far influenced by human action (Allendorf & Luikart, 2007).

Wildlife populations and their genetic differentiation are inevitably influenced by both present and past processes (Hewitt & Butlin, 1997). Climatic variations are the primary source of variability for animal and plant species and they could act as limiting factors. Indeed, climate fluctuations could affect individual physiology and metabolism or food availability and accessibility, having consequences on the demography of the species. Current geographical distribution of populations mainly depend on climate changes occurred in the past and the phylogeographical patterns observed in many European mammals are thought to be strongly related to Quaternary climatic fluctuations (Hewitt, 1999; 2004). The extensive ice cover of northern regions during ice ages temporarily confined thermophilic species to lower latitudes, whereas northward recolonizations took place during the milder interglacials. In particular, most of the current phylogeographical discontinuities are likely to have arisen during the last glaciation (c. 110,000–12,000 yr ago), as suggested also by ancient DNA studies (e.g. Hofreiter et al., 2004). During the Last Glacial Maximum (LGM; 23,000–16,000 yr ago), ice caps covered northern regions and the main mountain chains, and most of continental Europe was covered by permafrost (Vandenberghe et al., 2012). Iberia, Italy, the Balkans and the Caspian/Caucasus region acted as refugia for remnant populations and represented the source for the following recolonization of northern Europe for several species (Taberlet et al., 1998; Hewitt, 2004). In agreement with the

traditional 'southern refugia model', present-day genetic diversity is expected to decrease gradually from refugial to recolonized areas (Guo, 2012).

Instead, at a more localized geographical scale, the distribution of intraspecific genetic diversityis mostly influenced by landscape features (Manel et al., 2003). In fact, it was observed that the presence of natural barriers and/or anthropogenic infrastructures can limit gene flow, leading to population fragmentation and promoting genetic isolation among demes (Coulon et al., 2004, 2006; Castilho et al., 2010; Cushman et al., 2010; Frantz et al., 2012). Over the last century important phenomena like urbanization and the development of large networks of transport infrastructures have rapidly increased in Europe. Such transformations may have affected the spatial behaviour of many large mammals, like ungulates, and consequently their population genetic structure. Recently, several studies have been aimed at evaluating the effect of anthropogenic barriers on wildlife genetic structure (bighorn sheep *Ovis canadensis* Epps et al., 2005; roe deer *Capreolus capreolus* Coulon et al., 2006; carnivores Riley et al 2006; bobcat (*Lynx rufus*) Milion & Swanson 2007; small mammals Rico et al., 2007; red deer *Cervus elaphus* Šprem et al., 2013, wild boar *Sus scrofa* and red deer Frantz et al., 2012). Finally, much of the current population genetic variability is also influenced by human intervention through translocations, hybridization and hunting exploitation of species (Mamuris et al., 2001; Chazara et al., 2010; Goedbloed et al., 2013).

A new scientific discipline that has emerged in recent years and has given an innovative contribution in analyzing the population genetic structure in relation to environmental characteristics is the Landscape Genetics (Manel et al., 2003; Holderegger & Wagner, 2006). It attempts to provide on temporal and spatial scale and on individual and population level information on the interactions between environmental characteristics and microevolutionary processes (Storfer et al., 2007).

The Landscape Genetics tries to explain population substructure testing the effect of landscape elements on the observed genetic pattern, by implementing statistical and spatial analysis in GIS (Geographical Information Systems). Interpreting genetic data by spatially-explicit models allows us to test if the structuring can be dependent on the physical characteristics of the environment or, rather, on population parameters only (Piertney et al., 1998). Indeed, environmental characteristics may have a strong influence on the genetic make-up of population

acting on the dispersal ability of individuals (Coulon et al., 2004; Pérez-Espona et al., 2008; Cushman et al., 2010). Understanding the influences of landscape and habitat quality on gene flow and genetic variation can be of great importance to increase knowledge on the spatial behaviour and genetic status of animal species, indispensable in planning efficient management and conservation practices (DeSalle & Amato, 2004).

The fundamental distinction between the traditional Population Genetics and the Landscape Genetics is the integration of explicit tests of landscape heterogeneity in genetic variation within and between populations (Holderegger & Wagner, 2006). Thus, while population genetics studies the subdivision within a populations and the genetic diversity among the different genetic units, landscape genetics describes the spatial pattern of genetic variability, examining the causes of population substructure.

In the last years a high number of statistical approaches and targeted software have been developed to associate environmental variables to genetic data. Their goal is to find explanatory variables for a the genetic structure in a population, that can explain data better than a simple Isolation-By-Distance model (Wright, 1943; Balkenhol et al., 2009; Castihlo et al., 2010; Frantz et al., 2010; 2012; Huck et al., 2011). Due to the complexity of relationships, authors suggested the use of multiple tests and many environmental variables to best represent the landscape matrix in which populations are studied. The incorporation of environmental matrices in genetic models implies the knowledge of the species' ecology and behaviour, and generally passes through the creation of "resistance maps" describing the potential permeability of the landscape to animal movements (Ray, 2005; McRae, 2006; McRae et al., 2007, 2008). Such maps may incorporate information on land use, natural barriers (sea, lakes, rivers, mountains) and human infrastructures (roads, urban areas) that may influence dispersal and ultimately gene flow (Wang et al., 2009, van Strien et al., 2012; Weckworth et al., 2013). Such applications are useful to detect spatial genetic units and to identify outstanding barriers and possible ecological corridors for species management and conservation (Kuehn et al., 2007). These methods can be also useful to model the spread of specific genetic variants in a population, like in the case of introgressed genes deriving from human-caused hybridization events (Fitzpatrick & Shaffer, 2007; Mucci et al., 2012). Under these circumstances, the identification of barriers or of land resistance to gene flow can acquire a positive value for the conservation of the threatened population genetic integrity.

At a finer scale, the distribution of genes within a population is determined by the individual spatial behaviour and by inter-individual social relationships. Social organization and mating system and the influence of the environment on them are of utmost importance and can be studied with the contribution of genetic data (van Staaden, 1995; Gaillard et al., 2008; Rossiter et al., 2012).

On the basis of the different temporal and geographic scale, the influence of environmental variables on the diversity and genetic structure of a species can be assessed using different genetic markers.. Mitochondrial DNA (mtDNA) is used to identify genetic differentiation on large spatial and temporary scale, due to its characteristics. MtDNA, as opposed to the nuclear genome, presents a uniparental (maternal) inheritance and is not subject to recombination. Moreover, the low mutation rate makes it useful in studies of phylogeography and to identify genetic differences accumulated over a long time span. On the contrary, differences at fine spatial scale are best comprehended using more polymorphic molecular markers such as microsatellites (or Short Tandem Repeat STR). Opposite to mtDNA, they are neutral loci, and therefore not under selection, biparentally inherited, codominant, locus-specific and are subject to a high mutation rate Thanks to these features, microsatellites have a great potential to describe the genetic variability of subpopulations and to estimate inter-individual genetic relationships (i.e. relatedness).

In the present doctoral study, the genetic diversity of two species of European ungulates, wild boar (*Sus scrofa*) and roe deer (*Capreolus capreolus*) were investigated in relation to landscape features at different geographical scales. They are the most widespread ungulates in Europe, as well as important game species, thus subjected for years to human manipulation and management. The two species had a very similar demographic history. In fact, their current presence and distribution in Europe arises mainly from post-glacial recolonization events from southern Mediterranean refugia. In the last two centuries both species were subject to a dramatic contraction of its range and disappeared in certain areas. Such demographic decrease was mainly caused by overhunting, but also by the loss of suitable habitats, due to the spread of urban centres and agriculture, that has forced species to contract their range of distribution.

At the beginning of the XIX century, wild boar and roe deer were present in localized areas throughout Europe and only after the II World War, ungulate populations have begun their demographic recovery. This phenomenon was also enforced by human reintroduction and translocation events, inevitably influencing genetic variation of the two species (Apollonio & Putman, 2011).

In the **CHAPTER 1** we assessed the mitochondrial diversity of the wild boar across Europe, in relation to post- glacial dynamics. This project was conducted in collaboration with Prof. G. Bertorelle and Dr. S. Torres-Villaca of the University of Ferrara, Italy. Despite being a generalist with an omnivorous opportunistic behaviour, the wild boar is not ubiquitous in Europe, in fact is present from the South of Spain to the central Karelia (Danilov & Panchenko, 2012). The only limiting factor is the climate, having direct effects on the physiology and metabolism, and indirect effects on the availability and accessibility of food (Geisser & Reyer, 2005 Melis et al., 2006). In particular, the severity of winters affects the species' ability to move and to search for food. Consequently, due to the limited adaptation to cold climates, the geographical distribution of wild boar has suffered severe climatic oscillations during the Holocene and Pleistocene.

To identify traces of this climatic influence on the gene pool of the species we combined mitochondrial DNA genetic diversity coming from different European regions, with macroscopic past climate variations. Through the use of global climate models we projected climate into the past; consequently, it was possible to construct a map of suitability for the presence of the boar at the time of the last glacial maximum (~20,000 years ago) identifying the likely refuge areas from which the species would have re-colonized the entire continent. Then we tested whether current patterns of genetic variation can be explained by the estimated past suitability, despite recent human manipulation.

In **CHAPTER 2** the genetic structure of wild boar was considered at a regional scale. The isolated population of Sardinia was used as model species to evaluate the influence of environmental features (both natural and anthropogenic) on the local patterns of genetic diversity in wild boar. The Sardinian population originated during the Neolithic and is currently classified as a distinct subspecies (*Sus scrofa meridionalis*), on the basis of both morphological and genetic

evidences. Its long-lasting isolation has produced a relevant genetic differentiation of the Sardinian population, observed using different types of genetic markers by Scandura and colleagues (2008, 2009, 2011). Furthermore, Sardinian wild boar population was previously observed to be genetically divided into three main subpopulations (Scandura et al., 2011), highlighting the presence of isolation or limited gene flow among them. Genetic differences within the population were poorly explained by a simple isolation-by-distance model, indicating a more complex genetic pattern (Scandura et al., 2011), that could be explained by considering landscape features. In particular, the presence of high-traffic roads, fencing, urban and industrial areas, as well as relevant landscape breaks like water basins or rivers can severely affect gene flow in ungulate species (Worley et al., 2004; Coulon et al., 2006; Frantz et al., 2012; Ito et al., 2013). Specifically, wild boar is known to modify its activity and spatial patterns in relation to human disturbance. Nevertheless, thanks to its plasticity, a tendency to adapt to human presence and infrastructures is observed around urban centres (Cahill et al., 2012; Osashi et al., 2013). Therefore, keeping in mind the history of the species in the island, also characterized by appreciable levels of genetic introgression from domestic pigs and continental wild boar, aim of this study was to detect the limiting effects of landscape variables on wild boar gene flow. For this purpose we evaluated the land use and the presence of physical barriers, such as roads, which could have led over time to a genetic segregation within the Sardinian population.

Human intervention can affect the population genetic structure not only by changing the environment in which the species live (habitat fragmentation, land use changes, creation of road networks) but also through reintroductions or restocking events that implied the translocation of individuals from different geographical areas. In most cases, these events have led to a homogenization of the genetic variability with the disappearance of local genetic variants (Allendorf et al., 2001; Olden et al., 2004). In **CHAPTER 3** we assessed the impact of human action on the genetic integrity of the roe deer population inhabiting Central and Southern Italy and their role in shaping their current genetic structure. Indeed, Italian roe deer population experienced events of recolonization and translocation during the last century, after a drastic decline of the population in previous centuries (Perco & Calò, 1995). In the first XIX century roe deer was present only in the Eastern Alps, in a region of Central Italy, Maremma, and in a few small relict areas in

the Castelporziano estate (near Rome), and in Gargano and Orsomarso massifs in the south of the Italian peninsula (Perco & Calò, 1995). Such central-southern populations were considered as remnants of the endemic Italian subspecies (*C. c. italicus*; Festa 1925), characterized by both distinct morphometric (Montanaro et al., 2003) and genetic peculiarities (Lorenzini et al., 2002; Vernesi et al., 2002; Randi et al., 2004; Gentile et al., 2009; Mucci et al., 2012). Thus, reintroduction events were necessary to increase the small populations remained, and usually were realized with individuals coming from the Alps, the Balkans or Eastern Europe, leading to hybridization events between the European and Italian subspecies, especially in central Italy (Lorenzini et al 2002; Mucci et al 2012). The effectiveness of the reintroductions primarily derives from the ecological requirements of the populations from which the individuals are taken, which shall reflect those found in the area of reintroduction. This is not often the case, so it was possible to raise doubts on the spread of exotic genes where they faced the competition with locally adapted gene pools. In fact, although the two forms exhibit similar ecological requirements and similar social and spatial behaviour, it was observed the preference by native roe deer for woodland and scrubland/maquis, typical of the Italian peninsula (Focardi et al., 2009). Thus, genetic diversity was evaluated in three areas of Central Italy where it was possible to identify: i) a centre of diffusion of *C. c. italicus*, ii) an area of reintroduction or centre of diffusion of non-native roe deer, and iiI) an intermediate contact zone. Genetic patterns in each population were then evaluated in relation to reintroduction history and environmental features that could affect the success of the reintroductions and the spread of native versus non-native genes across the territory.

As remarked, environmental features may affect different biological aspects of species at different geographical scales, from the colonization of new habitats, individual dispersion, but also social and mating behaviour within a population. Indeed social behaviour, mating system and dispersal pattern are strongly affected by the distribution of resources in the environment, that is the amount, quality and distribution of food and the availability of sheltered or protected areas to defend against human disturbance and predation. Different reactions to similar environmental characteristics may also derive from the local density of individuals, ultimately producing a different genetic structure (Lott 1991). In roe deer, the fine-scale genetic structure is expected to

change in relation to landscape structure (Coulon et al., 2004), climatic conditions (Cagnacci et al., 2011), habitat structure (Rossi et al., 2001; Lamberti et al., 2006), population structure and density (Wahlström & Liberg 1995).

CHAPTER 4 is addressed to study the spatio-social behaviour and the resulting genetic structure in a localized roe deer population inhabiting the Alpe di Catenaia massif in Tuscany. These aspects were evaluated in relation to population density and environmental features of the area, also considering the species life cycle and testing possible differences between sexes.

References

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001). The problems with hybrids: setting conservation guidelines. Trends Ecol Evol 16: 613–622.

Allendorf FW, Luikart G (2007). Conservation and the Genetics of Populations. Blackwell Publishing.

Balkenhol N, Waits LP, Dezzani RJ (2009). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32: 818-830.

Cahill S, Llimona F, Cabañeros L, Calomardo, F (2012). Characteristics of wild boar (*Sus scrofa*) habituation to urban areas in the Collserola Natural Park (Barcelona) and comparison with other locations. Animal Biodiversity Conservation 35: 221–233.

Cagnacci F, Focardi S, Heurich M, Stache A, Hewison AJM, Morellet N, Kjellander P, Linnell JDC, Neteler M, Delucchi L, Ossi F, Urbano F (2011). Partial migration in roe deer: migratory and resident tactics are end points of a behavioural gradient determined by ecological factors. Oikos 120: 1790–1802

Focardi S, Montanaro P, La Morgia V, Riga F (2009). Piano d'azione per il capriolo italico (*Capreolus capreolus italicus*). Ministero dell'Ambiente 177 pp.

Castihlo CS, Marins-Sá LG, Benedet RC, Freitas TO (2010) Landscape genetics of mountain lions (*Puma concolor*) in southern Brazil. Mammalian Biology 76 (4):476-483.

Chazara O, Minvielle F, Roux D, Bed'hom B, Feve K, Coville JL, Kayang BB, Lumineau S, Vignal A, Boutin JM, Rognon X (2010). Evidence for introgressive hybridization of wild common quail (*Coturnix coturnix*) by domesticated Japanese quail (*Coturnix japonica*) in France. Conservation Genetics 11 (3):1051-1062.

Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004). Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. Molecular Ecology 13: 2841-2850.

Coulon A, Cosson JF, Morellet N, Angibault JM, Cargnelutti B, Galan M, Aulagnier S, Hewison AJM (2006). Dispersal is not female biased in a resource-defence mating ungulate, the European roe deer. Proceedings of the Royal Society B 273: 341-348.

Cushman SA, Lewis JS (2010). Movement behavior explains genetic differentiation in American black bears. Landscape Ecology 25:1613–1625.

DeSalle R, Amato G (2004). The expansion of conservation genetics. Nature Reviews Genetics 5 (9): 702-712.

Epps CW, Palsbøll PJ, Wehausen JD, Roderick GK, Ramey RR, McCullough DR (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters 8: 1029–1038.

Festa E (1925). Il capriolo dell'Italia central (*Capreolus capreolus italicus sbsp*. Nova). Bollettino del Museo di Zoologia e Anatomia Comparata dell'Università di Torino 7: 1-2.

Fitzpatrick BM, Shaffer HB (2007). Introduction history and habitat variation explain the landscape genetics of hybrid tiger salamanders. Ecological Applications 17:598–608

Frantz AC, Pope LC, Etherington TR, Wilson GJ, Burke T (2010). Using isolation-by-distance-based approaches to assess the barrier effect of linear landscape elements on badger (*Meles meles*) dispersal. Molecular Ecology 19:1663–1674.

Frantz AC, Bertouille S, Eloy MC, Licoppe A, Chaumont F, Flamand MC (2012) Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (*Cervus elaphus*), but not wild boars (*Sus scrofa*). Molecular Ecolology 21: 3445-3457.

Gaillard JM, Hewison AJM, Kjellander P, Pettorelli N, Bonenfant C, Van Moorter B, Liberg O, Andren H, Van Laere G, Klein,FAngibault JM, Coulon A, Vanpé C (2008). Population density and sex do not influence fine-scale natal dispersal in roe deer. Proceedings of the Royal Society B 275: 2025-2030

Geisser H, Reyer H-U (2005). The influence of food and temperature on population density of wild boar *Sus scrofa* in the Thurgau (Switzerland). Journal of Zoology of London 267: 89–96.

Gentile G, Vernesi C, Vicario S, Pecchioli E, Caccone A, Bertorelle G et al (2009). Mitochondrial DNA variation in roe deer (*Capreolus capreolus*) from Italy: evidence of admixture in one of the last *C.c.italicus* pure populations from central-southern Italy. Italian Journal of Zoology 76:16-27.

Goedbloed DJ, van Hooft P, Megens HJ, Langenbeck K, Lutz W, Crooijmans RP, van Wieren SE, Ydenberg RC, Prins HHT (2013). Reintroductions and genetic introgression from domestic pigs

have shaped the genetic population structure of Northwest European wild boar BMC Genetics Article in Press.

Guo Q (2012). Incorporating latitudinal and central–marginal trends in assessing genetic variation across species ranges. Molecular Ecology 21:5396–5403.

Hewitt GM (1999). Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68:87–112.

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London. Biological Sciences 359:183–195.

Hewitt GM, Butlin RK (1997). Causes and consequences of population structure. In: Krebs J, Davies, N, eds. Behavioural Ecology: An Evolutionary Approach. Oxford: Blackwell 350–372.

Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Münzel S, Pääbo S (2004). Lack of phylogeography in European mammals before the last glaciation. Proceedings of the National Academy of Sciences of USA 101:12963–12968.

Holderegger R, Wagner HH (2006). A brief guide to Landscape Genetics. Landscape Ecology 21:793–796.

Huck M, Jędrzejewski W, Borowik T, Jędrzejewska B, Nowak S, Mysłajek RW (2011). Analyses of least cost paths for determining effects of habitat types on landscape permeability: wolves in Poland. Acta Theriologica 56:91–101.

Ito TY, Lhagvasuren B, Tsunekawa A, Shinoda M, Takatsuki S, Buuveibaatar B, Chimeddorj B (2013). Fragmentation of the Habitat of Wild Ungulates by Anthropogenic Barriers in Mongolia. PLoS ONE 8(2): e56995. doi:10.1371/journal.pone.0056995.

Kuehn R, Hindenlang KE, Holzgang O, Senn J, Stoeckle B, Sperisen C (2007). Genetic Effect of Transportation Infrastructure on Roe Deer Populations (*Capreolus capreolus*). Journal of Heredity 98: 13–22.

Lamberti P, Mauri L, Merli E, Dusi S, Apollonio M (2006). Use of space and habitat selection by roe deer (*Capreolus capreolus*) in a Mediterranean coastal area: how does woods landscape affect the home range? Journal of Ethology 24(2):181-188

Lorenzini R, Lovari S, Masseti M (2002). The rediscovery of the Italian roe deer; genetic differentiation and management implications. Italian Journal of Zoology 69:367-379.

Lott DF (1991). Intraspecific variation in the social systems of wild vertebrates. Cambridge University Press, Cambridge

Mamuris Z, Sfougaris AI, Stamatis C (2001). Genetic structure of Greek brown hare (*Lepus europaeus*) populations as revealed by MtDBNA RFLP-PCR analysis: Implications for conserving genetic diversity. Biological Conservation, 101 (2):187-196.

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.

McRae BH (2005). Isolation by Resistance. Evolution 60(8):1551–1561.

McRae BH, Beier P (2007). Circuit theory predicts gene flow in plant and animal populations. PNAS 104(50): 19885-19890.

McRae HB, Dickson BG, Keitt TH, Shah VB (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89 (10):2712–2724.

Melis C, Szafrańska PA, Jędrzejewska B, Barton K (2006). Biogeographical variation in the population density of wild boar (*Sus scrofa*) in western Eurasia. J Biogeogr 33: 803–811.

Monmonier M (1973). Maximum-difference barriers: An alternative numerical regionalization method. Geographical Analysis. 3:245–261

Millions DG, Swanson BJ (2007). Impact of natural and artificial barriers to dispersal on the population structure of bobcats. Journal of Wildlife Management 71:96–102.

Montanaro P, de Marinis AM, Riga F, Focardi S, Toso S (2003). Variabilità craniometrica in alcune popolazioni italiane di capriolo. Paper presented at the Atti IV Congresso Italiano di Teriologia: Scientific Research and Conservation of Mammals in Italy, Riccione, 6-8 November.

Mucci N, Mattucci F, Randi E (2012). Conservation of threatened local gene pools: landscape genetics of the Italian roe deer (*Capreolus c. Italicus*) populations. Evolutionary Ecology Research 14: 897-920.

Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD (2004). Ecological and evolutionary consequences of biotic homogenization. TRENDS IN ECOLOGY AND EVOLUTION 19:1.

Osashi H, Saito M, Horie R, et al., (2013). Differences in the activity pattern of the wild boar *Sus scrofa* related to human disturbance. European Journal of Wildlife Research 59: 167-177.

Perco F, Calò CM (1995). The status of roe deer in Italy. In: Wotschikowsky U., cur., Proceeding of the 2nd European Roe Deer Meeting. Wildbiologische Gesellschaft, München.

Pérez-Espona S, Pérez-Barbería FJ, McLeod JE, Jiggins CD, Gordon IJ, Pemberton JM (2008). Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). Molecular Ecology 17: 981–996

Piertney SB, McColl ADC, Bacon PJ, Dallas JF (1998). Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. Molecular Ecology 7: 1645-1654.

Randi E, Alves PC, Carranza J, Milosevic-Zlatanovic S, Sfougaris A, Mucci N (2004). Phylogeography of roe deer (*Capreolus capreolus*) populations: the effects of historical genetic subdivisions and recent non-equilibrium dynamics. Molecular Ecology 13:3071-3083.

Ray N (2005). PATHMATRIX: a geographical information system tool to compute effective distances among samples. Molecular Ecology Notes 5:177–180.

Rico A, Kindlmann P, Sedláček (2007). Barrier effects of roads on movements of small mammals. Folia Zoologica 56 (1):1-12

Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne RK (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. Molecular Ecology 15:1733-1741.

Rossi L, Lamberti P, Mauri L, Apollonio M (2001). Male and female spatial behaviour of roe deer in a mountainous habitat during pre-rutting and rutting period. Journal Mountain Ecology 6(1): 1-6

Rossiter SJ, Zubaid A, Mohd-Adnan A, Struebig MJ, Kunz TH, Gopal S, Petit EJ, Kingston T (2012). Social organization and genetic structure: insights from codistributed bat populations. Molecular Ecology 21(3):647-61.

Scandura M, Iacolina L, Crestanello B et al., (2008). Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Molecular Ecology 17:1745-1762.

Scandura M, Iacolina L, Crestanello B, Pecchioli E, Bertorelle G, Apollonio M (2009). Occurrence of a private mtDNA clade in Italian wild boars. In: Proceedings of the 7th International Symposium on Wild Boar (*Sus scrofa*) and on Sub-order Suiformes (eds by Náhlik A, Tari T.), pp. 96-97. Sopron, Hungary.

Scandura M, Iacolina L, Cossu A, Apollonio M (2011). Effects of human perturbation on the genetic make-up of an island population: the case of the Sardinian wild boar. Heredity 106: 1012-1020.

Šprem N, Frantz AC, Cubric-Curik V, Safner T, Curik I (2013). Influence of habitat fragmentation on population structure of red deer in Croatia. Mammalian Biology 78: 290– 295.

Storfer A, Murphy MA, Evans JS, et al., (2007). Putting the 'landscape' in landscape genetics. Heredity 98: 128-142.

Van Staaden MJ (1995). Breeding tactics, social structure and genetic variation in mammals: problems and prospects. Acta Theriologica 3:165-182.

van Strien MJ, Keller D, Holderegger R (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. Molecular Ecology 21(16):4010–4023.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998). Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7:453–464.

Vandenberghe J, Renssen H, Roche DM, Goosse H, Velichko AA, Gorbunov A, Levavasseur G (2012). Eurasian permafrost instability constrained by reduced sea-ice cover. Quaternary Science Reviews 34:16–23.

Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G (2002). The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. Molecular Ecology 11:1285-1297.

Wahlström LM, Liberg O (1995). Patterns of dispersal and seasonal migration in roe deer (*Capreolus capreolus*). Journal of Zoology 235: 455-467

Wang IJ, Savage WK, Shaffer HB (2009). Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). Molecular Ecology 7:1365-74

Weckworth BV, Musiani M, DeCesare NJ, McDevitt AD, Hebblewhite M, Mariani S (2013). Preferred habitat and effective population size drive landscape genetic patterns in an endangered species. Proceedings of the Royal Society B 22 280 1769.

Worley K, Strobeck C, Arthur S, Carey J, Schwantje H, Veitch A, Coltman DW (2004). Population genetic structure of North American thinhorn sheep (ovis dalli). Molecular Ecology 13:2545–2556.

Wright S. (1943). Isolatin by Distance. Genetics, 28: 114-138.

Chapter 1

Mitochondrial phylogeography of the European wild boar: the effect of climate on genetic diversity and spatial lineage sorting across Europe

Sibelle T. Vilaça, Daniela Biosa, Frank Zachos, Laura Iacolina, Julia Kirschning, Paulo C. Alves, Ladislav Paule, Christian Gortazar, Zizzis Mamuris, Bogumiła Jędrzejewska, Tomasz Borowik, Vadim E. Sidorovich, Josip Kusak, Stefano Costa, Laurent Schley, Günther B. Hartl, Marco Apollonio, Giorgio Bertorelle and Massimo Scandura

Article In press: *Journal of Biogepgraphy*

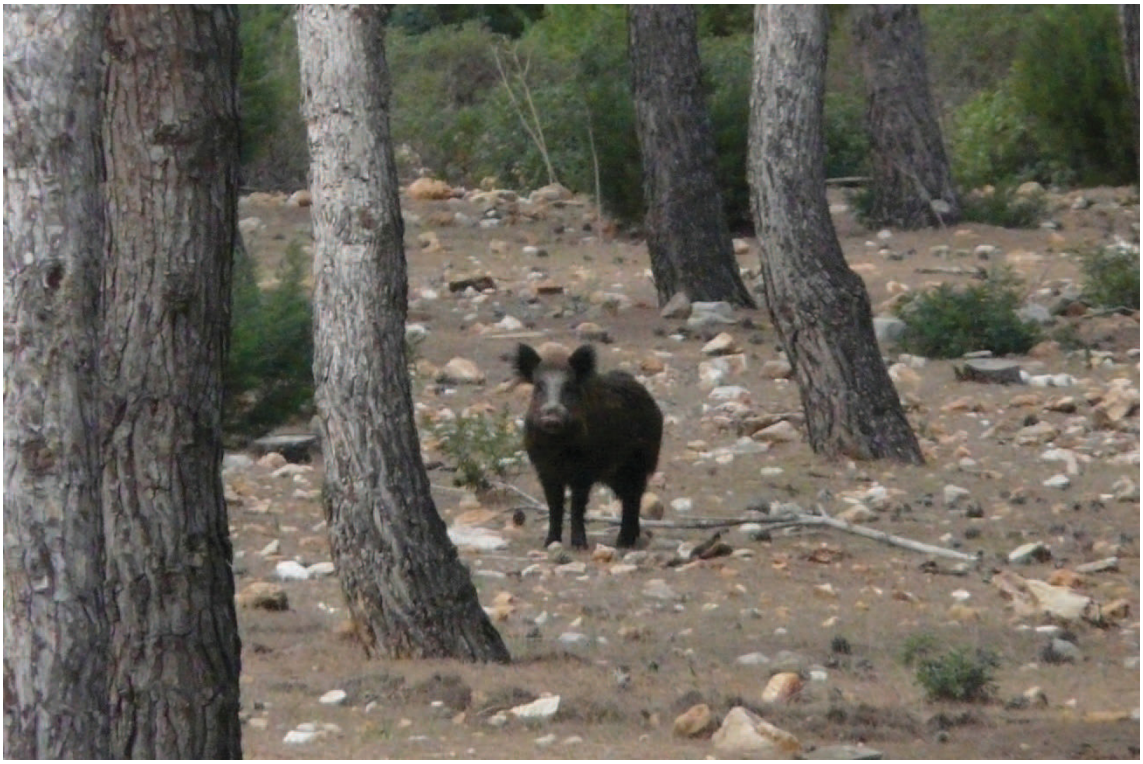


Foto di Alberto Addis

Mitochondrial phylogeography of the European wild boar: the effect of climate on genetic diversity and spatial lineage sorting across Europe

Sibelle T. Vilaça[1], Daniela Biosa[2], Frank Zachos[3,4], Laura Iacolina[2], Julia Kirschning[4], Paulo C. Alves[5,6], Ladislav Paule[7], Christian Gortazar[8], Zizzis Mamuris[9], Bogumiła Jędrzejewska[10], Tomasz Borowik[10], Vadim E. Sidorovich[11], Josip Kusak[12], Stefano Costa[13], Laurent Schley[14], Günther B. Hartl[4], Marco Apollonio[2], Giorgio Bertorelle[1†] and Massimo Scandura[2*†]

[1]Dipartimento di Biologia ed Evoluzione, Università di Ferrara, 44121 Ferrara, Italy, [2]Dipartimento di Scienze della Natura e del Territorio, Università di Sassari, 07100 Sassari, Italy, [3]Natural History Museum Vienna, A-1010 Vienna, Austria, [4]Zoological Institute, Christian-Albrechts-Universität zu Kiel, 24118 Kiel, Germany, [5]CIBIO, Faculdade de Ciências, Universidade do Porto, 4485-661 Vairão, Portugal, [6]Wildlife Biology Program, University of Montana, Missoula, MT59812, USA, [7]Faculty of Forestry, Technical University in Zvolen, 960 53 Zvolen, Slovakia, [8]SaBio IREC, CSIC-UCLM-JCCM, 13071 Ciudad Real, Spain, [9]Department of Biochemistry and Biotech, University of Thessaly, 41221, Larissa, Greece, [10]Mammal Research Institute, Polish Academy of Sciences, 17-230 Białowieża, Poland, [11]Institute of Zoology, National Academy of Sciences of Belarus, 220072 Minsk, Belarus, [12]Department of Biology, Faculty of Veterinary Medicine, University of Zagreb, 10000 Zagreb, Croatia, [13]Laboratorio Chimico della Camera di Commercio di Torino, 10127 Torino, Italy, [14]Administration de la Nature et des Forêts, L-2453, Luxembourg

*Correspondence: Massimo Scandura, DipNeT, via Muroni 25, 07100 Sassari, Italy.
E-mail: scandura@uniss.it

Abstract

Aim Climate changes in the past had a deep impact on the evolutionary history of many species and left genetic signatures that are often still detectable today. We investigated the geographical pattern of mitochondrial DNA diversity in the European wild boar (*Sus scrofa*). Our final aims were to clarify the influence of present and past climatic conditions, infer the geographical position of glacial refugia, and suggest post-glacial spatial dynamics.

Location Europe.

Methods D-loop sequences were obtained for 763 individuals from Portugal to western Russia. Phylogenetic, multivariate and interpolation methods were used to describe the genetic and geographical patterns. Climatic suitability during the Last Glacial Maximum (LGM) was predicted using MAXENT. The effect of present and past suitability on the observed patterns of diversity was evaluated by multiple linear regression.

Results We confirmed the existence of a ubiquitous mitochondrial clade in Europe (E1), an endemic clade in Italy (E2) and a few East Asian haplotypes (A), presumably introgressed from domestic pigs. No Near Eastern haplotypes were detected. Genetic divergence was not simply related to geographical distance. A clear south–north decreasing gradient of diversity was observed, with maximum levels in putative glacial refugia. Latitudinal variation in climatic conditions during the LGM was shown to be a good predictor of current genetic diversity. Moreover, an unexpected similarity between Iberia and eastern Europe was observed, while central European populations showed a higher affinity to the Italian gene pool.

Main conclusions The current distribution of mitochondrial genetic diversity was highly influenced by past climatic events, especially those related to the LGM, and is consistent with a major contribution of the Italian peninsula and the Balkans to the post-glacial recolonization of northern areas. More recent processes, such as restocking and extensive hunting, probably acted at rather local scales, without great impact on the global pattern of mitochondrial diversity.

Keywords

Climate change, genetic differentiation, glacial refugia, Last Glacial Maximum, mtDNA, phylogeography, *Sus scrofa*.

Introduction

The phylogeographical patterns observed in many European mammals are thought to be strongly related to Quaternary climatic fluctuations (Hewitt, 2004). The extensive ice cover of northern regions during ice ages temporarily confined thermophilic species to lower latitudes, whereas northward recolonizations took place during the milder interglacials. In particular, most of the current phylogeographical discontinuities are likely to have arisen during the last glaciation (*c.* 110,000–12,000 yr ago), as suggested also by ancient DNA studies (e.g. Hofreiter *et al.*, 2004). During the Last Glacial Maximum (LGM; 23,000–16,000 yr ago), ice caps covered northern regions and the main mountain chains, and most of continental Europe north of 45° N was covered by permafrost (Vandenberghe *et al.*, 2012). Iberia, Italy, the Balkans and the Caspian/Caucasus region acted as refugia for remnant populations and represented the source for the following recolonization of northern Europe for several species (Taberlet *et al.*, 1998; Hewitt, 2004).

In agreement with the traditional 'southerly refugia model', present-day genetic diversity is expected to decrease gradually from refugia to recolonized areas. A progressive loss of genetic diversity at increasing distance from refugia is expected if the recolonization was fast and if it followed a 'leading-edge expansion model', where only populations at the periphery of refugial regions contributed to repopulating the rest of the continent (Hewitt, 2004).

The Eurasian wild boar (*Sus scrofa* Linnaeus, 1758) is a temperate species currently widespread in Europe, where it is one of the most important game species (Apollonio *et al.*, 2010). Apart from human exploitation, climate is the main limiting factor for it, either through its effect on physiology and metabolism or through its indirect effect on food availability and accessibility (Geisser & Reyer, 2005; Melis *et al.*, 2006). First, harsh winters and hot summers represent a problem for the thermoregulation of piglets (Berg *et al.*, 2006). Second, highly energetic food, like acorns or agricultural crops, has to be consumed throughout the year in order to survive the winter (Schley & Roper, 2003), when frozen or very dry soils can limit access to underground food (Herrero *et al.*, 2004). As a consequence, under extreme climatic conditions, the species may have a high mortality rate.

Such a limited adaptation to severely cold conditions suggests that the geographical distribution of the wild boar in Europe was largely influenced by Pleistocene and Holocene climatic oscillations. Fossil remains indicate that the species survived during the LGM in Iberia, southern

France, the Italian peninsula and the Balkans (Sommer & Nadachowski, 2006). Genetic data, based on a numerically and geographically restricted sampling, support the view that Italy, and possibly the Balkans, acted as major genetic reservoirs (Larson *et al.*, 2005; Scandura *et al.*, 2008; Alves *et al.*, 2010; Alexandri *et al.*, 2012). When the climate became milder, the recolonization of the continent would have started from these southern refugia, with their contributions probably depending on the effective size of relict populations and on local environmental conditions.

Here, we investigate the distribution of wild boar mitochondrial DNA (mtDNA) lineages in Europe, considering as explanatory variables the geographical location of the samples and the present and past climatic conditions. We assembled a large dataset by adding 467 new sequences of European wild boar to previously published ones, thus filling important gaps in Central and Eastern Europe and covering the area from the Atlantic coast of Portugal to western Russia. This large dataset allowed us to address five specific questions. (1) How is genetic diversity geographically partitioned in Europe? (2) Does genetic diversity show a latitudinal cline, being higher in southern glacial refuges? (3) Does genetic diversity correlate with present and past climatic suitability? (4) Does genetic divergence increase with spatial distance? (5) Can past climatic changes and subsequent demographic and range fluctuations explain the observed geographical partitioning of mitochondrial lineages?

Materials and methods

Sampling

Tissue samples provided by local hunters were collected in 16 countries (Germany, Luxembourg, France, Portugal, Italy, Greece, Croatia, Bosnia–Herzegovina, Serbia, Slovakia, Romania, Bulgaria, Poland, Belarus, Ukraine and Russia). Areas that were under-represented in previous studies (Larson *et al.*, 2005; Scandura *et al.*, 2008; Alves *et al.*, 2010; Alexandri *et al.*, 2012) were specifically targeted. We also collected information on the historical distribution of the wild boar in most of Europe, and on how translocations, habitat fragmentation and overhunting might have influenced its natural distribution and demography (see Appendix S1 in Supporting Information).

Sequencing

We sequenced a total of 467 wild boar specimens from 36 locations. Sequences from Tunisia (*n* =

77), previously published by Hajji & Zachos (2011), only partly overlapped with the D-loop alignment and were therefore extended by re-sequencing. Total genomic DNA was isolated using a commercial DNA isolation kit (Sigma-Aldrich, St Louis, MO, USA; Qiagen, Hilden, Germany). Laboratory analyses consisted of amplifying the mitochondrial D-loop and sequencing a fragment of 411 base pairs (bp) with the BigDye Terminator 3.1 methodology (Applied Biosystems, Foster City, CA, USA) after purification with ExoSAP (ThermoFisher Scientific, Waltham, MA, USA). Detailed methodological procedures are described in Scandura *et al.* (2008).

Electropherograms were visually inspected, and base calls edited in FINCHTV 1.2 (Geospiza Inc, Seattle, WA, USA). By virtue of the quality of the electropherograms and the shortness of the region, most sequences were obtained with a single (forward) primer. Nonetheless, to ensure accuracy of nucleotide identification, a subset of samples was sequenced in the reverse direction, as were all individuals assigned to singleton haplotypes and all samples showing double peaks at any nucleotide position.

Phylogenetic analysis of D-loop sequences

The 467 new sequences were aligned using the CLUSTALX algorithm implemented in MEGA 4.0 (Tamura *et al.*, 2007) together with 632 sequences retrieved from GenBank. The downloaded sequences represented animals classified as wild *S. scrofa* from Europe (*n* = 361), North Africa (*n* = 77) and Asia (*n* = 194) (sequences are listed in Appendix S2).

The final dataset combined the published and newly produced sequences of the mitochondrial control region (411-bp region) and comprised sequences from 1099 wild boars from three continents. Haplotypes were collapsed in COLLAPSE 1.2 (Posada, 2011), and new haplotypes were deposited in GenBank (accession numbers KC608827–KC608847). A median-joining (MJ) network of haplotypes (Bandelt *et al.*, 1999) was then built in NETWORK 4.6 (Fluxus Technologies, Clare, UK).

The most appropriate model of nucleotide change was selected using JMODELTEST 0.1.1 (Posada, 2008), including a sequence of *Sus barbatus* as outgroup (GenBank accession number AJ314540). The best model, according to both the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), was the HKY model (Hasegawa *et al.*, 1985) with gamma-distributed (G) rate variation across sites.

Bayesian phylogenetic analyses were carried out in MRBAYES 3.2 (Ronquist & Huelsenbeck, 2003) using the HKY+G model of sequence evolution and two independent runs of four Markov chains (one cold and three heated) over 1,000,000 generations and sampling every 100 generations. The first 25% of the sampled trees and estimated parameters were discarded as burn-in. Plots of log-likelihood scores against generation times were used to identify the point at which log-likelihood values reached stationarity. The final consensus tree was drawn in MEGA.

Spatial differentiation and isolation by distance in Europe

To assess the spatial differentiation among European wild boar, we excluded all non-European sequences and those for which no detailed geographical information was available. Asian haplotypes found in wild boars sampled in Europe were excluded, because they were attributed to local introgression with domestic pigs (Scandura *et al.*, 2011). In total, 763 sequences were considered from 77 sites across 19 countries (Appendices S2 & S3a). Some sites with sample size below 10 were pooled. The choice of pooling was based on the geographical location and the genetic composition: only adjacent sites not separated by physical barriers and showing similar haplotypes were grouped. One population in southern Italy (ISal), having $n = 7$ (after removal of three Asian haplotypes), was kept separate, as its allelic composition differed markedly from the nearest populations. After pooling, the final number of populations used in the statistical analyses was 39 (Appendix S3b).

The genetic structure in Europe was analysed with spatial factor analysis (spFA; Frichot *et al*., 2012) in R (R Development Core Team, 2012). This analysis uses the geographical information (coordinates) to correct for the effects of spatial autocorrelation in the exploratory analysis of genetic data (allele frequencies). Compared to similar methods, spFA appears to better remove the distortion introduced by the decay of genetic similarity with geographical distance when the genetic structure is inferred from a principal components analysis (Frichot *et al.*, 2012).

The software ARLEQUIN 3.5 (Excoffier & Lischer, 2010) was used to run a Mantel test (Mantel, 1967) to investigate whether an isolation-by-distance (IBD) model could explain the geographical pattern observed in Europe. The occurrence of IBD was tested for 38 European populations (Sardinia was excluded) by looking at the correlation between spatial distances and linearized genetic distances ($\Phi_{ST} / (1 - \Phi_{ST})$). Geographical distances were computed using either

the linear Euclidean distance or the minimum land distance between two sampling areas, considering the sea and high elevations (> 2000 m a.s.l.) to be barriers for wild boar.

Genetic diversity

Haplotype diversity (*Hk*; Nei, 1987) and nucleotide diversity (π) were computed in ARLEQUIN for the 39 populations. As sample sizes differed considerably among areas, an unbiased estimate of allelic richness (AR) was also calculated with CONTRIB 1.0 (Petit *et al.*, 1998). To summarize the spatial distribution of genetic diversity, *Hk* values for 38 of the 39 populations were interpolated using the ordinary kriging (OK) method and the SPATIAL ANALYST extension in ARCGIS 10 (ESRI, Redlands, CA, USA). Ordinary kriging is a geostatistical interpolator method that creates a smooth surface even when sampling is spatially uneven. The sample from Russia – the northernmost sampling region in our study – was excluded from this analysis. When included, it generated an isolated point of very high diversity, deforming the global pattern in the north-eastern area. This population resulted from pooling different sampling sites within a radius of 540 km, all of which had small sample sizes. Such forced grouping, coupled with the fact that this area was possibly affected by the post-glacial expansion of more eastern (unsampled) populations, may have led to biased estimates of genetic diversity and we therefore excluded Russia from this analysis as a precautionary measure.

Fu's F_S (Fu, 1997) and Tajima's *D* (Tajima, 1989) statistics were calculated in ARLEQUIN. Significance was assessed by randomly generating samples under the hypothesis of selective neutrality and demographic stability. Excluding strong selective effects on the mtDNA region analysed, *P*-values smaller than the significance threshold of 0.00128 (following the Bonferroni approach for multiple testing) can be considered evidence of deviation from demographic stability.

Modelling the species' present and past (LGM) range

To evaluate the influence of past range variation on the present wild boar genetic diversity in Europe, we assessed the ecological suitability for the species at the time of the LGM, identifying putative refugia. We used the machine learning method based on maximum entropy implemented in the program MAXENT 3.3.3 (Phillips *et al.*, 2006) to predict the present wild boar distribution and that during the LGM. We used as presence data the geographical coordinates of the 77 sampling

sites of this study, the sites used in Melis *et al.* (2006), and those available from the Global Biodiversity Information Facility (GBIF) database. As the latter was largely biased due to over-representation of points in some areas (e.g. France), we sub-sampled the GBIF locations to obtain an even density of points across countries and to produce a good representation of the different environmental contexts. A total of 215 locations were used, covering all the parts of Europe relevant to our study.

As climate represents the driving factor influencing other environmental variables that affect wild boar presence (such as habitat type, water and food availability), climatic variables were used to construct the climate prediction models. Specifically, we used annual mean temperature (BIO1), temperature seasonality (BIO4), annual temperature range (BIO7), mean temperature of the warmest quarter (BIO10), mean temperature of the coldest quarter (BIO11), annual precipitation (BIO12), precipitation of the wettest quarter (BIO16), and precipitation of the driest quarter (BIO17). Current and LGM data were downloaded from WorldClim 1.4 (Hijmans *et al.*, 2005). Two different general circulation models were adopted for the LGM estimations – the Community Climate System Model (CCSM) and the Model for Interdisciplinary Research on Climate (MIROC). As snow cover is a crucial limiting factor for wild boar (Melis *et al.*, 2006), we also included two variables from the Stage Three Project (van Andel, 2002): snow depth (in centimetres); and the number of days per year with snow cover. Layers were cropped to span from latitude 68° N to 33.8° N and from longitude 12° W to 51.7° E. All layers were used in their original spatial resolution (2.5 arc-minutes) and projection (WGS84 datum). Models were run in MAXENT using default settings. To establish a threshold and transform the continuous maps into binary presence/absence maps we used the 10% of training presences as a threshold value. We ran 10 replicates, and the average across all runs was calculated. To evaluate model performance, we used the area under the curve of the receiver operating characteristic (AUC), which measures the ability of a prediction to discriminate presence from absence (Elith *et al.*, 2010) and ranges from 0.5 to 1. An AUC value of 0.5 indicates that the model has no predictive ability, whereas a perfect discrimination between suitable and unsuitable cells will achieve the maximum AUC, i.e. 1.0 (Morueta-Holme *et al.*, 2010). We also performed a multivariate similarity surface (MESS) analysis when projecting in MAXENT, following Elith *et al.* (2010). The MESS analysis allows novel climatic regions to be distinguished, by predicting values that fall outside the training range of

variables.

To validate the palaeoclimatic map, we used fossil records compiled by William Davies (available at the Stage Three Project website: http://www.esc.cam.ac.uk/research/research-groups/oistage3) and also by Sommer & Nadachowski (2006). We only considered archaeological sites where the presence of *S. scrofa* had been reported. Two time intervals were considered: LGM (23,000–16,000 yr ago) and older than 23,000 yr ago. The occurrence of wild boar remains in a site dating back to the glacial period was considered as indicative of its presence in a glacial refugium.

Correlation between genetic diversity and environmental variables

To understand how present and past habitat suitability in Europe can explain the detected pattern of genetic diversity across the continent, we compared the effects of five predictive variables (latitude; longitude; present suitability; LGM suitability according to the MIROC model; LGM suitability according to the CCSM model) on *Hk* and AR by multiple linear regression in R (R Development Core Team, 2012). Data for Russia were not considered in the models, because genetic diversity can be biased in this population (see above). Both single effects and joint effects of variables were tested and the most parsimonious model was selected on the basis of the Akaike information criterion (AIC), using the corrected formula for small sample sizes (AIC_c; Symonds & Moussalli, 2011).

Results

Mitochondrial DNA diversity and phylogeography

The MJ network based on the large alignment including the Asian and North African haplotypes (in total 1099 individuals and 87 different haplotypes, Appendix S2) confirmed three major groups (Fig. 1): an Asian clade (corresponding to clade A from Giuffra *et al*., 2000); a pan-European (clade E1) clade; and a clade found only in Italy (clade E2). Only seven Asian haplotypes were observed among the 828 European sequences: four in Italy and one each in Germany, Luxembourg and Belgium. The European and Asian clades were separated by six fixed mutations (at positions 15545, 15568, 15573, 15583 and 15732 in the reference mtDNA sequence; Ursing & Arnason, 1998).

The analysis of 763 European sequences resulted in 50 haplotypes, corresponding to two A

haplotypes, 41 E1 haplotypes and seven E2 haplotypes (Appendix S2). The Bayesian tree restricted to these haplotypes (Fig. 2a) gave high (≥ 95%) posterior probabilities to the three clades, and also suggested the existence of an additional subgroup, E1a, within E1, matching the A-side group reported in other studies (Larson *et al.*, 2007; Scandura *et al.*, 2011; Alexandri *et al.*, 2012). This subclade showed high frequencies in Italy, France, Germany, Austria and the north Adriatic (Fig. 2b); H022 and H023 were its most frequent haplotypes (matching A and BK, respectively, in Larson *et al*., 2005). Several private haplotypes occurred in Iberia, but the most common (H021, matching haplotype E in Larson *et al*., 2005) was shared with eastern populations. The derived E1a clade is very rare in both eastern and western areas, where the most basal haplotypes prevail (roughly matching the previously reported C-side group; Larson *et al.*, 2007; Scandura *et al.*, 2011) (Fig. 2b).

Besides the private haplogroup E2, Italian wild boar showed an exclusive E1 haplotype that was spread across the peninsula (H075). North Adriatic populations (from Croatia and north-eastern Italy) were dominated by a private haplotype (H083, subclade E1a), which differed by a single mutation from H023 (Fig. 1). Greece showed a large proportion of private sequences (45%, H128), matching haplotype G in Alexandri *et al*. (2012), which was reported to be common in this region. France and Germany had low numbers of haplotypes when compared to other European populations, whereas Sardinia had the largest number of haplotypes (14), most of which were private (64%).

In the spFA of wild boar populations (Fig. 3), factor 1 separated three geographical regions: Italy/Central Europe, Iberia, and the Balkans/Eastern Europe. A few exceptions to this pattern were observed: southern Portugal (population 21) appeared closer to eastern populations, while Bulgaria and Greece (populations 3 and 15, respectively) were intermediate between Iberian and Italy/Central Europe populations. The highest genetic distances were observed between Italy/Central Europe and eastern populations and not, as predicted by a simple geographical pattern, between the two disjointed groups separated by the largest geographical distances (Iberia and the Balkans/Eastern Europe). Sardinian wild boar (population 27) appeared intermediate between populations from mainland Italy and Austria.

No significant correlation was detected between linearized genetic distances and geographical distances (Euclidean distance, r= 0.0003, P = 0.74; least-cost distance, r = −0.033, P = 0.78).

The highest levels of genetic diversity were observed in Italy and Sardinia, mainly due to the presence of the divergent E2 haplotypes (Appendix S3b). The map of interpolated haplotype diversities throughout Europe based on 38 populations highlighted the pattern of higher diversity in southern regions, with a maximum in Greece, Italy and eastern Spain (Fig. 4). A similar pattern was observed for the distribution of allelic richness (results not shown). Fu's *Fs* and Tajima's *D* neutrality tests were not significant for the majority of the populations (Appendix S3b).

Present and past (LGM) range

Under the current distribution, the AUC values for the training and test data showed satisfactory values (0.886 and 0.817, respectively). AUC values above 0.8 are considered an excellent model prediction (Hosmer & Lemeshow, 2000). The MAXENT estimation for the present was consistent with the current distribution of the wild boar and was able to predict even newly colonized areas, such as Finland and Sweden. Regarding the LGM distribution, both models predicted the occurrence of the wild boar in Iberia, southern France, Italy and the Balkans. The CCSM model showed smaller areas of climatic suitability, especially in Italy and France, while the MIROC showed wider refugia for the wild boar (Fig. 5).

In fitting the model to the current wild boar distribution, it was possible to evaluate the relative contribution of variables to the presence/absence of the species and take their weight into account in constructing the predictive model. The most important variable was snow depth, followed by mean annual temperature. The standard deviation for the estimations was low, and most MESS values were positive (i.e. present in the training range). Negative MESS values were observed in northern areas, mainly because snow depth during the LGM could reach values currently not observed in Europe, and thus out of the training range of the model.

Predictions for the LGM were consistent with fossil records (Appendix S3c). The presence predicted by the MIROC model best reflected the fossil distribution, especially in southern France, where the model estimated a larger suitable area than the CCSM model.

Comparing the past suitability obtained with MAXENT with the current distribution of haplotype diversity across the sampled populations (Fig. 4), it emerges that areas with higher diversity correspond to areas of predicted presence during the LGM, which were located at lower latitudes. Among the considered models (including geographical coordinates and present and past

climatic suitability for wild boar as factors), the best-fitting model to explain haplotype diversity was that using latitude alone (Table 1), accounting for 17% of the overall variance (B = −0.02, t = −2.99, P = 0.006, adjusted R^2 = 0.172). It is, however, noteworthy that the best five models in Table 1 include geographical location and MIROC, the latter alone representing a good predictor of haplotype diversity (B = 0.38, t = 2.84, P = 0.007, adjusted R^2 = 0.160), as expected by virtue of its high correlation with latitude (R = −0.62, t = −4.78, P < 0.01). Conversely, current assessed suitability alone had no effect on haplotype diversity (B = −0.17, t = −0.62, P = 0.542, adjusted R^2 = −0.017).

Discussion

Mitochondrial DNA sequences from throughout Europe allowed us to reconstruct the wild boar's phylogeographical pattern and to identify the relevant factors that are likely to have generated it. Our main findings can be summarized as follows: (1) as a general pattern, mtDNA genetic diversity decreases northwards; (2) genetic diversity is better explained by climatic suitability during the Last Glacial Maximum than at the present day; (3) genetic and geographical distances are not correlated; (4) Italy and Central Europe show clear genetic similarities; (5) populations at the longitudinal extremes (Iberia and Eastern Europe) are genetically more similar than expected considering their geographical position; and (6) genetic introgression from domestic breeds into the wild boar lineages appears to be limited, due to low frequency of Asian haplotypes, restricted to a few localities.

South–north gradient of genetic diversity

Wild boar populations from northern areas, especially in central Europe, show very low genetic diversity. Southern areas, on the contrary, are more variable, and the Italian peninsula has the highest values of haplotype and nucleotide diversity. High levels of variation are also observed in Iberia (increasing from Portugal to eastern Spain) and in the Balkans, where Greece shows the highest values, in agreement with Alexandri *et al*. (2012). The Sardinian population shares the major mtDNA groups with the Italian mainland, but many haplotypes are exclusive to the island (see also Scandura *et al.*, 2008).

The existence of higher genetic diversity at lower latitudes suggests that southern areas

had an important role as genetic reservoirs during the last glaciation. This effect is confirmed by the correlation between haplotype diversity and the predicted climatic suitability for the species during the LGM. The current suitability, on the other hand, turned out to be a poor predictor of genetic variation, suggesting that an equilibrium situation (where high suitability corresponds to large populations and high genetic diversity) has not been reached after at least 15,000 years of warmer conditions.

An unusual phylogeographical pattern

The geographical distribution of clades and individual haplotypes from Iberia to the Balkans and western Russia shows some specific features not observed in other species. Excluding a few Asian sequences that are likely to be related to introgression events from domestic animals, only two major clades are observed, E1 and E2. E1 is widespread, whereas E2 is found only in the Italian peninsula and Sardinia. When the frequencies of single haplotypes and phylogenetic clades are analysed, a clear genetic affinity emerges between Italian and central European areas, and the eastern and western regions appear distinct but still showing some degree of genetic affinity.

In fact, the two longitudinal extremes in our sample (Iberia and Eastern Europe) share their most frequent haplotypes, a pattern already appreciated in other studies (Scandura *et al.*, 2008; Alexandri *et al*., 2012) but with a far smaller sample of sequences. The spFA also highlighted this similarity, showing few reciprocal mismatches between Iberian and East European populations and, more generally, supporting the genetic proximity of eastern and western areas, which is not compatible with a simple pattern of isolation-by-distance.

This global pattern also seems incompatible with what we know about recent translocations, which were common but probably only left a minor genetic signature at a local scale (e.g. Vernesi *et al*., 2003). Instead, we believe that the last glaciation and the subsequent recolonization processes from southern refugia were important factors in generating this geographical pattern.

Location and role of LGM refugia and post-glacial recolonization routes

The results obtained with MAXENT, integrated with fossil data (Sommer & Nadachowski, 2006; Appendix S3c), point to the following refuge areas: Balkans, Italy, Southern France and Iberia.

These areas are shared by many other species (Hewitt, 2004; Schmitt, 2007; Sommer & Zachos, 2009) including ungulates and, given their current genetic diversity, we interpret them as LGM genetic reservoirs. For example, both spatial predictive models (Stockwell & Peters, 1999) and fossil records (Banks *et al.*, 2008; Sommer *et al.*, 2008) suggest that, during the LGM, the red deer, *Cervus elaphus*, took refuge in the same areas identified for the wild boar, and the same is true for the second most widespread European ungulate, the roe deer, *Capreolus capreolus* (Sommer *et al.*, 2009).

But what happened in terms of connectivity and migration before, during, and after the last glaciation? The simplest hypothesis to justify the current geographical distribution of wild boar mtDNA lineages requires a stronger isolation of the Italian populations than the other European groups, probably occurring both before and during the last glaciation, and two major post-glacial colonization routes: one starting from the Italian and/or the south-western France refugium, and leading wild boars into central Europe, and another starting from the Balkans with wild boars colonizing the north-eastern regions. Conversely, the contribution of Iberian populations seems negligible.

The similarity between Eastern Europe and Iberia may therefore reflect the pre-LGM distribution, when Iberia, Central Europe and Eastern Europe might have formed a single, possibly panmictic group. This scenario is consistent with previous findings revealing weak phylogeographical structuring in pre-LGM populations of European mammals (Hofreiter *et al.*, 2004). On the other hand, the suggested lack of differentiation across Europe in the interglacial implies that the sharing of E1a sequences between Central Europe and Italy should be mostly due to gene flow from the peninsula, while more recent secondary contacts (natural recolonization or translocations from France) could have played a minor role.

We hypothesize that when the ice cap retreated, the recolonization of suitable regions by the southern remnant populations was driven by a density-dependent leading-edge pulse. Rapid expansions were sustained by dense populations, where the major source of migrants was represented by their northern portions (i.e. the edge). Such a mechanism could explain the genetic longitudinal discontinuity we observed in Europe. In particular, it can explain the glaring dissimilarity between populations across the Pyrenees, where none of the seven haplotypes observed in around 80% of the Iberian individuals was found in France, Luxembourg or Germany.

Accordingly, the recolonization of central Europe could have started from high-density populations inhabiting the Italian peninsula and southern France. This expansion presumably prevented further expansion from the Iberian refugium, possibly hosting less abundant populations. In this scenario, dispersal across the Alps would have been assured by the existence of several passes lying at medium elevations (< 2000 m a.s.l.), well accessible for this species. Simultaneously, wild boars from the Balkans would have recolonized north-eastern regions, with a minor contribution to Central Europe. As a result, the two extremes of Europe remained isolated, although they still share some of their pre-LGM diversity. Due to the leading-edge process, southern populations within the Italian and Balkan refugia did not contribute much to the recolonization of new suitable regions in the north, as suggested by their high diversity and genetic distinctiveness (for Greece, see also Alexandri *et al.*, 2012). Such a complex demographic and historical model can also explain the absence of isolation by distance, but it certainly needs further analyses and validation.

The suggested role of Italian refugial populations to the post-glacial recovery of wild boar in Europe may appear to be contradicted by the absence of the E2 clade north of the Alps. E2 haplotypes were (Larson *et al*., 2007) and remain common only in mainland Italy and Sardinia, although they were also detected in ancient specimens from Croatia (dated to around 11,000 years ago; Larson *et al*., 2007). Their current absence from the rest of Europe can be attributed to a low frequency in leading-edge populations at the time of the post-glacial population expansion, and to successive drift events, due to demographic oscillations. The moderate frequency of E2 we observed in northern Italy is more likely to be related to recent dispersal or translocation events, because the wild boar was extinct in this region at the beginning of the 20th century (Apollonio *et al.*, 1988).

Conclusions

As recently noted, a latitudinal gradient of intraspecific genetic diversity seems to be the rule in mammals (Adams & Hadly, 2013). In addition to greater species richness, low latitudes tend to show a higher genetic variation within species than regions closer to the pole (Guo, 2012). This pattern can be attributed to periodic global climate changes which led to repeated extinctions toward the poles, followed by natural recolonizations (Hewitt, 2004).

This general trend is observed also in the European wild boar, despite the fact that

translocation, restocking activities and extensive hunting have been common in this species in recent times. Our results are consistent with a prominent role for climatic and habitat oscillation during the Quaternary. Southern areas acted as genetic reservoirs in glacial times, and northern areas were mainly recolonized from Italian and French refugia in central Europe, and from the Balkans in Eastern Europe. Leading-edge expansion and density-dependent migration processes are also required to explain the complex mtDNA phylogeographical pattern we observed. Further studies on additional (nuclear) markers are needed to test our biogeographical reconstruction.

Acknowledgements

Sequencing activities were funded by the Sardinian Regional Government (L.R. 7/2007 Promozione della Ricerca Scientifica e dell'Innovazione Tecnologica in Sardegna). D.B. had a PhD grant financed by the European Social Fund, Operational Programme 2007/2013 – Axis IV – Human Capital Investment. S.T.V. was supported by a PhD Fellowship for International students of the University of Ferrara. B.J. and T.B. were supported by grants N N304 253935 from the Polish Ministry of Science and Higher Education and Research Potential in Conservation and Sustainable Management of Biodiversity – BIOCONSUS from the EC's 7th Framework Programme (contract no. 245737). Mobility between labs in Sassari and Kiel was financed by the Italy–Germany Vigoni cooperation program 2010–2011. We are grateful to R. Chirichella, A. Canu, G. Falceri, the Population and Conservation Genetics group at the University of Ferrara, and especially to A. Benazzo and S. Hoban, for their help in statistical analysis, English review, and insightful discussions. Thanks are due to CASPUR standard HPC grant (2011) and CINECA-ISCRA HPC class C grant (2010) for computational resources awarded to S.T.V.

References

Adams, R. & Hadly, E. (2013) Genetic diversity within vertebrate species is greater at lower latitudes. *Evolutionary Ecology*, **27**, 133–143.

van Andel, T.H. (2002) The climate and landscape of the middle part of the Weichselian glaciation in Europe: the Stage 3 Project. *Quaternary Research*, **57**, 2–8.

Alexandri, P., Triantafyllidis, A., Papakostas, S., Chatzinikos, E., Platis, P., Papageorgiou, N., Larson, G., Abatzopoulos, T.J. & Triantaphyllidis, C. (2012) The Balkans and the colonization of Europe: the post-glacial range expansion of the wild boar, *Sus scrofa*. *Journal of Biogeography*, **39**, 713–723.

Alves, P.C., Pinheiro, I., Godinho, R., Vicente, J., Gortázar, C. & Scandura, M. (2010) Genetic diversity of wild boar populations and domestic pig breeds (*Sus scrofa*) in South-western Europe. *Biological Journal of the Linnean Society*, **101**, 797–822.

Apollonio, M., Randi, E. & Toso, S. (1988) The systematics of the wild boar (*Sus scrofa* L.) in Italy. *Bollettino di Zoologia*, **3**, 213–221.

Apollonio, M., Andersen, R. & Putman, R. (eds) (2010) *European ungulates and their management in the 21st century*. Cambridge University Press, Cambridge, UK.

Bandelt, H.J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Banks, W.E., d'Errico, F., Peterson, A.T., Kageyama, M. & Colombeau, G. (2008) Reconstructing ecological niches and geographic distributions of caribou (*Rangifer tarandus*) and red deer (*Cervus elaphus*) during the Last Glacial Maximum. *Quaternary Science Reviews*, **27**, 2568–2575.

Berg, F., Gustafson, U. & Andersson, L. (2006) The uncoupling protein 1 gene *(UCP1)* is disrupted in the pig lineage: a genetic explanation for poor thermoregulation in piglets. *PLoS Genetics*, **2**, e129.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Excoffier, L. & Lischer, H.E.L. (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.

Frichot, E., Schoville, S., Bouchard, G., François, O. (2012) Correcting principal component maps for effects of spatial autocorrelation in population genetic data. *Frontiers in Genetics*, **3**, 254.

Fu, Y-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Geisser, H. & Reyer, H.-U. (2005) The influence of food and temperature on population density of wild boar *Sus scrofa* in the Thurgau (Switzerland). *Journal of Zoology*, **267**, 89–96.

Giuffra, E., Kijas, J.M.H., Amarger, V., Carlborg, Ö., Jeon, J.-T. & Andersson, L. (2000) The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics*, **154**, 1785–1791.

Guo, Q. (2012) Incorporating latitudinal and central–marginal trends in assessing genetic variation across species ranges. *Molecular Ecology*, **21**, 5396–5403.

Hajji, G.E. & Zachos, F.E. (2011) Mitochondrial and nuclear DNA analyses reveal pronounced genetic structuring in Tunisian wild boar *Sus scrofa*. *European Journal of Wildlife Research*, **57**, 449–456.

Hasegawa, M., Kishino, H. & Yano, T-A. (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Herrero, J., Couto, S., Rosell, C. & Arias, P. (2004) Preliminary data on the diet of wild boar living in a Mediterranean coastal wetland. *Galemys*, **16**, 115–123.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hofreiter, M., Serre, D., Rohland, N., Rabeder, G., Nagel, D., Conard, N., Münzel, S. & Pääbo, S. (2004) Lack of phylogeography in European mammals before the last glaciation. *Proceedings of the National Academy of Sciences USA*, **101**, 12963–12968.

Hosmer, D.W., Lemeshow, S. (2000) *Applied logistic regression*, 2nd edn. John Wiley & Sons, New York.

Larson, G., Dobney, K., Albarella, U., Fang, M-Y., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L. & Cooper, A. (2005)

Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, **307**, 1618–1621.

Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., Vigne, J.-D., Edwards, C.J., Schlumbaum, A., Dinu, A., Bălăçsescu, A., Dolman, G., Tagliacozzo, A., Manaseryan, N., Miracle, P., van Wijngaarden-Bakker, L., Masseti, M., Bradley, D.G. & Cooper, A. (2007) Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences USA*, **104**, 15276–15281.

Mantel, N.A. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Melis, C., Szafrańska, P.A., Jędrzejewska, B. & Bartoń, K. (2006) Biogeographical variation in the population density of wild boar (*Sus scrofa*) in western Eurasia. *Journal of Biogeography*, **33**, 803–811.

Morueta-Holme, N., Fløjgaard, C. & Svenning, J.-C. (2010) Climate change risks and conservation implications for a threatened small-range mammal species. *PLoS ONE*, **5**, e10360.

Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Petit, R.J., El Mousadik, A. & Pons, O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Posada, D. (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Posada, D. (2011) *Collapse 1.2: describing haplotypes from sequence alignments. Website http://www.softpedia.com/get/Science-CAD/Collapse.shtml (accessed on 13 November 2013).*

R Development Core Team (2012) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Scandura, M., Iacolina, L., Crestanello, B., Pecchioli, E., Di Benedetto, M.F., Russo, V., Davoli, R., Apollonio, M. & Bertorelle, G. (2008) Ancient vs. recent processes as factors shaping the

genetic variation of the European wild boar: are the effects of the last glaciation still detectable? *Molecular Ecology*, **17**, 1745–1762.

Scandura, M., Iacolina, L. & Apollonio, M. (2011) Genetic diversity in the European wild boar *Sus scrofa*: phylogeography, population structure and wild × domestic hybridization. *Mammal Review*, **41**, 125–137.

Schley, L. & Roper, T.J. (2003) Diet of wild boar *Sus scrofa* in Western Europe, with particular reference to consumption of agricultural crops. *Mammal Review*, **33**, 43–56.

Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

Sommer, R.S. & Nadachowski, A. (2006) Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Review*, **36**, 251–265.

Sommer, R.S. & Zachos, F.E. (2009) Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. *Journal of Biogeography*, **36**, 2013–2020.

Sommer, R.S., Zachos, F.E., Street, M., Jöris, O., Skog, A. & Benecke, N. (2008) Late Quaternary distribution dynamics and phylogeography of the red deer (*Cervus elaphus*) in Europe. *Quaternary Science Reviews*, **27**, 714–733.

Sommer, R.S., Fahlke, J.M., Schmölcke, U., Benecke, N. & Zachos, F.E. (2009) Quaternary history of the European roe deer *Capreolus capreolus*. *Mammal Review*, **39**, 1–16.

Stockwell, D. & Peters, D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, **13**, 143–158.

Symonds, M. & Moussalli, A. (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, **65**, 13–21.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Ursing, B.M. & Arnason, U. (1998) The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). *Journal of Molecular Evolution*, **47**, 302–306.

Vandenberghe, J., Renssen, H., Roche, D.M., Goosse, H., Velichko, A.A., Gorbunov, A. & Levavasseur, G. (2012) Eurasian permafrost instability constrained by reduced sea-ice cover. *Quaternary Science Reviews*, **34**, 16–23.

Vernesi, C., Crestanello, B., Pecchioli, E., Tartari, D., Caramelli, D., Hauffe, H. & Bertorelle, G. (2003) The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*): a microsatellite analysis. *Molecular Ecology*, **12**, 585–595.

Table 1 Comparison of the performance of multiple linear regression models explaining gene diversity in European wild boar (*Sus scrofa*) populations on the basis of their geographical position (Lat., latitude; Long., longitude), and of present (Current) and past (Last Glacial Maximum: MIROC or CSSM) environmental suitability. The Russian population was excluded for reasons explained in the text. Environmental suitability was assessed in MAXENT on the basis of presence-only data and eleven climatic variables. Climatic data at LGM were estimated on the basis of the MIROC and CSSM global climatic models.

Model predictors	AIC_c	Δ_i	W_i	cum. W_i	adj. R^2
Lat.	−12.091	0	0.154	0.154	0.172
MIROC	−11.574	0.517	0.119	0.273	0.160
Lat. + Long. + MIROC	−11.531	0.559	0.116	0.390	0.223
Lat. + MIROC	−11.440	0.651	0.111	0.501	0.189
Lat. + Long.	−11.026	1.065	0.090	0.591	0.180
MIROC + Current	−10.681	1.410	0.076	0.668	0.172
Lat. + Long. + MIROC + Current	−10.319	1.772	0.064	0.731	0.233
Long. + MIROC	−9.991	2.100	0.054	0.915	0.157
Lat. + Current	−9.645	2.446	0.045	0.830	0.149
Lat. + CCSM	−9.602	2.489	0.044	0.875	0.148
Lat. + MIROC + Current	−9.429	2.662	0.041	0.915	0.179
Lat. + Long. + Current	−8.518	3.573	0.026	0.941	0.159

AIC_c, AIC corrected for small sample sizes; Δ_i, difference in AIC_c between model *i* and the best model; W_i, weight of model *i*; cum. W_i, cumulative weight of model *i* and the upper models; adj. R^2, adjusted R^2.

Figure Legends

Figure 1 Median-joining network of mitochondrial D-loop haplotypes (411 bp) observed in 1099 wild boar (*Sus scrofa*) from Europe, Asia and North Africa. Circle size is proportional to haplotype frequencies; border colour refers to haplogroups (see Fig. 2); fill colour denotes geographical distribution; numbers indicate haplotype codes (abbreviated to save space: e.g. 29 denotes haplotype H029). Connections with more than one nucleotide change are identified by transverse bars. Major clades (A, E1 and E2) are delimited by dashed lines. A sequence of *Sus barbatus* was used as outgroup.

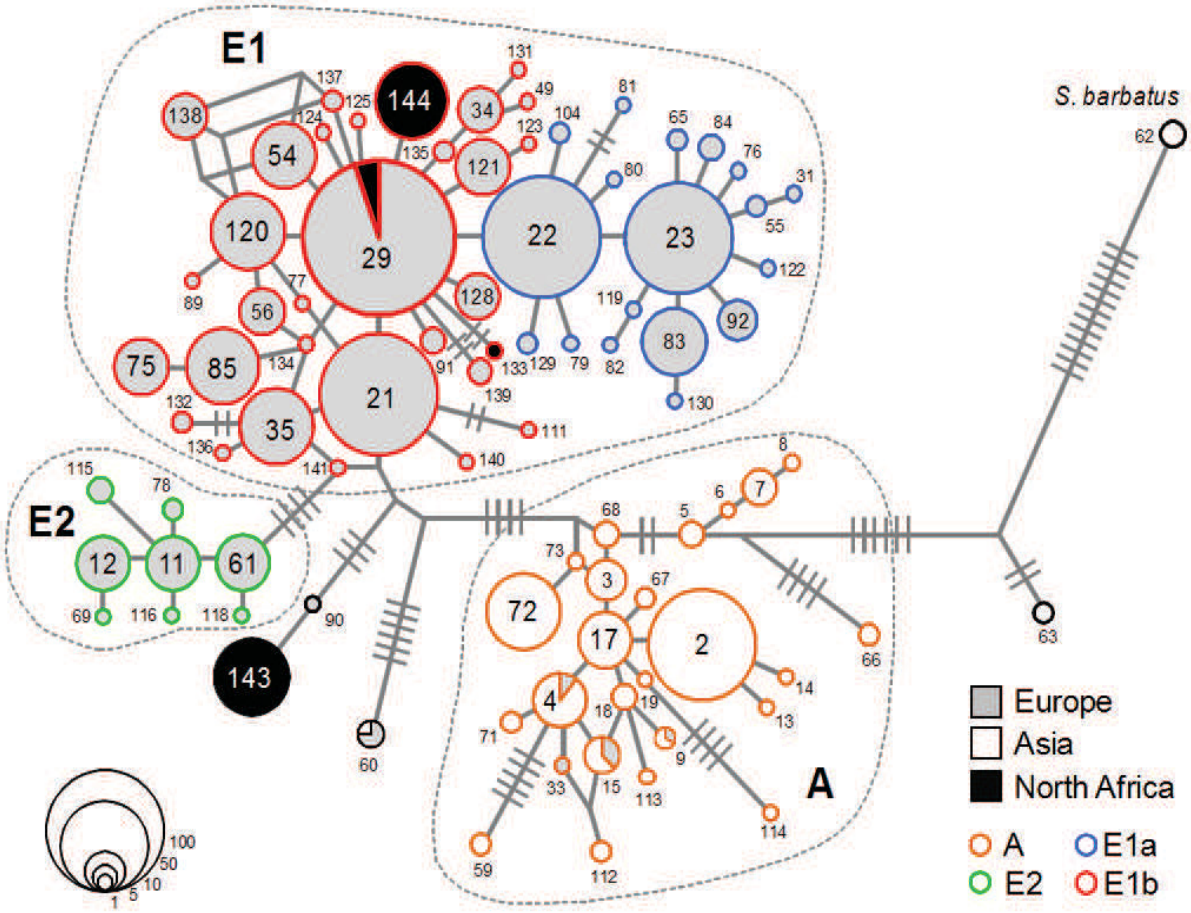


Figure 2 Mitochondrial DNA haplogroups observed in 763 European wild boar (*Sus scrofa*). (a) Bayesian tree of the 50 mitochondrial D-loop haplotypes. The tree is rooted using a homologous sequence of *Sus barbatus*. Branch colours represent different haplogroups. Posterior probabilities and major clades are indicated on internodes. The E1 haplotypes not belonging to the well supported E1a clade were grouped into one haplogroup (E1c) for convenience. E1a and E1c roughly correspond to the haplogroups appearing in previous papers and referred to as A-side and C-side, respectively (Larson *et al.*, 2007; Scandura *et al.*, 2011; Alexandri *et al.*, 2012). (b) Map showing the frequency of the three haplogroups (E1a, E1c and E2) in each of the 39 European populations (numbers indicate population codes).

Fig. 2a)

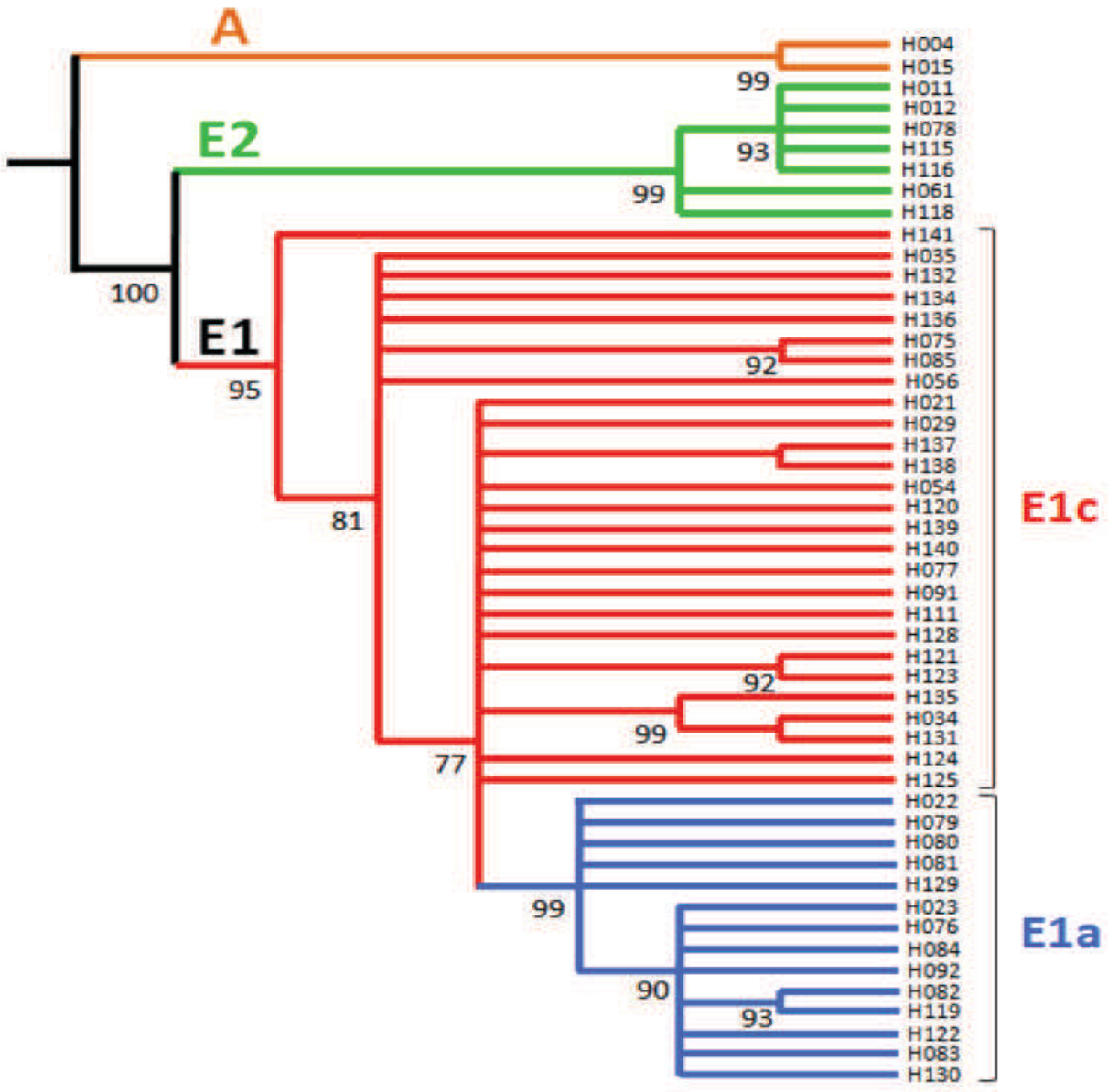

A
E2
E1
100
95
99
93
99
92
81
77
99
92
99
90
93
H004
H015
H011
H012
H078
H115
H116
H061
H118
H141
H035
H132
H134
H136
H075
H085
H056
H021
H029
H137
H138
H054
H120
H139
H140
H077
H091
H111
H128
H121
H123
H135
H034
H131
H124
H125
H022
H079
H080
H081
H129
H023
H076
H084
H092
H082
H119
H122
H083
H130
E1c
E1a

Fig. 2b)

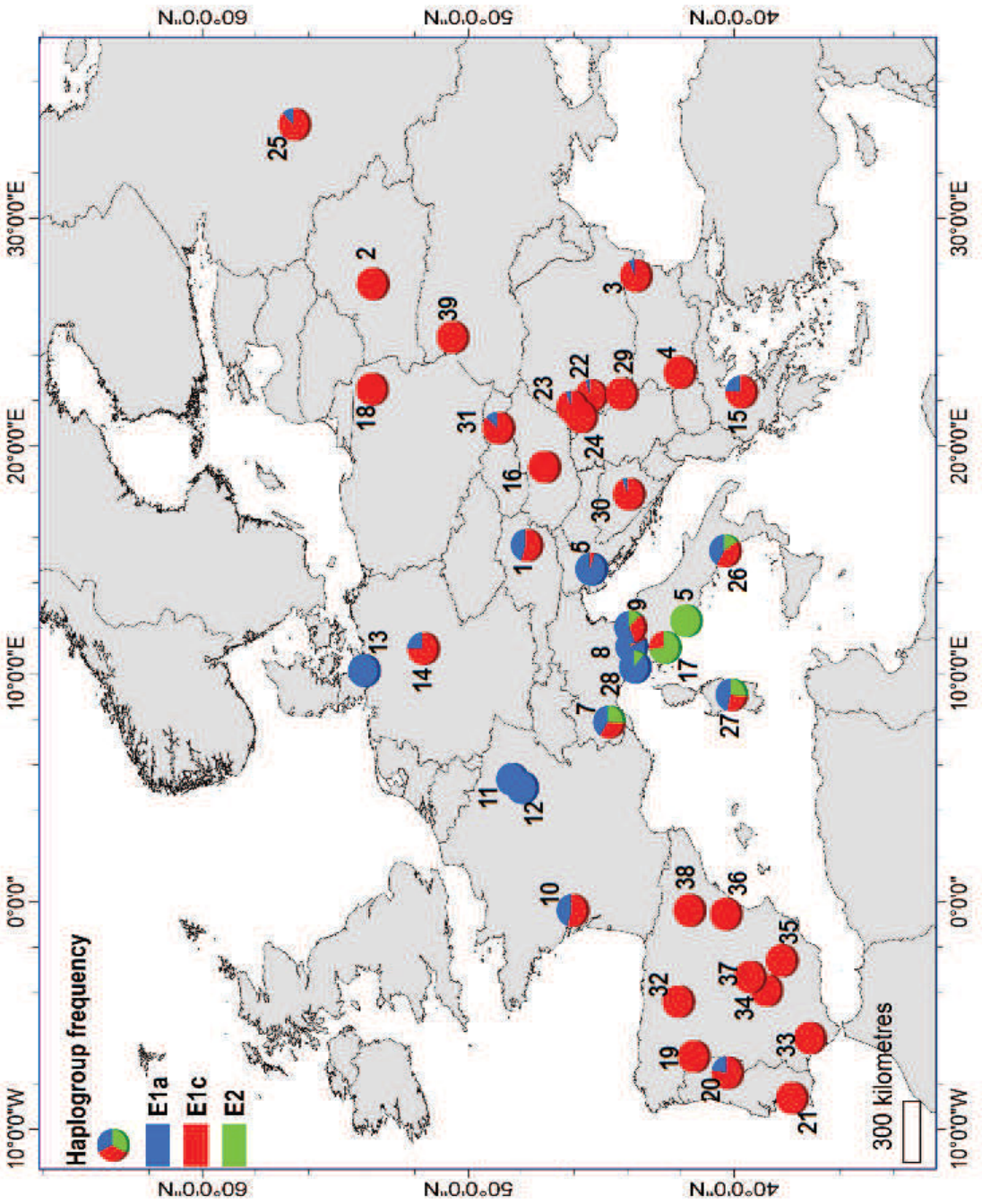

60°0'0"N
50°0'0"N
40°0'0"N
30°0'0"E
20°0'0"E
10°0'0"E
0°0'0"
10°0'0"W
Haplogroup frequency
E1a
E1c
E2
300 kilometres
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

Figure 3 Bidimensional plot of the European wild boar populations (*Sus scrofa*) sampled in this study, obtained by spatial factor analysis (spFA). Black dots represent populations located in Italy and central Europe, grey stars represent populations in Iberia, and black triangles represent populations in the Balkans and Eastern Europe. Numbers indicate population codes.

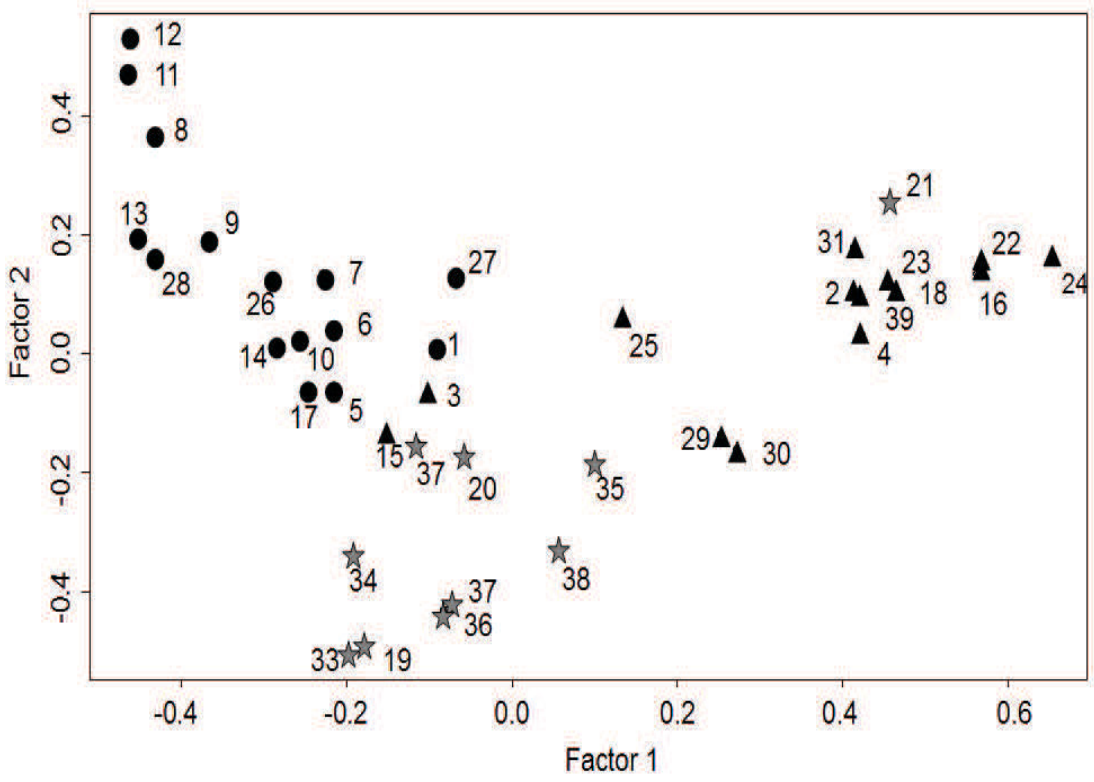


Figure 4 Estimated distribution of mitochondrial (D-loop) haplotype diversity (*Hk*) in wild boar (*Sus scrofa*) across Europe, resulting from the interpolation of 38 different sampling populations (Russia excluded) using the kriging method in ArcGIS 10. Bluish colours correspond to low values of *Hk*, while reddish colours represent high values.

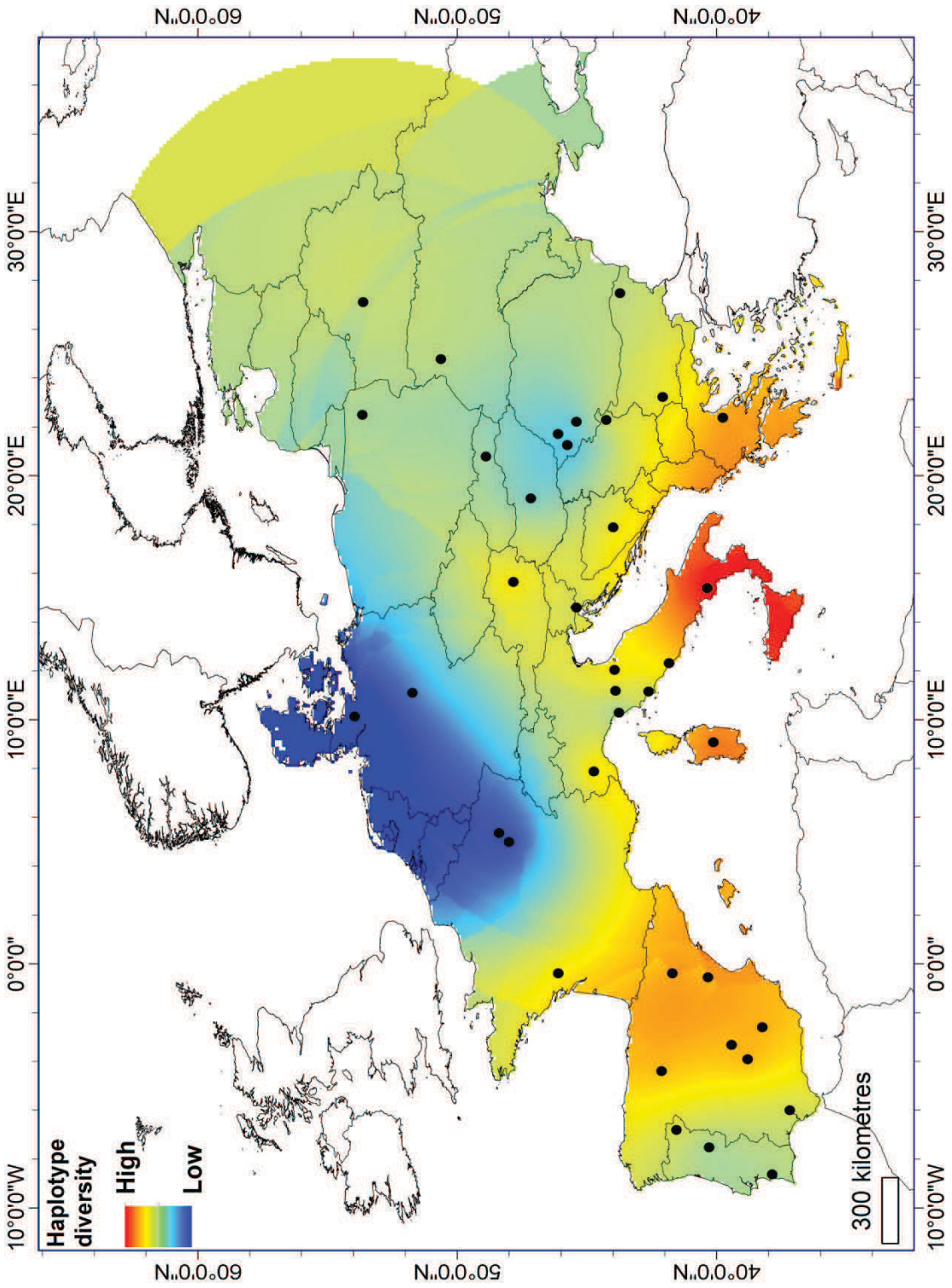

60°0'0"N
50°0'0"N
40°0'0"N
10°0'0"W
0°0'0"
10°0'0"E
20°0'0"E
30°0'0"E
Haplotype diversity
High
Low
300 kilometres

Figure 5 Distribution range of wild boar (*Sus scrofa*) in Europe, as predicted on the basis of climatic suitability data estimated for the present time and for the Last Glacial Maximum (LGM). (a) Current predicted distribution; (b) LGM distribution based on the MIROC model; (c) LGM distribution based on the CCSM model.


a
b
c
Unsuitable
Suitable

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Current status of wild boar in different European countries and comparison with the early 20th century.

Appendix S2 Information on samples and sequences used for phylogenetic and phylogeographical analyses.

Appendix S3 Additional information and calculations on wild boar distribution and mtDNA diversity in Europe.

Appendix S1 Current status of wild boar in different European countries and comparison with the early 20th century. n/a, data not available.

Country	Early 20th century	Early 21st century	Source
Austria	nearly absent	Widespread, more common in the east (annual hunting bag 30,000 head)	Reimoser & Reimoser (2010)
Baltic countries	limited to the southern regions	Widespread	Andersone-Lilley *et al.* (2010)
Belgium	present	Almost ubiquitous	Casaer & Licoppe (2010)
Croatia	present	Ubiquitous in forested areas (population size 18,200 head); hybridization with domestic pigs in the eastern part of the country	Kusak & Krapinec (2010)
Czech Republic	absent	Widespread after natural immigration (probably from Hungary)	Bartoš *et al.* (2010)
Denmark	absent	Very limited presence (fenced animals and incursions from Germany)	Andersen & Holthe (2010)
Finland	n/a	Some hundreds of animals dispersing from Russia	Ruusila & Kojola (2010)
France	present	Widespread (annual hunting bag 400,000 head), illegal introductions of possible hybrids	Maillard *et al.* (2010)
Germany	present	Widespread in the country, except for the alpine region in the South (annual hunting bag 500,000 head)	Wotschikowsky (2010)
Great Britain	absent	Some populations originating from animals that have escaped from enclosures, possibly crossbred with domestic pigs	Putman (2010)
Greece	present	Widely distributed but declining; reintroductions and restocking reported	Papaioannou (2010)
Hungary	present	Widespread, but more abundant in the north and west (annual hunting bag: about 90,000 head)	Csányi & Lehoczki (2010)
Ireland	absent	Few animals escaped from enclosures	Carden (2012)

Country	Early 20th century	Early 21st century	Source
Italy	present	Widely distributed (annual hunting bag > 110,000 heads); recovery after reintroductions and restocking	Apollonio *et al.* (2010)
Macedonia	n/a	Widespread (annual hunting bag > 700 head)	Stojanov *et al.* (2010)
Netherlands	absent	Population originating from reintroduced Eastern European animals and immigrants from Germany	van Wieren & Groot Bruinderink (2010)
Norway	absent	Occasional immigrants from Sweden	Andersen *et al.* (2010)
Poland	present	More abundant in western regions (annual hunting bag 136,000 animals in total)	Wawrzyniak *et al.* (2010)
Portugal	restricted to mountains and some fenced areas	Widespread (annual hunting bag about 8000 head)	Vingada *et al.* (2010)
Romania	present	Widespread, more common in the west (about 60,000 head estimated); hybridization in the Danube delta	Micu *et al.* (2010)
Serbia	n/a	Widespread (annual hunting bag > 2000 head)	Paunović *et al.* (2010)
Slovakia	present	Almost widespread (annual hunting bag: 15,000–20,000 head)	Find'o & Skuban (2010)
Slovenia	rare	Patchy distribution; population originating from a fenced stock with animals imported from Germany (annual hunting bag 7000 head)	Adamic *et al.* (2010)
Spain	present (from north-east to south-west)	Widespread (annual hunting bag 150,000 head)	Carranza (2010)
Sweden	absent	Population originating from Central European animals that have escaped from enclosures, possibly cross-bred with domestic pigs	Liberg *et al.* (2010)
Switzerland	absent	Expanding population arising from immigration from Germany (about 5000 head killed annually)	Imesch-Bebié *et al.* (2010)

REFERENCES

Adamic, M. & Jerina, K. (2010) Ungulates and their management in Slovenia. *European ungulates and their*

management in the 21st century (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 507–526. Cambridge University Press, Cambridge, UK.

Andersen, R. & Holthe, V. (2010) Ungulates and their management in Denmark. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 71–85. Cambridge University Press, Cambridge, UK.

Andersen, R., Lund, E., Solberg, E.J. & Sæther, B.-E. (2010) Ungulates and their management in Norway. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 14–36. Cambridge University Press, Cambridge, UK.

Andersone-Lilley, Ž., Balčiauskas, L., Ozoliņš, J., Randveer, T. & Tõnisson, J. (2010) Ungulates and their management in the Baltics (Estonia, Latvia and Lithuania). *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 103–128. Cambridge University Press, Cambridge, UK.

Apollonio, M., Ciuti, S., Pedrotti, L. & Banti, P. (2010) Ungulates and their management in Italy. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 475–506. Cambridge University Press, Cambridge, UK.

Bartoš, L., Kotrba, R. & Pintíř, J. (2010) Ungulates and their management in the Czech Republic. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 243–261. Cambridge University Press, Cambridge, UK.

Carden, R.F. (2012) *Review of the natural history of wild boar (*Sus scrofa*) on the island of Ireland.* Northern Ireland Environment Agency, Belfast; National Parks and Wildlife Service, Dublin; National Museum of Ireland, Dublin.

Carranza, J. (2010) Ungulates and their management in Spain. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 419–440. Cambridge University Press, Cambridge, UK.

Casaer, J. & Licoppe, A. (2010) Ungulates and their management in Belgium. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 184–200. Cambridge University Press, Cambridge, UK.

Csányi, S. & Lehoczki, R. (2010) Ungulates and their management in Hungary. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 291–318. Cambridge University Press, Cambridge, UK.

Find'o, S. & Skuban, M. (2010) Ungulates and their management in Slovakia. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 262–290. Cambridge University Press, Cambridge, UK.

Imesch-Bebié, N., Gander, H. & Schnidrig-Petrig, R. (2010) Ungulates and their management in Switzerland. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 357–391. Cambridge University Press, Cambridge, UK.

Kusak, J. & Krapinec, K. (2010) Ungulates and their management in Croatia. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 527–539. Cambridge University Press, Cambridge, UK.

Liberg, O., Bergström, R., Kindberg, J. & von Essen, H. (2010) Ungulates and their management in Sweden. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 37–70. Cambridge University Press, Cambridge, UK.

Maillard, D., Gaillard, J.-M., Hewison, M., Ballon, P.,

Duncan, P., Loison, A., Toïgo, C., Baubet, E., Bonenfant, C., Garel, M. & Saint-Andrieux, C. (2010) Ungulates and their management in France. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 441–474. Cambridge University Press, Cambridge, UK.

Micu, I., Náhlik, A., Neguş, S., Mihalache, I. & Szabó, I. (2010) Ungulates and their management in Romania. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen & R. Putman), pp. 319–337. Cambridge University Press, Cambridge, UK.

Papaioannou, H. (2010) Ungulates and their management in Greece. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 540–562. Cambridge University Press, Cambridge, UK.

Paunović, M., Ćirović, D. & Linnell, J.D.C. (2010) Ungulates and their management in Serbia. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 563–571. Cambridge University Press, Cambridge, UK.

Putman, R. (2010) Ungulates and their management in Great Britain and Ireland. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 129–164. Cambridge University Press, Cambridge, UK.

Reimoser, F. & Reimoser, S. (2010) Ungulates and their management in Austria. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 338–356. Cambridge University Press, Cambridge, UK.

Ruusila, V. & Kojola, I. (2010) Ungulates and their management in Finland. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 86–102. Cambridge University Press, Cambridge, UK.

Stojanov, A., Melovski, D., Ivanov, G. & Linnell, J.D.C. (2010) Ungulates and their management in Macedonia. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 572–577. Cambridge University Press, Cambridge, UK.

Vingada, J., Fonseca, C., Cancela, J., Ferreira, J. & Eira, C. (2010) Ungulates and their management in Portugal. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 392–418. Cambridge University Press, Cambridge, UK.

Wawrzyniak, P., Jędrzejewski, W., Jędrzejewska, B. & Borowik, T. (2010) Ungulates and their management in Poland. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 223–242. Cambridge University Press, Cambridge, UK.

van Wieren, S.E. & Groot Bruinderink, G.W.T.A. (2010) Ungulates and their management in the Netherlands. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 165–183. Cambridge University Press, Cambridge, UK.

Wotschikowsky, U. (2010) Ungulates and their management in Germany. *European ungulates and their management in the 21st century* (ed. by M. Apollonio, R. Andersen and R. Putman), pp. 201–222. Cambridge University Press, Cambridge, UK.

Appendix S2 List of mtDNA sequences (partial control region, 411 bp) used for phylogenetic and phylogeographical analyses. For new sequences matching already published haplotypes a reference GenBank accession number is indicated; new haplotypes (KC608827–KC608847) are shaded. Asterisks mark original Tunisian sequences from Hajji & Zachos (2010) that were extended in the present study to be fully aligned with the other sequences in the dataset. No. ind., number of individuals; GenBank, GenBank accession number.

Specimen/Isolate	No. ind.	Country	Location	Continent	Haplotype	GenBank	Source
WB56	1	Japan		Asia	H002	AB015084	Okumura *et al.* (2001)
WB55	7	Japan		Asia	H003	AB015085	Okumura *et al.* (2001)
WB72	2	Japan		Asia	H004	AB015086	Okumura *et al.* (2001)
WB1	3	Japan	Ryukyu	Asia	H005	AB015087	Okumura *et al.* (2001)
WB2	1	Japan	Ryukyu	Asia	H006	AB015088	Okumura *et al.* (2001)
WB6	7	Japan	Ryukyu	Asia	H007	AB015089	Okumura *et al.* (2001)
WB10	1	Japan	Ryukyu	Asia	H008	AB015090	Okumura *et al.* (2001)
European wild boar 1	2	Italy		Europe	H011	AB015094	Okumura *et al.* (2001)
European wild boar 3	1	Italy		Europe	H012	AB015095	Okumura *et al.* (2001)
Tokushima.5001	1	Japan	Tokushima	Asia	H013	AB041467	Okumura *et al.* (2001)
Gunma.4	14	Japan		Asia	H004	AB041468	Okumura *et al.* (2001)
WB7	5	Japan	Miyazaki	Asia	H002	AB041469	Okumura *et al.* (2001)
ù	11	Japan	Miyazaki	Asia	H002	AB041470	Okumura *et al.* (2001)
WB12	2	Japan	Miyazaki	Asia	H002	AB041471	Okumura *et al.* (2001)
WB43	1	Japan	Miyazaki	Asia	H014	AB041472	Okumura *et al.* (2001)
WB49	3	Japan	Miyazaki	Asia	H002	AB041473	Okumura *et al.* (2001)
WB4	1	Italy		Europe	H033	AB059651	Okumura *et al.* (2001)
Jpn.Wild Boar	1	Japan		Asia	H015	AB059652	Okumura *et al.* (2001)
EWB1	1	Poland		Europe	H034	AF136555	Giuffra *et al.* (2000)
EWB2	1	Poland		Europe	H029	AF136556	Giuffra *et al.* (2000)
EWB3	1	Italy		Europe	H012	AF136563	Giuffra *et al.* (2000)
Asian WB10	1	Japan		Asia	H002	AF136564	Giuffra *et al.* (2000)
Asian WB11	1	Japan		Asia	H017	AF136565	Giuffra *et al.* (2000)
GL72	1	Japan		Asia	H004	AY884634	Larson *et al.* (2005)
GL83	1	India		Asia	H019	AY884643	Larson *et al.* (2005)
GL81	1	China		Asia	H015	AY884642	Larson *et al.* (2005)
Finland.36	1	Finland		Europe	H034	AF535163	Gongora *et al.* (2003)
Finland.41	1	Finland		Europe	H049	AF535164	Gongora *et al.* (2003)
Ssc1	1	Bulgaria		Europe	H029	AJ314542	Randi *et al.* (2002)
Ssc2	1	Japan		Asia	H003	AJ314543	Randi *et al.* (2002)
Ssc3	1	Italy	Sardinia	Europe	H022	AJ314544	Randi *et al.* (2002)
SWB1	1	Spain		Europe	H054	AY232868	Alves *et al.* (2003)
SWB2	1	Spain		Europe	H054	AY232869	Alves *et al.* (2003)
SWB3	1	Spain		Europe	H054	AY232870	Alves *et al.* (2003)
SWB4	1	Spain		Europe	H055	AY232871	Alves *et al.* (2003)

Specimen/Isolate	No. ind.	Country	Location	Continent	Haplotype	GenBank	Source
SWB5	1	Spain		Europe	H055	AY232872	Alves *et al.* (2003)
SWB6	1	Spain		Europe	H021	AY232873	Alves *et al.* (2003)
SWB7	1	Spain		Europe	H056	AY232874	Alves *et al.* (2003)
GL232	1	China	Amoy	Asia	H009	AY884707	Larson *et al.* (2005)
GL135	1	India	Monghyr, Bengal	Asia	H059	AY884674	Larson *et al.* (2005)
GL162	1	India	Kashmir, Woolar Lake	Asia	H059	AY884689	Larson *et al.* (2005)
GL46	1	India	Kashmir, Valley of Kashmir	Asia	H059	AY884612	Larson *et al.* (2005)
GL52	1	Spain	Burgos, Palacios de la Sierra	Europe	H089	AY884616	Larson *et al.* (2005)
GL59	1	Iran	Hamadan	Asia	H060	AY884622	Larson *et al.* (2005)
GL142	1	Armenia	Vaik region	Europe	H060	AY884680	Larson *et al.* (2005)
GL236	1	Armenia		Europe	H060	AY884710	Larson *et al.* (2005)
GL271	1	Armenia		Europe	H060	AY884727	Larson *et al.* (2005)
GL63	1	Germany	Harz Mountains	Europe	H022	AY884626	Larson *et al.* (2005)
GL65	1	Italy	Sardinia, Sinnai	Europe	H061	AY884628	Larson *et al.* (2005)
GL70	1	Malaysia	Johore	Asia	H062	AY884632	Larson *et al.* (2005)
GL88	1	Indonesia	Rhio archipelago, Pulo Jombol	Asia	H062	AY884648	Larson *et al.* (2005)
GL101	1	Indonesia	Rhio archipelago, Pulo Penjait	Asia	H062	AY884659	Larson *et al.* (2005)
GL138	1	Malaysia	Johore, Mount Austin	Asia	H062	AY884676	Larson *et al.* (2005)
GL71	1	France	Côte d'Or, Is-sur-Tille	Europe	H065	AY884633	Larson *et al.* (2005)
GL73	1	Morocco	Ougda, Taforalt	Africa	H029	AY884635	Larson *et al.* (2005)
GL84	1	Indonesia	Nias Samasama	Asia	H063	AY884644	Larson *et al.* (2005)
GL99	1	Indonesia	Nias Feliwaa	Asia	H063	AY884657	Larson *et al.* (2005)
GL107	1	Germany	Czech / German border	Europe	H023	AY884664	Larson *et al.* (2005)
GL108	1	Germany	WFG Wriesen	Europe	H022	AY884665	Larson *et al.* (2005)
GL109	1	Germany	Rathsdorf	Europe	H029	AY884666	Larson *et al.* (2005)
GL110	1	Corsica		Europe	H022	AY884667	Larson *et al.* (2005)
GL111	1	Italy	Sardinia	Europe	H022	AY884668	Larson *et al.* (2005)
GL112	1	the Netherland	Het Loo	Europe	H022	AY884669	Larson *et al.* (2005)
GL113	1	Macedonia		Europe	H021	AY884670	Larson *et al.* (2005)
GL143	1	Frace	Corsica	Europe	H083	AY884681	Larson *et al.* (2005)
GL144	1	Italy	Sardinia, Tempio Pausania	Europe	H065	AY884682	Larson *et al.* (2005)
GL146	1	China	Hunan Yochow	Asia	H018	AY884683	Larson *et al.* (2005)
GL78	1	China	Shensi, Yen-An-Fu	Asia	H066	AY884639	Larson *et al.* (2005)
GL64	1	China	Shanxi Fen Chow	Asia	H066	AY884627	Larson *et al.* (2005)
GL147	1	China	Shanxi Tai-Yuan-Fu	Asia	H066	AY884684	Larson *et al.* (2005)
GL190	1	Italy	Sardinia	Europe	H061	AY884690	Larson *et al.* (2005)
GL192	1	Russia	Vladivostock	Asia	H009	AY884692	Larson *et al.* (2005)
GL191	1	Russia	Vladivostock	Asia	H067	AY884691	Larson *et al.* (2005)
GL220	1	France	Corsica	Europe	H090	AY884696	Larson *et al.* (2005)
GL221	1	Spain	Asturias	Europe	H029	AY884697	Larson *et al.* (2005)
GL222	1	Portugal	Estremadra	Europe	H091	AY884698	Larson *et al.* (2005)
GL60	1	Burma	Tenasserim, Bok Pyin	Asia	H068	AY884623	Larson *et al.* (2005)
GL67	1	Burma	Tenasserim, Kisseraing Islan	Asia	H068	AY884629	Larson *et al.* (2005)
GL68	1	Thailand	Trang Trong	Asia	H068	AY884630	Larson *et al.* (2005)
GL86	1	Malaysia	Johore, Muar	Asia	H018	AY884646	Larson *et al.* (2005)
GL237	1	Russia	South-eastern Siberia	Asia	H002	AY884711	Larson *et al.* (2005)

Specimen/Isolate	No. ind.	Country	Location	Continent	Haplotype	GenBank	Source
GL240	1	Burma	Tenasserim, Champang	Asia	H068	AY884712	Larson *et al.* (2005)
GL242	1	Spain	Aragon	Europe	H091	AY884714	Larson *et al.* (2005)
GL244	1	Italy	Maremma	Europe	H069	AY884716	Larson *et al.* (2005)
GL245	1	Italy	Maremma	Europe	H011	AY884717	Larson *et al.* (2005)
GL246	1	Italy	Maremma	Europe	H011	AY884718	Larson *et al.* (2005)
GL247	1	Italy	Maremma	Europe	H012	AY884719	Larson *et al.* (2005)
GL248	1	Italy	Maremma	Europe	H012	AY884720	Larson *et al.* (2005)
GL249	1	Italy	Maremma	Europe	H012	AY884721	Larson *et al.* (2005)
GL250	1	Italy	Maremma	Europe	H011	AY884722	Larson *et al.* (2005)
GL251	1	Italy	Maremma	Europe	H011	AY884723	Larson *et al.* (2005)
GL252	1	Italy	Near Florence	Europe	H022	AY884724	Larson *et al.* (2005)
GL284	1	France	Corsica	Europe	H022	AY884728	Larson *et al.* (2005)
GL285	1	France	Corsica	Europe	H022	AY884729	Larson *et al.* (2005)
GL287	1	France	Corsica	Europe	H023	AY884731	Larson *et al.* (2005)
GL288	1	Italy	Sardinia	Europe	H022	AY884732	Larson *et al.* (2005)
GL289	1	Italy	Sardinia, Urzulei	Europe	H023	AY884733	Larson *et al.* (2005)
LNSar88	1	Italy	Sardinia, Limbara	Europe	H023	AY884795	Larson *et al.* (2005)
LCorsica82	1	France	Corsica	Europe	H023	AY884796	Larson *et al.* (2005)
LeifFrance	1	France		Europe	H092	AY884815	Larson *et al.* (2005)
leucomystax.1	2	Japan		Asia	H071	D42171	Okumura *et al.* (2001)
leucomystax.2	6	Japan		Asia	H002	D42172	Okumura *et al.* (2001)
leucomystax.3	13	Japan		Asia	H017	D42173	Okumura *et al.* (2001)
leucomystax.4	33	Japan		Asia	H072	D42174	Okumura *et al.* (2001)
leucomystax.5	35	Japan		Asia	H002	D42175	Okumura *et al.* (2001)
leucomystax.6	2	Japan		Asia	H017	D42176	Okumura *et al.* (2001)
leucomystax.7	1	Japan		Asia	H002	D42177	Okumura *et al.* (2001)
leucomystax.8	1	Japan		Asia	H073	D42178	Okumura *et al.* (2001)
domesticus.1	1	Germany		Europe	H004	D42180	Okumura *et al.* (2001)
domesticus.2	3	Japan		Asia	H015	D42182	Okumura *et al.* (2001)
domesticus.3	2	Japan		Asia	H018	D42183	Okumura *et al.* (2001)
riukiuanus	1	Japan		Asia	H005	D42184	Okumura *et al.* (2001)
78	1	Belgium		Europe	H031	DQ379234	Fang *et al.* (2006)
81	1	Belgium		Europe	H054	DQ379235	Fang *et al.* (2006)
83	1	Belgium		Europe	H054	DQ379236	Fang *et al.* (2006)
89	1	Belgium		Europe	H054	DQ379237	Fang *et al.* (2006)
96	1	Belgium		Europe	H104	DQ379238	Fang *et al.* (2006)
105	1	Belgium		Europe	H104	DQ379239	Fang *et al.* (2006)
113	1	Belgium		Europe	H009	DQ379240	Fang *et al.* (2006)
114	1	Belgium		Europe	H021	DQ379241	Fang *et al.* (2006)
S10	1	France		Europe	H092	DQ379261	Fang *et al.* (2006)
DB2	1	China		Asia	H112	DQ379262	Fang *et al.* (2006)
DB4	1	China		Asia	H112	DQ379263	Fang *et al.* (2006)
DB3	1	China		Asia	H067	DQ379264	Fang *et al.* (2006)
DB5	1	China		Asia	H018	DQ379265	Fang *et al.* (2006)
HN8	1	China		Asia	H113	DQ379266	Fang *et al.* (2006)
HN9	1	China		Asia	H114	DQ379267	Fang *et al.* (2006)

Specimen/Isolate	No. ind.	Country	Location	Continent	Haplotype	GenBank	Source
haplotype_X	32	Tunisia		Africa	H143	KC608846	Hajji & Zachos (2010)*
haplotype_Y	32	Tunisia		Africa	H144	KC608847	Hajji & Zachos (2010)*
haplotype_Z	3	Tunisia		Africa	H029	EU362490	Hajji & Zachos (2010)*
D250	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D251	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D253	1	Morocco	Atlas, Rabat	Africa	H133	KC608841	present study
D254	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D255	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D258	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D292	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D293	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study
D294	1	Morocco	Atlas, Rabat	Africa	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
S4	1	France	Chizè	10	H023	DQ379232	Fang *et al.* (2006)
Sm10	1	France	Chizè	10	H023	DQ379233	Fang *et al.* (2006)
mitS1	1	France	Chizè	10	H085	DQ379242	Fang *et al.* (2006)
mitS2	1	France	Chizè	10	H085	DQ379243	Fang *et al.* (2006)
mitS3	1	France	Chizè	10	H111	DQ379244	Fang *et al.* (2006)
Sm1	1	France	Chizè	10	H085	DQ379245	Fang *et al.* (2006)
Sm2	1	France	Chizè	10	H085	DQ379246	Fang *et al.* (2006)
Sm3	1	France	Chizè	10	H085	DQ379247	Fang *et al.* (2006)
Sm4	1	France	Chizè	10	H085	DQ379248	Fang *et al.* (2006)
Sm5	1	France	Chizè	10	H085	DQ379249	Fang *et al.* (2006)
Sm6	1	France	Chizè	10	H085	DQ379250	Fang *et al.* (2006)
Sm7	1	France	Chizè	10	H085	DQ379251	Fang *et al.* (2006)
Sm8	1	France	Chizè	10	H085	DQ379252	Fang *et al.* (2006)
Sm9	1	France	Chizè	10	H022	DQ379253	Fang *et al.* (2006)
S1	1	France	Chizè	10	H092	DQ379254	Fang *et al.* (2006)
S2	1	France	Chizè	10	H092	DQ379255	Fang *et al.* (2006)
S3	1	France	Chizè	10	H092	DQ379256	Fang *et al.* (2006)
S6	1	France	Chizè	10	H092	DQ379257	Fang *et al.* (2006)
S7	1	France	Chizè	10	H092	DQ379258	Fang *et al.* (2006)
S8	1	France	Chizè	10	H092	DQ379259	Fang *et al.* (2006)
S9	1	France	Chizè	10	H092	DQ379260	Fang *et al.* (2006)
S10	1	France	Chizè	10	H092	DQ379261	Fang *et al.* (2006)
AR16	1	Italy	Arezzo province	9	H075	EU362409	Scandura *et al.* (2008)
AR17	1	Italy	Arezzo province	9	H023	EU362410	Scandura *et al.* (2008)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
AR18	1	Italy	Arezzo province	9	H075	EU362411	Scandura *et al.* (2008)
AR20	1	Italy	Arezzo province	9	H022	EU362412	Scandura *et al.* (2008)
AR21	1	Italy	Arezzo province	9	H011	EU362413	Scandura *et al.* (2008)
AR22	1	Italy	Arezzo province	9	H022	EU362414	Scandura *et al.* (2008)
AR23	1	Italy	Arezzo province	9	H022	EU362415	Scandura *et al.* (2008)
AR24c	1	Italy	Arezzo province	9	H011	EU362416	Scandura *et al.* (2008)
FO1R	1	Italy	Forlì province	9	H075	EU362417	Scandura *et al.* (2008)
FO2	1	Italy	Forlì province	9	H023	EU362418	Scandura *et al.* (2008)
FO3R	1	Italy	Forlì province	9	H023	EU362419	Scandura *et al.* (2008)
FO4R	1	Italy	Forlì province	9	H023	EU362420	Scandura *et al.* (2008)
FO5	1	Italy	Forlì province	9	H022	EU362421	Scandura *et al.* (2008)
FO6	1	Italy	Forlì province	9	H075	EU362422	Scandura *et al.* (2008)
FO7	1	Italy	Forlì province	9	H075	EU362423	Scandura *et al.* (2008)
FO8	1	Italy	Forlì province	9	H022	EU362424	Scandura *et al.* (2008)
FO9	1	Italy	Forlì province	9	H075	EU362425	Scandura *et al.* (2008)
FO10	1	Italy	Forlì province	9	H023	EU362426	Scandura *et al.* (2008)
SI1	1	Italy	Siena province	8	H023	EU362427	Scandura *et al.* (2008)
SI2	1	Italy	Siena province	8	H023	EU362428	Scandura *et al.* (2008)
SI3	1	Italy	Siena province	8	H023	EU362429	Scandura *et al.* (2008)
SI4	1	Italy	Siena province	8	H023	EU362430	Scandura *et al.* (2008)
SI5	1	Italy	Siena province	8	H023	EU362431	Scandura *et al.* (2008)
SI6c	1	Italy	Siena province	8	H023	EU362432	Scandura *et al.* (2008)
SI7c	1	Italy	Siena province	8	H023	EU362433	Scandura *et al.* (2008)
SI9	1	Italy	Siena province	8	H023	EU362434	Scandura *et al.* (2008)
SA1	1	Italy	Salerno province	26	H076	EU362435	Scandura *et al.* (2008)
SA2R	1	Italy	Salerno province	26	H023	EU362436	Scandura *et al.* (2008)
SA3R	1	Italy	Salerno province	26	H077	EU362437	Scandura *et al.* (2008)
SA4	1	Italy	Salerno province	26	H075	EU362438	Scandura *et al.* (2008)
SA5R	1	Italy	Salerno province	26	H023	EU362439	Scandura *et al.* (2008)
SA6R	1	Italy	Salerno province	26	H015	EU362440	Scandura *et al.* (2008)
SA7R	1	Italy	Salerno province	26	H015	EU362441	Scandura *et al.* (2008)
SA8R	1	Italy	Salerno province	26	H075	EU362442	Scandura *et al.* (2008)
SA9	1	Italy	Salerno province	26	H012	EU362443	Scandura *et al.* (2008)
SA10R	1	Italy	Salerno province	26	H015	EU362444	Scandura *et al.* (2008)
FI27	1	Italy	Firenze province	8	H023	EU362445	Scandura *et al.* (2008)
FI29	1	Italy	Firenze province	8	H023	EU362446	Scandura *et al.* (2008)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
FI30	1	Italy	Firenze province	8	H023	EU362447	Scandura *et al.* (2008)
FI31	1	Italy	Firenze province	8	H022	EU362448	Scandura *et al.* (2008)
FI33c	1	Italy	Firenze province	8	H022	EU362449	Scandura *et al.* (2008)
FI34c	1	Italy	Firenze province	8	H022	EU362450	Scandura *et al.* (2008)
FI36	1	Italy	Firenze province	8	H075	EU362451	Scandura *et al.* (2008)
FI37	1	Italy	Firenze province	8	H022	EU362452	Scandura *et al.* (2008)
FI40	1	Italy	Firenze province	8	H012	EU362453	Scandura *et al.* (2008)
FI47c	1	Italy	Firenze province	8	H075	EU362454	Scandura *et al.* (2008)
MR1	1	Italy	Maremma National Park	17	H011	EU362455	Scandura *et al.* (2008)
MR2	1	Italy	Maremma National Park	17	H011	EU362456	Scandura *et al.* (2008)
MR3	1	Italy	Maremma National Park	17	H011	EU362457	Scandura *et al.* (2008)
MR4	1	Italy	Maremma National Park	17	H011	EU362458	Scandura *et al.* (2008)
MR5	1	Italy	Maremma National Park	17	H011	EU362459	Scandura *et al.* (2008)
MR6	1	Italy	Maremma National Park	17	H011	EU362460	Scandura *et al.* (2008)
MR8	1	Italy	Maremma National Park	17	H075	EU362461	Scandura *et al.* (2008)
MR9	1	Italy	Maremma National Park	17	H011	EU362462	Scandura *et al.* (2008)
MR11	1	Italy	Maremma National Park	17	H075	EU362463	Scandura *et al.* (2008)
MR12	1	Italy	Maremma National Park	17	H075	EU362464	Scandura *et al.* (2008)
MR14R	1	Italy	Maremma National Park	17	H011	EU362465	Scandura *et al.* (2008)
SR2	1	Italy	San Rossore	28	H022	EU362466	Scandura *et al.* (2008)
SR6c	1	Italy	San Rossore	28	H022	EU362467	Scandura *et al.* (2008)
SR10	1	Italy	San Rossore	28	H022	EU362468	Scandura *et al.* (2008)
SR11	1	Italy	San Rossore	28	H011	EU362469	Scandura *et al.* (2008)
SR12	1	Italy	San Rossore	28	H022	EU362470	Scandura *et al.* (2008)
SR15c	1	Italy	San Rossore	28	H022	EU362471	Scandura *et al.* (2008)
SR16	1	Italy	San Rossore	28	H022	EU362472	Scandura *et al.* (2008)
SR17c	1	Italy	San Rossore	28	H022	EU362473	Scandura *et al.* (2008)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
SR18c	1	Italy	San Rossore	28	H022	EU362474	Scandura *et al.* (2008)
SR20c	1	Italy	San Rossore	28	H022	EU362475	Scandura *et al.* (2008)
CP1	1	Italy	Castelporziano	5	H011	EU362476	Scandura *et al.* (2008)
CP6R	1	Italy	Castelporziano	5	H078	EU362477	Scandura *et al.* (2008)
CP8	1	Italy	Castelporziano	5	H012	EU362478	Scandura *et al.* (2008)
CP10R	1	Italy	Castelporziano	5	H012	EU362479	Scandura *et al.* (2008)
CP13	1	Italy	Castelporziano	5	H012	EU362480	Scandura *et al.* (2008)
CP14R	1	Italy	Castelporziano	5	H012	EU362481	Scandura *et al.* (2008)
CP15R	1	Italy	Castelporziano	5	H012	EU362482	Scandura *et al.* (2008)
CP16R	1	Italy	Castelporziano	5	H012	EU362483	Scandura *et al.* (2008)
CP20R	1	Italy	Castelporziano	5	H078	EU362484	Scandura *et al.* (2008)
CP26R	1	Italy	Castelporziano	5	H078	EU362485	Scandura *et al.* (2008)
SS1	1	Italy	Sardinia	27	H061	EU362486	Scandura *et al.* (2008)
SS4	1	Italy	Sardinia	27	H022	EU362487	Scandura *et al.* (2008)
SS16	1	Italy	Sardinia	27	H022	EU362488	Scandura *et al.* (2008)
SS33	1	Italy	Sardinia	27	H022	EU362489	Scandura *et al.* (2008)
SS36	1	Italy	Sardinia	27	H029	EU362490	Scandura *et al.* (2008)
SS37c	1	Italy	Sardinia	27	H029	EU362491	Scandura *et al.* (2008)
SS38	1	Italy	Sardinia	27	H029	EU362492	Scandura *et al.* (2008)
SS42	1	Italy	Sardinia	27	H079	EU362493	Scandura *et al.* (2008)
SS43	1	Italy	Sardinia	27	H080	EU362494	Scandura *et al.* (2008)
SS45	1	Italy	Sardinia	27	H081	EU362495	Scandura *et al.* (2008)
SS48c	1	Italy	Sardinia	27	H022	EU362496	Scandura *et al.* (2008)
SS50	1	Italy	Sardinia	27	H082	EU362497	Scandura *et al.* (2008)
GO1	1	Italy	Gorizia province	6	H083	EU362498	Scandura *et al.* (2008)
GO4	1	Italy	Gorizia province	6	H083	EU362499	Scandura *et al.* (2008)
GO5	1	Italy	Gorizia province	6	H083	EU362500	Scandura *et al.* (2008)
GO6	1	Italy	Gorizia province	6	H083	EU362501	Scandura *et al.* (2008)
GO11R	1	Italy	Gorizia province	6	H083	EU362502	Scandura *et al.* (2008)
GO13R	1	Italy	Gorizia province	6	H083	EU362503	Scandura *et al.* (2008)
GO15R	1	Italy	Gorizia province	6	H083	EU362504	Scandura *et al.* (2008)
GO18R	1	Italy	Gorizia province	6	H083	EU362505	Scandura *et al.* (2008)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
GO19	1	Italy	Gorizia province	6	H083	EU362506	Scandura *et al.* (2008)
SP2	1	Spain	Coto Doñana	3	H021	EU362507	Scandura *et al.* (2008)
				3			
SP3	1	Spain	Coto Doñana	3	H021	EU362508	Scandura *et al.* (2008)
				3			
SP6	1	Spain	Coto Doñana	3	H021	EU362509	Scandura *et al.* (2008)
				3			
SP8	1	Spain	Coto Doñana	3	H021	EU362510	Scandura *et al.* (2008)
				3			
SP9	1	Spain	Coto Doñana	3	H021	EU362511	Scandura *et al.* (2008)
				3			
SP10	1	Spain	Coto Doñana	3	H021	EU362512	Scandura *et al.* (2008)
				3			
SP12	1	Spain	Coto Doñana	3	H021	EU362513	Scandura *et al.* (2008)
				3			
SP13	1	Spain	Coto Doñana	3	H021	EU362514	Scandura *et al.* (2008)
				3			
SP15	1	Spain	Coto Doñana	3	H021	EU362515	Scandura *et al.* (2008)
				3			
FR2R	1	France	Haute-Marne	1	H023	EU362516	Scandura *et al.* (2008)
				1			
FR3	1	France	Haute-Marne	1	H023	EU362517	Scandura *et al.* (2008)
				1			
FR6R	1	France	Haute-Marne	1	H023	EU362518	Scandura *et al.* (2008)
				1			
FR8	1	France	Haute-Marne	1	H023	EU362519	Scandura *et al.* (2008)
				1			
FR10R	1	France	Haute-Marne	1	H023	EU362520	Scandura *et al.* (2008)
				1			
FR11R	1	France	Haute-Marne	1	H022	EU362521	Scandura *et al.* (2008)
				1			
FR13R	1	France	Haute-Marne	1	H023	EU362522	Scandura *et al.* (2008)
				1			
FR15	1	France	Haute-Marne	1	H023	EU362523	Scandura *et al.* (2008)
				1			
FR16	1	France	Haute-Marne	1	H023	EU362524	Scandura *et al.* (2008)
				1			
FR20R	1	France	Haute-Marne	1	H023	EU362525	Scandura *et al.* (2008)
				1			
AS1	1	Austria	Lower Austria	1	H084	EU362526	Scandura *et al.* (2008)
AS3	1	Austria	Lower Austria	1	H029	EU362527	Scandura *et al.* (2008)
AS4	1	Austria	Lower Austria	1	H035	EU362528	Scandura *et al.* (2008)
AS5	1	Austria	Lower Austria	1	H084	EU362529	Scandura *et al.* (2008)
AS6	1	Austria	Lower Austria	1	H084	EU362530	Scandura *et al.* (2008)
AS10	1	Austria	Lower Austria	1	H084	EU362531	Scandura *et al.* (2008)
AS11	1	Austria	Lower Austria	1	H035	EU362532	Scandura *et al.* (2008)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
AS12	1	Austria	Lower Austria	1	H035	EU362533	Scandura *et al.* (2008)
AS13	1	Austria	Lower Austria	1	H023	EU362534	Scandura *et al.* (2008)
AS14	1	Austria	Lower Austria	1	H084	EU362535	Scandura *et al.* (2008)
PO8	1	Poland	Giżycko	1	H029	EU362536	Scandura *et al.* (2008)
				8			
PO11	1	Poland	Giżycko	1	H085	EU362537	Scandura *et al.* (2008)
				8			
PO12	1	Poland	Giżycko	1	H029	EU362538	Scandura *et al.* (2008)
				8			
PO14	1	Poland	Giżycko	1	H085	EU362539	Scandura *et al.* (2008)
				8			
PO16	1	Poland	Giżycko	1	H085	EU362540	Scandura *et al.* (2008)
				8			
PO24	1	Poland	Giżycko	1	H085	EU362541	Scandura *et al.* (2008)
				8			
PO25	1	Poland	Giżycko	1	H029	EU362542	Scandura *et al.* (2008)
				8			
PO27	1	Poland	Giżycko	1	H029	EU362543	Scandura *et al.* (2008)
				8			
UN1R	1	Hungary		1	H029	EU362544	Scandura *et al.* (2008)
				6			
UN2	1	Hungary		1	H029	EU362545	Scandura *et al.* (2008)
				6			
UN6	1	Hungary		1	H029	EU362546	Scandura *et al.* (2008)
				6			
UN9R	1	Hungary		1	H029	EU362547	Scandura *et al.* (2008)
				6			
UN10R	1	Hungary		1	H029	EU362548	Scandura *et al.* (2008)
				6			
UN15R	1	Hungary		1	H029	EU362549	Scandura *et al.* (2008)
				6			
UN17R	1	Hungary		1	H029	EU362550	Scandura *et al.* (2008)
				6			
UN18	1	Hungary		1	H029	EU362551	Scandura *et al.* (2008)
				6			
UN24	1	Hungary		1	H035	EU362552	Scandura *et al.* (2008)
				6			
UN28	1	Hungary		1	H035	EU362553	Scandura *et al.* (2008)
				6			
Jav01Ses	1	Portugal	Serra da Estrela	2	H056	HM747208	Alves *et al.* (2010)
				0			
Jav14Ses	1	Portugal	Serra da Estrela	2	H056	HM747208	Alves *et al.* (2010)
				0			
Jav15Ses	1	Portugal	Serra da Estrela	2	H021	HM747209	Alves *et al.* (2010)
				0			

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jav21Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav27Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav28Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav29Ses	1	Portugal	Serra da Estrela	20	H029	HM747202	Alves *et al.* (2010)
Jav30Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav31Ses	1	Portugal	Serra da Estrela	20	H029	HM747202	Alves *et al.* (2010)
Jav33Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav34Ses	1	Portugal	Serra da Estrela	20	H021	HM747209	Alves *et al.* (2010)
Jav35Ses	1	Portugal	Serra da Estrela	20	H056	HM747208	Alves *et al.* (2010)
Jav36Cer	1	Portugal	Cercal	21	H029	HM747205	Alves *et al.* (2010)
Jav37Cer	1	Portugal	Cercal	21	H029	HM747205	Alves *et al.* (2010)
Jav38Cer	1	Portugal	Cercal	21	H029	HM747205	Alves *et al.* (2010)
Jav39Cer	1	Portugal	Cercal	21	H029	HM747205	Alves *et al.* (2010)
Jav40Cer	1	Portugal	Cercal	21	H029	HM747205	Alves *et al.* (2010)
Jav41Rmo	1	Portugal	Reguengos de Monsaraz	21	H029	HM747205	Alves *et al.* (2010)
Jav42Rmo	1	Portugal	Reguengos de Monsaraz	21	H029	HM747205	Alves *et al.* (2010)
Jav43Rmo	1	Portugal	Reguengos de Monsaraz	21	H029	HM747205	Alves *et al.* (2010)
Jav44Rmo	1	Portugal	Reguengos de Monsaraz	21	H029	HM747205	Alves *et al.* (2010)
Jav45Rmo	1	Portugal	Reguengos de Monsaraz	21	H029	HM747205	Alves *et al.* (2010)
Jav83Bra	1	Portugal	Bragança	19	H021	HM747201	Alves *et al.* (2010)
Jav84Bra	1	Portugal	Bragança	19	H029	HM747202	Alves *et al.* (2010)
Jav85Bra	1	Portugal	Bragança	19	H021	HM747203	Alves *et al.* (2010)
Jav86Bra	1	Portugal	Bragança	19	H120	HM747204	Alves *et al.* (2010)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jav87Bra	1	Portugal	Bragança	19	H120	HM747204	Alves *et al.* (2010)
Jav89Bra	1	Portugal	Bragança	19	H120	HM747204	Alves *et al.* (2010)
Jav47Vgu	1	Portugal	Serpa, Vale do Guadiana	21	H029	HM747205	Alves *et al.* (2010)
Jav50Vgu	1	Portugal	Serpa, Vale do Guadiana	21	H029	HM747205	Alves *et al.* (2010)
Jav54Vgu	1	Portugal	Mértola, Vale do Guadiana	21	H029	HM747205	Alves *et al.* (2010)
Jav56Vgu	1	Portugal	Mértola, Vale do Guadiana	21	H029	HM747205	Alves *et al.* (2010)
Jav59Alg	1	Portugal	Aljezur, Algarve	21	H023	HM747198	Alves *et al.* (2010)
Jav61Alg	1	Portugal	Aljezur, Algarve	21	H023	HM747198	Alves *et al.* (2010)
Jav65Alg	1	Portugal	Vila do Bispo, Algarve	21	H023	HM747198	Alves *et al.* (2010)
Jav67Alg	1	Portugal	Vila do Bispo, Algarve	21	H023	HM747198	Alves *et al.* (2010)
Jve01Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve02Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve46Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve47Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve48Jae	1	Spain	Jaén	35	H029	HM747202	Alves *et al.* (2010)
Jve49Jae	1	Spain	Jaén	35	H021	HM747215	Alves *et al.* (2010)
Jve50Jae	1	Spain	Jaén	35	H021	HM747215	Alves *et al.* (2010)
Jve66Jae	1	Spain	Jaén	35	H021	HM747215	Alves *et al.* (2010)
Jve67Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve68Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve69Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve70Jae	1	Spain	Jaén	35	H054	HM747211	Alves *et al.* (2010)
Jve71Jae	1	Spain	Jaén	35	H029	HM747202	Alves *et al.* (2010)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jve72Jae	1	Spain	Jaén	35	H029	HM747202	Alves *et al.* (2010)
Jve73Jae	1	Spain	Jaén	35	H029	HM747202	Alves *et al.* (2010)
Jve74Jae	1	Spain	Jaén	35	H029	HM747202	Alves *et al.* (2010)
Jve11Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve12Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve13Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve14Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve15Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve34Cre	1	Spain	Ciudad Real	34	H054	HM747211	Alves *et al.* (2010)
Jve35Cre	1	Spain	Ciudad Real	34	H138	HM747210	Alves *et al.* (2010)
Jve36Cre	1	Spain	Ciudad Real	34	H138	HM747210	Alves *et al.* (2010)
Jve37Cre	1	Spain	Ciudad Real	34	H138	HM747210	Alves *et al.* (2010)
Jve38Cre	1	Spain	Ciudad Real	34	H120	HM747212	Alves *et al.* (2010)
Jve39Cre	1	Spain	Ciudad Real	34	H138	HM747210	Alves *et al.* (2010)
Jve40Cre	1	Spain	Ciudad Real	34	H120	HM747212	Alves *et al.* (2010)
Jve82Cre	1	Spain	Ciudad Real	34	H021	HM747203	Alves *et al.* (2010)
Jve83Cre	1	Spain	Ciudad Real	34	H021	HM747203	Alves *et al.* (2010)
Jve84Cre	1	Spain	Ciudad Real	34	H021	HM747203	Alves *et al.* (2010)
Jve85Cre	1	Spain	Ciudad Real	34	H021	HM747203	Alves *et al.* (2010)
Jve16Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve17Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve18Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve19Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jve20Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve21Tol	1	Spain	Toledo	37	H054	HM747211	Alves *et al.* (2010)
Jve22Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve23Tol	1	Spain	Toledo	37	H054	HM747211	Alves *et al.* (2010)
Jve24Tol	1	Spain	Toledo	37	H138	HM747210	Alves *et al.* (2010)
Jve25Tol	1	Spain	Toledo	37	H029	HM747202	Alves *et al.* (2010)
Jve26Tol	1	Spain	Toledo	37	H120	HM747212	Alves *et al.* (2010)
Jve76Sev	1	Spain	Sevilla	33	H139	HM747213	Alves *et al.* (2010)
Jve77Sev	1	Spain	Sevilla	33	H139	HM747213	Alves *et al.* (2010)
Jve78Sev	1	Spain	Sevilla	33	H139	HM747213	Alves *et al.* (2010)
Jve79Sev	1	Spain	Sevilla	33	H139	HM747213	Alves *et al.* (2010)
Jve80Sev	1	Spain	Sevilla	33	H139	HM747213	Alves *et al.* (2010)
Jve86Ast	1	Spain	Asturias	32	H120	HM747204	Alves *et al.* (2010)
Jve87Ast	1	Spain	Asturias	32	H120	HM747204	Alves *et al.* (2010)
Jve88Ast	1	Spain	Asturias	32	H021	HM747209	Alves *et al.* (2010)
Jve89Ast	1	Spain	Asturias	32	H021	HM747209	Alves *et al.* (2010)
Jve90Ast	1	Spain	Asturias	32	H021	HM747209	Alves *et al.* (2010)
Jve91Ast	1	Spain	Asturias	32	H029	HM747206	Alves *et al.* (2010)
Jve92Ast	1	Spain	Asturias	32	H140	HM747214	Alves *et al.* (2010)
Jve94Ast	1	Spain	Asturias	32	H029	HM747206	Alves *et al.* (2010)
Jve95Ast	1	Spain	Asturias	32	H029	HM747206	Alves *et al.* (2010)
Jve102Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve104Pyr	1	Spain	Pyrenees	32	H085	HM747199	Alves *et al.* (2010)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jve105Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve106Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve110Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve111Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve112Pyr	1	Spain	Pyrenees	32	H021	HM747201	Alves *et al.* (2010)
Jve115Veb	1	Spain	Valle del Ebro	38	H021	HM747201	Alves *et al.* (2010)
Jve116Veb	1	Spain	Valle del Ebro	38	H021	HM747201	Alves *et al.* (2010)
Jve117Veb	1	Spain	Valle del Ebro	38	H021	HM747201	Alves *et al.* (2010)
Jve119Veb	1	Spain	Valle del Ebro	38	H021	HM747201	Alves *et al.* (2010)
Jve120Veb	1	Spain	Valle del Ebro	38	H029	HM747206	Alves *et al.* (2010)
Jve124Veb	1	Spain	Valle del Ebro	38	H029	HM747206	Alves *et al.* (2010)
Jve125Veb	1	Spain	Valle del Ebro	38	H029	HM747206	Alves *et al.* (2010)
Jve127Veb	1	Spain	Valle del Ebro	38	H021	HM747201	Alves *et al.* (2010)
Jve143Gua	1	Spain	Guadalajara	36	H120	HM747204	Alves *et al.* (2010)
Jve145Gua	1	Spain	Guadalajara	36	H120	HM747204	Alves *et al.* (2010)
Jve144Gua	1	Spain	Guadalajara	36	H120	HM747204	Alves *et al.* (2010)
Jve146Gua	1	Spain	Guadalajara	36	H120	HM747204	Alves *et al.* (2010)
Jve147Gua	1	Spain	Guadalajara	36	H120	HM747204	Alves *et al.* (2010)
Jve128Mun	1	Spain	Montes Universales	36	H021	HM747201	Alves *et al.* (2010)
Jve129Mun	1	Spain	Montes Universales	36	H021	HM747201	Alves *et al.* (2010)
Jve131Mun	1	Spain	Montes Universales	36	H021	HM747201	Alves *et al.* (2010)
Jve132Mun	1	Spain	Montes Universales	36	H021	HM747201	Alves *et al.* (2010)
Jve133Mun	1	Spain	Montes Universales	36	H021	HM747201	Alves *et al.* (2010)

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
Jve134Bur	1	Spain	Burgos	32	H120	HM747204	Alves *et al.* (2010)
Jve135Bur	1	Spain	Burgos	32	H120	HM747204	Alves *et al.* (2010)
Jve136Bur	1	Spain	Burgos	32	H021	HM747201	Alves *et al.* (2010)
Jve138Bur	1	Spain	Burgos	32	H120	HM747204	Alves *et al.* (2010)
Jve140Bur	1	Spain	Burgos	32	H120	HM747204	Alves *et al.* (2010)
Jve141Bur	1	Spain	Burgos	32	H021	HM747201	Alves *et al.* (2010)
Jve148Alb	1	Spain	Albacete	36	H029	HM747202	Alves *et al.* (2010)
Jve151Alb	1	Spain	Albacete	36	H021	HM747207	Alves *et al.* (2010)
Jve155Alb	1	Spain	Albacete	36	H120	HM747204	Alves *et al.* (2010)
Jve157Alb	1	Spain	Albacete	36	H021	HM747207	Alves *et al.* (2010)
Jve159Alb	1	Spain	Albacete	36	H029	HM747202	Alves *et al.* (2010)
JvFr26Mad	1	France	Madine	10	H023	HM747198	Alves *et al.* (2010)
JvFr27Mad	1	France	Madine	10	H023	HM747198	Alves *et al.* (2010)
JvFr28Pre	1	France	Prémian	10	H085	HM747199	Alves *et al.* (2010)
JvFr32Pre	1	France	Prémian	10	H085	HM747199	Alves *et al.* (2010)
JvFr28bPre	1	France	Prémian	10	H085	HM747200	Alves *et al.* (2010)
JvFr32bPre	1	France	Prémian	10	H085	HM747200	Alves *et al.* (2010)
JvAust01	1	Austria	Vienna	1	H035	HM747196	Alves *et al.* (2010)
JvAust07	1	Austria	Vienna	1	H035	HM747196	Alves *et al.* (2010)
JvAust07b	1	Austria	Vienna	1	H141	HM747197	Alves *et al.* (2010)
D82	1	Spain	Lugar Nuevo	38	H054	HM747211	present study
D83	1	Spain	Lugar Nuevo	38	H054	HM747211	present study
D84	1	Spain	Lugar Nuevo	38	H054	HM747211	present study
F1	1	France	Arc-en-Barrois	12	H023	EU362410	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
F2	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F6	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F7	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F8	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F9	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F10	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F11	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F12	1	France	Arc-en-Barrois	12	H023	EU362410	present study
F13	1	France	Arc-en-Barrois	12	H023	EU362410	present study
POL46	1	Poland	Dębniki	18	H085	DQ379242	present study
POL47	1	Poland	Płaska	18	H029	EU362490	present study
POL49	1	Poland	Jałowo	18	H134	KC608842	present study
POL50	1	Poland	Łomża	18	H029	EU362490	present study
POL51	1	Poland	Leman	18	H029	EU362490	present study
POL52	1	Poland	Nowogród	18	H029	EU362490	present study
POL53	1	Poland	Nowogród	18	H029	EU362490	present study
POL54	1	Poland	Szczebra	18	H029	EU362490	present study
POL55	1	Poland	Głęboki Bród	18	H029	EU362490	present study
SS21	1	Italy	Sardinia	27	H029	EU362490	present study
SS25	1	Italy	Sardinia	27	H022	EU362412	present study
SS30	1	Italy	Sardinia	27	H022	EU362412	present study
SS53	1	Italy	Sardinia	27	H022	EU362412	present study
SS67	1	Italy	Sardinia	27	H061	EU362486	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
SS71	1	Italy	Sardinia	27	H029	EU362490	present study
SS74	1	Italy	Sardinia	27	H022	EU362412	present study
SS75	1	Italy	Sardinia	27	H029	EU362490	present study
SS80	1	Italy	Sardinia	27	H061	EU362486	present study
SS81	1	Italy	Sardinia	27	H023	EU362410	present study
SS84	1	Italy	Sardinia	27	H023	EU362410	present study
SS89	1	Italy	Sardinia	27	H115	KC608827	present study
SS95	1	Italy	Sardinia	27	H022	EU362412	present study
SS100	1	Italy	Sardinia	27	H061	EU362486	present study
SS102	1	Italy	Sardinia	27	H029	EU362490	present study
SS108	1	Italy	Sardinia	27	H029	EU362490	present study
SS18	1	Italy	Sardinia	27	H022	EU362412	present study
SS35	1	Italy	Sardinia	27	H116	KC608828	present study
SS66	1	Italy	Sardinia	27	H029	EU362490	present study
SS78	1	Italy	Sardinia	27	H023	EU362410	present study
SS91	1	Italy	Sardinia	27	H118	KC608829	present study
SS98	1	Italy	Sardinia	27	H022	EU362412	present study
SS112	1	Italy	Sardinia	27	H119	KC608830	present study
SS138	1	Italy	Sardinia	27	H061	EU362486	present study
SS139	1	Italy	Sardinia	27	H061	EU362486	present study
SS146	1	Italy	Sardinia	27	H022	EU362412	present study
SS155	1	Italy	Sardinia	27	H023	EU362410	present study
SS161	1	Italy	Sardinia	27	H023	EU362410	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
SS166	1	Italy	Sardinia	27	H023	EU362410	present study
SS168	1	Italy	Sardinia	27	H091	AY884698	present study
SS174	1	Italy	Sardinia	27	H022	EU362412	present study
SS180	1	Italy	Sardinia	27	H022	EU362412	present study
SS181	1	Italy	Sardinia	27	H022	EU362412	present study
SS183	1	Italy	Sardinia	27	H022	EU362412	present study
SS190	1	Italy	Sardinia	27	H061	EU362486	present study
SS191	1	Italy	Sardinia	27	H061	EU362486	present study
SS197	1	Italy	Sardinia	27	H061	EU362486	present study
SS23	1	Italy	Sardinia	27	H029	EU362490	present study
SS32	1	Italy	Sardinia	27	H029	EU362490	present study
SS44	1	Italy	Sardinia	27	H029	EU362490	present study
SS46	1	Italy	Sardinia	27	H029	EU362490	present study
SS62	1	Italy	Sardinia	27	H061	EU362486	present study
SS76	1	Italy	Sardinia	27	H023	EU362410	present study
SS79	1	Italy	Sardinia	27	H091	AY884698	present study
SS86	1	Italy	Sardinia	27	H029	EU362490	present study
SS111	1	Italy	Sardinia	27	H023	EU362410	present study
SS125	1	Italy	Sardinia	27	H023	EU362410	present study
SS128	1	Italy	Sardinia	27	H022	EU362412	present study
SS225	1	Italy	Sardinia	27	H061	EU362486	present study
SS234	1	Italy	Sardinia	27	H029	EU362490	present study
SS237	1	Italy	Sardinia	27	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
SS244	1	Italy	Sardinia	27	H029	EU362490	present study
SS248	1	Italy	Sardinia	27	H022	EU362412	present study
SS251	1	Italy	Sardinia	27	H029	EU362490	present study
SS259	1	Italy	Sardinia	27	H022	EU362412	present study
SS261	1	Italy	Sardinia	27	H022	EU362412	present study
SS268	1	Italy	Sardinia	27	H029	EU362490	present study
SS269	1	Italy	Sardinia	27	H022	EU362412	present study
SS270	1	Italy	Sardinia	27	H022	EU362412	present study
SS276	1	Italy	Sardinia	27	H029	EU362490	present study
SS279	1	Italy	Sardinia	27	H061	EU362486	present study
SS286	1	Italy	Sardinia	27	H061	EU362486	present study
SS306	1	Italy	Sardinia	27	H115	KC608827	present study
SS313	1	Italy	Sardinia	27	H115	KC608827	present study
SS319	1	Italy	Sardinia	27	H115	KC608827	present study
SS339	1	Italy	Sardinia	27	H022	EU362412	present study
SS340	1	Italy	Sardinia	27	H022	EU362412	present study
SS344	1	Italy	Sardinia	27	H129	KC608837	present study
SS347	1	Italy	Sardinia	27	H061	EU362486	present study
SS349	1	Italy	Sardinia	27	H061	EU362486	present study
SS350	1	Italy	Sardinia	27	H061	EU362486	present study
GR01	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR02	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR03	1	Greece	Olympos, Thessaly	15	H021	EU362507	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
GR04	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR05	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR06	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR07	1	Greece	Olympos, Thessaly	15	H128	KC608836	present study
GR08	1	Greece	Kissavos, Thessaly	15	H128	KC608836	present study
GR09	1	Greece	Kissavos, Thessaly	15	H128	KC608836	present study
GR10	1	Greece	Veria, Macedonia	15	H128	KC608836	present study
GR11	1	Greece	Pelio, Magnesia	15	H128	KC608836	present study
GR13	1	Greece	Serres, Macedonia	15	H021	EU362507	present study
GR14	1	Greece	Serres, Macedonia	15	H021	EU362507	present study
GR15	1	Greece	Serres, Macedonia	15	H021	EU362507	present study
GR16	1	Greece	Serres, Macedonia	15	H021	EU362507	present study
GR17	1	Greece	Serres, Macedonia	15	H021	EU362507	present study
GR18	1	Greece	Serres, Macedonia	15	H022	EU362412	present study
GR19	1	Greece	Almiros, Magnesia	15	H022	EU362412	present study
GR20	1	Greece	Almiros, Magnesia	15	H128	KC608836	present study
GR21	1	Greece	Petritsi, Macedonia	15	H029	EU362490	present study
GR22	1	Greece	Petritsi, Macedonia	15	H022	EU362412	present study
GR23	1	Greece	Petritsi, Macedonia	15	H022	EU362412	present study
GR24	1	Greece	Petritsi, Macedonia	15	H129	KC608837	present study
GR25	1	Greece	Petritsi, Macedonia	15	H022	EU362412	present study
GR26	1	Greece	Petritsi, Macedonia	15	H029	EU362490	present study
GR27	1	Greece	Petritsi, Macedonia	15	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
GR28	1	Greece	Vrynena, Magnesia	15	H128	KC608836	present study
GR29	1	Greece	Vrynena, Magnesia	15	H128	KC608836	present study
GR30	1	Greece	Petritsi, Macedonia	15	H022	EU362412	present study
D241	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D242	1	Serbia	Djerdap, Bor	29	H021	EU362507	present study
D243	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D244	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D245	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D246	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D247	1	Serbia	Djerdap, Bor	29	H029	EU362490	present study
D181	1	Serbia	Crni Lug, Zemun, Belgrade	30	H029	EU362490	present study
D182	1	Serbia	Crni Lug, Zemun, Belgrade	30	H029	EU362490	present study
D183	1	Serbia	Crni Lug, Zemun, Belgrade	30	H029	EU362490	present study
D184	1	Serbia	Crni Lug, Zemun, Belgrade	30	H029	EU362490	present study
D185	1	Serbia	Crni Lug, Zemun, Belgrade	30	H029	EU362490	present study
D85	1	Serbia	Plavna, Bor	29	H021	EU362507	present study
D86	1	Serbia	Plavna, Bor	29	H029	EU362490	present study
D87	1	Serbia	Plavna, Bor	29	H021	EU362507	present study
D88	1	Serbia	Plavna, Bor	29	H035	EU362528	present study
D89	1	Serbia	Plavna, Bor	29	H021	EU362507	present study
D90	1	Serbia	Sremska Mitrovica, Vojvodina	30	H029	EU362490	present study
D91	1	Serbia	Sremska Mitrovica, Vojvodina	30	H022	EU362412	present study
D92	1	Serbia	Sremska Mitrovica, Vojvodina	30	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
D93	1	Serbia	Sremska Mitrovica, Vojvodina	30	H029	EU362490	present study
D94	1	Serbia	Sremska Mitrovica, Vojvodina	30	H029	EU362490	present study
D65	1	Bosnia		30	H029	EU362490	present study
D66	1	Bosnia		30	H125	KC608835	present study
D59	1	Bosnia	Glasinac	30	H021	EU362507	present study
D212	1	Bosnia	Hadžići	30	H021	EU362507	present study
D214	1	Bosnia	Maglaj	30	H021	EU362507	present study
D215	1	Bosnia	Priluk	30	H021	EU362507	present study
D216	1	Bosnia	Živinice	30	H021	EU362507	present study
D217	1	Bosnia	Kakanj	30	H021	EU362507	present study
D218	1	Bosnia	Kladanj	30	H021	EU362507	present study
D262	1	Romania	Arad	22	H035	EU362528	present study
D264	1	Romania	Târgu Secuiesc	23	H132	KC608840	present study
D265	1	Romania	Craiova	29	H121	KC608831	present study
D268	1	Romania	Craiova	29	H035	EU362528	present study
D269	1	Romania	Târgu Secuiesc	23	H132	KC608840	present study
D273	1	Romania	Craiova	29	H124	KC608834	present study
D274	1	Romania	Târgu Secuiesc	23	H035	EU362528	present study
D278	1	Romania	Craiova	29	H021	EU362507	present study
D279	1	Romania	Târgu Secuiesc	23	H029	EU362490	present study
D283	1	Romania	Craiova	29	H029	EU362490	present study
D284	1	Romania	Târgu Secuiesc	23	H035	EU362528	present study
LX1	1	Luxembourg	Rumelange, Esch-sur-Alzette	11	H004	D42180	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
LX2	1	Luxembourg	Eschweiler, Wiltz	1	H022	EU362412	present study
				1			
LX3	1	Luxembourg	Vianden, Vianden	1	H023	EU362410	present study
				1			
LX4	1	Luxembourg	Junglinster, Grevenmacher	1	H023	EU362410	present study
				1			
LX5	1	Luxembourg	Mersch,Mersch	1	H023	EU362410	present study
				1			
LX6	1	Luxembourg	Boevange-sur-Attert, Mersch	1	H023	EU362410	present study
				1			
LX7	1	Luxembourg	Tuntange, Mersch	1	H022	EU362412	present study
				1			
LX8	1	Luxembourg	Stadtbremidus, Remich	1	H023	EU362410	present study
				1			
LX9	1	Luxembourg	Waldbredimus, Remich	1	H023	EU362410	present study
				1			
CR01	1	Croatia	Istria	6	H130	KC608838	present study
CR02	1	Croatia	Istria	6	H083	EU362498	present study
CR03	1	Croatia	Gorski Kotar	6	H023	EU362410	present study
CR04	1	Croatia	Gorski Kotar	6	H023	EU362410	present study
CR05	1	Croatia	Istria	6	H083	EU362498	present study
CR06	1	Croatia	Gorski Kotar	6	H083	EU362498	present study
CR07	1	Croatia	Lokve	6	H083	EU362498	present study
CR08	1	Croatia	Istria	6	H083	EU362498	present study
CR09	1	Croatia	Lika	6	H083	EU362498	present study
CR10	1	Croatia	Gorski Kotar	6	H083	EU362498	present study
CR11	1	Croatia	Gorski Kotar	6	H083	EU362498	present study
CR12	1	Croatia	Gorski Kotar	6	H023	EU362410	present study
CR13	1	Croatia	Gorski Kotar	6	H023	EU362410	present study
CR14	1	Croatia	Lokve	6	H083	EU362498	present study
CR15	1	Croatia	Gorski Kotar	6	H083	EU362498	present study
CR16	1	Croatia	Karlovac	6	H083	EU362498	present study
CR17	1	Croatia	Lika	6	H029	EU362490	present study
BU01	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU02	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU03	1	Bulgaria	Ludogorie	3	H123	KC608833	present study
BU04	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU05	1	Bulgaria	Ludogorie	3	H122	KC608832	present study
BU06	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU07	1	Bulgaria	Ludogorie	3	H029	EU362490	present study
BU08	1	Bulgaria	Ludogorie	3	H035	EU362528	present study
BU09	1	Bulgaria	Ludogorie	3	H035	EU362528	present study
BU10	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU11	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU12	1	Bulgaria	Ludogorie	3	H121	KC608831	present study
BU13	1	Bulgaria	Rila	4	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
BU14	1	Bulgaria	Rila	4	H029	EU362490	present study
BU15	1	Bulgaria	Rila	4	H121	KC608831	present study
BU16	1	Bulgaria	Rila	4	H021	EU362507	present study
BU17	1	Bulgaria	Rila	4	H121	KC608831	present study
BU18	1	Bulgaria	Rila	4	H029	EU362490	present study
BU19	1	Bulgaria	Rila	4	H029	EU362490	present study
BU20	1	Bulgaria	Rila	4	H029	EU362490	present study
BU21	1	Bulgaria	Rila	4	H121	KC608831	present study
BU22	1	Bulgaria	Rila	4	H029	EU362490	present study
BU23	1	Bulgaria	Rila	4	H029	EU362490	present study
BU24	1	Bulgaria	Rila	4	H029	EU362490	present study
TO01	1	Italy	Cuneo province	7	H012	EU362443	present study
TO02	1	Italy	Cuneo province	7	H012	EU362443	present study
TO03	1	Italy	Cuneo province	7	H012	EU362443	present study
TO04	1	Italy	Cuneo province	7	H029	EU362490	present study
TO06	1	Italy	Cuneo province	7	H029	EU362490	present study
TO07	1	Italy	Cuneo province	7	H022	EU362412	present study
TO08	1	Italy	Cuneo province	7	H085	DQ379242	present study
TO09	1	Italy	Torino province	7	H120	HM747204	present study
TO10	1	Italy	Torino province	7	H022	EU362412	present study
TO11	1	Italy	Torino province	7	H023	EU362410	present study
TO12	1	Italy	Torino province	7	H023	EU362410	present study
TO13	1	Italy	Torino province	7	H012	EU362443	present study
TO14	1	Italy	Torino province	7	H012	EU362443	present study
TO15	1	Italy	North-eastern Piedmont	7	H075	EU362409	present study
TO16	1	Italy	North-eastern Piedmont	7	H075	EU362409	present study
TO17	1	Italy	Torino province	7	H023	EU362410	present study
TO18	1	Italy	North-eastern Piedmont	7	H023	EU362410	present study
TO19	1	Italy	North-eastern Piedmont	7	H022	EU362412	present study
TO20	1	Italy	North-eastern Piedmont	7	H022	EU362412	present study
PT01	1	Portugal	Serra da Lousã, Lousã	20	H056	HM747208	present study
PT04	1	Portugal	Serra da Lousã, Lousã	20	H056	HM747208	present study
PT05	1	Portugal	Morais, Macedo de Cavaleiros	19	H120	HM747204	present study
PT06	1	Portugal	Morais, Macedo de Cavaleiros	19	H120	HM747204	present study
PT07	1	Portugal	Morais, Macedo de Cavaleiros	19	H120	HM747204	present study
PT08	1	Portugal	Morais, Macedo de Cavaleiros	19	H120	HM747204	present study
PT09	1	Portugal	Morais, Macedo de Cavaleiros	19	H120	HM747204	present study
PT10	1	Portugal	Junqueira, Vimioso	19	H120	HM747204	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
PT11	1	Portugal	Junqueira, Vimioso	19	H120	HM747204	present study
PT12	1	Portugal	Junqueira, Vimioso	19	H021	EU362507	present study
PT13	1	Portugal	Junqueira, Vimioso	19	H120	HM747204	present study
PT14	1	Portugal	Junqueira, Vimioso	19	H120	HM747204	present study
PT15	1	Portugal	Junqueira, Vimioso	19	H021	EU362507	present study
PT16	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT17	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT18	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT19	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT20	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT21	1	Portugal	Quintela, Bragança	19	H056	HM747208	present study
PT22	1	Portugal	Quintela, Bragança	19	H021	EU362507	present study
PT23	1	Portugal	Alvaredos, Vinhais	19	H021	EU362507	present study
PT24	1	Portugal	Alvaredos, Vinhais	19	H021	EU362507	present study
PT26	1	Portugal	Alvaredos, Vinhais	19	H120	HM747204	present study
PT27	1	Portugal	Alvaredos, Vinhais	19	H021	EU362507	present study
PT28	1	Portugal	Palaçoulo, Miranda do Douro	19	H120	HM747204	present study
PT30	1	Portugal	Palaçoulo, Miranda do Douro	19	H021	EU362507	present study
PT31	1	Portugal	Palaçoulo, Miranda do Douro	19	H120	HM747204	present study
PT32	1	Portugal	Palaçoulo, Miranda do Douro	19	H120	HM747204	present study
PT33	1	Portugal	Palaçoulo, Miranda do Douro	19	H021	EU362507	present study
PT34	1	Portugal	Palaçoulo, Miranda do Douro	19	H021	EU362507	present study
PT36	1	Portugal	Palaçoulo, Miranda do Douro	19	H021	EU362507	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
PT37	1	Portugal	Palaçoulo, Miranda do Douro	1	H021	EU362507	present study
				9			
BEL02	1	Belarus	Studenka	2	H035	EU362528	present study
BEL03	1	Belarus	Skripky	2	H035	EU362528	present study
BEL04	1	Belarus	Beshenkovichy	2	H029	EU362490	present study
BEL05	1	Belarus	Ushachy	2	H029	EU362490	present study
BEL10	1	Belarus	Gytkovivchy	2	H029	EU362490	present study
BEL11	1	Belarus	Jazwinskoje forest	2	H029	EU362490	present study
BEL12	1	Belarus	Paszukowskoje forest	2	H029	EU362490	present study
BEL13	1	Belarus	Belanskoje forest	2	H034	AF136555	present study
BEL18	1	Belarus	Tolochyn	2	H029	EU362490	present study
BEL19	1	Belarus	Luban	2	H029	EU362490	present study
BEL20	1	Belarus	Iv'e	2	H029	EU362490	present study
BEL33	1	Belarus	Vologyn	2	H035	EU362528	present study
BEL34	1	Belarus	Molodechno	2	H035	EU362528	present study
BEL35	1	Belarus	Petrykov	2	H029	EU362490	present study
BEL36	1	Belarus	Vileyka	2	H035	EU362528	present study
BEL37	1	Belarus	Vileyka	2	H029	EU362490	present study
BEL38	1	Belarus	Volkovysk	2	H029	EU362490	present study
BEL39	1	Belarus	Gitkovichy	2	H029	EU362490	present study
BEL40	1	Belarus	Bereza	2	H034	AF136555	present study
RUS06	1	Russia	Novgorod region	2	H029	EU362490	present study
				5			
RUS07	1	Russia	Novgorod region, Chvojnya district	2	H035	EU362528	present study
				5			
RUS08	1	Russia	Novgorod region, Chvojnya district	2	H034	AF136555	present study
				5			
RUS14	1	Russia	Pskov region, Pskov district	2	H029	EU362490	present study
				5			
RUS15	1	Russia	Pskov region, Pskov district	2	H022	EU362412	present study
				5			
RUS17	1	Russia	Pskov region, Pskov district	2	H022	EU362412	present study
				5			
RUS22	1	Russia	Smolensk region, Hyslabychy district	2	H034	AF136555	present study
				5			
RUS23	1	Russia	Smolensk region, Viazma district	2	H035	EU362528	present study
				5			
RUS24	1	Russia	Smolensk region, Viazma district	2	H035	EU362528	present study
				5			
RUS25	1	Russia	Smolensk region, Glinka district	2	H029	EU362490	present study
				5			
RUS26	1	Russia	Smolensk region, Yarcevo district	2	H029	EU362490	present study
				5			
RUS27	1	Russia	Smolensk region, Gagaryn district	2	H029	EU362490	present study
				5			

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
RUS28	1	Russia	Smolensk region, Glinka district	25	H035	EU362528	present study
RUS30	1	Russia	Smolensk region, Viazma district	25	H131	KC608839	present study
RUS31	1	Russia	Smolensk region, Pochynok district	25	H029	EU362490	present study
RUS32	1	Russia	Smolensk region, Pochynok district	25	H034	AF136555	present study
RUS43	1	Russia	Moskov region, Rostov district	25	H035	EU362528	present study
RUS45	1	Russia	Moskov region, Rostov district	25	H035	EU362528	present study
RLip01	1	Romania	Lipova	23	H029	EU362490	present study
RLip02	1	Romania	Lipova	23	H029	EU362490	present study
RLip03	1	Romania	Lipova	23	H029	EU362490	present study
RLip04	1	Romania	Lipova	23	H029	EU362490	present study
RLip05	1	Romania	Lipova	23	H029	EU362490	present study
RLip06	1	Romania	Lipova	23	H029	EU362490	present study
RLip07	1	Romania	Lipova	23	H029	EU362490	present study
RLip08	1	Romania	Lipova	23	H029	EU362490	present study
RLip09	1	Romania	Lipova	23	H029	EU362490	present study
RLip10	1	Romania	Lipova	23	H029	EU362490	present study
RLip11	1	Romania	Lipova	23	H029	EU362490	present study
RPla01	1	Romania	Plateau	23	H029	EU362490	present study
RPla02	1	Romania	Plateau	23	H029	EU362490	present study
RCar01	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar02	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar03	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar04	1	Romania	Caransebeş	22	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
RCar05	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar06	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar07	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar08	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar09	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar10	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar11	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar12	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar13	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar14	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar15	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar16	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar17	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar18	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar19	1	Romania	Caransebeş	22	H029	EU362490	present study
RCar20	1	Romania	Caransebeş	22	H029	EU362490	present study
RTim01	1	Romania	Timişoara	24	H029	EU362490	present study
RTim02	1	Romania	Timişoara	24	H029	EU362490	present study
RTim03	1	Romania	Timişoara	24	H029	EU362490	present study
RTim04	1	Romania	Timişoara	24	H029	EU362490	present study
RTim05	1	Romania	Timişoara	24	H029	EU362490	present study
RTim06	1	Romania	Timişoara	24	H029	EU362490	present study
RTim07	1	Romania	Timişoara	24	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
RTim08	1	Romania	Timişoara	24	H029	EU362490	present study
RTim09	1	Romania	Timişoara	24	H029	EU362490	present study
RTim10	1	Romania	Timişoara	24	H029	EU362490	present study
RTim11	1	Romania	Timişoara	24	H029	EU362490	present study
RTim12	1	Romania	Timişoara	24	H029	EU362490	present study
RTim13	1	Romania	Timişoara	24	H029	EU362490	present study
RTim14	1	Romania	Timişoara	24	H029	EU362490	present study
RTim15	1	Romania	Timişoara	24	H029	EU362490	present study
RTim16	1	Romania	Timişoara	24	H029	EU362490	present study
RTim17	1	Romania	Timişoara	24	H029	EU362490	present study
RTim18	1	Romania	Timişoara	24	H029	EU362490	present study
RTim19	1	Romania	Timişoara	24	H029	EU362490	present study
RTim20	1	Romania	Timişoara	24	H029	EU362490	present study
RTim21	1	Romania	Timişoara	24	H029	EU362490	present study
RTim22	1	Romania	Timişoara	24	H029	EU362490	present study
RTim23	1	Romania	Timişoara	24	H029	EU362490	present study
RTim24	1	Romania	Timişoara	24	H029	EU362490	present study
RTim25	1	Romania	Timişoara	24	H029	EU362490	present study
RTim26	1	Romania	Timişoara	24	H029	EU362490	present study
RTim27	1	Romania	Timişoara	24	H029	EU362490	present study
RTim28	1	Romania	Timişoara	24	H029	EU362490	present study
RTim29	1	Romania	Timişoara	24	H029	EU362490	present study
RTim30	1	Romania	Timişoara	24	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
RTim31	1	Romania	Timişoara	24	H029	EU362490	present study
PLSud01	1	Poland		18	H029	EU362490	present study
PLSud02	1	Poland		18	H029	EU362490	present study
PLSud03	1	Poland		18	H029	EU362490	present study
PLSud04	1	Poland		18	H029	EU362490	present study
PLSud05	1	Poland		18	H029	EU362490	present study
UPla01	1	Ukraine	Plain	39	H029	EU362490	present study
UPla02	1	Ukraine	Plain	39	H029	EU362490	present study
UPla03	1	Ukraine	Plain	39	H029	EU362490	present study
UPla04	1	Ukraine	Plain	39	H029	EU362490	present study
UPla05	1	Ukraine	Plain	39	H029	EU362490	present study
UPla06	1	Ukraine	Plain	39	H029	EU362490	present study
UPla07	1	Ukraine	Plain	39	H029	EU362490	present study
UPla08	1	Ukraine	Plain	39	H029	EU362490	present study
UMts01	1	Ukraine	Mountains	39	H029	EU362490	present study
UMts02	1	Ukraine	Mountains	39	H029	EU362490	present study
UMts03	1	Ukraine	Mountains	39	H029	EU362490	present study
SLDif01	1	Slovakia		31	H029	EU362490	present study
SLDif02	1	Slovakia		31	H029	EU362490	present study
SLDif03	1	Slovakia		31	H029	EU362490	present study
SLDif04	1	Slovakia		31	H029	EU362490	present study
SLDif05	1	Slovakia		31	H029	EU362490	present study
SLDif06	1	Slovakia		31	H029	EU362490	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
SLDif07	1	Slovakia		3	H029	EU362490	present study
				1			
SLDif08	1	Slovakia		3	H029	EU362490	present study
				1			
SLDif09	1	Slovakia		3	H029	EU362490	present study
				1			
SLDif10	1	Slovakia		3	H029	EU362490	present study
				1			
SLDif11	1	Slovakia		3	H029	EU362490	present study
				1			
RPla03	1	Romania	Plateau	2	H022	EU362412	present study
				3			
RCar21	1	Romania	Caransebeş	2	H022	EU362412	present study
				2			
GSwh01	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh02	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh03	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh04	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh05	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh06	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh07	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh08	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh09	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh10	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh11	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh12	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh13	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh14	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh15	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			
GSwh16	1	Germany	Schleswig-Holstein	1	H022	EU362412	present study
				3			

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
GSwh17	1	Germany	Schleswig-Holstein	13	H022	EU362412	present study
GSwh18	1	Germany	Schleswig-Holstein	13	H022	EU362412	present study
GSwh19	1	Germany	Schleswig-Holstein	13	H022	EU362412	present study
GSwh20	1	Germany	Schleswig-Holstein	13	H022	EU362412	present study
GHar01	1	Germany	Harz	14	H022	EU362412	present study
GHar02	1	Germany	Harz	14	H022	EU362412	present study
GHar03	1	Germany	Harz	14	H022	EU362412	present study
GHar04	1	Germany	Harz	14	H022	EU362412	present study
GHar05	1	Germany	Harz	14	H022	EU362412	present study
SLDif12	1	Slovakia		31	H135	KC608843	present study
SLDif13	1	Slovakia		31	H135	KC608843	present study
RPla04	1	Romania	Plateau	23	H035	EU362528	present study
RTim32	1	Romania	Timişoara	24	H035	EU362528	present study
RTim33	1	Romania	Timişoara	24	H035	EU362528	present study
RTim34	1	Romania	Timişoara	24	H035	EU362528	present study
SLDif14	1	Slovakia		31	H035	EU362528	present study
UPla09	1	Ukraine	Plain	39	H136	KC608844	present study
SLDif15	1	Slovakia		31	H023	EU362410	present study
SLDif16	1	Slovakia		31	H023	EU362410	present study
RCon01	1	Romania	Constanța	3	H121	KC608831	present study
RCon02	1	Romania	Constanța	3	H121	KC608831	present study
RCon03	1	Romania	Constanța	3	H121	KC608831	present study
RCon04	1	Romania	Constanța	3	H121	KC608831	present study
RCon05	1	Romania	Constanța	3	H121	KC608831	present study
RCon06	1	Romania	Constanța	3	H121	KC608831	present study
GHar06	1	Germany	Harz	14	H085	DQ379242	present study

Specimen	No. ind.	Country	Location	Pop. ID	Haplotype	GenBank	Source
GHar07	1	Germany	Harz	14	H085	DQ379242	present study
GHar08	1	Germany	Harz	14	H085	DQ379242	present study
GHar09	1	Germany	Harz	14	H085	DQ379242	present study
GHar10	1	Germany	Harz	14	H085	DQ379242	present study
GHar11	1	Germany	Harz	14	H085	DQ379242	present study
GHar12	1	Germany	Harz	14	H085	DQ379242	present study
GHar13	1	Germany	Harz	14	H085	DQ379242	present study
GHar14	1	Germany	Harz	14	H085	DQ379242	present study
GHar15	1	Germany	Harz	14	H085	DQ379242	present study
GHar16	1	Germany	Harz	14	H085	DQ379242	present study
GHar17	1	Germany	Harz	14	H085	DQ379242	present study
GHar18	1	Germany	Harz	14	H085	DQ379242	present study
GHar19	1	Germany	Harz	14	H085	DQ379242	present study
GHar20	1	Germany	Harz	14	H085	DQ379242	present study
RCar22	1	Romania	Caransebeş	22	H137	KC608845	present study
RCar23	1	Romania	Caransebeş	22	H137	KC608845	present study
UMts04	1	Ukraine	Mountains	39	H034	AF136555	present study
UMts05	1	Ukraine	Mountains	39	H034	AF136555	present study
UMts06	1	Ukraine	Mountains	39	H034	AF136555	present study
UMts07	1	Ukraine	Mountains	39	H034	AF136555	present study
UMts08	1	Ukraine	Mountains	39	H034	AF136555	present study

REFERENCES

Alves, E., Ovilo, C., Rodriguez, M. C. & Silio, L. (2003) Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. *Animal Genetics*, **34**, 319–324.

Alves, P.C., Pinheiro, I., Godinho, R., Vicente, J., Gortázar, C. & Scandura, M. (2010) Genetic diversity of wild boar populations and domestic pig breeds (*Sus scrofa*) in South-western Europe. *Biological Journal of the Linnean Society*, **101**, 797–822.

Fang M, Berg F, Ducos A, Andersson L (2006) Mitochondrial haplotypes of European wild boars with 2*n* = 36 are closely related to those of European domestic pigs with 2*n* = 38. *Animal Genetics*, **37**, 459–464.

Giuffra, E., Kijas, J. M. H., Amarger, V. Carlborg, Ö., Jeon, J.-T. & Andersson L. (2000) The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics*, **154**, 1785–1791.

Gongora, J., Peltoniemi, O. A. T., Tammen, I., Raadsma, H. & Moran, C. (2003) Analyses of possible domestic pig contribution in two populations of Finnish farmed wild boar. *Acta Agriculturae Scandinavica A*, **53**, 161–165.

Hajji, G.E. & Zachos, F.E. (2011) Mitochondrial and nuclear DNA analyses reveal pronounced genetic structuring in Tunisian wild boar *Sus scrofa*. *European Journal of Wildlife Research*, **57**, 449–456.

Larson, G., Dobney, K., Albarella, U., Fang, M-Y., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L. & Cooper, A. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, **307**, 1618–1621.

Okumura, N. , Kurosawa, Y., Kobayashi, E., Watanobe, T., Ishiguro, N., Yasue, H. & Mitsuhashi, T. (2001) Genetic relationship amongst the major non-coding regions of mitochondrial DNAs in wild boars and several breeds of domesticated pigs. *Animal Genetics*, **32**, 139–147.

Randi, E., D'Huart, J. P., Lucchini, V. & Aman, R. (2002) Evidence of two genetically deeply divergent species of waterhog, *Phacochoerus africanus* and *P. aethiopicus* (Artiodactyla: Suiformes) in East Africa. *Mammalian Biology*, **67**, 91–96.

Scandura, M., Iacolina, L., Crestanello, B., Pecchioli, E., Di Benedetto, M.F., Russo, V., Davoli, R., Apollonio, M. & Bertorelle, G. (2008) Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? *Molecular Ecology*, **17**, 1745–1762.

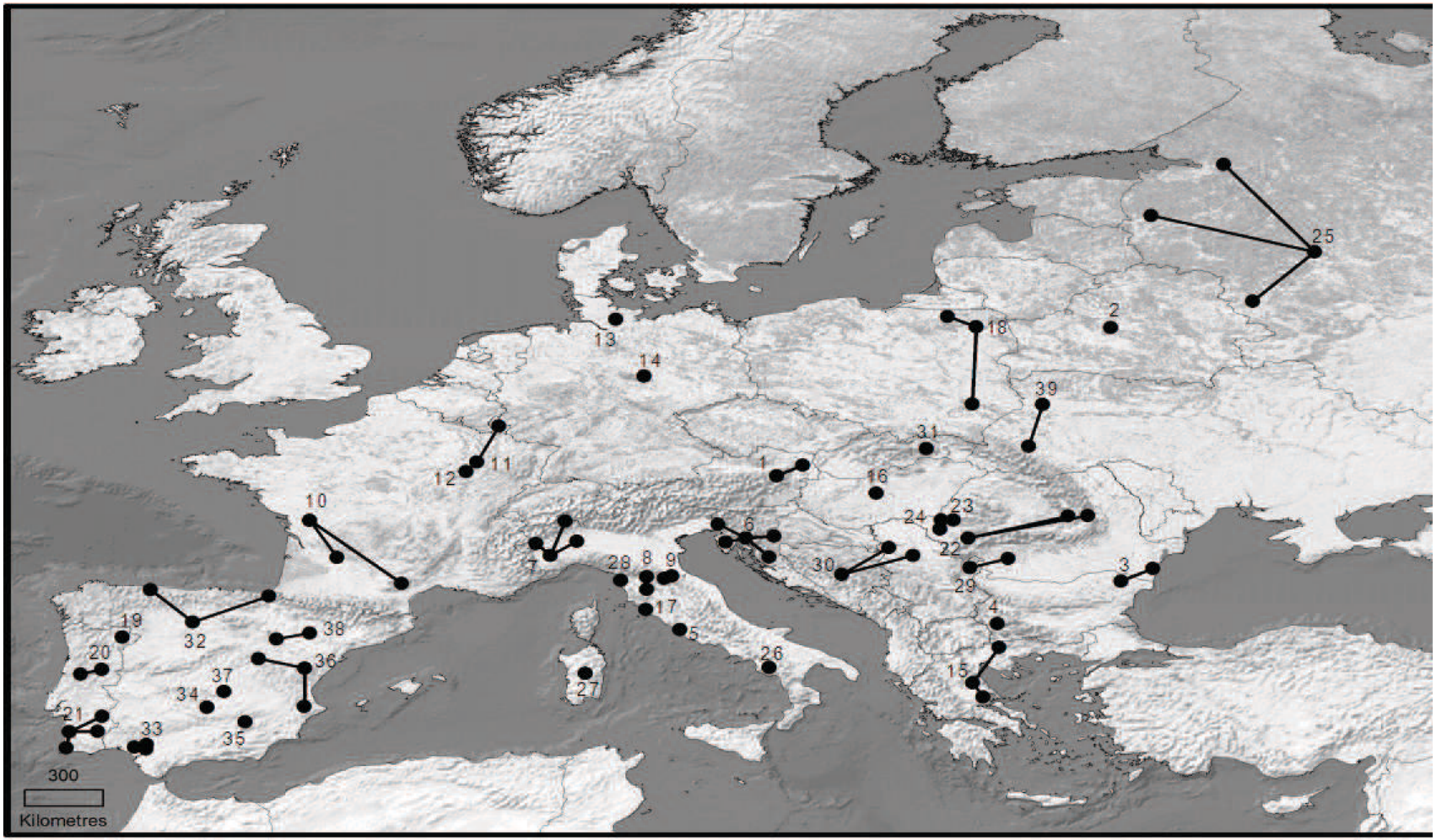
Appendix S3 Additional information and calculations on wild boar distribution and mtDNA diversity in Europe.

TABLE OF CONTENTS

REFERENCES

Wilson, D.E. & Reeder, D.A.M. (2005) *Mammal species of the world: a taxonomic and geographic reference*. Johns Hopkins University Press, Baltimore, MD.

Appendix S3a Map showing the 77 sites (black dots) where wild boar were sampled. Numerical codes refer to the 39 populations used for calculations (see Appendices S2 & S3b). Sites that were grouped into one population are connected by solid lines (see text for a description of the procedure used).


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
300
Kilometres

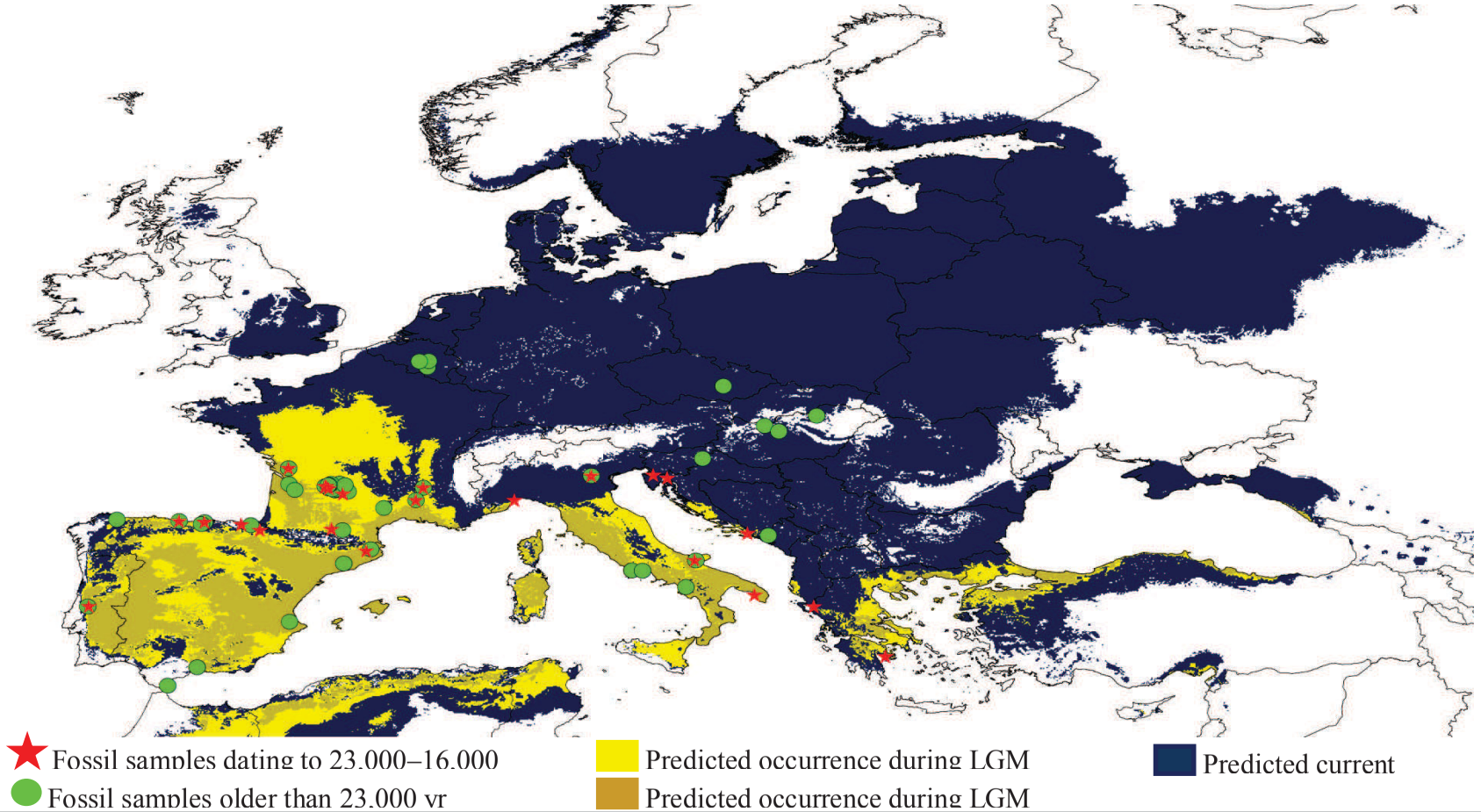
Appendix S3b Genetic diversity observed in the 39 European populations. Pop., population number; subsp., putative subspecies, following Wilson & Reeder (2005); *n*, sample size of each population; *S*, number of polymorphic sites; Ts, number of transitions; Tv, number of transversions; *H*, number of haplotypes; *Hk*, haplotype diversity; AR, allelic richness; π, nucleotide diversity; *k*, mean number of pairwise differences. Numbers in bold refer to statistically significant values at the neutrality tests Tajima's *D* and Fu's F_S.

Pop.	Area	Subsp.	*n*	*S*	Ts	Tv	*H*	*Hk*	AR	π	*k*	Tajima's *D*	Fu's F_S
1	Lower Austria	*scrofa*	13	6	5	1	5	0.744	2.606	0.007	2.821	**1.694**	0.795
2	Belarus	*attila*	19	4	4	0	3	0.550	1.546	0.003	1.216	0.185	1.648
3	Black Sea	*lybicus*	18	6	5	1	5	0.484	1.807	0.003	1.229	−1.183	−0.735
4	Western Bulgaria	*lybicus*	12	2	2	0	3	0.530	1.538	0.001	0.576	−0.382	−0.362
5	Castelporziano	*majori*	10	1	1	0	3	0.600	1.692	0.002	1.000	**1.303**	0.477
6	North Adriatic	*scrofa*?	26	3	2	1	4	0.397	1.279	0.001	0.554	−0.965	−1.098
7	Piedmont	*majori*?	19	12	11	1	7	0.860	3.641	0.011	4.421	1.047	1.080
8	Central Tuscany	*majori*?	18	11	10	1	4	0.595	1.922	0.005	2.176	−1.166	1.865
9	Northern Apennines	*majori*?	18	12	11	1	4	0.765	2.508	0.010	3.786	0.410	**3.969**
10	Western France	*scrofa*	28	8	7	1	5	0.669	2.139	0.007	2.698	0.965	2.246
11	Eastern France–Luxembourg	*scrofa*	18	1	1	0	2	0.294	0.798	0.004	1.544	0.022	0.463
12	Arc-en-Barrois	*scrofa*	10	0	0	0	1	0.000	0.000	0.000	0.000	0.000	0.000
13	Schleswig-Holstein	*scrofa*	20	0	0	0	1	0.000	0.000	0.000	0.000	0.000	0.000
14	Harz	*scrofa*	20	3	2	1	2	0.395	0.917	0.003	1.184	1.069	3.433
15	Greece	*lybicus*	29	3	2	1	5	0.727	2.498	0.003	1.300	1.455	0.028
16	Hungary	*attila*?	10	2	2	0	2	0.356	0.933	0.002	0.711	0.019	1.523
17	Maremma	*majori*	11	9	9	0	2	0.436	0.976	0.010	3.927	1.169	6.822

Pop.	Area	Subsp.	*n*	*S*	Ts	Tv	*H*	*Hk*	AR	π	*k*	Tajima's *D*	Fu's F_S
18	Eastern Poland	*scrofa*	22	2	2	0	3	0.437	1.204	0.002	0.784	0.970	0.761
19	Northern Portugal	*scrofa*	36	3	3	0	4	0.567	1.376	0.003	1.083	1.112	0.786
20	Central Portugal	*scrofa*	14	3	3	0	3	0.484	1.538	0.003	1.143	0.647	1.145
21	Southern Portugal	*scrofa*	18	2	1	1	2	0.366	0.892	0.002	0.732	0.632	2.082
22	Romanian Carpathians	*attila*	24	4	3	1	4	0.308	1.091	0.001	0.409	−1.690	**−1.854**
23	Lipova	*attila*	20	5	2	3	4	0.500	1.689	0.003	1.268	−0.306	0.556
24	Timişoara	*attila*	34	2	2	0	2	0.166	0.511	0.001	0.332	−0.636	0.953
25	Western Russia	*attila*	18	5	4	1	5	0.778	2.777	0.005	1.993	0.933	0.506
26	Salerno	*majori*?	7	14	13	1	5	0.905	4.000	0.013	5.238	−0.459	0.384
27	Sardinia	*meridionalis*	83	20	19	1	14	0.815	3.281	0.010	4.086	0.022	−0.326
28	San Rossore	*majori*	10	9	8	1	2	0.200	0.700	0.004	1.800	**−1.901**	3.672
29	Eastern Serbia	*lybicus*?	17	3	3	0	5	0.713	2.450	0.002	0.971	−0.110	−1.414
30	Western Serbia–Bosnia	*lybicus*?	19	3	1	2	4	0.614	1.720	0.002	0.702	−0.496	−0.859
31	Slovakia	*scrofa*	16	5	4	1	4	0.525	1.838	0.002	0.950	−1.218	−0.374
32	Northern Spain	*scrofa*	22	5	5	0	5	0.684	2.272	0.003	1.208	−0.355	−0.508
33	Coto Doñana	*scrofa*	14	2	2	0	2	0.495	0.990	0.004	1.484	**1.554**	**3.641**
34	Ciudad Real	*scrofa*	16	4	4	0	4	0.767	2.551	0.004	1.800	1.543	1.137
35	Jaén	*scrofa*	16	2	2	0	3	0.658	1.820	0.002	0.858	1.085	0.668
36	Eastern Spain	*scrofa*	15	2	2	0	3	0.648	1.726	0.003	1.048	**1.850**	1.029
37	Toledo	*scrofa*	11	3	3	0	4	0.600	2.164	0.003	1.273	0.830	−0.228
38	Aragon	*scrofa*	11	2	2	0	3	0.709	1.952	0.002	0.982	1.339	0.551

Pop.	Area	Subsp.	*n*	*S*	Ts	Tv	*H*	*Hk*	AR	π	*k*	Tajima's *D*	Fu's F_S
39	Ukraine	*attila*	17	4	4	0	3	0.522	1.371	0.003	1.235	−0.169	1.556

Appendix S3c Wild boar distribution as predicted by MAXENT on the basis of current and past (i.e. LGM, CCSM and MIROC models) climatic conditions. Fossil records are represented by red stars or green dots. Predicted current suitable areas are represented by dark blue colour, while predicted suitable areas during the LGM are represented by yellow (predicted by only one model) or dark yellow (predicted by both models).



Chapter 2

Isolation by distance, by barrier and by resistance have shaped a sharp genetic structure in a island wild boar population.

Daniela Biosa, Antonello Canu, Laura Iacolina, Marco Apollonio, Massimo Scandura

Manuscript

Foto Alberto Addis

Isolation by distance, by barrier and by resistance have shaped a sharp genetic structure in a island wild boar population.

Daniela Biosa[1], Antonello Canu[1], Laura Iacolina[1], Marco Apollonio[1], Massimo Scandura[1]

1 Dipartimento di Scienze della Natura e del Territorio, Università di Sassari, via Muroni 25, 07100 Sassari, Italy

Abstract

Genetic diversity within animal populations can be evenly distributed or, more often, be structured by the simple effect of distance or by the existence of breaks in landscape connectivity.

We studied the genetic differentiation within the wild boar (*Sus scrofa*) population in Sardinia island (Italy), and tested it for the existence of isolation by distance (IBD) and isolation by barrier (IBB), after the exclusion of the possible role of local genetic introgression from continental boars and domestic pigs. A total of 368 Sardinian wild boar samples were analysed with a set of 16 microsatellites. Signals of genetic introgression were revealed by a Bayesian cluster analysis including a total of 328 reference wild boars and domestic pigs. After removal of putative introgressed individuals (above a threshold of 10%), genetic structure in the population was confirmed by different statistical approaches. Blind (STRUCTURE and PSMIX) and spatially explicit methods (GENELAND and STRUCTURE) supported a sharp partition into three discrete subpopulations. A significant IBD pattern was detected, taking land use into account (least-coast distance) the pattern increased. In addition, genetic discontinuities between subpopulations were also explained by the presence of the main motorway (S.S. 131), crossing the island from north to south.

Interestingly, this physical barrier S.S. 131and areas of low permeability to wild boar have limited gene flow in the island, protecting western subpopulations from the spread of exotic genes mostly introgressed in the eastern subpopulation. This study reveals how human-transformed landscapes can strongly impact the genetic connectivity even in large-sized and highly mobile animal species.

Key words: *Sus scrofa meridionalis*, Sardinia, landscape genetics, microsatellites, gene flow, road impact

Introduction

Connectivity and habitat fragmentation can have a strong impact on the onset of genetic differentiation between and within populations of the same species. Indeed, the presence of barriers leads to a disjunction and, sometimes, a complete isolation of portions of the population, leading to genetic drift and, consequently, to possible changes in the genetic composition. Usually, the presence of barriers and absence of corridors promotes the limitation of gene flow, the reduction of genetic diversity and the increase of inbreeding (Balkenhol & Waits 2009). In the last decade numerous analytical approaches have been developed to infer microevolutionary processes driven by environmental fragmentation and human infrastructures, giving rise to a field named 'landscape genetics' (Manel et al., 2003; Storfer et al., 2010). Briefly, landscape genetics analysis consists in correlating genetic variation with environmental characteristics to quantify the effect of the environment on the population genetic structure (Storfer et al., 2007). These kind of approaches have been widely adopted to evaluate the influence of natural, but even more, anthropogenic barriers on the gene flow of animal and plant species. Since urban and suburban development and road network extension are among the primary causes of habitat fragmentation, this analysis can be of help in planning management practices for species conservation (Holderegger & Di Giulio 2010).

For this purpose, several studies on the landscape genetics of large mammals have been conducted (e.g. roe deer *Capreolus capreolus* Coulon et al., 2006; red deer *Cervus elephus* Pérez-Espona et al., 2008; mountain lion *Puma concolor* Castilho et al., 2011; giant panda *Ailuropoda melanoleuca* Zhu et al., 2010; tiger *Panthera tigris* Sharma et al., 2013; caribou *Rangifer tarandus* Weckworth et al., 2013). Such studies, using different statistical approaches, indicated that landscape features could explain a large proportion of genetic differentiation that is not explained by spatial distance only.

Important phenomena like urbanization and the development of large networks of transport infrastructures have rapidly increased in Europe. Such expansion may have affected the spatial behaviour of many large mammals, like ungulates, and consequently their population genetic structure. Although their social relevance and diffusion, only few studies have been aimed at evaluating the effect of anthropogenic barriers on ungulates genetic structure (bighorn sheep *Ovis canadensis* Epps et al., 2005; roe deer Coulon et al., 2006; Kuehn et al., 2007; Hepenstrick et al., 2012; red deer Šprem et al., 2013, wild boar *Sus scrofa* and red

deer Frantz et al., 2012). Due to scarcity of information, it is difficult to establish the real impact of such barriers, since they could have various levels of permeability depending on the species behavioural characteristics. For example Frantz et al. (2012) showed how the presence of a motorway could affect differently two ungulate species in Belgium. It seemed to act as a barrier for the red deer, leading to a genetic differentiation between subpopulations inhabiting the opposite sides of the road, while it seemed not to be a disturbing factor for the wild boar.

The wild boar is one of the most important and widespread ungulate game species in Europe. It shows an opportunistic behaviour and is adaptable to almost any type of environment. Climate represents the main limiting factor for wild boar through its effect on physiology and metabolism or through its indirect effect on food availability and accessibility (Geisser & Reyer, 2005; Melis et al., 2006).

The wild boar is considered a sedentary species with a quite small-scaled use of the space (Keuling et al., 2008), regardless of the habitat occupied. Dispersal is male-biased, females are philopatric and form matriarchal social groups, while adult males stay mostly isolated. The available data indicate that wild boar dispersal takes place between 11 and 16 months of age and covers limited distances (mostly < 20 km, Briedermann, 1990; Truvè & Lemel, 2003; Keuling et al., 2010). Dispersal patterns are influenced by various factors such as population density, habitat structure and quality, and climate (Dardaillon & Bougnon, 1987; Keuling et al., 2010). In addition, human activities can have an impact on different aspects of the species ecology and behaviour. For instance, wild boar is known to modify its activity and spatial patterns in relation to human disturbance. If undisturbed wild boar tends to be active during the day, while under hunting pressure and high human disturbance it shifts its activity to nocturnal (Boitani et al., 1994; Podgórski et al., 2013). Nevertheless, thanks to its plasticity, a tendency to adapt to human presence and infrastructures is observed around urban centres (Cahill et al., 2012; Osashi et al., 2013).

Our study is focused on the wild boar population inhabiting Sardinia, a scarcely populated island and still underdeveloped in terms of main infrastructure, if compared to other regions of continental Europe. Sardinian wild boar is considered a dwarf form of the European wild boar, which originated during the Neolithic, and it is currently classified as a distinct subspecies (*Sus scrofa meridionalis* Major 1883), on the basis of both morphological and genetic evidences. Its long-lasting isolation has produced a relevant genetic differentiation of the Sardinian

population, observed using different types of genetic markers by Scandura and colleagues (2008, 2009, 2011) and Iacolina and colleagues (submitted). Furthermore, Scandura et al., (2011) detected appreciable levels of genetic introgression from domestic pigs and continental wild boar, and a relevant genetic structure. The population came out to be divided into three subpopulations: one in the east of the island, one in the north-west and the last in the south-west. The authors concluded that the sharp east-west genetic differentiation could not be explained by isolation-by-distance only and rather argued that landscape features could play an important role.

In the present study, we analysed the Sardinian wild boar population, increasing the number of individuals and genetic markers, with the aim to evaluate the genetic structure suggested in the previous study in relation to natural and anthropogenic environmental variables that could act as a barrier, preventing genetic shuffling among subpopulations. Both isolation by distance (IBD) and isolation by barrier (IBB) were tested using different landscape genetics approaches and considering various environmental features, as suggested by Balkenhol et al. (2009) and Frantz et al. (2012).

Materials and Methods

Study area

The study area is represented by the Sardinia island, the second biggest in the Mediterranean sea for extension (24,100 km^2) and inhabitants (1,640,379). Nonetheless, population density is relatively low for Europe (around 68 inhabitants/km^2) and people are concentrated in the five main cities and along the coast, while the interior is characterized by small villages and large uninhabited areas.

The climate is mediterranean-temperate at low elevations and along the coast, and continental in the inland at higher elevations. Temperature is mild and relatively constant during the year (on average 18°C, ranging between a mean of 7°C in winter and 25°C in summer). Precipitations mostly fall during autumn and winter, being more frequent in the northern and western sides of the island. Annual precipitations range from less than 400 mm in the dry south to almost 1500 mm in the eastern mountains.

The island is relatively dry, and major rivers have mostly the features of streams in summer. A single small natural lake (Lake Baratz, 0.6 km^2) and some tens of artificial basins (the

biggest one being Lake Omodeo, 29 km^2) are also present, as well as a number of ponds and lagoons along the coast. Because of the island size , habitats may differ greatly. Coasts (1,849 km long) are generally high and rocky, interspersed by a number of sandy shores. Mountains occupy only the 13.6% of the territory and are mainly concentrated in the central-eastern part of the island. The main mountainous massifs are the Gennargentu, in the central-eastern side (reaching 1,834 m a.s.l.) and the chains of Marghine and Goceano, crossing the island from north-east to south-west. Plateaus and flatlands occupy the 18,5% of the territory, the main flatland being represented by the Campidano Plain in the south-west, a human-modified landscape dominated by cultivations, especially cereal crops, fruit trees and vineyards.

The wild boar is a game species widespread in the island, where it occurs in various habitats, it is rare only in the Campidano plain, due to its ecological plasticity. Estimates of population size are weak and affected by large confidence intervals. Local densities were derived from hunting data of 168 hunting grounds spread throughout the island and ranged between 0 and 33 culled heads/km^2. On the basis of habitat suitability analyses higher densities are expected to occur in the central and northern part of the island (Autonomous Region of Sardinia 2012).

To evaluate the impact of the land use on the dispersal of wild boar and its ultimate effect on the genetic structure, we used the 4th Level CORINE Land Cover at scale 1:50,000 (Heymann et al., 1994). We divided the island into 6451 cells of 2x2 km and for each cell we calculated the percentage surface of each land cover category. Cover categories were pooled according their effectiveness to act as barrier to the species' movements and to the ecological requirements of the wild boar (Tab. 1). Considering the lack of information regarding the Sardinian population, we relied on studies in Mediterranean areas that have stressed the importance of seasonal availability of food and water, and the presence of refuge areas (Boitani et al., 1994; Massei et al., 1997; Fernandez-Llario & Carranza, 2000; Focardi et al., 2008). In addition to such natural variables, also the distribution of human activities and infrastructures has a strong impact on the presence of the wild boar. Road and railway networks are in most cases crossed by wild boar during their movements, but if intensively frequented and associated to permanent shields may become an effective barrier, limiting the species' dispersal. Only one main road with the mentioned features occurs in Sardinia: the S.S. 131 'Carlo Felice', a superficial motorway with 4 lanes and with very few crossing points for wildlife. Its trail was firstly set in

the XIX century, but it was paved in its present form in the last 50 years. It crosses the island from north to south connecting the two major cities, Cagliari and Sassari.

As a result of our spatial analysis, we predicted that the main landscape features that can restrict the movement of the species in the island are represented by the Campidano plain and the S.S. 131 motorway.

Sampling and microsatellite genotyping

A total of 368 wild boar samples were obtained from all over Sardinia by local hunters during the period 2001-2011 (Fig. 1a). Muscle or hair samples were collected from hunted animals and stored, respectively, in absolute ethanol or frozen until analysis. Sampling locations were mapped using ArcGIS v. 10 (ESRI, Redlands, CA, USA). Accuracy of spatial information differed among samples: punctual GPS locations were available for some animals only, whereas for most of them hunters reported either the municipality or the hunting ground where the animal was shot (i.e. polygons in the range 26-547 km^2, median size 79 km^2). In the last case, geographical coordinates of the geometric centre of the area were used for statistical analysis.

DNA was extracted using Genelute kit (Sigma-Aldrich, St Louis, MO, USA) for tissue samples and Instagene Matrix (Bio-Rad, Hercules, CA, USA) for hair samples, and then stored at -20°C. All samples were genotyped with a panel of 16 microsatellites (STR; Short Tandem Repeat): S090, SW122, SW2532, S355, SW1492, SW461, IGF1, SW951, SW2021, SW2496, S026, S215, SW72, SW857, S155 and SW24 (details at www.thearkdb.org). Each PCR was performed in a 10 µL reaction volume, containing 3 µL of DNA solution, 0.5 U of Taq DNA polymerase (Euroclone, Pero, Italy), 1× PCR buffer (Euroclone), 2.5 mM $MgCl_2$, 100 µM of each dNTP and 2 pmol of each primer. The forward primer of each pair was labelled with an ABI fluorescent dye (6-FAM, HEX or NED). The amplification profile was set up with an initial step of denaturation at 95 °C for 3 min, followed by 35 cycles of 92 °C for 45 s, Ta (54–62 °C) for 45 s, and 72 °C for 30 s. A further extension step of 72 °C for 10 min concluded the reaction. Amplicons were sized using capillary electrophoresis in an ABI PRISM 3100 and 3730XL Avant automatic sequencer (Applied Biosystems, Carlsbad, CA, USA) by the BMR-Genomics sequencing service (Padua, Italy). Peak Scanner software v. 1.0 (Applied Biosystems) was used to analyze electrophoretic data.

In order to evaluate possible signatures of genetic introgression, genotypes of the Sardinian wild boars were compared with 214 reference wild boars from different European

countries and with 114 domestic pigs, including commercial and local free-ranging individuals (Tab. S1).

Microsatellite data analysis

In order to detect evidences of null alleles, stuttering or large allele dropout, data were checked with MICRO-CHECKER 2.2.3 (van Oosterhout et al., 2004).

Deviations from Hardy-Weinberg equilibrium (HWE) and linkage equilibrium (LE) were tested in the Sardinian population using GENEPOP v. 4.2 (Raymond & Rousset, 1995). Tests for HWE employed the Markov chain method proposed by Guo and Thompson (1992), with the following chain parameters: 10000 dememorizations, 100 batches and 10000 iterations. Deviations from LE were tested for each pair of loci. Significance levels were lowered, accounting for the number of multiple tests by the sequential Bonferroni procedure (Rice 1989).

The occurrence of imported exotic boars as long with the signature of genetic introgression from continental populations (Italian peninsula or central Europe) and from domestic pigs was recently proven in Sardinia by Scandura and coauthors (2011). Notably, they showed the effect of such gene flow on the inference of the population genetic structure. On account of those findings, in the present study we preliminarily screened all Sardinian genotypes in order to remove non-negligible distortions in allele frequencies that were attributable to human-mediated introgressive hybridization or to introduction of exotic boars.

With this aim in mind, we implemented a Bayesian cluster analysis in STRUCTURE v. 2.3.4 (Pritchard et al., 2000; Falush et al., 2003, 2007; Hubisz et al., 2009). Firstly, we performed 10 independent Monte Carlo Markov chain (MCMC) runs simulating a number of subpopulations (K) ranging between 1 and 10, with the following settings: admixture model, no population information, correlated allele frequencies, 200,000 burn-in and 200,000 iterations of data collection. Then, we selected the value of K that maximized the discrimination among the following four groups: Sardinian wild boar (Sardinian WB), mainland Italy wild boar (Italian WB), other European wild boars (European WB) and domestic pigs. Accordingly we assessed the nature of each individual sampled in Sardinia, in relation to the possible occurrence of gene introgression from the other wild and domestic populations. The degree of admixture was individually evaluated by referring to the Q-values obtained in the run with highest likelihood at

the selected K-value. To be conservative, only individuals showing >90% membership to the Sardinian cluster (Q_{Sar}) were retained for further analyses. In so doing we admit the erroneous exclusion of non-introgressed individuals that can be misclassified, but we expect that real immigrants and first-generation hybrids will be correctly classified ($Q_{Sar} < 0.90$) and excluded (see also Frantz et al., 2013).

The pruned dataset of Sardinian wild boars was hence used to infer population structure. In agreement with Balkenhol et al., (2009) and Frantz et al., (2012), suggesting the use of multiple approaches when seeking signals of population structure, two different Bayesian clustering algorithms (STRUCTURE v. 2.3.4 and GENELAND v.4.0.3, Guillot et al., 2008, 2012) and a general maximum likelihood method (PSMIX, Wu et al., 2006) were implemented.

Two procedures were adopted in STRUCTURE: a 'blind' simulation, neglecting any prior information, and an 'informed' analysis, where samples were attributed 'a priori' to one of the subpopulations identified in the previous study (Scandura et al., 2011), were run. In the latter, the attribution to a subpopulation was based only on the geographic position of samples. Those collected on the border between two subpopulations were omitted. In total, 165 individuals were ascribed to the Eastern (ES), 50 to the North-West (NWS) and 62 to the South-West (SWS) subpopulation.

The first ('blind') analysis was performed with the following settings: K = 1-10, 10 independent Monte Carlo Markov chain runs for each K, admixture model, no population information, correlated allele frequencies, 200,000 + 200,000 iterations (burn-in + data collection). The optimal K-value was chosen according to the ΔK statistics developed by Evanno et al., (2005).

In the second ('informed') analysis the same settings were used, except for the incorporation of population information in the prior (usepopinfo function) and the number of K, which was fixed to 3 (equal to the number of expected subpopulations). Thereby we forced the algorithm to calculate the assignment probability to three inferred clusters, which were expected to correspond to the three assumed subpopulations.

To visualize the genetic population structure across geographic space we used the inverse distance weighted (IDW) interpolation in Spatial Analyst tool in ArcGis (as in Vandergast et al., 2011) to interpolate Q-values. The interpolated surface is obtained by assigning values to unknown points from a weighted sum of values of known points (Shepard, 1968). In this way

each point is influenced by the values of nearest points and less by more distant points. The tool creates a raster surface of genetic Q-values characterizing the 3 different subpopulations in relation to their geographical positions.

A further 'informed' analysis was implemented in GENELAND, a Bayesian clustering algorithm that use both geographic and genetic information for each individual to estimate the number of subpopulations, to assign each individual to the subpopulation of origin and to identify possible migrants between subpopulations. The method used by the program to identify spatial patterns is in the Poisson-Voronoi tessellation (Muche 2005). It assumes that there is an unknown number of polygons that approximate the true partition into subpopulations across space. The area covered by each subpopulation can be approximated by the union of polygons. We determined the membership of each sample to one population by running the algorithm 10 times, with the following settings: K = 3, 100,000 MCMC iterations, with thinning of 50 and 500 burn-in, admixture model, correlated allele frequencies, amount of uncertainty on spatial coordinates equal to 500 metres. The run providing the highest average posterior probability was considered for population membership.

Additionally, we used PSMix (Population Stratification inference via Mixture model; Wu et al., 2006), an R package based on maximum likelihood method using the Expectation-Maximization (EM) algorithm (Dempster et al., 1977). Contrary to Bayesian approaches, the EM algorithm computes the confidence intervals of the estimates via resampling methods (bootstrap statistics) and overcomes the main drawbacks of the more used Bayesian methods: the dependence of estimates on the prior adopted and the great computational effort (Tang et al., 2005). Program settings were: K = 3 with 100,000 iterations in the EM algorithm, admixture model and a convergence criterion eps = 1 x e^{-10}. Results consisted in estimates of individual membership to the three inferred clusters.

Finally, we performed a FCA (Factorial Correspondence Analysis) in GENETIX v. 4.05 (Belkhir et al., 2004) to detect differences among non-introgressed genotypes in relation to their assigned subpopulation. Genotypes were finally plotted in a three-dimensional space on the basis of their genetic proximity.

Isolation by distance and by barrier

As the presence of a sharp isolation by distance (IBD) pattern in the population can induce to overestimate the real number of genetic clusters (Frantz et al., 2009; Guillot 2009; Schwartz and McKelvey 2009), we used GenAlEx 6.4 (Peakall & Smouse 2005) to evaluate the occurrence of IBD in the Sardinian population by Mantel test (Mantel 1967). A genetic distance matrix and an Euclidean distance matrix were used, the latter calculated from Universal Transverse Mercator geographic coordinates.

Euclidean distances were often used to describe geographic distances among sample locations but they do not reflect the real routes that species use during their movements. In fact, individual movements are influenced by different landscape elements that are not taken into account when using linear distances. Accordingly, to account for the influence of landscape features on gene flow in the island, we calculated the least-cost path distance (LCD) between sample locations on the basis of a map of land cover resistance weights that were assigned with regard to the habitat permeability to wild boar and its use of the habitat (feeding, refuge), using Spatial Analyst Tool in ArcGIS (Fig. 1b). The LCD is an "effective distance", calculated as the shortest distance between two points, corrected by considering the cost of the species movement across different habitat patches. To quantify this cost, we associated to each cell a weight to create a resistance map (the higher the weight, the greater the cost to move through that cell; Tab. 1). Finally, for every pair of individuals, the LCD was calculated and the obtained path plotted in ArcGIS (Fig. 1b). The matrix of pairwise LCD was then used as distance matrix to perform a Mantel test.

In addition to the effect of land cover on gene flow, we also evaluated the influence of physical barriers represented by conspicuous infrastructures. In particular, we considered the S.S. 131 'Carlo Felice'. In order to test isolation by barrier (IBB) due to the presence of the motorway, we produced two different matrices to be compared with the genetic matrix. Firstly, we produced a binary model considering the presence/absence of the barrier between two samples (1 = presence; 0 = absence). Secondly, we considered the number of times an animal should cross the barrier to reach a different area in the island. In this case the matrix included three values: 0 = no crossing, 1 = single crossing and 2 = double or multiple crossing. Stretches

of sea separating peripheral islands from the major one were considered a conspicuous barrier as well.

To evaluate the relative role of Euclidean distances, least-cost path and the presence of physical barriers, the four different spatial distance matrices were used to perform Partial Mantel tests (Smouse et al., 1986) using the vegan package (Oksanen et al., 2007) in R. Significant correlations were determined by the calculation of Pearson product-moment correlation using a permutation test with 999 replicates. Monte Carlo P-values were calculated to determine the significance of partial Mantel tests.

Finally, we also evaluated the presence of IBB by detecting barriers to gene flow with the software BARRIER v. 2.2 (Manni et al., 2004), which implements the Monmonier's (1973) maximum difference algorithm. The software detects edges associated with the highest genetic diversity, statistically tested by resampled bootstrap matrices of molecular data. The graphical pattern of the genetic boundaries is computed by the Delaunay triangulation (Brassel and Reif 1979), the fastest triangulation method for connecting a set of localities on a map with a set of triangles. It is the dual structure of Voronoi diagrams, which implies that all points within a polygon are closest to its centroids (the location of the sampled population) than to any other polygon, that determines which samples are adjacent. The pruned Sardinian STRs dataset was divided into 16 sampling areas on the basis of the geographic proximity of samples and a F_{ST} distance matrix was calculated among areas to detect the most conspicuous genetic boundaries.

Results

The total number of different alleles in the Sardinian sample was 154 , with a range of 6-16 and an average of 9.63 ± 3.18 (standard deviation, SD) per locus. Missing alleles represented 2.17 % of the dataset. MICRO-CHECKER did not find scoring errors in the dataset or evidence of large allele dropout. Not surprisingly, a significant deviation from Hardy-Weinberg equilibrium was detected in the overall population (all loci $P<0.01$) and several pairs of loci resulted in linkage disequilibrium (51/120 at $\alpha=0.05$ and 45/120 at $\alpha=0.01$, significance corrected for multiple tests).

Identification of immigrant/introgressed individuals

Similarly to the results obtained by Scandura et al. (2011), the Bayesian analysis in STRUCTURE sharply distinguished the four main populations (Sardinian WB, Italian WB, European WB, and domestic pigs) in the overall sample of 696 individuals. However, this result was not achieved at K = 4, but when K was equal to 5. Each of the four populations was univocally assigned to one of the inferred clusters (with Q>0.9), with the exception of Sardinian wild boars which were mainly assigned to two clusters (I and III), corresponding to a cumulative average membership of 0.886 (Fig. S1 Suppl. Mat.). Hence, to conservatively assess which individual was a possible recent immigrant or affected by introgression, we applied the threshold of 0.9 to the sum of Q-values referred to the two 'Sardinian' clusters (Q_{I+III}).

Some degree of introgression was found in the Sardinian sample: 75.8% of individuals (n=279) were attributed to the Sardinian population with $Q_{I+III} > 0.90$, 12 samples were attributed to the domestic pig cluster with a $Q_V > 0.80$, while 36 individuals showed $0.80 < Q_{I+III} < 0.90$. The remaining wild boars (n=41) showed intermediate membership proportions and could not be assigned to any population (all Q-values < 0.8, Fig. 2). For further analyses we thus removed from the dataset 89 possibly introgressed individuals (=12+36+41) with the addition of other two individuals, that, because of uncertainty on the death place, could not be attributed *a-priori* to any subpopulation. Consequently the final pruned dataset was composed by 277 individuals.

Genetic structure of the Sardinian population

The 'blind' Bayesian analysis in STRUCTURE on the pruned dataset confirmed the results obtained by Scandura and coauthors (2011), detecting K = 3 as the most likely partition (Evanno's method) identifying three subpopulations in the island: one in the north-west (NWS), one in the south-west (SWS), and one including all eastern Sardinia (ES, Fig. 3). In total 160 individuals were assigned (on the basis of the highest Q-value) to the ES cluster, 58 to the NWS, and 59 to the SWS. Only 17 individuals (6.1%) were assigned to a different population from those expected on the basis of the sampling site. The observed pattern of population differentiation seemed to identify the presence of abrupt genetic discontinuity in coincidence with the S.S. 131 and the Campidano plain.

The 'informed' analysis at fixed K=3, as expected, greatly improved the allocation of samples to the respective sampling area. All samples out of two (99.3%) were assigned to the expected subpopulation, equalling the number of mismatches obtained in GENELAND (Tab. 2). PSMIX showed a similar partition into subpopulations to that obtained with STRUCTURE without any prior geographical information. In this case, 23 samples (8.3%) were not attributed to the expected subpopulation (Tab. 2).

The consistency of results obtained by different analytical approaches points to a sharp structuring in the island with a limited ongoing gene flow between subpopulations. Pairwise differentiation was maximum for NWS-SWS pair (F_{ST} = 0.161) and minimum for ES-NWS (F_{ST} = 0.089), as confirmed by the FCA plot (Fig S2 Suppl. Mat). As expected (see Scandura et al., 2011), HWE and LE analyses within subpopulations revealed a sensitive reduction of significance with respect to the global population (HWE: 3/16 deviations at =0.01 in ES, no deviation in NWS and 4/16 deviations at =0.01 in SWS; LE: 1/120 loci pair at =0.05 in ES, 2/120 at =0.05 in NWS, 4/120 at =0.01 in SWS, significance corrected for multiple tests).

Influence of distance and landscape on gene flow

Mantel tests and correlograms based on Euclidean distance matrices showed the presence of a weak degree of IBD in the population (R_{xy} = 0.099, P < 0.001; Fig. 4a), suggesting a more complex genetic pattern in the island. Replicated Mantel tests using the least-cost matrix provided stronger correlations (R_{xy} = 0.337, P < 0.001, Fig. 4b), similar to those obtained using the binary barrier matrix (R_{xy} = 0.365, P < 0.001) and the matrix accounting for the number of crossings (R_{xy} = 0.337, P < 0.001). Such results were confirmed by partial Mantel tests, where genetic distances were well explained by least-cost and barrier matrices, even when controlling for the effect of geographical position (Tab. 3).

In accordance with results of the other methods, the barrier predictor analysis using Monmonier's maximum difference algorithm identified two major barriers (Fig. 5). The first barrier separated sampling locations belonging to the SWS subpopulation from the rest of samples, emphasizing its strong genetic differentiation. The second barrier separated the NWS subpopulation from sampling sites occupying the south-west and the east part of the island. Again, both putative barriers matched quite well the extension of the motorway S.S. 131.

Discussion

This study clearly reveals how the joint effect of distance and landscape features can generate genetic discontinuities across a large mammal population. IBD and IBB were detected in the Sardinian wild boar population, the latter being associated to the effect of unsuitable habitats and infrastructures.

In the present study we have enlarged the dataset used in Scandura et al., (2011) by including a larger sample of Sardinian wild boars (from 210 to 368) and increasing the number of autosomal markers (from 10 to 16). Nonetheless, new results confirmed the partition into the same three subpopulations (ES, NWS and SWS) that had been previously detected. Three different statistical approaches (STRUCTURE, GENELAND and PSMIX) gave full support to such genetic structure, thus accomplishing recommendations given by Balkenhol and coauthors (2009) and Frantz and coauthors (2012).

A signature of recent gene introgression from continental wild boars and domestic pigs was also confirmed. Specifically, gene introgression seemed to affect mainly the eastern subpopulation, while the northwestern and southwestern ones were marginally interested. In total, almost 25% of individuals sampled in the island were recognized as putative hybrids and their exclusion from population structure analyses prevented the confounding effect possibly arising from the local occurrence of exogenous alleles.

An IBD pattern was observed in the population and was evident even neglecting landscape features (i.e. using Euclidean distances between sampling sites). Nonetheless, the presence of a sharp genetic structure suggested the existence of barriers to gene flow in the island. IBD was observed in other European wild boar populations at a local scale, whereas it appeared to be absent at a continental scale (Scandura et al., 2008). Both Frantz et al., (2009) and Goedbloed et al., (2013) found a IBD pattern in populations of Central-Western Europe. Interestingly, and in agreement with our results, in both studies IBD was retained also in presence of genetic introgression from alien or domestic sources (see also Frantz et al., 2013).

As expected, the genetic diversity within the Sardinian population was better explained when landscape features, i.e. land use and the main motorway (S.S. 131), were included in statistical correlations. This result suggests that the presence of unsuitable habitats and man-made infrastructures can effectively limit wild boar movements in the island. Particularly, the

sudden genetic differentiation between western and eastern subpopulations seemed to occur in conjunction with the motorway S.S. 131 (Fig. 3).

When the assignment tests were performed using the geographical information on the individuals, the majority of individuals (> 96%) was associated with the respective sampling subpopulation. In contrast, in analysis based only on the genetic information the number of misassigned individuals slightly increased and included especially individuals sampled near the motorway. Such result would suggest that wild boar could disperse across the motorway, but at a very low rate.

The impact of the motorway seemed to be stronger in the north, where a continuity of land use is observed and a higher degree of connectivity is therefore expected on the basis of environmental conditions only (Fig. 1b). This prediction is supported by the records of road accidents along the S.S. 131. During the last 10 years (2001-2011, Autonomous Region of Sardinia, unpublished data) through the complete extent of the S.S. 131, 106 accidents involving wild boars were recorded, mostly concentrated in the area separating NWS and ES subpopulations. On the opposite, the reduced gene flow between SWS and ES is more likely an effect of the high resistance to the species given by the Campidano plain (Fig. 1b). In this area, the motorway connects two major urban centres (Cagliari and Oristano), crossing a lowland characterized by important suburban and industrial surfaces as well as agricultural crops, thus being fairly unsuitable for the species movements. Hence, in this case a pattern of isolation by resistance (McRae 2006) would prevail.

Finally, subpopulations NWS and SWS appeared reciprocally isolated (i.e., no recent gene flow), probably due to the breaking presence of the urban area of Oristano, delimited in the west by the coast and in the east by the S.S. 131.

Notably, such barriers in Sardinia have also prevented the spread of introgressed genes from the eastern subpopulation to the rest of the island and probably safeguarded the genetic integrity of the SWS subpopulation (seemingly the purest one). This is a very interesting case in a conservation viewpoint, as only negative effects are typically associated to anthropogenic barriers.

But can the current landscape, which has been mostly influenced by recent transformations, explain the high genetic differentiation observed across the island?

Actually, although a long time lag is usually expected between a causal factor and a detectable genetic effect, simulations have proved that a limited number of generations (as small as 15) can be sufficient for the genetic signature of a landscape barrier to become detectable (Landguth et al., 2010). Accordingly, several studies exploring genetic discontinuities linked to linear barriers have documented the relevant effect of infrastructures built only 10-40 years before (Epps et al., 2005; Pérez-Espona et al., 2008; Hepenstrick et al., 2012). This time span is similar to that elapsed from the enlargement of the S.S. 131, thus confirming that its role is likely to have been substantial.

However, unlike our study, Frantz et al., (2012) found that a motorway had no influence on the genetic structure of a wild boar population in Belgium. In contrast to S.S. 131, that road has many subways and tunnels that can allow the passage of wildlife. Despite the lack of population sub-structure between the two sides of the road, the authors did not exclude that the road acts as a barrier, but that other factors (like a large N_e) could have masked it. Likewise, Vassant et al., (1993) argued that wild boars were not impacted by the motorway A5 in France; as marked individuals regularly crossed the road taking advantage of wildlife corridors. In our case the Sardinian motorway is almost devoid of corridors that facilitate movements of the local fauna for the entire stretch of about 200 kilometres.. Unfortunately, up to now nothing is known about spatial behaviour and habitat preferences of the Sardinian subspecies and we cannot exclude differences from the continental counterparts. Future projects should be addressed to combine population genetics and ecology to test the actual impact of land use and roads on the spatial patterns of *Sus scrofa meridionalis*.

Our results have direct consequences on the management of wild species in Sardinia. Given the motorway has seemingly an important role as barrier to gene flow in the wild boar population, it could as well represent a cause of fragmentation for other animals, promoting local genetic isolation. However, the effect on other species should be tested by targeted studies, as the same infrastructure might have different impact on different species (see for instance Frantz et al., 2012).

Finally, possible long-term detrimental effects (small population size, inbreeding, genetic drift) of habitat fragmentation should be carefully evaluated in the wild boar, in order to promote a sustainable management of its genetic resources.

Acknowledgement

The first and second authors PhD grant was financed by the Sardinia Region (European Social Fund, Operational Programme 2007/2013 – Axis IV - Human Capital Investment).

References

Autonomous Region of Sardinia (2012). Carta delle Vocazioni Faunistiche della Sardegna - sez. Ungulati (aggiornamento).

Balkenol N, Waits LP (2009). Molecular road ecology: Exploring the potential of genetics for investigating transportation impacts on wildlife. Molecular Ecology 18: 4151-4164.

Balkenhol N, Waits LP, Dezzani RJ (2009). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32: 818-830.

Belkhir K (2004). Page WEB de GENETIX - Version 4.05 pour Windows TM [Online]. Available at http://kimura.univmontp2.fr/genetix/

Boitani L, Mattei L, Nonis D, Corsi F (1994). Spatial and activity patterns of wild boar in Tuscany, Italy. Journal of Mammalogy 75: 600–612

Brassel KE, Reif D (1979). A procedure to generate Thiessen polygons. Geographical. Anaalysis. 325: 31–36.

Briedermann L (1990). Schwarzwild. 2nd edn. Neumann-Neudamm Melsungen.

Cahill S, Llimona F, Cabañeros L, Calomardo, F (2012). Characteristics of wild boar (*Sus scrofa*) habituation to urban areas in the Collserola Natural Park (Barcelona) and comparison with other locations. Animal Biodiversity Conservation 35: 221–233.

Castilho CS, Marins-Sá LG, Benedet RC, Freitas TO (2011). Landscape genetics of mountain lions (*Puma concolor*) in southern Brazil. Mammalian Biology 76: 476–483.

Coulon A, Guillot G, Cosson J-F, Angibault JMA, Aulagnier S, Cargnelutti B, Galan M, Hewison AJM (2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. Molecular Ecology 15: 1669–1679.

Dardaillon M, Bougnon G (1987). The influence of some environmental characteristics on the movement of wild boar *Sus scrofa*. Biology Behaviour 12: 82-92.

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistic Society B 39: 1–38.

Epps CW, Palsbøll PJ, Wehausen JD, Roderick GK, Ramey RR, McCullough DR (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters 8: 1029–1038

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software Structure: a simulation study. Molecular Ecology 14: 2611–2620.

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

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.

Fernandez-Llario P and Carranza J (2000). Reproductive performance of the wild boar in a Mediterranean ecosystem under drought conditions. Ethology, Ecology and Evolution 12: 335–343.

Focardi S, Gaillard JM, Ronchi F, Rossi S (2008). Survival of wild boars in a variable environment: unexpected life-history variation in an unusual ungulate. Journal of Mammalogy 89: 1113–1123.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? Journal of Applied Ecology 46: 493–505.

Frantz AC, Bertouille S, Eloy MC, Licoppe A, Chaumont F, Flamand MC (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (*Cervus elaphus*), but not wild boars (*Sus scrofa*). Molecular Ecology 21: 3445-3457.

Frantz AC, Zachos FE, Kirschning J, Cellina S, Bertouille S, Mamuris Z, Koutsogiannouli EA, Burke T (2013). Genetic evidence for introgression between domestic pigs and wild boars (*Sus scrofa*) in Belgium and Luxembourg: a comparative approach with multiple marker systems. Biological Journal of Linnean Society 110: 104–115.

Geisser H, Reyer H-U (2005). The influence of food and temperature on population density of wild boar *Sus scrofa* in the Thurgau (Switzerland). Journal of Zoology London 267: 89–96.

Goedbloed DJ, van Hooft P, Megens H-J, Langenbeck K, Lutz W, Crooijmans RPMA, van Wieren SE, Ydenberg RC, PrinsHHT (2013). Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar. BMC Genetics 14: 43

Guillot G (2009). On the inference of spatial structure from population genetics data. Bioinformatics 25: 1796-1801. Guillot G, Santos F, Estoup A (2008). Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics 24 (11): 1406-1407.

Guillot G, Santos F, Estoup A (2012). Population genetics analysis using R and the Geneland program [Documentation file]. Available at http://www2.imm.dtu.dk/~gigu/Geneland/

Guo SW, Thompson EA (2002). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361-372.

Hepenstrick D, Thiel D, Holderegger R, Gugerli F (2012). Genetic discontinuities in roe deer (*Capreolus capreolus*) coincide with fenced transportation infrastructure. Basic and Applied Ecology 13: 631–638.

Heymann, Y, Steenmans, Ch, Croisille, G, Bossard, M (1994). CORINE Land Cover. Technical Guide. Luxembourg (Office for Official Publications of European Communities)

Holderegger R, Di Giulio M (2010). The genetic effects of roads: A review of empirical evidence. Basic and Applied Ecology 11: 522–531

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

Keuling O, Stier N, Roth M (2008). Annual and seasonal space use of different age classes of female wild boar *Sus scrofa*. European Journal of Wildlife Research 54: 403–412.

Keuling O, Lauterbach K, Stier N, Roth M (2010). Hunter feedback of individually marked wild boar *Sus scrofa* L.: dispersal and efficiency of hunting in northeastern Germany. European Journal of Wildlife Research 56: 159-167.

Kuehn R, Hindenlang KE, Holzgang O, Senn J, Stoeckle B, Sperisen C (2007). Genetic Effect of Transportation Infrastructure on Roe Deer Populations (*Capreolus capreolus*). Journal of Heredity 98: 13–22.

Landguth EL, Cushman SA, Schwartz MK, McKelvey, Murphy M, Luikart G (2010). Quantifying the lag time to detect barriers in landscape genetics. Molecular Ecology 19: 4179–4191.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003). Landscape genetics: combining landscape ecology and population genetics. TREE 18: 189-197.

Manni F, Guérard E, Heyer E (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Human Biology 76: 173–190.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209–220.

Massei G, Genov PV, Staines B, Gorman ML (1997). Factors influencing home range and activity of wild boar (*Sus scrofa*) in a Mediterranean coastal area. J Zool 242: 411–423.

McRae BH (2006) Isolation by resistance. Evolution 60: 1551-1561.

Melis C, Szafrańska PA, Jędrzejewska B, Barton K (2006). Biogeographical variation in the population density of wild boar (Sus scrofa) in western Eurasia. J Biogeogr 33: 803–811.

Monmonier M (1973). Maximum-difference barriers: An alternative numerical regionalization method. Geographical Analalysis. 3:245–261

Muche L (2005). The Poisson-Voronoi tessellation: relationships for edges. Advance in Applied Probability 37: 279-296.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PG, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2011). Package 'vegan' version 2.0-2. Available at http://cran.r-project.org/web/packages/vegan/index.html

Osashi H, Saito M, Horie R, et al., (2013). Differences in the activity pattern of the wild boar *Sus scrofa* related to human disturbance. European Journal of Wildlife Research 59: 167-177.

Peakall R, Smouse PE (2005). GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295.

Pérez-Espona S, Pérez-Barbería FJ, McLeod JE, Jiggins CD, Gordon IJ, Pemberton JM (2008). Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). Molecular Ecology 17: 981–996

Podgórski T, Baś G, Jędrzejewska B, Sönnichsen L, Śniezko S, Jędrzejewski W, Okarma H (2013). Spatiotemporal behavioral plasticity of wild boar (*Sus scrofa*) under contrasting conditions of human pressure: primeval forest and metropolitan area. Journal of Mammalogy 94: 109–119.

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

Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86: 248-249.

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

Scandura M, Iacolina L, Crestanello B et al., (2008) Ancient vs. recent processes as factors shaping the genetic variation of the European wild boar: are the effects of the last glaciation still detectable? Molecular Ecology 17: 1745-1762.

Scandura M, Iacolina L, Crestanello B, Pecchioli E, Bertorelle G, Apollonio M (2009) Occurrence of a private mtDNA clade in Italian wild boars. In: Proceedings of the 7th International Symposium on Wild Boar (*Sus scrofa*) and on Sub-order Suiformes (eds by Náhlik A, Tari T.), pp. 96-97. Sopron, Hungary.

Scandura M, Iacolina L, Cossu A, Apollonio M (2011). Effects of human perturbation on the genetic make-up of an island population: the case of the Sardinian wild boar. Heredity 106: 1012-1020.

Schwartz MK, McKelvey KS (2009). Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conservation Genetics 10: 441-452.

Sharma S, Dutta T, Maldonado JE, Wood TC, Panwar HS, Seidensticker J (2013). Forest corridors maintain historical gene flow in a tiger metapopulation in the highlands of central India. Proceedings of the Royal Society B 280: 20131506.

Shepard D (1968). A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 ACM National Conference. pp. 517-524.

Smouse P, Long J, Sokal R (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology 35: 627-632.

Šprem N, Frantz AC, Cubric-Curik V, Safner T, Curik I (2013). Influence of habitat fragmentation on population structure of red deer in Croatia. Mammalian Biology 78: 290–295.

Storfer A, Murphy MA, Evans JS, et al., (2007). Putting the 'landscape' in landscape genetics. Heredity 98: 128-142.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010). Lanscape genetics: Where are we now? Molecular Genetics 19: 3496-3514.

Tang H, Peng J, Wang P, Risch NJ (2005) Estimation of individual admixture: analytical and study design considerations. Genetic Epidemiology 28: 289–301.

Truvè J, Lemel J (2003). Timing and distance of natal dispersal for wild boar Sus scrofa in Sweden. Wildlife Biology 9: 51-57.

Vandergast AG, Perry WM, Lugo RV, Hathaway SA (2011). Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. Molecular Ecology Res 11: 158-161.

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

Vassant J, Brandt S, Jullien JM (1993). Influence du passage de l'autoroute A5 sur les populations cerf et sanglier du Massif d'Arc-en-Barrois. 2^{e} et dernière partie: cas de la population sanglier. Bulletin Mensuel de l'Office National de la Chasse 184: 24–33.

Weckworth BV, Musiani M, DeCesare NJ, McDevitt AD, Hebblewhite M, Mariani S (2013). Preferred habitat and effective population size drive landscape genetic patterns in an endangered species. Proceedings of the Royal Society B 280: 20131756.

Wu B, Liu N, Zhao H (2006). PSMIX: an R package for population structure inference via maximum likelihood method. BMC Bioinformatics 7: 317.

Zhu L, Zhan X, Meng T, Zhang S, Wei F (2010). Landscape features influence gene flow as measured by cost-distance and genetic analyses: A case study for giant pandas in the Daxiangling and Xiaoxiangling Mountains. BMC Genetics 11: 72.

Table 1. Description of land cover categories used to describe the Sardinia environment. Each category was associated to a specific weight, assigned on the basis of the permeability to the species and the use of the habitat by wild boar (feeding, refuge).

Land use category	Description	Categories weight
Anthropic areas	Urbanized and industrialized areas	100
Wetland	Rivers, water basins, swamps	80
Open areas	Nude land, rocks	60
Cultivated areas	Arable land, fruit and olive trees, vineyards	30
Pastures	Meadows dedicated to livestock	20
Shrubs	Unused land with low and bushy vegetation	10
Maquis	Typical mediterranean maquis	2
Forest	Deciduous, coniferous and mixed forests	0

Table 2. Results of the genetic assignment of Sardinian wild boars to the three subpopulations on the basis of different statistical methods. Stru1- 'blind' Bayesian assignment in Structure (no population information), Stru2 – 'informed' Bayesian assignment in Structure (geographical information), GL – Bayesian assignment in Geneland (geographical information), PM – maximum likelihood assignment in PSMix (no population information). For Structure the area of assignment is that corresponding to the cluster receiving the highest probability.

		Area of assignment											
		ES				**NSW**				**SWS**			
		Stru 1	Stru 2	GL	PM	Stru 1	Stru 2	GL	PM	Stru 1	Stru 2	GL	PM
Sampling area	**ES**	**94%**	**99%**	**100 %**	**89 %**	5%	1%	0%	9%	1%	0%	0%	2%
	NWS	4%	2%	4%	2%	**96%**	**98%**	**96 %**	**98 %**	0%	0%	0%	0%
	SWS	5%	0%	0%	2%	3%	0%	0%	3%	**92%**	**100 %**	**100 %**	**95 %**

Table 3. Results of partial Mantel tests for correlation between genetic distance and different spatial distances in the Sardinian wild boar population. Geo = Euclidean distance matrix, Bar = matrix of presence/absence (1/0) of physical barrier, also accounting for the number of barrier crossings (2/1/0), LeastC = least-cost distance matrix, which incorporates different costs for each land use category.

Partial Mantel Test						
			Barrier 0/1		Barrier 0/1/2	
Matrix 1	Matrix 2	Controlled variable	*R*	*P*	*R*	*Pval*
Gen	Geo	Bar	0.081	0.001*	0.076	0.001*
Gen	Geo	LeastC	0.075	0.001*		
Gen	Bar	Geo	0.361	0.001*	0.331	0.001*
Gen	Bar	LeastC	0.193	0.001*	0.145	0.001*
Gen	LeastC	Bar	0.123	0.001*	0.143	0.001*
Gen	LeastC	Geo	0.331	0.001*		

Legends to figures

Figure 1. Maps of Sardinia showing information used in landscape genetic analysis. (a) Map showing the composition of land use categories, the distribution of sample locations and the position of the main roads. (b) Map showing cell resistance weights used to calculate least cost paths among sampling sites (in red).

Figure 1

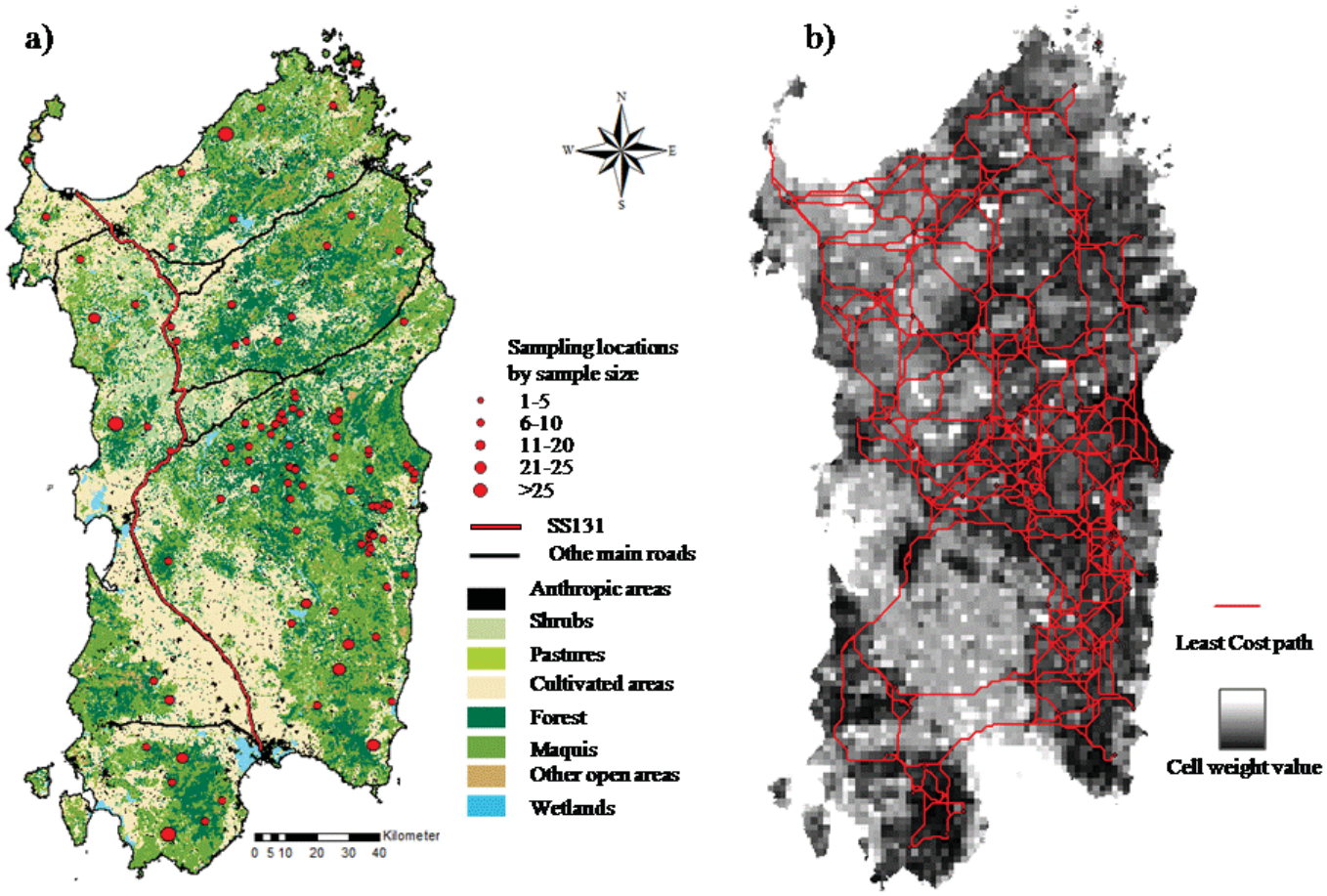


Figure 2. Assignment proportions obtained by STRUCTURE for each of the 368 Sardinian wild boar microsatellite genotypes. Individuals are represented by thin vertical lines, showing the membership (Q) to the clusters inferred by the program (colored segments). Membership to clusters I and III (in yellow), both exclusive to the Sardinian population, were pooled. Only individuals that were univocally assigned to the Sardinian component ($Q_{I+III} \geq 0.9$, i.e. left to the solid black line) were identified as non-introgressed members of the Sardinian population and used for the inference of population structure (n = 277).

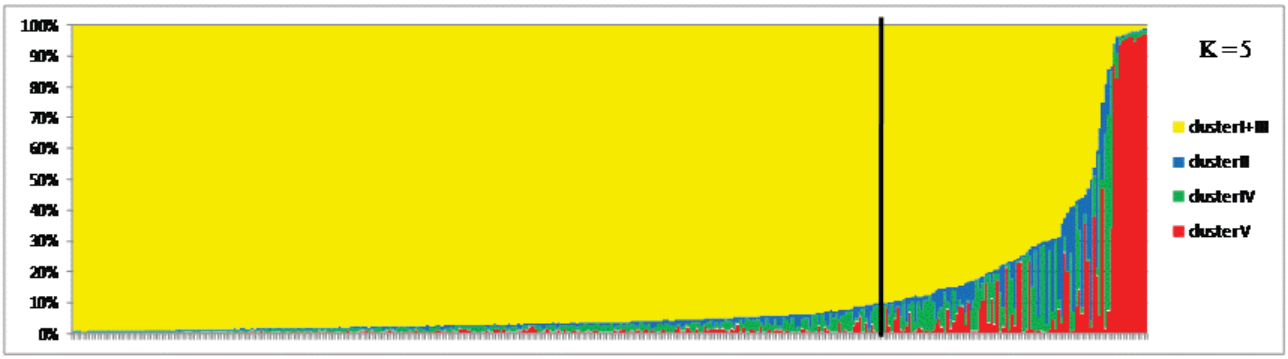


Figure 3. Genetic structure of the Sardinian wild boar population inferred by STRUCTURE. (a) Log-likelihood values for the different K values in the simulation and corresponding outcome of the Evanno's method. Three subpopulations were identified: north-west (NWS), south-west (SWS) and eastern Sardinia (ES). (b) Graphical representation of the membership of wild boars to each of the three subpopulations through the Inverse Distance Weight (IDW) interpolation method.

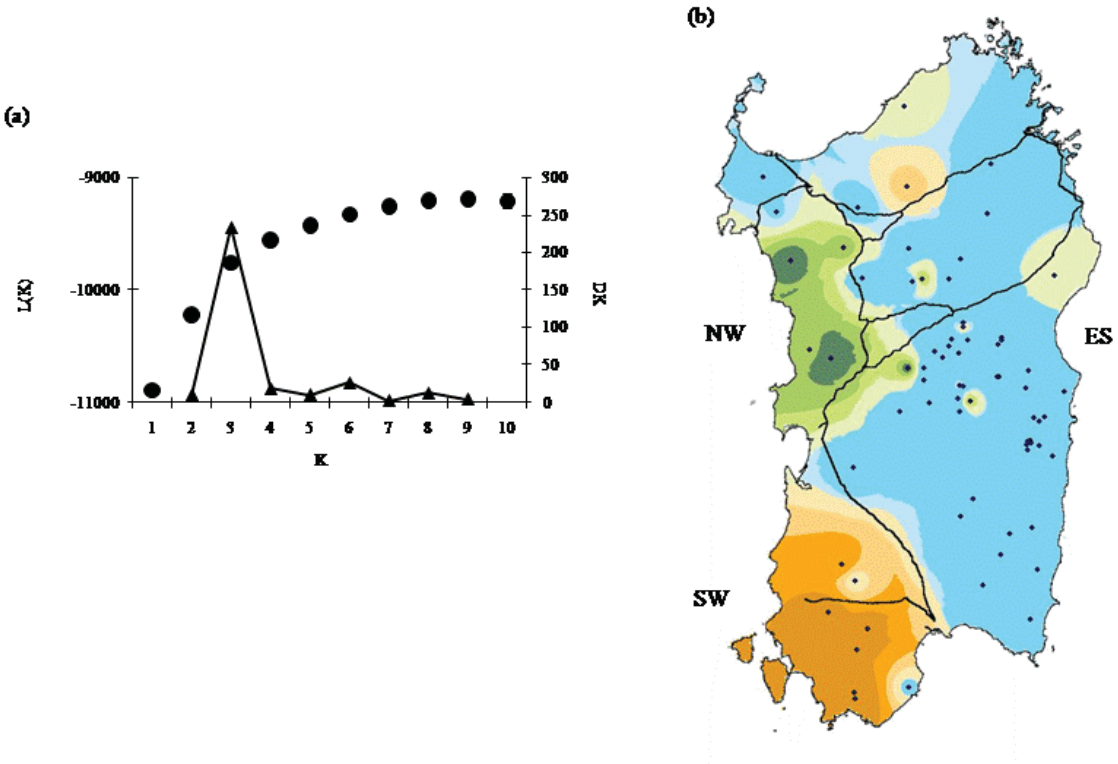


Figure 4. Correlation between genetic and geographic distance in the Sardinian wild boar population, as resulting from Mantel test in Genalex. Putative hybrids are excluded (purged dataset, n = 277). In the upper plot (a) geographic distances are the Euclidean distances between sampling sites; in the lower plot (b) they are represented by the least-cost distance.

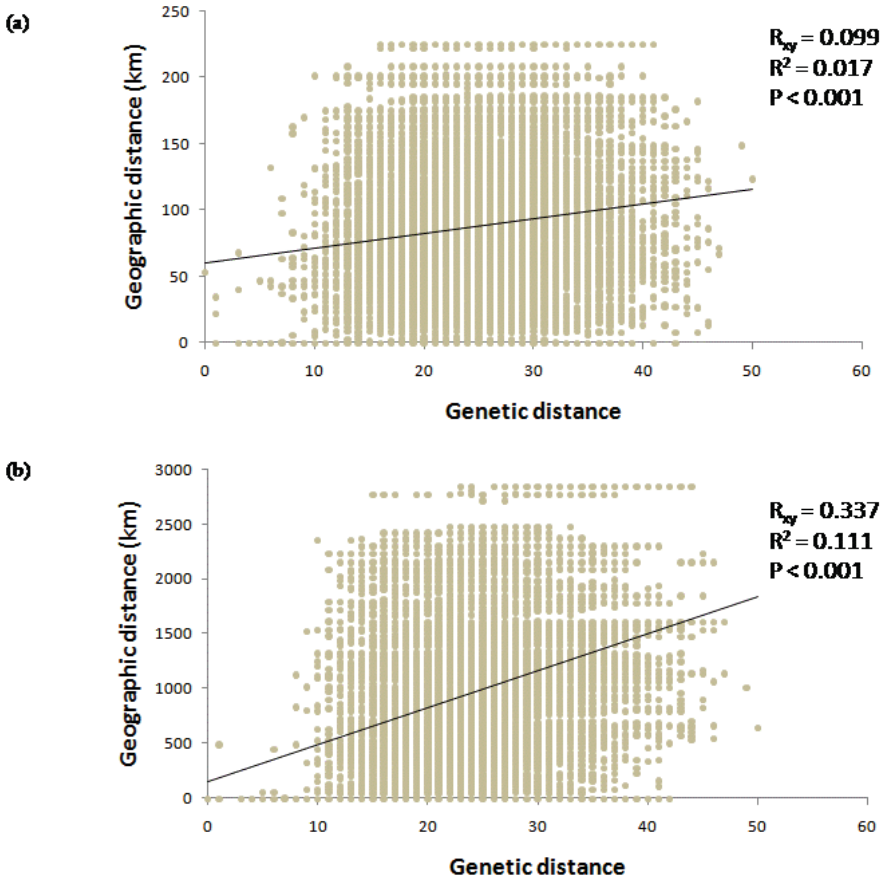
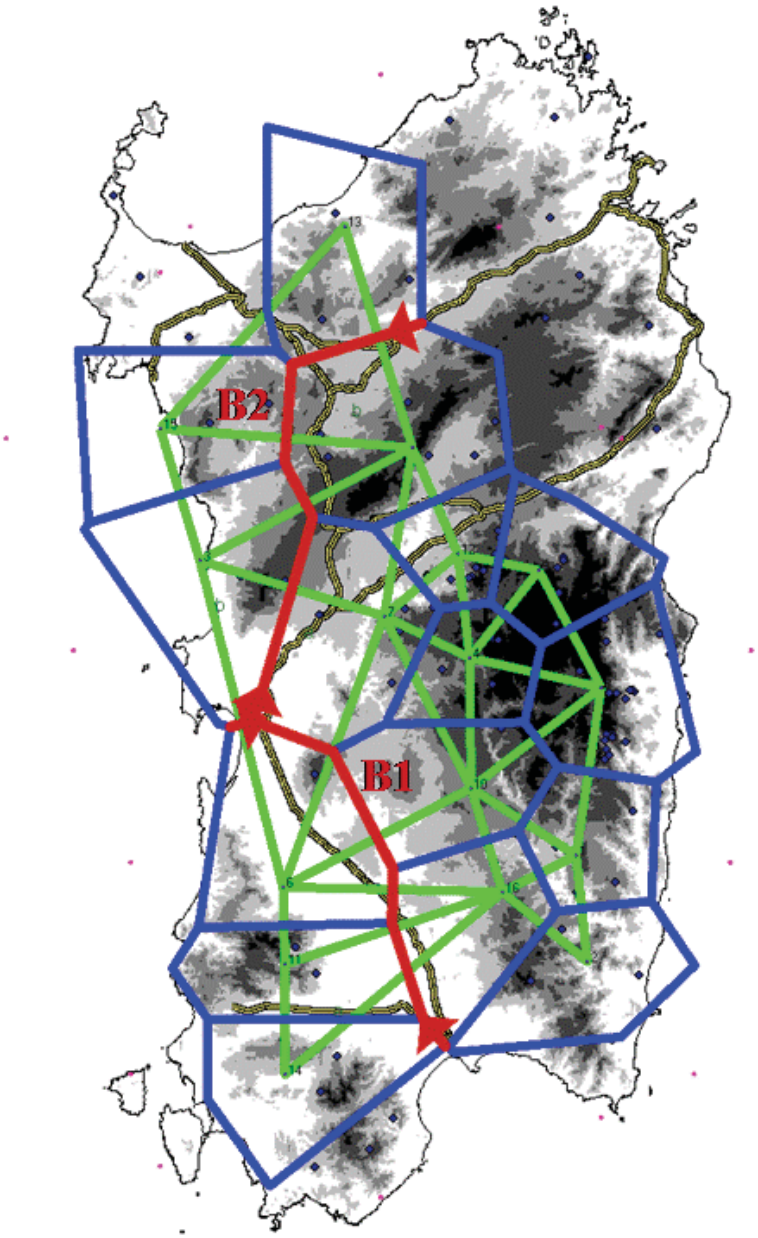


Figure 5. Barriers to gene flow identified by the Monmonier's algorithm in the Sardinian wild boar population. Major barriers are represented by red lines (B1 and B2). Green and blue polygons represent respectively the Delaunay triangulation linking sampling sites and the Voronoi tessellation used by the program BARRIER.



Supplementary Materials

Table S1. List of samples of continental wild boar and domestic pig used as reference for assignment tests.

Reference population	Country/breed	Nr of individuals
Italian WB	Central-Southern Italy	75
European WB	Spain	15
	France	15
	Luxembourg	10
	Austria	13
	North-East Italy	19
	Poland	43
	Belarus	24
Domestic pigs	Commercial breeds	16
	Sardinian free-ranging	98

Figure S1. Assignment proportions to the five clusters (I-V) identified by Structure for the four populations in the global dataset (n=698), namely: Italian wild boar (WB-Ita), European wild boar (WB), domestic pigs (DP) and Sardinian wild boar (WB-Sar).

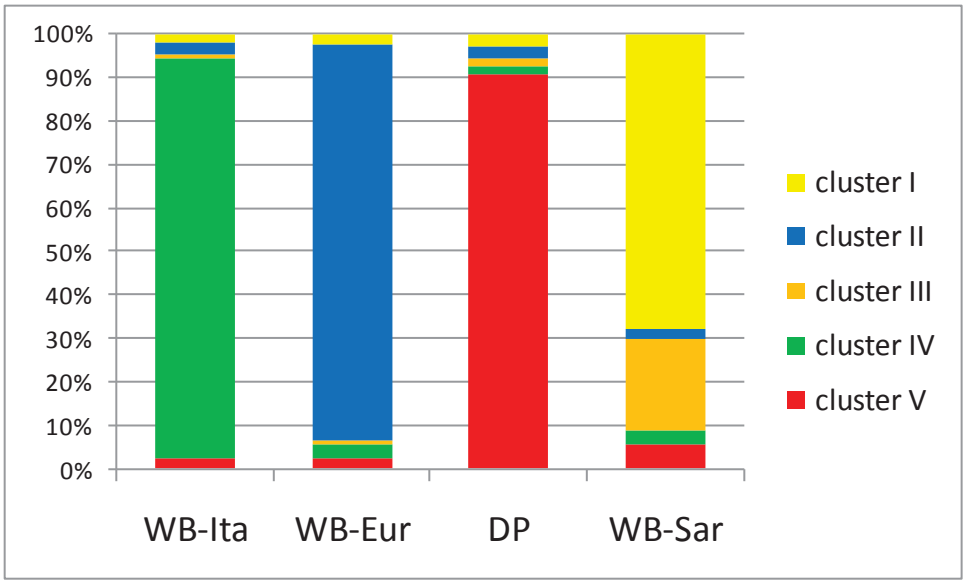
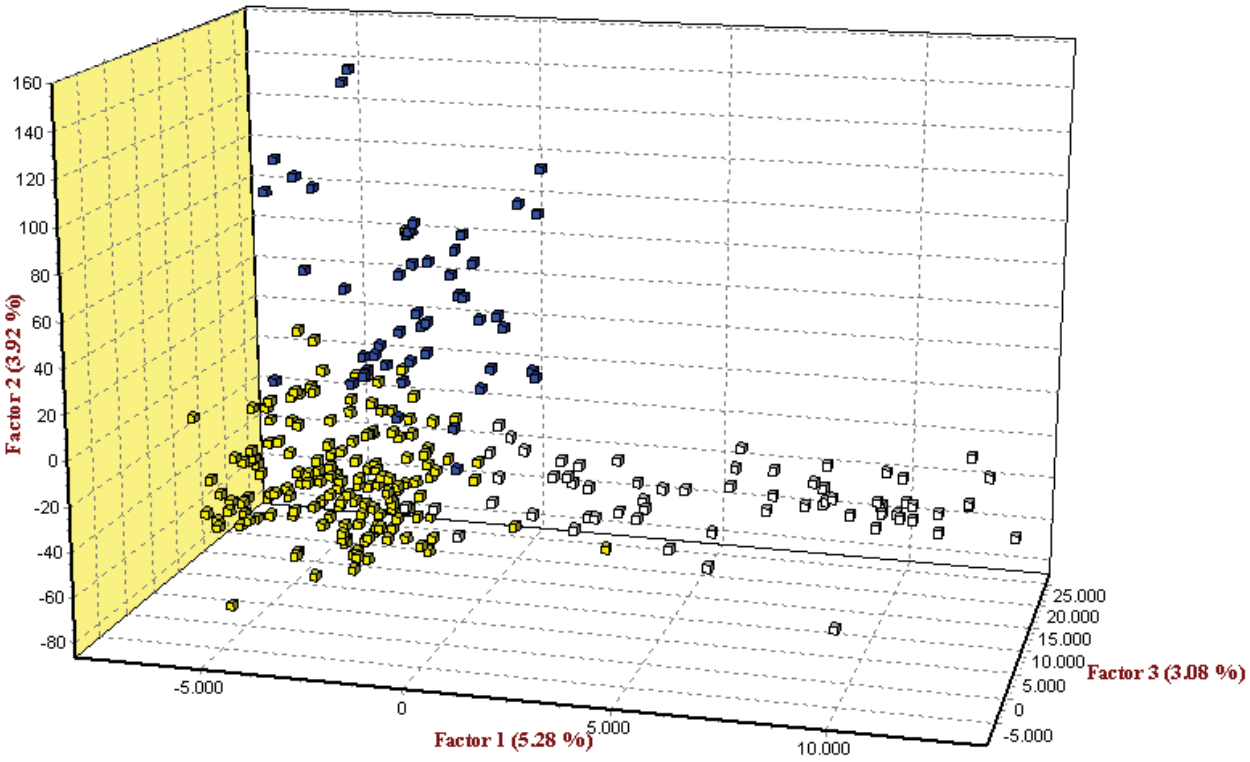


Figure S2. Factorial Correspondence Analysis (FCA) plot of multilocus genotypes belonging to the three subpopulations (ES – yellow, NWS – blue, SWS – white).



Chapter 3

Genetic competition between native and non-native lineages: a genetic survey of the endemic roe deer in three areas of central Italy

Daniela Biosa, Massimo Scandura, Siriano Luccarini, Francesco Nonnis Marzano, James Tagliavini, Marco Apollonio

Manuscript

Foto: Paolo Bongi

Genetic competition between native and non-native lineages: a genetic survey of the endemic roe deer in three areas of central Italy

Daniela Biosa[1], Massimo Scandura[1], Siriano Luccarini[1], Francesco Nonnis Marzano[2], James Tagliavini[2], Marco Apollonio[1]

[1] Dipartimento di Scienze della Natura e del Territorio, Università di Sassari, via Muroni 25, 07100 Sassari, Italy

[2] Dipartimento di Biologia Evolutiva e Funzionale, Università di Parma, Via G.P. Usberti 23A, 43124 Parma, Italy

Abstract

Reintroduction of game species and restocking of depressed populations have interested many areas of Europe during the last century. In some cases, such operations have threatened the endemic diversity of remnant populations, creating a general concern for their conservation. The Italian peninsula hosts relict populations of an endemic roe deer (*Capreolus capreolus italicus*), which are studied since the early 2000s. Growing evidences testified its occurrence in a larger range than previously assumed, but the genetic status of these populations is worrisome, because of emerging signals of admixture with human-released exotic stocks.

We present a study on three areas in central Italy (provinces of Arezzo, Pisa and Parma), where a contact between indigenous and reintroduced roe deer was suspected on the basis of available historical data. Mitochondrial (control region) and nuclear (11 microsatellites) data were gathered for 304 individuals, partitioned within each study area among three subareas: a source of *italicus* (A), a source of non-*italicus* (C) and an intermediate contact zone (B). Results revealed that the native genetic component was distributed along a decreasing gradient moving from A to C in all three areas and coherently between mitochondrial and nuclear markers. High frequency of Italian haplotypes was found in the lower province of Pisa, but the most effective diffusion of native haplotypes was observed in Parma, apparently due to a high forest cover and connectivity. Where an outstanding landscape barrier was present (the Arno river in Arezzo, B), it seemingly limited the dispersal of *C. c. italicus* northwards generating a detectable genetic discontinuity.

Our results reveal the presence of a patchy genetic composition at the northern border of the Italian subspecies and a varying pattern of genetic admixture. The appreciable genetic differentiation between the extremes of our sampling areas, after many generations from human releases can suggest that a certain resistance to genetic homogenization does take place. Further investigations are recommended to elucidate which conditions may favor the resilience of this endemic taxon.

Keywords: *Capreolus capreolus italicus*, genetic admixture, introgression, mitochondrial DNA, microsatellites, landscape genetics, conservation

Introduction

In the last century, reintroduction and restocking activities, carried out relocating animals for hunting or conservation purposes, have affected several species in Europe, in particular game species (Randi 2005, Barbanera et al. 2010, Linnel and Zachos 2011; Champagnon et al. 2012a). Such anthropogenic modifications have led to a considerable increase in the encounter rate between forms evolved in different geographical and ecological contexts. The consequence of such contacts is often the introgressive hybridization between a native and an exotic stock, with loss of local variants (Allendorf et al. 2001) and an overall genetic homogenization among different populations (Olden et al. 2004), both causing a huge impact on the global biodiversity. Furthermore, the release of exotic conspecifics could generate changes in morphological and behavioural traits of the native form, which can cause meaningful ecological problems, such as an increased predation impact on endemic species, competition with native animals, transmission of new diseases (Leighton 2002; Ryan et al. 2009, Consuegra et al. 2011, Champagnon et al. 2010 and 2012b). Human-mediated hybridization can negatively impact the genetic variability of species causing taxa extinctions, genetic mixing and flattening genetic peculiarities among populations (Rhymer & Simberloff 1996, Olden et al. 2004), but it can also confer fitness advantages to introgressed individuals (e.g. 'hybrid vigour' due to heterosis, Keller & Taylor 2010), thus favouring the recovery of declining populations. Further, in some cases the protection of these hybrids may represent the only way to preserve the genetic characteristics of one or both parental populations (Allendorf et al. 2004).

The effects of translocations followed by intraspecific hybridization have been documented in several game species in Europe, such as wild boar *Sus scrofa* (Goedbloed et al. 2013), red deer *Cervus elaphus* (Frantz et al. 2006, Haanes et al. 2013), brown hare *Lepus europaeus* (Mamuris et al. 2001), and some birds (common quail *Coturnix coturnix* Chazara et al. 2010; mallard *Anas platyrhynchos* Champagnon et al. 2013). These studies have stressed the importance of prioritizing the knowledge and maintenance of genetic diversity, investigating and monitoring the effects of human-induced intraspecific hybridization and implementing adequate management policies to limit its detrimental impact on native populations (see also Laikre et al. 2010).

The roe deer (*Capreolus capreolus*) is the most widespread ungulate in Europe and is strongly managed throughout its distribution range (Apollonio et al 2010). As other game

species, at the beginning of the last century it had gone extinct in several regions of Europe, but subsequently it recovered almost everywhere (Apollonio et al 2010). Similarly to the rest of Europe, in Italy the disappearance of the roe deer began in the XVI century with range and population sizes that progressively declined, reaching a critical point in the XIX century, when the species crashed in central-southern Italy and Sicily, where it died out. After the II World War, the roe deer was present only in the Eastern Alps, in a region of Central Italy (Maremma), and in a few small relict areas in the Castelporziano estate (near Rome), and in Gargano and Orsomarso massifs in the south of the Italian peninsula (Perco & Calò 1995). Such central-southern populations were considered as remnants of the endemic Italian subspecies (*C. c. italicus*; Festa 1925), characterized by both distinct morphometric (Montanaro et al. 2003) and genetic peculiarities at both mitochondrial and nuclear markers (Lorenzini et al. 2002; Vernesi et al. 2002; Randi et al. 2004; Gentile et al. 2009; Mucci et al. 2012). In the last decades, the relict Italian subspecies was rediscovered also in southern Tuscany (Lorenzini et al. 2002), and genetic signals of its presence were found in a wider portion of central and northern Apennines (Randi et al. 2004). Nonetheless, genetic studies also revealed a mixture in these area between individuals bearing *italicus* haplotypes and others bearing different European haplotypes. Similarly, Gentile et al. (2009) detected the presence in the Orsomarso population of individuals geneticallyrelated to those found in some populations of northern and central Apennines and not classified as *C. c. italicus*.

In fact, due to the disappearance of roe deer in most of the Italian peninsula, reintroduction and restocking operations were necessary to restore the original range and sustain the small remnant populations. Reintroductions were carried out mainly in the western Alps and in the Apennines, and involved stocks from the eastern Italian Alps, central Europe and the Balkans (Lorenzini et al. , 1993, 1997; Randi et. al 1998; Lorenzini et al. 2002; Vernesi et al. 2002). Such practices have unavoidably led to hybridization between the *C. c. capreolus* and *C. c. italicus* subspecies, especially in central Italy (Randi et al. 2004; Mucci et al 2012). The consequent introgression of exotic genes into the native gene pool of the Italian roe deer have raised concern for the genetic integrity of this relict subspecies and further studies focused on patterns of current admixture have been recommended (Gentile et al. 2009).

Besides the abovementioned morphological and genetic differences it was argued that the two forms have different ecological requirements, due to the different environmental

contexts where they historically occurred and to which they have supposedly adapted. In fact, although the two forms exhibit similar food habits and similar social and spatial behaviour, a preference by native roe deer for deciduous woodland and scrubland-maquis, typical of the Italian peninsula, was noticed (Focardi et al. 2009). Such ecological difference could have influenced the race between the competing gene pools in the contact zones, ultimately affecting the success of reintroductions and of the following recolonization of the area.

In the present study, we considered three areas of Italy, in which we could identify: *i*) an area of expansion of the native *C. c. italicus*, *ii*) an area where exotic *C. c. capreolus* had been reintroduced over the last 50 years, and *iii*) an intermediate contact zone between the two subspecies. The intermediate area was considered as an arena where the mutual introgression of the two gene pools was evaluated and put into relation with the environmental features and the reintroduction history, in terms of time elapsed since the reintroduction event and number of released individuals Both biparentally- (microsatellites) and uniparentally-inherited (mitochondrial DNA, mtDNA) genetic markers were used.

Materials and methods

Sampling design and study areas

In order to identify the genetic structure of each roe deer population and patterns of hybridization between native and exotic forms, we considered three study areas in central Italy (Fig. 1) and adopted the following sampling scheme within each of them: a sub-area of natural spread of the italic roe deer (A), a sub-area of European roe deer reintroduction (C), and an intermediate contact zone between the two roe deer forms (B). For each area we aimed at collecting 25 samples for sub-area A, 50 for B and 25 for C. A total of 284 tissue samples of hunted roe deer were finally collected from the province of Arezzo (AR, n = 103), the province of Pisa (PI, n = 93) and the province of Parma and neighbouring territories (PR, n = 88). Further 20 samples from the northern part of the province of Grosseto (GR) were used as control group, having been sampled in a ‘core’ *italicus* area (Mucci et al. 2012). For each area, we collected documented information about events of restocking and translocation, comprehending approximate time and site of releases, as well as number and origin of released individuals (Tab. 1).

Area AR: historical records for this area indicates that roe deer never disappeared in Foreste Casentinesi (sub-area AR-C) during the XIX-early XX century, but its presence after the II World War was so small to justify restocking operations (Mazzarone et al. 2000). Already in 1933 and 1955 two resocking with very limited number of roe deer of Alpine origin were performed (Crudele 1988) A further release interested in 1959 the Forest of Vallombrosa (north-east border of sub-area C), where a group of roe deer were translocated from Foreste Casentinesi in order to recolonize the Pratomagno massif (Mattioli et al 1995). In the rest of the province the species was absent for a long time (no area was mentioned by Perco 1981) and reappeared in the 1980s, thanks to the natural expansion from the close province of Siena to the hilly areas at south-west (sub-area A) and from contemporary expansion of the northern restocked population (Mattioli L. pers. comm.). Delimitation of the three sub-areas followed previous genetic data, showing the prevalence of European haplotypes in the mountains and valleys in the north of the province (Casentino) and the occurrence of *italicus* populations in neighbouring areas of the province of Siena (Mucci et al. 2012). The intermediate area includes Valdarno and Val di Chiana, two alluvial valleys, considered as possible contact zones for the two gene pools (sub-area Sr-B). This subarea is crossed by the Arno river, which may act as a barrier to gene flow.

Area PI: roe deer has recolonized this area over the last decades through the combined effect of a natural recovery of native populations in the south (in continuity with Maremma) and of the spread of exotic stocks reintroduced into a private estate (Miemo) in the central part of the province. This reintroduction occurred in the mid 1960s and involved the release of around 20 heads coming from ex Jugoslavia (Baldacci M. pers. comm.). For our purposes, and according to results presented by Mucci and colleagues (2012), we considered the Castelnuovo Val di Cecina hunting district in the south, where roe deer was hunted and present in the, ‘60s and the ‘70s, as an *italicus* diffusion area (sub-area PI-A), and the Miemo hunting estate in the north of the province as a putative source of exotic genes (sub-area PI-C). The upper Val di Cecina, in the middle, was interpreted as the zone where the two gene pools came into contact (sub-area PI-B).

Area PR: in this portion of Apennines roe deer was almost extinct at the end of the XIX century. In Boschi di Carrega, a private estate then converted to a regional park, the local population died out at the end of the XIX century, and was restored by introducing 5-6 roe deer from Balkans in the early XX century (Parco Regionale Boschi di Carrega, pers. comm.). Reintroductions of doubtful success occurred in the '30s in the Alto-Parmense (Mattioli 1994), but most meaningful reintroductions were realized between 1950 and 1970 by the Corpo Forestale dello Stato, with the release of 28 roe deer of underreported origin in Alta Garfagnana (sub-area PR-A). In this area additional information report on reintroductions occurring between 1967 and 1972 with 16 individuals coming from Capalbio (southern province of Grosseto) and 27 roe deer from breeding stations in the Eastern Alps,. The former showed better adaptation and were used for further restocking of local reserves whereas the latter died out (Corpo Forestale dello Stato, pers. comm.). Hills and mountains comprised between Garfagnana (sub-area PR-A) and the Boschi di Carrega Regional Park (sub-area PR-C) were interpreted as contact zones between two divergent gene pools (sub-area PR-B).

Environmental features of the nine subareas were obtained in ArcGIS v. 10 (ESRI Italy) on the basis of Digital Terrain Model (Istituto Geografico Militare, Italy) and IV Level Corine Landcover at scale 1:50 000 (Heymann et al. 1994) and are shown in Table 1.

Mitochondrial DNA and microsatellite genotyping

Tissue samples were stored in absolute ethanol or frozen until DNA extraction, obtained using Genelute kit (Sigma-Aldrich, St Louis, MO, USA).

For 252 samples the sequence of the mtDNA control region was obtained using primers developed by Randi et al. (1998) LcapPro (5'CGT CAG TCT CAC CAT CAA CCC CCA AAG C-3') and HcapPhe (5'-GGG AGA CTC ATC TAG GCA TTT TCA GTG-3'). Reactions were performed with the following amplification conditions: 35 cycles of 92 °C for 1 min, 62 °C for 1 min and 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min. PCR products were purified by Exo/SAP digestion and the fragment of interest was directly sequenced using the forward primer LcapPro and the BigDye Terminator kit version 3.1 (Applied Biosystems). Fragments were finally purified in columns loaded with Sephadex G-50 and run in an ABI PRISM 3130xl Avant automatic sequencer (Applied Biosystems). The size of the target region was also selected in

order to maximize the possible alignment that could be created with published sequences available in GenBank.

All samples were genotyped with a panel of 11 polymorphic autosomal microsatellites: Roe1, Roe6, Roe8, Roe9 (Fickel & Reinsch 2000), NVRTH16, NVRTH21, NVRTH24 (Roed & Midthjell 1998), ILSTS011 (Kemp et al 1995), OarFCB304 (Talbot et al 1996), BMC1009 (Kappes et al 1997) and RT1 (Wilson et al 1997). They were amplified by three multiplexed (multiplex A: Roe1, Roe8, Roe9; multiplex B: RT1, NVRTH21, BMC1009; multiplex C: Roe6, NVRTH16, ILSTS011) and two single PCRs (NVRTH24 and OarFCB304). PCR conditions are available upon request. Fluorescently-labeled PCR products were sized by BMR Genomics (Padua, Italy) using capillary electrophoresis on a ABI PRISM automatic sequencer (Applied Biosystems).

Statistical analysis

To check and editing the obtained mtDNA sequences, electropherograms were imported into FinchTV 1.4.0 (Geospiza Inc.). Subsequently, a multiple-sequence alignment was created using the function ClustalW in Mega 5.2.2 (Kumar et al. 2008) and adding *Capreolus capreolus* control region sequences downloaded by GenBank. For phylogenetic purposes a sequence of *C. pygargus* was included as outgroup (GenBank accession code JQ906130,). Hereafter, the aligned sequences were cut to obtain a fragment matching that used by Randi et al. 2004 and Mucci et al. 2012, having a total length of 704 bp. In total we obtained an alignment of 443 sequences, 191 downloaded from GeneBank and 252 obtained from this study. Unique haplotypes were identified with Collapse 1.2 (D. Posada).

In order to identify native haplotypes (i.e., *italicus*, hereafter IT), we followed Randi et al. (2004) and Mucci et al. (2012), who considered as diagnostic a nucleotide deletion at position 103 in the alignment, replaced in all other European haplotypes by a thymine or cytosine. On the basis of this discrimination we calculated the frequency of *italicus* and non-*italicus* (hereafter EU) haplotypes in the three sub-areas of each province. Thereby, we could evaluate their distribution within each area.

A median-joining (MJ) network of haplotypes (Bandelt et al., 1999) was built in NETWORK 4.6 (Fluxus Technologies Ltd.

Microsatellite multilocus genotypes were used in Genetix 4.05 (Belkhir et al. 2001) to calculate observed (H_O) and expected heterozygosity (H_E), mean number of alleles per locus (k) and the allele frequencies in the sample. F_{IS} was also calculated and its significance tested by permutations. Such analyses were performed for the complete dataset of each area and taking into account the subdivision into three sub-areas (A, B and C). MicroChecker 2.2.3 (Van Oosterhout et al 2004) was used to check for possible genotyping errors, allelic dropout, or null alleles at the 11 loci, while deviations from the Hardy-Weinberg equilibrium (HWE) and from linkage equilibrium (LE) were assessed by Genepop 4 (Raymond & Rousset 1995). Tests for HWE employed the Markov chain method proposed by Guo and Thompson (1992), with the following chain parameters: 10000 dememorizations, 100 batches and 20000 iterations. Deviations from LE were tested for each pair of loci. Significance levels were lowered, accounting for the number of multiple tests by the sequential Bonferroni procedure (Rice, 1989).

Genetic diversity across samples was assessed by the Factorial Correspondence Analysis (FCA) in Genetix. The procedure provides a method to check for similarities among samples by making no a priori assumption of group membership. We performed the FCA in the three areas together, also including the reference sample from Grosseto, and within each area to verify genetic differences among the three sub-areas.

The fine genetic structure within each population was inferred using the individual-based approach implemented in STRUCTURE v 2.3 (Pritchard et al 2000). The algorithm uses Bayesian model-based clustering to attribute each individual to the most likely suorce population, thus being able to identify possible immigrants and gene introgression.

The software STRUCTURE calculates the individuals' membership proportion (Q) to each of the inferred clusters. STRUCTURE was performed by 10 independent Monte Carlo Markov Chain simulations having a number of putative populations (K) ranging between 2 and 15, using the following settings: admixture model, no population information, correlated allele frequencies, 200 000 burn-in and 200 000 iterations of data collection. The analysis was run adding the phenotypic information relative to the mtDNA haplotype beared by the animal (0 = non-*italicus* and 1 = *italicus*). This information is not used by STRUCTURE in the computation but is used as descriptive attribute to calculate cumulative Q proportion for each phenotypic class.

Furthermore, to verify the actual presence of two different gene pools within the same area, we run STRUCTURE with data of single provinces with the same settings used in the previous analysis, but using a fixed number of clusters, K = 2. Under the admixture model, the algorithm calculates the individuals' membership proportion (Q) to each of the two cluster. Individuals having more than 75% of membership ($Q \geq 0.75$) to one cluster were assigned to the corresponding population. Furthermore, we calculated the cumulative proportion of membership of each pre-defined population as mean Q. Additionally, to establish the degree of admixture between the two genetic clusters we also evaluated the number of individuals with $Q \geq 0.75$ for either cluster in the different sub-areas. In this way, we were able to identify the proportion of individuals that presented intermediate genotypes and evaluate the rate of hybridization between the two roe deer subspecies.

Results

Mitochondrial DNA

Thirty five different mtDNA haplotypes (704bp) were found in our sample, 23 of which did not match any published haplotype. Their alignment with the 166 haplotypes used in Mucci et al. (2012) allowed us to discriminate those that could be attributed to the *italicus* subspecies, thanks to the indel at position 103. Accordingly, we identified 14 IT haplotypes (11 novel) carried by 115 individuals (46 %) in the total sample, whereas 137 roe deer were bearing one of the 21 different sequence (hereafter EU, 12 novel) belonging to the Central and East clades (18 and 3 haplotypes, respectively; Tab. 2). The MJ network confirmed the partition into three major clades (Central, West and East), but barely resolved divergence within the Central clade (Fig. S1 in Supplementary Materials). In any case, the group of IT sequences clustered into a single sub-clade, diverging from the main clade by the indel at position 103, but also by a mutation at position 204.

The frequency of IT haplotypes decreased moving from A to C in all three study areas, reflecting the occurrence of two different gene pools at the extreme sub-areas and thus confirming historical information (Tab. 1). The area with the highest frequency of *italicus* mtDNA was PR (54 %) followed by PI (51 %) and AR (20 %). The AR area hosted a very low native diversity with only two different IT haplotypes, despite the 7 of PI and 8 of PR. The highest percentage of EU haplotypes resulted in Casentino (AR-C, 94 %). As concerns contact

zones, almost half of individuals showed a IT-mtDNA in PR-B (50%) and PI-B (41%), while only one fourth in AR-B (24%). As expected the GR population hosted prevalently native haplotypes, with a frequency of IT-mtDNA equal to 89%.

The most common IT haplotypes was H15 (n = 76), previously found in Central Italy (Siena, Grosseto, Massa; Randi et al. 2004, Gentile et al. 2009) and widespread in our three study areas and in GR. Similarly, H43 (n = 15) was detected in PI and GR, while it had been recorded in Siena (Gentile et al. 2009). Noticeably, none of the sampled individuals showed haplotypes characterizing the isolated populations of Castelporziano, Gargano and Orsomarso, historically identified as relict areas of pure Italian roe deer populations (Festa 1925). EU haplotypes were mostly represented by H13 (n = 55) and H18 (n = 29), previously detected in populations of the Italian Alps and Apennines and belonging to the Central European clade (Randi et al, 2004).

All detected EU haplotypes belonged to the Central clade, apart three sequences falling in the East clade (H16, H23 and H193) and detected in PR and in PI-C. While the former is widespread in Europe, the East clade characterizes roe deer populations inhabiting the Balkans, but it occurs also in East and Central Europe, and in northern Italy (Randi et al. 2004). East-clade haplotypes as well as two EU-Central-clade sequences (H41 and H48) can be interpreted as non-indigenous mitochondrial sequences.

Autosomal microsatellites

The proportion of missing data in the complete STR dataset was 4.6%. No evidence of genotyping errors due to dropout or null alleles was detected by MicroChecker. A total of 108 different alleles (5-14 alleles per locus, k = 9.82) were found in the overall population. Average observed heterozygosity (H_O = 0.597) was lower than expected heterozygosity (H_E = 0.701), suggesting an excess of homozygotes that was confirmed by a significant deviation from HWE ($P < 0.01$ for all loci but Roe1 and BMC1009). Significant deviations from HWE were also found the three study areas separately. Pairwise Fst values calculated among areas, including the reference GR population, revealed a high level of differentiation for all pairs (range Fst = 0.086 - 0.144) except PI-GR (Fst = 0.031), which are geographically very close. Pairwise Fst calculated within areas showed values comprised between 0.002 (AR-A – AR-B) and 0.105 (PI-A – PI-C), highlighting a general trend of major divergence between A and C in every area and confirming

the intermediate diversity of subareas B (Tab 3). Such pattern of genetic differentiation was confirmed also by FCA and the Bayesian cluster analysis.

The FCA including all genotypes produced three clouds of points that matched the three study areas. GR individuals clustered with the neighbouring PI population (Fig. 2), and the closer genotypes of the other two populations were represented by AR-A and PR-A respectively. This position suggested that native individuals occupied the central part of the graph.

Bayesian cluster analysis on the global dataset, provided a clear genetic differentiation among the three provinces at K = 3 (Fig. 3a). In agreement with FCA, the Gr population clustered with individuals coming from PI area. At K = 7 the overall sample was partitioned in a way that that showed the genetic composition of the populations within each area. In each of the three, two different components were recognizable, one prevailing in subarea A and the other in C, with some mixing in the intermediate subarea (Fig 3b). Most homogeneous groups appeared subareas C in AR and PR, and subarea A in PI.

A finer analysis within each area remarked (at K = 2) an opposite assignment in individuals belonging to subareas A and C, confirming their genetic differentiation and suggesting an intuitive association of one cluster (IT) to the native lineage (Fig. 4). Indeed, between 64% and 72% in the three subareas A were assigned to the IT cluster with Q ≥ 0.75 and no individual was assigned at the same threshold to the alternative cluster (Tab. 4).Interestingly, while in PI and Pr the contact zone (subarea B) was mostly represented by admixed individuals (only 17% and 36% of individuals, respectively, with $Q_{I/II} \geq 0.75$), in AR it appeared as a mix of individuals with different ancestry (90% of individuals showing $Q_{I/II} \geq 0.75$, 46% assigned to cluster IT and 44% to cluster II, Tab.4), suggesting a strong population structure. Moreover, in Casentino (AR-C) all individuals were fully assigned to cluster II. Average assignment proportions to the native cluster (Q_{IT}) in the nine subareas correlated with the frequency of IT mtDNA haplotypes (arcsin-trasformation, $R^2 = 0.567$, $p = 0.019$).

Furthermore, frequency of IT haplotypes and high average membership to the IT cluster tended to increase with forest cover and with average size of wood patches (Tab. 1). The gradual distribution of mitochondrial and nuclear alleles and the high degree of admixture in contact zones in PI and PR would not subtend an outstanding role of landscape elements in shaping the genetic structure of these populations. On the contrary, the genetic pattern observed in AR suggests a discontinuity across the B zone, that can be due to the Arno river.

Actually the frequency of genotypes ascribed to the IT cluster is more than two-fold higher south than north of the river (69% vs. 29%, respectively).

We finally remark the occurrence of 10 mtDNA sequences (3 haplotypes) belonging to the East clade: all but one were detected in PR (5 in PR-A, 3 in PR-B and 1 in PR-C), none in AR.

Discussion

Genetic dynamics following a contact between two divergent gene pools are not easy to predict. In the present study we assessed the genetic composition of roe deer populations in three areas of central Italy: two of them were located at the northern border of the reported native range of the Italian subspecies *C. c. Italicus,* one was situated further north and resulted from a reintroduction of the native subspecies. Each study area was specifically selected to model a situation of contact between expanding native nuclei (coming from relict populations or reintroduction) and reintroduced/restocked populations of different origin.

Results confirm our previous assumptions based on historical information and clearly prove that genetic divergence between native and introduced lineages was maintained after decades of admixture opportunity.

As expected, in all three areas we found out a clear genetic gradient with a decreasing frequency of *italicus* genotypes at both mitochondrial and nuclear markers moving from south (subareas A) to north (subareas C). Nonetheless, differences among areas were observed.

The highest frequency of IT haplotypes was detected in upper Val di Cecina (PI-A), an area in ecological continuity with northern Maremma, where the native Italian roe deer have never gone extinct (Ghigi 1917). This is in agreement with observed frequencies in the north of Grosseto province (95% in our sampled site) and in the next areas of the province of Siena (Mucci et al. 2012). Yet, a surprisingly high frequency was detected in Garfagnana as well (PR-A), an area well north to the historical *italicus* border but where roe deer from Maremma had been translocated in the late 1960s. Genetic data seem to support the success of this introduction, apparently sustained by further translocations from the initial stock to other sites in the area (A. Vanoni pers. comm.). Nonetheless, the presence of an eastern haplotype (H16) in PR-A is likely the legacy of additional releases with northern roe deer (e.g. imported captive animals from eastern Alps) and indicative of a possible local admixture of diverging gene pools.

Although the ecological requirements of the Italian subspecies are poorly known (Focardi et al. 2009) and no specific ecological data was considered in this study, we can argue that the presence of large extensions of deciduous forests in Garfagnana (see Tab. 1) may have represented an advantage for the spread of the native lineage. This could also explain the great diffusion of the *italicus* gene pool in PR-B, a huge area covering the forested mountains in between the provinces of Reggio Emilia and Parma. We cannot exclude that undocumented releases of animals from southern populations have occurred in this area or nearby, but the gradient of IT frequency (A > B > C) and haplotype sharing between the two areas would suggest a gene flow from PR-A to PR-B.

On the opposite extreme of *italicus* presence is Casentino (AR-C). In this area, only one roe deer exhibited an IT haplotype. The massive presence of haplotypes ascribed to the Central clade agrees with previous evidences for the province of Arezzo (Lorenzini et al.,1993, 2002, Vernesi et al. 2002, Randi et al. 2004, Gentile et al. 2009). Nonetheless the origin of this roe deer population is still questionable. Lorenzini and colleagues (2002) noticed relationships with alpine populations, but the two most frequent haplotypes (H12 and H13) have not been observed abroad (only in Alps and Apennines, Randi et al. 2004, Gentile et al. 2009). Therefore, either an alpine origin through restocking or the recovery of a native depressed ('northern') lineage, differing from *C. c. italicus*, seem plausible explanations.

Noticeably, in the province, an appreciable genetic discontinuity was detected in AR-B. The presence of the Arno river can effectively limit roe deer dispersal, preventing the spread of IT genotypes northwards (see Fig. 1). The barrier-effect of the Arno river was proposed by Vernesi et al. (2002) to justify the lack of admixture between divergent gene pools in central Italy. This evidence was then contradicted by the finding of exotic alleles in populations of the province of Siena and Grosseto (Gentile et al. 2009, Mucci et al. 2012) and by our results. Anyway, a possible role of natural barriers and anthropogenic infrastructures should be carefully taken into account in the future management of the Italian roe deer.

Finally, the high consistency between mitochondrial and nuclear markers is rather uncommon in studies on mammal species (Lawson Handley & Perrin 2007). This is due to the widespread occurrence of male-biased dispersal (Greenwood 1980), mediating gene flow of nuclear but not mitochondrial alleles. On the co ntrary our data are in full agreement with

recent studies pointing to a lack of sex bias in roe deer dispersal (Coulon et al. 2006, Gaillard et al. 2008, Bonnot et al. 2010, Debeffe et al. 2012).

Conclusions

Our data support the existence of a patchy genetic composition of roe deer populations in central Italy, due to replicated and extremely localized releases that were conducted in the past. Such operations have put in a competition native and non-native gene pools, whose relative success was determined by several factors including time of the introduction, number of released individuals, habitat and landscape matrix (included relevant natural or artificial barriers). However, it should not be neglected a possible role of the species' ecology, which may differ among populations with a different genetic background, especially if they have experienced spatial segregation.

Once again, these results reveal that, although at a wide (continental) scale the phylogeography of large ungulates seems to be coherent with major demographic processes like over-glacial dynamics (roe deer Randi et al. 2004, red deer Skog et al. 2009, wild boar Scandura et al. 2008), at a local scale the genetic structure seems to be largely affected by more recent human manipulations and by landscape features.

This contribution to the knowledge of the ongoing hybridization dynamics involving the Italian roe deer could represent a starting step to further investigate the genetic structure of admixed populations by modelling the effect of landscape in spreading native/exotic genes.

At last, we recommend to pay attention in future sampling of Italian roe deer for genetic studies. Due to the observed fine-scale differentiation, comprehensive sampling would be necessary to represent an administrative area, like a province . For instance, the absence of *italicus* haplotypes in the samples from Arezzo analysed in previous studies (Lorenzini et al. 2002, Vernesi et al. 2002, Mucci et al. 2012) was probably due to having analysed samples from the north of the province only.

Acknowledgements

We are indebted with M. Corradi (Parco Regionale Boschi di Carrega), E. Mori (University of Siena), L. Mattioli (Provincial Administration of Arezzo) Baldacci (AFV Miemo), for sample

collection and historical information. We also thank L. Iacolina, C.S. Soffiantini and I. La Fata for their help in laboratory analyses.

The first author's PhD grant was financed by the Sardinia Region (European Social Fund, Operational Programme 2007/2013 – Axis IV - Human Capital Investment).

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.

Allendorf FW, Leary RF, Hitt NP, Knudsen KL, Lundquist LL, Spruell P (2004). Intercrosses and the U.S. Endangered Species Act: Should hybridized populations be included as westslope cutthroat trout? Conservation Biology 18: 1203–1213.

Apollonio M, Andersen R, Putman R (2010). European Ungulates and their Management in the 21st Century. Cambridge University Press.

Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16 (1):37-48.

Barbanera F, Pergams ORW, Guerrini M, Forcina G, Panayides P, Dini F (2010) Genetic consequences of intensive management in game birds. Biological Conservation 143:1259-1268.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) Genetix, logiciel sous Windows TM pour la génétique des populations, Laboratoire Génome, Populations, Interactions. CNRS UPR 9060, Université de Montpellier II, Montpellier (France).

Bonnot N, Gaillard J-M, Coulon A, Galan M, Cosson J-F, Delorme D, Klein F, Hewison M (2010). difference between the sexes in fine-scale spatial genetic structure of roe deer. PlosOne 5(12): e14436.

Champagnon J, Guillemain M, Elmberg J, Folkesson K, Gauthier-Clerc M (2010) Changes in Mallard Anas platyrhynchos bill morphology after 30 years of supplemental stocking. Bird Study 57 (3): 344-351.

Champagnon J, Elmberg J, Guillemain M, Gauthier-Clerc M, Lebreton J.-D. (2012a). Conspecifics can be aliens too: A review of effects of restocking practices in vertebrates. J Nat Cons 20: 231-241.

Champagnon J, Guillemain M, Elmberg J, Massez G, Cavallo F, Gauthier-Clerc M (2012b) Low survival after release into the wild: assessing “the burden of captivity” on Mallard physiology and behaviour. European Journal of Wildlife Research 58:255–267.

Champagnon, J, Crochet, P.A , Kreisinger J , Zížková D , Gauthier-Clerc M , Massez G , Söderquist P, Albrecht T , Guillemain M (2013). Assessing the genetic impact of massive restocking on wild mallard. Animal Conservation, 16 (3):295-305.

Chazara O, Minvielle F, Roux D, Bed'hom B, Feve K, Coville JL, Kayang BB, Lumineau S, Vignal A, Boutin JM, Rognon X (2010). Evidence for introgressive hybridization of wild common quail (*Coturnix coturnix*) by domesticated Japanese quail (*Coturnix japonica*) in France. Conservation Genetics 11 (3):1051-1062.

Consuegra S, Phillips N, Gajardo G, De Leaniz CG (2011). Winning the invasion roulette: Escapes from fish farms increase admixture and facilitate establishment of non-native rainbow trout. Evolutionary Applications 4:660-671.

Coulon A, Cosson J, Morellet N, Angibault J, Cargnelutti B, Galan M, Aulagnier S, Hewison A (2006). Dispersal is not female biased in a resource-defence mating ungulate, the European roe deer. Proceeding of thr Royal Society B 273:341-348.

Crudele G 1988 La fauna. In: Le foreste di Campigna-Lama nell'Appennino Tosco-romagnolo. M Padula and G Crudele Eds. Regione Emilia Romagna, Bologna. Pp. 325-401.

Debeffe L, Morellet N, Cargnelutti B, Lourtet B, Bon R, Gaillard J-M, Hewison AJM (2012). Condition-dependent natal dispersal in a large herbivore: heavier animals show a greater propensity

to disperse and travel further. Journal of Animal Ecology 81: 1327–1337.

Festa E. (1925). Il capriolo dell'Italia central (*Capreolus capreolus italicus sbsp*. Nova). Bollettino del Museo di Zoologia e Anatomia Comparata dell'Univarsità di Torino 7: 1-2.

Fickel J, Reinsch A, (2000) Microsatellite markers for the European roe deer (*Capreolus capreolus*). Molecular Ecology 9(7): 994-995.

Frantz AC, Pourtois JT, Heuertz M, Schley L, Flamand MC, Krier A, Bertouille S, Chaumont F, Burke T (2006). Genetic structure and assignment tests demonstrate illegal translocation of red deer (*Cervus elaphus*) into a continuous population. Molecular Ecology 15 (11):3191-3203.

Focardi S, Montanaro P, La Morgia V, Riga F (a cura di) (2009). Piano d'azione nazionale per il capriolo italico (*Capreolus capreolus italicus*). Quad Cons Natura nr. 31, Ministero Ambiente – ISPRA. Pp. 1-172.

Gaillard JM, Hewison AJM, Kjellander P, Pettorelli N, Bonenfant C, Van Moorter B, Liberg O, Andren H, Van Laere G, Klein F, Angibault JM, Coulon A, Vanpé C (2008). Population density

and sex do not influence fine-scale natal dispersal in roe deer. Proceedings of th Royal Society B 275: 2025–2030.

Gentile G, Vernesi C, Vicario S, Pecchioli E, Caccone A, Bertorelle G et al (2009). Mitochondrial DNA variation in roe deer (*Capreolus capreolus*) from Italy: evidence of admixture in one of the last C.c.italicus pure populations from central-southern Italy. Italian Journal of Zoology 76:16-27.

Ghigi A (1917). I Mammiferi d'Italia considerati nei loro rapporti con l'agricoltura. Riv. Sc. Nat. - Natura, VIII: 126.

Goedbloed DJ, van Hooft P, Megens H-J, Langenbeck K, Lutz W, Crooijmans RP, van Wieren SE, Ydenberg RC, Prins HHT (2013). Reintroductions and genetic introgression from domestic pigs have shaped the genetic population structure of Northwest European wild boar. BMC Genetics 43.

Guo SW, Thompson EA (2002). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361-372.

Greenwood PJ (1980). Mating systems, philopatry and dispersal in birds and mammals. Animal Behavior 28: 1140–1162.

Haanes H, Rosvold J, Røed KH (2013). Non-indigenous introgression into the Norwegian red deer population. Conservation Genetics 14: 237-242.

Hasegawa, M., Kishino, H. & Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22: 160-174.

Heymann Y, Steenmans Ch, Croisille G, Bossard M (1994). CORINE Land Cover. Technical Guide. Luxembourg (Office for Official Publications of European Communities).

Kappes SM, Keel JW, Stone RT, McGraw RA, Sonstegard TS, Smith TPL, Lopez-Corrales NL, Beattie CW (1997) A second-generation linkage map of the bovine genome. Genome Research 7:235–249.

Keller SR, Taylor DR (2010). Genomic admixture increases fitness during a biological invasion. Journal of Evolutionary Biology 23: 1720-1731.

Kemp SJ, Hishida O, Wambugu J (1995) A panel of polymorphic bovine, ovine and caprine microsatellite markers. Animal Genetics 26: 299-306.

Kumar S, Nei M, Dudley J, Tamura K. 2008. MEGA: a biologist centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinformatics 9:299–306

Laikre L, Allendorf FW, Aroner LC, Baker CS, Gregovich DP, Hansen MM, Jackson JJ, Kendall KC, McKelvey K, Neel MC, Olivieri I, Ryman N, Schwartz MK, Short Bull R, Stetz JB,

Tallmon DA, Taylor BL, Vojta CD, Waller DM, Waples RS (2010) Neglect of genetic diversity in implementation of the convention on biological diversity. Conservation Biology 24:86–88

Lawson Handley LJ, Perrin N (2007). Advances in our understanding of mammalian sex-biased dispersal. Molecular Ecology 16: 1559–1578.

Leighton FA (2002). Healts risk assessement of the translocation of wild animals. Rev. sci. tech. Off. int. Epiz 21(1): 187-195.

Linnel J., Zachos F. 2011. Status, distribution patterns of European ungulates: genetics, population history and conservation in " Putman R., Apollonio M., Andersen R. Eds Ungulate management in Europe" Cambridge University Press 12-53

Lorenzini, R., Patalano, M., Apollonio, M., and Mazzarone, V. (1993). Genetic variability of roe deer Capeolus capreolus in Italy: electrophoretic survey on populations of different origin. Acta theriologica. 38 (2):141-151.

Lorenzini R, Burrini L, Mazzoni della Stella R (1997). Biochemical genetic differentiation in some roe deer populations of Tuscany, central Italy. Italian Journal of Zoology 64:239-244.

Lorenzini R, Lovari S, Masseti M (2002). The rediscovery of the Italian roe deer; genetic differentiation and management implications. Italian Journal of Zoology 69:367-379.

Mamuris Z , Sfougaris AI , Stamatis C (2001b). Genetic structure of Greek brown hare (*Lepus europaeus*) populations as revealed by MtDBNA RFLP-PCR analysis: Implications for conserving genetic diversity. Biological Conservation, 101 (2) : 187-196.

Mattioli S., 1994. Le reintroduzioni di Cervidi in Italia, con particolare riferimento

all'Appennino. In: Inglisa M., cur., La fauna appenninica e la sua conservazione. Airone e Corpo Forestale dello Stato, Caramanico.

Mazzarone V, Lovari C, Mattioli L, Apollonio M, Siemoni N, Orlandi L, Castellani, Gualazzi L, Trinca M, Turchini I, Baldini G, Borchi S, Toni C (2000) Gli ungulati delle Foreste Casentinesi: dieci anni di monitoraggio 1988-1997. Ed. Regione Toscana, Firenze. Pp. 24-27.

Montanaro P, de Marinis AM, Riga F, Focardi S, Toso S (2003). Variabilità craniometrical in alcune popolazioni italiane di capriolo. Paper presented at the Atti IV Congresso Italiano di Teriologia: Scientific Research and Conservation of Mammals in Italy, Riccione, 6-8 November.

Mucci N, Mattucci F, Randi E (2012). Conservation of threatened local gene pools: Landscape genetics of the Italian roe deer (*Capreolus c. italicus*) populations. Evolutionary Ecology Research 14:897-920.

Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD (2004). Ecological and evolutionary consequences of biotic homogenization. Trends in Ecology and Evolution 19:18-24.

Perco F (1981). Capriolo, *Capreolus capreolus* Linnaeus, 1758. In: Pavan M. (a cura di), Distribuzione e biologia di 22 specie di mammiferi in Italia. Collana del Progetto Finalizzato "Promozione della qualità dell'ambiente". AQ/1/142-164. Consiglio Nazionale delle Ricerche, Roma, 145-152.

Perco F, Calò CM (1995). The status of roe deer in Italy. In: Wotschikowsky U., cur., Proceedings of the 2nd European Roe Deer Meeting. Wildbiologische Gesellschaft, München.

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

Rambaut 2012 FigTree v. 1.4.0. http://tree.bio.ed.ac.uk/software/figtree/

Randi E (2005). Management of wild ungulate populations in Italy: Captive-breeding, hybridisation and genetic consequences of translocations. Veterinary Research Communications 29 (Suppl. 2): 71-75.

Randi E, Pierpaoli M, Danilkin A (1998). Mitochondrial DNA polymorphism in populations of Siberian and European roe deer (*Capreolus pygarus* and *C. capreolus*). Heredity 80:429-443.

Randi E, Alves PC, Carranza J, Milosevic-Zlatanovic S, Sfougaris A, Mucci N (2004). Phylogeography of roe deer (*Capreolus capreolus*) populations: the effects of historical genetic subdivisions and recent non-equilibrium dynamics. Molecular Ecology 13:3071-3083.

Raymond M & Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86: 248-249.

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

Roed KH &Midthjell L (1998) Microsatellites in reindeer, *Rangifer tarandus*, and their use in other cervids. Molecular Ecology 7: 1773-1776.

Ryan ME, Johnson JR, Fitzpatrick BM (2009). Invasive hybrid tiger salamander genotypes impact native amphibians. PNAS 106: 11166-11171.

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

Scandura M, Iacolina, L, Crestanello B, Pecchioli E, Di Benedetto MF, Russo V, Davoli R, Apollonio M, Bertorelle G (2008). Ancient vs. recent processes as factors shaping the genetic

variation of the European wild boar: are the effects of the last glaciation still detectable? Molecular Ecology 17: 1745-1762.

Skog A, Zachos FE, Rueness EK, Feulner PGD, Mysterud A, Langvatn R, Lorenzini R, Hmwe SS, Lehoczky I, Hartl GB, Stenseth NC, Jakobsen KS (2009). Phylogeography of red deer (*Cervus elaphus*) in Europe. Journal of Biogeography 36:66-77.

Talbot J, Haigh J, Plante Y (1996) A parentage evaluation test in North American elk (*Wapiti)* using microsatellites of bovine and bovine origin. Animal Genetics 27: 117-119.

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

Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G (2002). The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. Molecular Ecology 11:1285-1297.

Wilson GA, Strobeck C, Wu L, Coffin JW (1997). Characterization of microsatellite loci in caribou *Rangifer tarandus*, and their use in other artiodactyls. Molecular Ecology 6 :697–699.

Table 1. Study areas, subareas, environmental features and genetic composition.

Area	Sub-area	Recent history	Surface	Elevation	Wood	Shrubs	Open	WP	H'	N_{mt}	mt_{IT}	N_{nuc}	Q_{IT}
AR	A	No release. Natural recolonization by expansion of neighbouring populations of the province of Siena	289.9	337	43 %	4 %	46 %	3.02	2.51	23	26.1 %	28	69.4 %
	B	No release. Natural recolonization from neighbouring areas.	782.5	417	45 %	7 %	36 %	2.55	3.08	34	23.5 %	50	49.8 %
	C	Never disappeared. Restocking with individuals from Alps in the 1930s and in 1950-1964.	584.5	787	64 %	10 %	23 %	3.18	2.70	17	5.9 %	25	6.5 %
PI	A	No release. Natural recovery of native populations in continuity with province of Grosseto	94.2	439	51 %	13 %	35 %	3.69	2.54	21	95.2 %	22	79.4 %
	B	No release. Natural recolonizarion from neighbouring areas	626.0	243	35 %	8 %	53 %	3.14	2.74	41	41.5 %	48	47.7 %
	C	Reintroduction of ca. 20 individuals from ex-Jugoslavia in the 1960s	31.6	365	85 %	3 %	12 %	4.48	1.25	19	21.1 %	23	34.9 %
PR	A	Reintroductions from multiple sources between 1950 and 1972, the most effective one being that of 16 individuals from Capalbio (southern Tuscany) in 1967-1969	504.4	870	79 %	4 %	14 %	8.81	1.73	23	73.9 %	25	75.9 %
	B	Releases of doubtful success in the 1930s. Natural recolonization from neighbouring areas.	1786.1	802	51 %	8 %	37 %	2.65	2.21	34	50.0 %	36	41.6 %
	C	Reintroductions at the end of XIX century with individuals from Balkans	68.3	191	12 %	12 %	63 %	0.70	2.45	22	40.9 %	27	23.2 %

Surface = extension of study subareas in km^2
Elevation = mean elevation expressed in m a.s.l.
Wood - percentage of surface covered by woods (deciduous/coniferous)
Shrubs = percentage of surface covered by bushes, shrubs or maquis
Open = percentage of surface covered by crops, pastures or meadows
WP = mean extension of wood patches
H' = Shannon's index of land use diversity
Nmt = number of roe deer mtDNA sequences obtained
mt_{IT} = frequency of mtDNA *italicus* haplotypes
Nnuc = number of roe deer microsatellite genotypes obtained
Q_{IT} = average assignment proportion to the Italian cluster resulting from Bayesian analysis in Structure

Table 2. Variables sites and haplotypes detected in the roe deer study populations.

Haplotype	Polymorphic sites	Clade	Geographic Range	AR A	AR B	AR C	PI A	PI B	PI C	PR A	PR B	PR C	GR	TOT	GenBank acc. code	Haplotype Gen
	` . . . . . . . . . . 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 3 3 4 4 4 4 5 6 6 6 6 7`															
	`3 3 3 3 5 5 6 6 7 8 0 3 4 9 0 0 0 2 2 2 3 4 4 6 6 9 1 7 0 0 1 4 4 6 6 6 8 0`															
	`0 3 4 6 5 6 1 7 4 4 3 2 1 5 0 1 4 0 4 5 9 3 4 1 7 2 2 7 0 4 2 4 4 3 4 9 2 4`															
H1	`G T T T T A A A T T T A T C T C G C G T A G T C T C A A T C G T T G A A A C`	Central														
H12	`. . . . . . . . . . . G C . C . A . . . . . . . . T . . . . . . . A G G G .`	Central	Ita(RM)	3	4	3								10	AY625743	L12
H13	`. . . . . . . . . . . G . . C . A . . . . . . . . T . . . . . . . A G . G .`	Central	Ita (Alp, AR,FO,FL)	13	20	13		2	1	1	2	3		55	AY625744	L11
H16	`. . . . . . . . C . . . . . C . . . . . . . . . C . . G . T . C . . G G G .`	East	Ita(CAlp, Ape), Ger							5	2	1		8	AY625747	L19
H18	`. . . . . . . . . . . G . . C . . . . . G . . . . T . . . . . . . . G G . .`	Central	Ita(Alp, Ape)					13	4		8	4		29	AY625749	
H23	`. . . . . . . . C . . . . . C . . . . . . . . . C . . G . T . C . . . G G .`	East	Ita(CAlp), Ser						1					1	AY625754	L20
H41	`. . . . . . . . . . . G . . C . . . . . G . . . . T . . . . . . . . G . . .`	Central	Ita(SI,GR), Ger										1	1	AY625772	L14/L15
H48	`. . . . . . . . . . . G . . C . . . . . G . . . C T . . . . . . . . G G . .`	Central	Ita(CAlp), Ser, Ger								2	2		4	AY625779	L17
H52	`. . . . . . . . . . . G . . C . . . . . G . . . . T . . . . . C . . G G G .`	Central						6	3				1	10	AY625783	
H58	`. . . . . . . . . . . G . . C . . . . . . . . . . T . . . . . . . A G . G .`	Central			1									1	AY625799	
H177	`. . . . . G . . . A . G . . C . . . . . G . . . C T . . . . . . . . G G . .`	Central										1		1	new	
H180	`. . . . . . . . . A . G . . C . . . . . G . . . . T . . . . . . . . G G . .`	Central									1	1		2	new	
H181	`. . . . . . . . . A . G . . C . . . . . G . . . C T . . . . . . . . G G . .`	Central										1		1	new	
H183	`T . G . G G . G . A . G . . C . A . . . . . . . . T . . . . . . . A G . G .`	Central			1									1	new	
H185	`. . . . . . . . . . . G . . C . A A . . . . A . . T . . A . . . . A G . G .`	Central		1										1	new	
H187	`. . . . . . . . . . . G . . C . . . . . G . . . . T . . . . . C . . . G G .`	Central					1		1					2	new	
H188	`. . . . . . . . . A . G . . C . . . . . G . . . . T . . . . . C . . G G G .`	Central						3	2					5	new	
H189	`. . . . . . . . . . . G C . C . A . . . . A A . . T . . . . . . . A G G G .`	Central							1					1	new	
H190	`A G . . . . . . . . . G . . C . . . . . G . . . . T . . . . . . . . G G . .`	Central							1					1	new	
H191	`. . . . . . . . . . . G . . C . . . . . G . . . . T . . . . A . . . G G . .`	Central							1					1	new	
H192	`. . G . G G . . . A . G . . C . . . . . G . . . . T . . . . . . . . G G . .`	Central									1			1	new	
H193	`. . . . . . . . C A . . . . C . . . . . . . . . C . . G . T . C . . G G G .`	East									1			1	new	
H15	`. . . . . . . G . A - . . . C C A . . . G . . . . . . . . . . . . . G . . .`	Central	Ita (SI,GR,MS)	5	8	1	11	5	1	14	10	7	14	76	AY625746	L5
H40	`. . . . . . . . . . - . . . C . A . . . G . . . . . G . . . . . . A G . G .`	Central	Ita (Ape)										1	1	AY625771	
H43	`. . . . . . . . . . - . . . C . A . . C G . . . . . . . . . . . . . G . G .`	Central	Ita (SI)				4	8	2				1	15	AY625774	L4
H176	`. . . . . . . . . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central						1		3	3			7	new	
H178	`. . G . G G G G . A - . . . C . A . . . G . . . . . . . . . . . G . G . . .`	Central										1		1	new	
H179	`. . . . . . . . . . - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central										1		1	new	
H182	`. . G . . G . . . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central		1							1			2	new	
H186	`. . . . . . . . . A - . . . C . A . . C G . . . . . . . . . . . . . G . G .`	Central					2	3	1					6	new	
H194	`. . . C . . . . . A - . . G C . A . A . G . . G . . . . . . . . . . G . . .`	Central									1			1	new	
H195	`. . . . . G . . . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central					1							1	new	
H196	`. . G . . G . G . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central					1							1	new	
H197	`T . G . . G . G . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central					1							1	new	
H199	`. . G . . . . . . A - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central									1			1	new	
H200	`. . . . . . . G . . - . . . C . A . . . G . . . . . . . . . . . . . G . . .`	Central									1			1	new	
			TOTAL	23	34	17	21	41	19	23	34	22	18	252		
			C. c. italicus (IT)	6	8	1	20	17	4	17	17	9	16	115		
			% IT	26.1%	23.5%	5.9%	95.2%	41.5%	21.1%	73.9%	50.0%	40.9%	88.9%	45.6%		
			East	0	0	0	0	0	1	5	3	1	0	10		
			Central	23	34	17	21	41	18	18	31	21	18	242		

Table 3. Fst values between subareas of each population. In each study area the highest divergence was observed between the two extreme subareas A and C (*).

a

Arezzo			
	AR-A	AR-B	AR-C
AR-A	-	-	-
AR-B	0.002	-	-
AR-C	0.050*	0.028	-

b

Pisa			
	PI-A	PI-B	PI-C
PI-A	-	-	-
PI-B	0.064	-	-
PI-C	0.105*	0.007	-

c

Parma			
	PR-A	PR-B	PR-C
PR-A	0		
PR-B	0.037	0	
PR-C	0.057*	0.011	0

Table 4. Individual assignment (Q ≥ 0.75) to one of the two clusters identified in each area.

Area	Subarea	N	cluster I (IT) Q ≥ 0.75	%	cluster II Q ≥ 0.75	%	cluster I/II Q ≥ 0.75	%
Arezzo	A	28	18	64%	5	18%	23	82%
	B	50	23	46%	22	44%	45	90%
	C	25	0	0%	25	100%	25	100%
Pisa	A	22	14	64%	0	0%	14	64%
	B	48	3	6%	5	10%	8	17%
	C	23	0	0%	7	30%	7	30%
Parma	A	25	18	72%	0	0%	18	72%
	B	36	2	6%	11	31%	13	36%
	C	27	0	0%	16	59%	16	59%

Legend Figures

Figure. 1 Map of the three roe deer study areas (Arezzo, AR – Pisa, PI – Parma, PR). Each area is partitioned into subareas A, B and C, respectively hosting native, reintroduced and admixed populations. Administrative boundaries (region – white, province – black line) and the Arno river are highlighted.

Figure 2. Factorial Correspondence Analysis (FCA) of roe deer genotypes from three study areas: AR – Arezzo (red), PI – Pisa (green), PR – Parma (blue) and the reference population in the province of Grosseto (GR, orange). Genotypes of subareas A are represented as squares, genotypes of subareas B and C as dots.

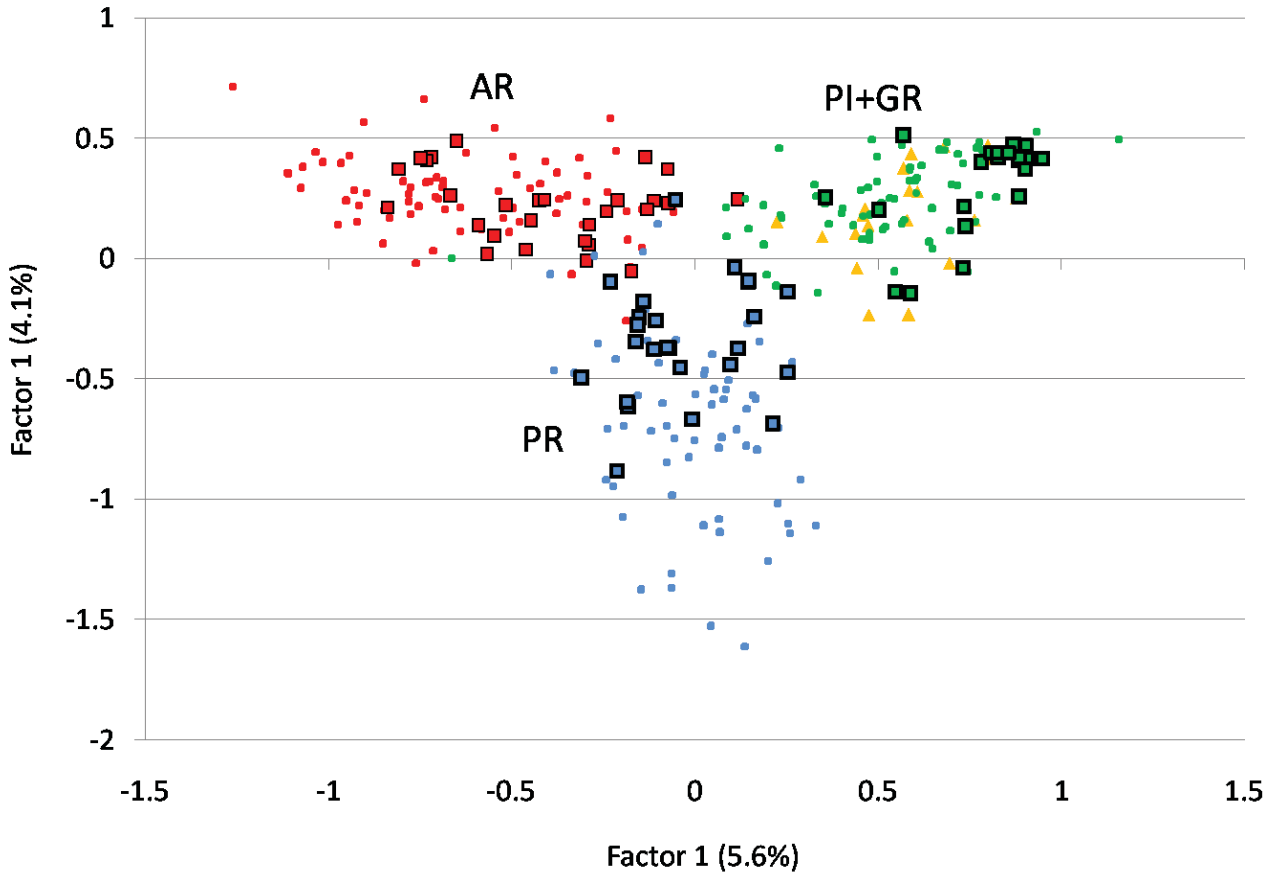


Figure 3. Bayesian cluster analysis of roe deer genotypes from three study areas (AR – Arezzo, PI – Pisa, PR – Parma) and the reference population in the province of Grosseto (GR). (a) Partition at K = 3. Each cluster corresponded to a study population. (b) Partition at K = 7. Each area is partitioned between two or more clusters. Subareas A and C are mostly distinguished.

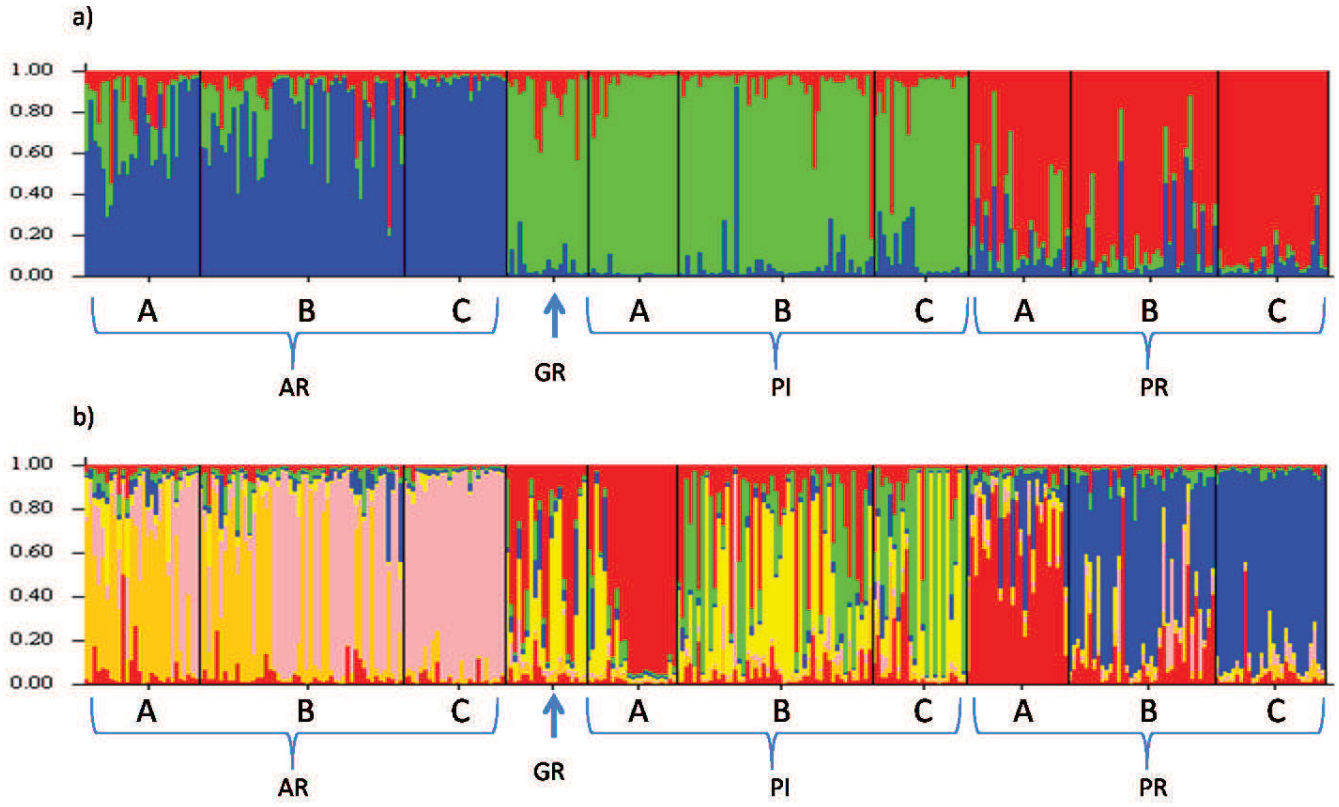


Figure 4. Bayesian cluster analysis within each province at K = 2. Results showed a clear identification of two main clusters roughly matching the A and C subareas.

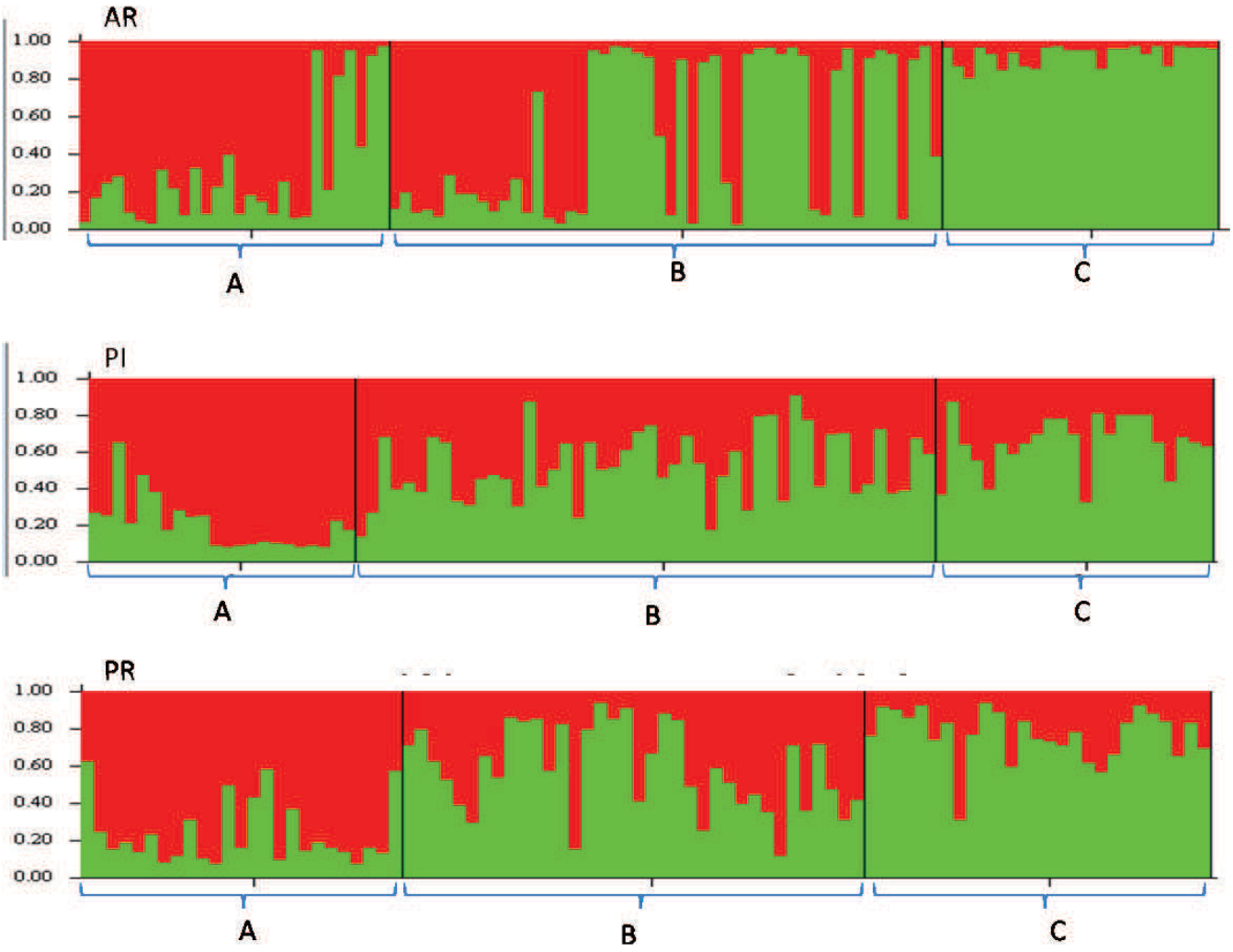
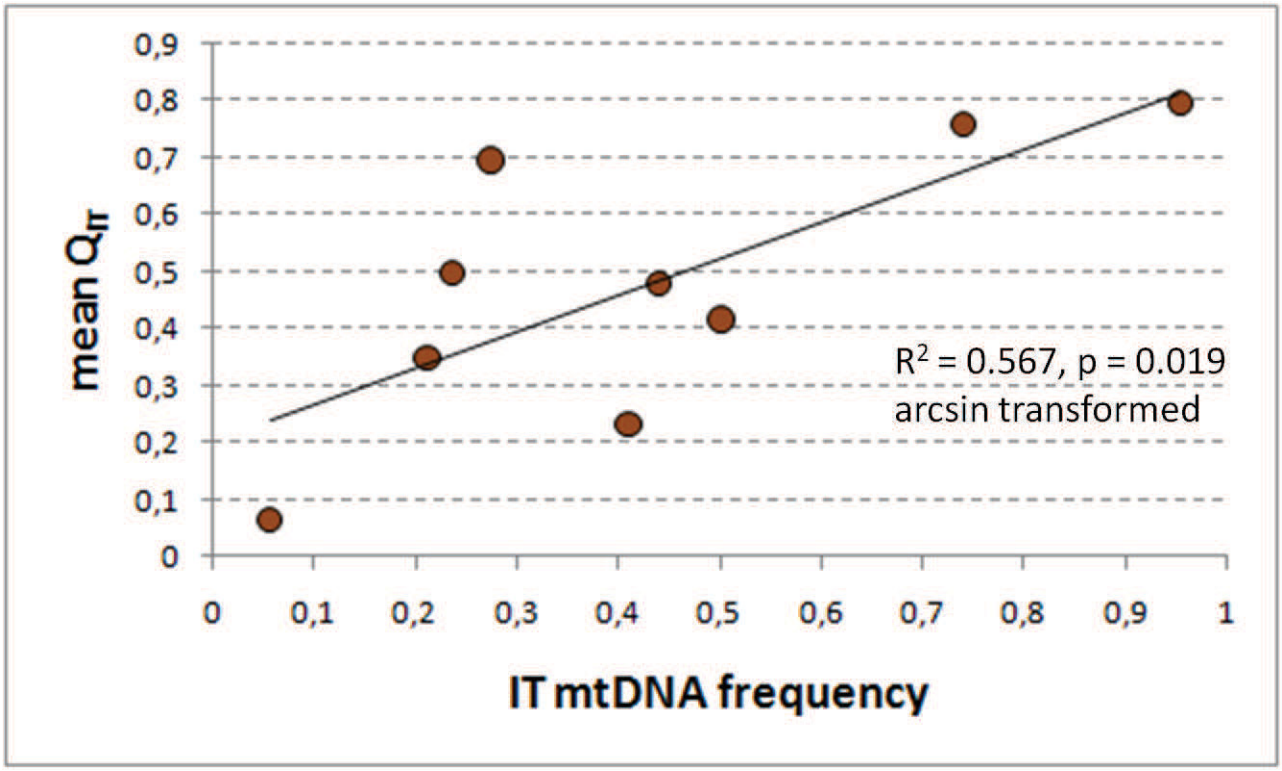
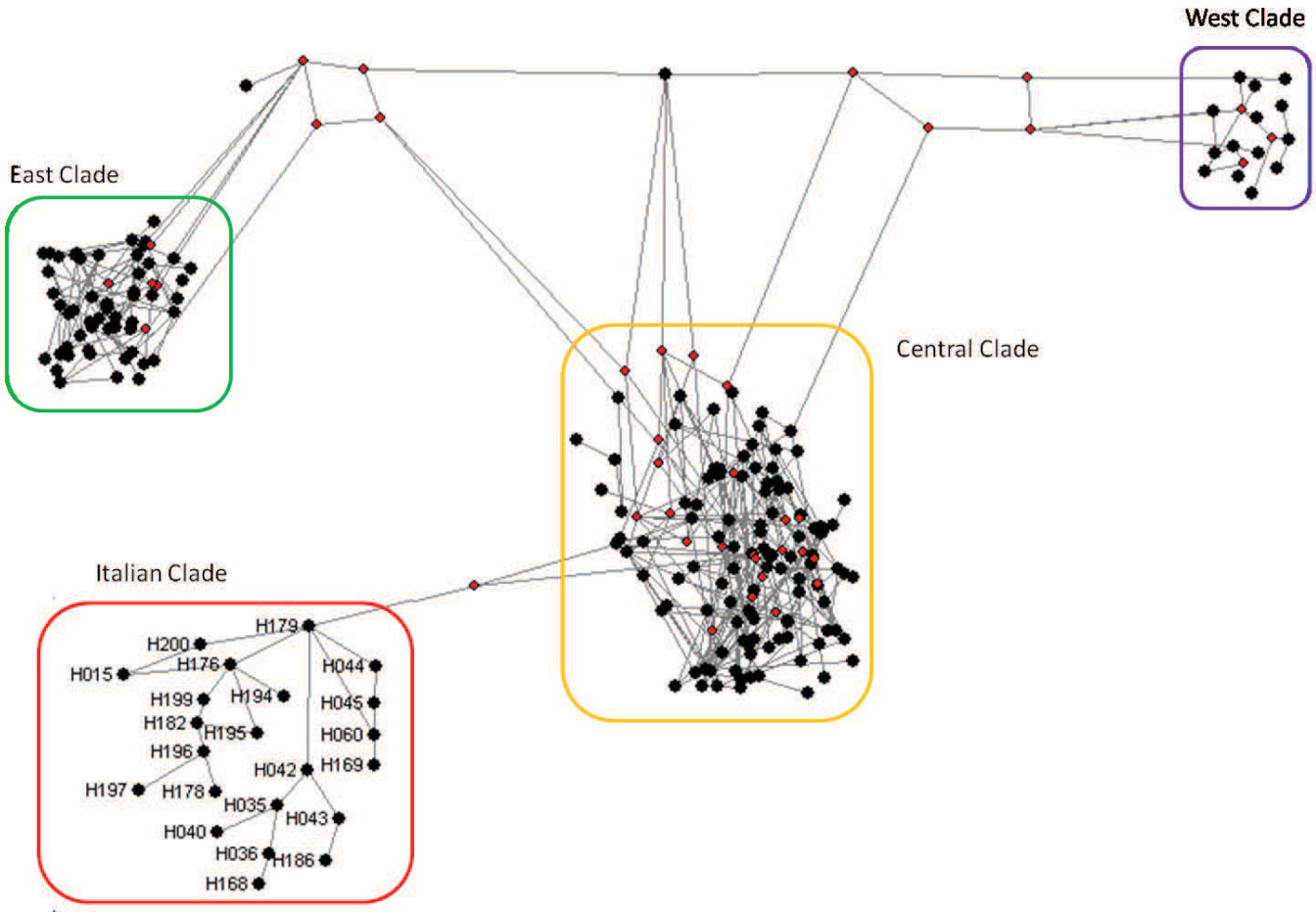


Figure 5. Correlation between frequency of italicus mitochondrial haplotypes and average assignment proportion to the Italian cluster across the nine subareas considered in this study.



Supplementary Material

Figure S1. Median-Joining network of European roe deer mitochondrial control region sequences (704 bp). The three major clades are coded according to Randi et al. (2004) as West, Central and East. Sequences attributed to the subspecies Capreolus capreolus italicus are shown as separate clade, but actually they are phylogenetically included in the Central clade (see Randi et al. 2004 and Mucci et al. 2012).



Chapter 4

Spatial pattern and genetic relatedness in a mountain roe deer population

Daniela Biosa, Nicoletta Sica, Nives Pagon, Stefano Grignolio, Massimo Scandura, Marco Apollonio

Manuscript

Spatial pattern and genetic relatedness in a mountain roe deer population

Daniela Biosa[1], Nicoletta Sica[1], Nives Pagon[1], Stefano Grignolio[1], Massimo Scandura[1], Marco Apollonio[1]

1 Department of Science for Nature and Environmental Resources, University of Sassari, Via Muroni 25, I-07100, Sassari, Italy. Phone: +39 079228628. Fax: +39 079228665.

Abstract

In this study we investigated genetic relatedness in a sample of 69 radio-collared roe deer of a high density population inhabiting a densely forested area in the Italian Apennines from 2002 to 2010. We assessed summer and winter home range by means of Kernel 90%. We used both centre of deer home range and overlap between inter and intra sexual pairs to look at population genetic structure along a spatial gradient. We calculated pairwise Queller and Goodnight's relatedness coefficients after genotyping individuals at 11 microsatellites and correlated them with spatial patterns arising from individual home range position and overlap. Genetic and spatial analyses, based on individual positions, did not reveal any population structure, neither in the complete dataset, nor in the two sexes separately. Mantel test did not detect significant correlation between relatedness and distance between home range centroids in summer, with the exception for the year 2003, whereas we found a negative correlation in some winter seasons (2002, 2003, 2004). Nevertheless, when we took into account overlap extent among individuals and genetic relatedness, we observed a different pattern: in general both sexes preferably seemed to share their home range with relatives of the same sex during the whole year, but this is not true for inter-sex dyads. In these case population organization during the summer rutting period and the winter grouping phase changes as inter-sex pairs during summers appeared to modify their home range overlap to limit the meeting with relatives. Our results confirm previous studies suggesting the absence of a sex-biased dispersal in roe deer, but they also point to the existence of a high gene flow across the continuous forest habitat which prevents the onset of population structuring at the local scale. However we found that behavioural mechanisms exist allowing the maintenance of genetic diversity in the population thanks to spatial tactics adopted during the mating season.

Key words: *Capreolus capreolus*, social organization, microsatellites, spatial behaviour, genetic structure.

Introduction

Social animals exhibit a heterogeneous intraspecific variation in their social structure, which is influenced by ecological and demographic constraints (Lott 1991). Social behaviour, mating system and dispersal pattern are strongly affected by the distribution of resources in the environment, which determines the amount, quality and distribution of food and the availability of shelters to protect against predation. But they also depend on population parameters as density and sex ratio, which determine the levels of intraspecific competition (both for food and for mates). Nonetheless, individual responses to ecological pressures may vary depending on internal factors (like sex, age, rank, body condition) resulting in a spectrum of individual strategies, whose combination will finally produce the socio-spatial organization of a population (Vehrencamp & Bradbury 1984). Although social behaviour and mating systems seem to be closely associated, extensive studies have shown that they are not always inter-dependent (Lott 1991). Consequently, observed patterns of social association between sexes could not reflect exclusively the mating system in the population but rather arise as a combination of individual strategies (Vehrencamp & Bradbury 1984; Davies 1991). Hence, distinct populations of the same species may react differently to similar environmental constraints, ultimately showing a different genetic structure (Lott 1991).

Several authors tried to disentangle the link between social and mating systems, dispersal patterns, and gene flow in numerous mammals species (Greenwood 1980, Lawson et al. 2007, Clutton-Brock & Lukas 2012), revealing that sex-biased dispersal is widespread in mammals and that female philopatry is the rule, resulting from a complex combination of evolutionary forces (Lawson et al. 2007, Dobson 2013). This common pattern leads to a divergent fine-scale genetic structure between sexes. In this frame, the role of social interactions, and especially the evolutionary advantages of a kin-based spatial association, have been remarked (Lehman & Keller 2006).

Philopatry and social viscosity (delayed dispersal) may induce individuals to live in kin clusters where animals share a large proportion of alleles by descent (Dobson 1982; Fowler 2005). Furthermore, social relationships between directly neighbouring individuals in a network can be stronger among related individuals (Wolf and Trillmich 2008). Cooperation among related males may help access to territory and mates in resource-defence systems (e.g. red grouse, Watson et al. 1994). For example, adult males of Hoffmann's two-toed sloths

(*Choloepus hoffmani*), one of the most sedentary terrestrial mammals on Earth, are segregated and defend some territory in core areas, although home ranges of adjacent males overlap in areas of low use (Peery and Pauli 2012). Neighboring adult males of this species are more closely related than expected. It was argued that the need for a male to defend a territory and mates would be reduced if the competitor was a relative, suggesting that kin selection could shape the mating system of sloths (Peery and Pauli 2012). Even from a female viewpoint, Nituch and colleagues (2008) showed that closely related ewes (*Ovis aries*) formed closer spatial associations than less-related and unrelated pairs.

To investigate on the influence of social and mating behaviour as well as of dispersal on the genetic structure of a mammal population we choose roe deer (*Capreolus capreolus*) as study case. We considered this species a good model because it is an income breeder with low sexual dimorphism, and a peculiar mating system adopted. The roe deer does not accumulate energy in form of fat reserves during the year but tends to use resources in proportion to metabolic needs (Andersen et al. 2000). Unlike the majority of cervids, the roe deer shows a very slight sexual dimorphism (Geist 1988) and very limited differences in spatial behaviour between sexes (Gaillard et al. 2008; Cagnacci et al. 2011), that help to evaluate the effect of social behaviour elements free from differential physiological constraints typical of strongly dimorphic ungulates. The roe deer mating strategy has been described as an “always stay” and a “low risk – low gain” mating strategy (Coulon et al. 2006; Vanpé et al. 2009), which does not conform to the expected pattern of a conventional resource-defence polygyny described by Greenwood (1980). Actually, levels of polygyny in roe deer are low (Vanpé et al. 2008), thus suggesting that a mating system based on resource-defence by males is not the primary selective force driving the evolution of dispersal patterns in this species (Coulon et al. 2006). Rate and extent of dispersal are region-dependent (Coulon et al. 2004), but common patterns are observed. Natal dispersal (sensu Howard 1960) generally occurs in individuals aged 10 months to 3 years during spring-early summer and often covers a few kilometres only (Linnel et al. 1998). High similarity in dispersal patterns between sexes was observed (Coulon et al. 2006, Gaillard et al. 2008, Bonnot et al. 2010). In both sexes a variable percentage of individuals is philopatric and shows a high fidelity to the natal range across the year and during lifetime, whilst the rest of the population either disperse early in their life or shows seasonal migration

(Wahlstrom &Liberg 1995, Cagnacci et al., 2011). Beyond natal dispersal and seasonal migration, short-range movements are also observed during the mating season, which were classified as 'breeding excursions' (San Josè & Lovari 1998, Lovari et al. 2008). The decision whether to disperse or not was observed to be condition-dependent, the attitude to disperse being stronger in heavier animals of both sexes (Wahlstrom & Liberg 1995, Debeffe et al. 2012). Threshold body weight for dispersing animals approximated 14 Kg in agricultural landscape in France (Debeffe et al. 2012) and 15 kg in forest habitat in Sweden (Wahlstrom & Liberg 1995). Therefore, body mass of individuals is a key factor in roe deer populations: while body mass of fawns in winter provides a good proxy of habitat quality (Gaillard et al. 2008), body mass of yearlings and adults is a good proxy of individual quality (Toigo et al. 2006) and represents a good predictor of life-history traits like natal dispersal (Debeffe et al 2012), survival (Gaillard et al. 2000), litter size (Hewison and Gaillard 2001) and reproductive success (Vanpé et al. 2010).

Finally, roe deer have adapted to a variety of habitats, consequently developing a high variability in the social structure and spatial behaviour among populations (Hewison et al. 1998 , 2001; Cargnelutti et al. 2002). Accordingly, the fine-scale genetic structure is expected to change in relation to landscape structure (Coulon et al. 2004), climatic conditions (Cagnacci et al. 2011), habitat structure (Rossi et al. 2001; Lamberti et al. 2006), population structure and density (Wahlström & Liberg 1995)

In this broad frame and given that roe deer show quite different levels of sociality and space occupancy during the territorial and the non-territorial period (Kurt 1966; Vincent et al. 1983; Danilkin & Hewison 1996), we specifically examined the socio-spatial structure of the study population in summer and compared it to that present in winter. Spatial behaviour during summer is indeed relevant to the social system of roe deer (Kjellander et al. 2004), as males are solitary and territorial on account of the approaching rutting season (July-August - Henning 1962; Bramley 1970), while female spatial behaviour is influenced by the social bond with fawns and by the active search for a partner during the oestrus (Lovari et al. 2008). Moreover, in summer, home ranges are smaller and more stable in both sexes and migratory animals tend to return to their natal site (Cagnacci et al. 2011). For these reasons, we supposed that the presence of philopatry in both sexes may increase the population viscosity due to a low sex-biased dispersal. Such conditions could be reflect in the genetic structure of the population resulted in a low relationship between relatedness and geographic distances.

Materials and Methods

Study area

The study was carried out between 2002 and 2010 in Alpe di Catenaia, a mountainous area located in Central Italy, which extends over an area of ca. 120 km^2. A protected area, Oasi Alpe di Catenaia (OAC), covering 28 km^2, is located within the study area. Elevation ranges from 330 to 1,514 m a.s.l., where snow occasionally falls between November and April. The study area is covered mainly by mixed forests (83%), consisting of beech (*Fagus sylvatica*) at altitudes higher than 900 m, and of Turkey oak (*Quercus cerris*), chestnut (*Castanea sativa*), black pine (*Pinus nigra*), silver fir (*Abies alba*) and Douglas-fir (*Pseudotsuga menziesii*) at lower altitudes. Shrubs and pastures cover around 16% of the area, while cultivated fields and urban areas are few and located outside OAC.

Roe deer and wild boar (*Sus scrofa*) are the only ubiquitous ungulates in the study area, whereas the presence of red deer (*Cervus elaphus*) is occasional. Natural predators are grey wolf (*Canis lupus*) and red fox (*Vulpes vulpes*), each consuming different roe deer age classes (Bassi et al. 2012). In the OAC hunting activities are permanently banned, whereas in the rest of the study area wild boar hunting with hounds and roe deer hunting from fixed high seats are permitted. The roe deer hunting season is open from 1st August to 15th September and from 15th January to 30th of March (for more information see Grignolio et al. 2011). Roe deer density is routinely assessed every year by drive census, following the methodology described in Mattioli et al. (1995) and Davis et al. (2012). The overall density of roe deer was then estimated as the mean across the different blocks and incremented accounting for the estimated number of newborns (assuming: sex ratio = 0.5, proportion of reproductive females = 0.812, average litter size = 1.905; unpublished data of the Provincial Administration of Arezzo). Between 2002 and 2010 post-birth roe deer density in the study area averaged 39.1 ± 5.0 heads/km^2 (mean ± SD), ranging annually between 33 and 48 heads/km^2. Mean live body weights of roe deer culled outside OAC area shown in Table 1 (unpublished data of the Provincial Administration of Arezzo).

Spatial behaviour data collection

Between March 2002 and March 2010, 69 roe deer (38 females and 31 males) were captured by means of vertical drop nets, sampled (hairs) for genetic analysis and fitted with

VHF radio-collars (Televilt, Sweden) before being released. Radio-collared deer were monitored by discontinuous radio-tracking obtaining individual locations by the triangulation method, on the basis of three bearings. We distributed the locations uniformly during the 24 hours on a monthly basis and separated consecutive locations by an interval of at least 12 hours, in order to avoid temporal and spatial autocorrelation (White & Garrott 1990). Further details on the field procedures adopted are described in Grignolio et al. (2011).

Genetic analysis

DNA was extracted from hair (at least 10 bulbs) using Instagene Matrix (Bio-Rad, Hercules, California, USA) and then stored at -20°C. The samples were genotyped with a panel of eleven polymorphic microsatellites: Roe1, Roe6, Roe8, Roe9 (Fickel & Reinsch 2000), NVRTH16, NVRTH21, NVRTH24 (Roed & Midthjell 1998), ILSTS011 (Kemp et al. 1995), OarFCB304 (Talbot et al. 1996), BMC1009 (Kappes et al. 1997) and RT1 (Wilson et al. 1997). They were amplified by three multiplexed (multiplex A: Roe1, Roe8, Roe9; multiplex B: RT1, NVRTH21, BMC1009; multiplex C: Roe6, NVRTH16, ILSTS011) and two single PCRs (NVRTH24 and OarFCB304). PCR conditions are available upon request. Fluorescently-labeled PCR products were sized by capillary electrophoresis on ABI PRISM 3730 automatic sequencer (Applied Biosystems).

We used Genetix 4.05 (Belkhir et al. 2001) to assess the level of genetic variability in the sampled population, through the estimation of observed (H_O) and expected heterozygosity (H_E), the mean number of alleles per locus (k) and the allele frequencies in the sample. F_{IS} was also calculated for the overall population and its significance tested by permutations. MicroChecker 2.2.3 (Van Oosterhout et al. 2004) was used to check for possible genotyping errors, allelic dropout, or null alleles at the 11 loci. MicroChecker 2.2.3 (Van Oosterhout et al. 2004) was used to check for possible genotyping errors, allelic dropout, or null alleles at the 11 loci, while deviations from the Hardy-Weinberg equilibrium (HWE) and from linkage equilibrium were assessed by Genepop 4 (Raymond & Rousset 1995). Significance of tests was corrected for multiple comparisons with the Bonferroni sequential procedure (Rice 1989).

The matrix of pairwise relatedness among the sampled individuals was calculated with GenAlEx 6.4 software (Peakall & Smouse 2005). In order to improve the calculation of relatedness coefficients we enlarged the dataset by including other 64 roe deer genotypes from

surrounding areas, so as to correct for possible distortions in allele frequencies in the monitored sample. As coefficient of relatedness we estimated the r_{xy} statistic (Queller & Goodnight 1989), ranging between -1 and +1, that quantifies the probability of sharing two alleles by descent. Highly positive relatedness values are expected for pairs that are more closely related than for random pairs in the population.

Spatial behaviour and genetic structure

Summer (1st June – 31th August, mating season) and winter (1st December – 28th February) home ranges were estimated using Kernel method 90% by means of Ranges VI software (Kenward et al. 2003).Each home range was obtained by at least 18 fixes per season (on average 24 fixes) evenly distributed among the three months. We excluded summer 2005 and 2006, and winter 2005, 2007 and 2009 from the analysis because all monitored animals had an insufficient number of fixes. Kernel centroids were obtained for each individual home range, and Euclidean distances between each pair of roe deer centroids were estimated.

Correlations between genetic relatedness and geographical distance were performed by Mantel test in GenAlEx 6.4. The test compares the correlation between the observed spatial (geographical coordinates of the centroid of each home range) and genetic (pairwise relatedness) matrices, to correlations obtained by 9,999 random permutations of the matrix columns and rows (Smouse et al. 1986). This computation was performed both jointly and separately for the two sexes, in order to test the hypothesis of sex-biased dispersal. All considered seasons had ≥3 monitored deer.

For each dyad of individuals (male-female, male-male and female-female), we calculated the extent of seasonal home range overlap using the Intersect tool in ArcGis 9.3 (ESRI, Environmental Systems Research Institute Inc. Redlands, California) and the Minta's index (1992; 1993) as follows:

$$Minta's\ index\ (\mathrm{MI}) = \sqrt{\frac{\mathrm{HR(AB)}}{\mathrm{HR(A)}} \times \frac{\mathrm{HR(AB)}}{\mathrm{HR(B)}}} \times 100$$

where HR(AB) is the area shared by individuals A and B, while HR(A) and HR(B) are the home ranges of individuals A and B, respectively (Minta 1992). Thereby, MI represents the mean potential interaction probability and may range between 0 and 100, with 100 indicating full overlap. On the basis of this parameter, we classified seasonal roe deer dyads into three categories representing different extents of overlap: O_0: MI = 0; O_1: 0.1≤ MI ≤ 40; O_2: MI >40. We assessed whether genetic relatedness increased with the extent of home range overlap by means of non-parametric randomization tests in R 2.15.1 (R Development Core Team 2011). Specifically, we calculated the significance of the difference between means of the two compared classes by computing the number of times a higher difference was obtained in 1,000 permuted datasets. This calculation was performed by sex and by season to verify different patterns of spatial and social organization. Furthermore, we tested whether buck and does tend to overlap their summer home range with unrelated potential partners, by comparing relatedness in male-female dyads belonging to the O_2 class (the highest overlap class) between summer and winter. Significance was estimated by randomization test in R.

Results

A total of 11,061 individual positions was recorded during a period of eight years. The mean home range size (± SE) in summer was 94.70 ± 15.90 ha and did not show significant differences between sexes (males: 87.25 ± 16.90 ha, females: 98.15 ± 21.20 ha; two-sample t-test: t =1.979, d.f. = 120, p = 0.412). In winter, home range size tends to decrease (69.18 ± 58.13 ha), again without a significant difference between males (67.38 ± 39.53 ha) and females (70.25 ± 67.13 ha; two-sample t-test: t =1.991; d.f. = 76; p = 0.813).

In total, 133 multi-locus genotypes were obtained (69 monitored + 64 non-monitored roe deer). The proportion of missing data was 2.93%. No evidence of allelic dropout or null alleles was detected by MicroChecker. A total of 71 different alleles (3-10 alleles per locus, k = 6.45) were found in the population. Average observed heterozygosity (H_O = 0.581) and expected heterozygosity (H_E = 0.616) were calculated. The overall F_{IS} in the population was positive (0.056) and significant (when tested by permutation in Genetix), in accordance with the presence of significant deviations from HWE across loci (χ^2=infinity; d.f.=22; p<0.001). None of the microsatellite loci showed a significant excess of heterozygotes. Similarly, no linkage

disequilibrium was found between any pair of microsatellites, when significance was corrected for multiple comparisons, except for pairs RT1-NVRTH16 and RT1-OarFCB304, that showed a p-value < 0.0001 (55 tests, $\alpha_{5\%}$=0.00091). The average relatedness coefficient in the sample of monitored roe deer was 0.098 ± 0.215 (mean ± SD, N=2,347 pairs). The overall distribution of r_{xy} values fitted a normal distribution (Kolmogorov-Smirnov normality test: p = 0.774; Fig. 1a) and did not differ between bucks and does (two-sample t-test: t = 2.101; d.f. = 18; p = 0.965; mean r_{xy}= 0.097 ± 0.214 among females and 0.087 ± 0.212 among males; Fig. 1b).

In summer most correlation coefficients (R) between relatedness and spatial distance were negative (Tab. 2), but no significant correlation was found (Mantel test, p> 0.05 for the overall population, and for the two sexes separately), with the single exception of summer 2003 in males (p = 0.036). On the contrary, negative correlation was highly significant in winter for the overall population only in 2003 (p = 0.009), and for females in the years 2002 (p = 0.050), 2003 (p = 0.033) and 2004 (p = 0.030). No significant correlation was found instead among males in winter.

In summer, the extent of overlap in roe deer dyads sharing a portion of their home range was on average 37.4 ± 12.5 % and no difference was detected between sexes across years (two-sample t-test: t = 2.44; d.f. = 6; p = 0.90). Limiting to intra-sex comparisons, dyads of does shared on average 30.5 ± 15.8% (n=92 pairs) of their summer home range, whereas dyads of bucks (n=38) averaged 30.1 ± 26.4% overlap. A single female summer home range was observed to overlap with as many as 6 monitored males, and the same maximum was observed in the opposite way round. As expected, overall overlap increased in winter (66.7 ± 10.9 %, for the overall sample, n = 238 pairs), averaging 55.1 ± 9.5 % and 41.1 ± 29.9 % for dyads of females and males, respectively. Also in winter the two sexes did not show a significant difference (two-sample t-test: t = 2.77; d.f. = 4; p-value = 0.64)

In both sexes, relatedness was higher between individuals with overlapping home ranges (classes O_1 and O_2) than between O_0 dyads (Tab. 3). This pattern was more evident, and statistically significant, in winter (randomization test, p-value = 0.020). In inter-sex dyads sharing a large part of their home range (O_2 class only), relatedness was low in summer (0.072± 0.03) and doubled in winter (0.154 ± 0.03), and this was the highest value observed (p = 0.046).

The relatedness of female dyads was always higher than male dyads for classes O_0 and O_1, while it was the opposite for the O_2 class both during summer and winter.

Discussion

Our results suggested the occurrence of a similar pattern of genetic relatedness between sexes within a forest roe deer population in the Italian Apennines. In particular, we detected the absence of correlation between genetic and spatial distance among males and females both in summer (mating period) and in winter (grouping period). Our data was carried out on a limited spatial scale (an area of about 30 km^2) and did not reveal any difference in spatial behaviour between the two sexes, according to similar results reported in previous studies on this species (Coulon et al. 2006; Gaillard et al. 2008; Bonnot et al. 2010; Debeffe et al. 2012) that proved, at different scales, the absence of sex-biased roe deer dispersal in mosaic landscapes in France and found evidence of similar dispersal behaviour in the two sexes.

The model on dispersal proposed by Wahlström and Liberg (1995) suggested that dispersal in roe deer depends on population density and fawns' body conditions. Thereby, no difference in dispersal rate is expected between males and females at high densities if fawns reach a critical weight in winter (around 15 kg in their study area in Scandinavia, 14 kg in a fragmented habitat in France; Debeffe et al. 2012). Despite high roe deer density in our study area, fawns body weight was high and beyond the threshold reported for the other studied populations. Hence, our data fit Wahlström and Liberg's prediction, suggesting similar dispersal patterns between sexes. We can infer that the environmental quality of our study area is elevated and enables juveniles of both sexes to quickly grow, giving them the opportunity to choose whether to disperse or not.

Actually, habitat conditions and high density in the population could promote the use of different spatial organization strategies to limit inbreeding and to shuffle genes in the population.

Coulon et al. (2004) stressed that, in investigating correlations between genetic relatedness and spatial distances, landscape features should be taken into account, as the degree of habitat fragmentation can affect individual dispersal and, consequently, the genetic structure of the whole population. Roe deer in our study area inhabited a continuous forest

without barrier and this might favour gene flow and account for the lack of genetic structure observed in the study population.

It is important to notice that, unlike the above-mentioned studies, we calculated pairwise distances between animals from home ranges centroids and not from geographic coordinates of captures or kills. The latter, indeed, can barely represent the geographic position of animals and refer to a specific time of the year (in most cases the winter). Accordingly, we expected more pronounced differences between sexes on account of their dissimilar spatial behaviour during the mating season. Instead, when we compared relatedness with different degrees of home range overlap, we observed that relatedness tended to increase with overlap, in both sexes. Such tendency is more pronounced during winters probably because of scarcity of food resources and the consequent formation of small groups, usually composed by relatives, in the most suitable areas (Danilkin & Hewison 1996).

During the breeding period, we found a similar pattern for intra-sex dyads, with a general preference in sharing their home range with relatives. For both sexes this may be explained by the presence of some level of philopatry for this species, even it was detected in different European populations only for females (Vincent et al. 1995; Danilkin & Hewison 1996; Andersen et al. 1998). Respecting males, summer represents the period when aggressive interactions among bucks are more frequent due to territoriality. This can be especially true in a high density area, as Alpe di Catenaia. Our results suggested that also in this period males prefer to share their home range with close relatives, also showing a certain degree of philopatry. This behavioural pattern (kin tolerance) could be a strategy to reduce the competition for mating and the number of aggressive interactions between related males and be finally associated to the inclusive fitness of males (Brown & Brown 1993). Our findings are in compliance with the mating system adopted by roe deer: “always stay” and a “low risk – low gain” mating strategy (Coulon et al. 2006; Vanpé et al. 2009). Indeed, if neighbours of a territorial male are relatives, we may assume that he will reduce the level of aggression and, consequently, the risks associated with competitive interactions with them, while maintaining an unaltered chance of mating for the kin cluster.

Finally, during the rutting period, males and females seem to be distributed in the space to minimize the chances to meet relatives of the opposite sex. The values of inter-sex kinship turn out to be much lower than intra-sex ones during summers, but not during winters. Such

genetic pattern observed within the population may be due to the presence of seasonal migration and reproductive excursions with the ultimate aim to prevent high level of inbreeding. Yet, seasonal migration and natal dispersal are not the only spatial strategies involved in inbreeding avoidance by females. Indeed, previous studies carried out on the Apennine population of Casentinesi Forest National Park, characterized by similar environmental features (Lamberti et al 2001, 2004; Rossi et al 2001), identified different spatial behaviours in females during the rutting period. They found stationary and migratory females; the latter showed summer home ranges much larger than the stationary ones. Our findings on spatial behaviour and different inter-sex genetic relatedness patterns between summer and winter indicate that females are likely to adopt different behavioural patterns to increase the encounter probability with unrelated males during the rut. This factor, together with the low level of polygamy (Vanpè et al 2008), can play a role in the maintenance of high levels of genetic variability

An apparent contradiction was the absence of correlation between genetic and spatial distances, but association between genetic relatedness and home range overlap emphasizes the need to use a multiple approach to evaluate spatial behaviour in relation to genetic relatedness in an animal population. Most studies in order to relate a spatial reference to individual genetic data took into account a single individual position in the space that can range from the observation/culling spot to the weighted centre of the home range (Gaillard et al. 2008; Bonnot et al. 2012), just a limited number of examples of the use of home range overlap among individuals are available (Jesmer et al. 2011; Innes et al. 2012). We decided to look at both kinds of data in order to consider also the probabilities of interaction between individuals, based on the shared use of space. Individual deer may live quite closer and avoid any home range overlap or, on the contrary, show an extended one. The use of different spatial information proved useful to face this problem. Further analyses on larger spatial scales in forested landscapes are needed to confirm the spatial pattern and the genetic structure observed, while taking into account the role of landscape features in shaping different patterns of genetic variation.

Acknowledgments

We would like to thank, P. Bongi, E. Bertolotto, S. Luccarini, E. Donaggio, E. Panozzi, B. Pizza, G. Meacci and the hunters of the URCA association, and all other volunteers contributing to field activities. We are also grateful to R. Chirichella for her support in GIS analyses and G. Falcer for linguistic revision. We are also indebted to P. Banti, G. Chianucci, P. Pedone and L. Mattioli for logistic support.

This study was funded by Tuscany Region, Arezzo Province and Italian Ministry for the University and Research (PRIN 2005). The first author's PhD grant was financed by the Sardinia Region (European Social Fund, Operational Programme 2007/2013 – Axis IV - Human Capital Investment).

References

Apollonio M, Festa Bianchet M, Mari F, Mattioli S, Sarno B (1992) To lek or not to lek – Mating strategies of male fallow deer.Behav Ecol 3: 25-31

Bassi E, Donaggio E, Marcon A, Scandura M, Apollonio M (2012) Trophic niche overlap and wild ungulate consumption by red fox and wolf in a mountain area in Italy. MammalBiol doi:10.1016/j.mambio.2011.12.002

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) Genetix, logiciel sous Windows TM pour la génétique des populations, Laboratoire Génome, Populations, Interactions. CNRS UPR 9060, Université de Montpellier II, Montpellier (France)

Bideau E, Vincent JP, Quéré JP, Angibault JM (1983) Occupation de l'espace chez le chevraeuil (*Capreoluscapreolus L*.). I. Cas des mâles. ActaOecol, OecolAppl 4: 163-184

Bonnot N, Gaillard JM, Coulon A, Galan M, Cosson JF, Delorme D, Klein F, Hewison AJH (2010) No Difference between the Sexes in Fine-Scale Spatial Genetic Structure of Roe Deer. PLoS ONE 5(12): e14436

Börger L, Franconi N, De Michele G, Gantz A, Meschi F, Manica A, Lovari S, Coulson T (2006) Effects of sampling regime on the mean and variance of home range size estimates. J Anim Ecol 75: 1393-1405

Bramley PS (1970) The evolution of diversity in avian territorial systems. The Wilson Bulletin 76: 160-168

Brown GE, Brown JA (1993) Social dynamics in salmonid fishes: do kin make better neighbours? Animal Behaviour 45(5): 863–871

Cagnacci F, Focardi S, Heurich M, Stache A, Hewison AJM, Morellet N, Kjellander P, Linnell JDC, Neteler M, Delucchi L, Ossi F, Urbano F (2011) Partial migration in roe deer: migratory and resident tactics are end points of a behavioural gradient determined by ecological factors. Oikos 120: 1790–1802

Cargnelutti B, Reby D, Desneux L, Angibault JM, Joachim J, Hewison AJM (2002) Space use by roe deer in a fragmented landscape. Some preliminary results. Revue d'Ecologie (Terre Vie) 57: 29-37

Clutton-Brock TH (1989) Mammalian mating systems. Proceedings of the Royal Society of London B 236: 339-372

Clutton-Brock TH, Green D, Hiraiwa-Hasegawa M, Albon SD (1988) Passing the buck – Resource defence, lek breeding and mate choice in fallow deer. Behav Ecol Sociobiol, 23: 281-296

Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. Mol Ecol 13: 2841-2850

Coulon A, Cosson JF, Morellet N, Angibault JM, Cargnelutti B, Galan M, Aulagnier S, Hewison AJM (2006) Dispersal is not female biased in a resource-defence mating ungulate, the European roe deer. Proceedings of the Royal Society B 273: 341-348

Danilkin A, in association with Hewison AJM (1996) Behav Ecol of Siberian and European Roe deer. Chapman and Hall, London

Davies NB (1991). Mating systems. In: Behav Ecol: An Evolutionary Approach (Krebs J.R. & Davies, N.B, eds.). Blackwell Scientific, Oxford, pp 263-299

Davis ML, Stephens, PA, Willis, SG, Bassi E, Marcon A, Donaggio E, Capitani C, Apollonio M (2012) Prey selection by an apex predator: the importance of sampling uncertainty. In Press.

Debeffe L, Morellet N, Cargnelutti B, Lourtet B, Bon R, Gaillard JM, Hewison AJM (2012). Condition-dependent natal dispersal in a large herbivore: heavier animals show a greater propensity to disperse and travel further. Journal of Animal Ecology 81: 1327-1337

Dingle, H. and V. A. Drake. 2007. What is migration? BioScience 57:113–121

Emlen ST, Oring LW (1977) Ecology, sexual selection and the evolution of mating systems. Science 197: 215-233

Fickel J, Reinsch A, (2000) Microsatellite markers for the European roe deer (*Capreoluscapreolus*). Mol Ecol 9(7): 994-995

Fischer F, Linsenmair KE (1999) The territorial system of the Kob antelope (*Kobuskobkob*) in the Comoe National Park, Côte d'Ivorie. African Journal of Ecology 37(4):386-399

Gaillard JM, Hewison AJM, Kjellander P, Pettorelli N, Bonenfant C, Van Moorter B, Liberg O, Andren H, Van Laere G, Klein,FAngibault JM, Coulon A, Vanpé C (2008) Population density and sex do not influence fine-scale natal dispersal in roe deer. Proceedings of the Royal Society B 275: 2025-2030

Geist V (1988) How markets for wildlife meat and parts, and the sale of hunting privileges, jeopardize wildlife conservation.ConservBiol 2: 15-26

Gosling LM (1986) The evolution of the mating strategies in male antilopes. In Ecological Aspect of Social Evolution: Birds and Mammals (edsRubestein, D.I. and Wrangham, R.W.). Princeton University Press, Princeton pp. 244-281

Greenwood PJ (1980) Mating system, philopatry and dispersal in birds and mammals. AnimBehav 28: 1140-1162

Griggio M, Biard C, Penn DJ, HoiH. (2012). Female house sparrows "count on" male genes: experimental evidence for MHC-dependent mate preference in birds. BMC Evolutionary Biology 2011 11: 44

Grignolio S, Merli E, Bongi P, Ciuti S, Apollonio M (2011) Effects of hunting with hounds on a non-target species living on the edge of a protected area. BiolConserv 144: 641–649

Henning R (1962) Uber das Revierverhalten des Rehbocke.Zeitschrift fur Jagdwissenschaft 8: 61-81

Hewison AJM, Vincent JP, Reby D (1998) Social organisation of European roe deer. In: Andersen, R, Duncan, P, Linnell, J.D.C. (Eds). The European roe deer: The biology of success. Oslo, Norway: Scandinavian University Press, pp. 189-219

Hewison AJM, Vincent JP, Joachim J, Angibault JM, Cargnelutti B, Cibien C (2001) The effects of woodlands fragmentation and human activity on roe deer distribution in agricultural landscapes. Can J Zool 79: 679-689

Hood GM (2010) PopTools version 3.2.5. Available on the internet. URL http://www.poptools.org

Kappes SM, Keel JW, Stone RT, McGraw RA, Sonstegard TS, Smith TPL, Lopez-Corrales NL, Beattie CW (1997) A second-generation linkage map of the bovine genome. Genome Research 7: 235–249

Kemp SJ, Hishida O, Wambugu J (1995) A panel of polymorphic bovine, ovine and caprine microsatellite markers.Anim Genet 26: 299-306

Kenward RE, South AB, Walls SS(2003). Ranges VI vl.2: for the analysis of tracking and location data. Anatrack Ltd., Whareham (UK).

Kjellander P, Hewison AJM, Liberg O, Angibault JM, Bideau E, Cargnelutti B (2004) Experimental evidence for density-dependence of home-range size in roe deer (*Capreoluscapreolus L.*): a comparison of two long-term studies. Oecologia 139: 478-485

Kjellander P, Gaillard JM and Hewison AJM (2006). Density-dependent responses of fawn cohort body mass in two contrasting roe deer populations. Oecologia (Berlin) 146:521–530

Kurt F (1966) Feldbeobachtungen und Versucheuber das Revierverhalten der Rehbocke (*Capreoluscapreolus L.*). Revue Suisse de Zoologie 73:408-421

Jordan WC and Bruford MW (1998). New perspectives on mate choice and the MHC. Heredity 81: 239–245

Lamberti P, Rossi I, Mauri L, Apollonio M (2001). Alternative use of space strategies of female roe deer (*Capreoluscapreolus*) in a mountainous habitat. Italian Journal of Zoology 68: 69-73

Lamberti P, Mauri L, Apollonio M (2004). Two distinct patterns of spatial behaviour of female roe deer (*Capreoluscapreolus*) in a mountainous habitat. Ethology Ecology & Evolution 16: 41-53

Lamberti P, Mauri L, Merli E, Dusi S, Apollonio M (2006) Use of space and habitat selection by roe deer (*Capreoluscapreolus*) in a Mediterranean coastal area: how does woods landscape affect the home range? J Ethol 24(2): 181-188

Lehmann, L., Keller, L.The evolution of cooperation and altruism - A general framework and a classification of models

(2006) Journal of Evolutionary Biology, 19 (5), pp. 1365-1376.

Linnell JDC, Andersen R (1998) Territoriality fidelity and tenure in roe deer bucks. ActaTheriol 43(1): 67-75

Linnell JDC, Wahlström LK, Gaillard JM (1998) From birth to independence: birth, growth, neonatal mortality, hiding behaviour and dispersal. In: Andersen, R, Duncan, P, Linnell, J.D.C. (Eds). The European roe deer: The biology of success. Oslo, Norway: Scandinavian University Press. Pp. 257-283

Lott DF (1991) Intraspecific variation in the social systems of wild vertebrates. Cambridge University Press, Cambridge

Lovari S, Bartolommei P, Meschi F, Pezzo F (2008). Going out of mate: excursion behaviour of female roe deer. Ethology 114: 886-896

Mattioli L, Apollonio M, Mazzarone V, Centofanti E (1995) Wolf food habits and wild ungulate availability in the ForesteCasentinesi National Park, Italy. ActaTheriol 40 (4) 387-402

Minta SC (1992). Tests of spatial and temporal interaction among animals. Ecol Appl 2: 178–188

Minta SC (1993). Sexual differences in spatio-temporal interaction among badgers.Oecologia 96:402-409

Moore J, Ali R (1984) Are dispersal and inbreeding avoidance related? AnimBehav 32: 94-112

Nefolt RJC, Thirgood SJ (1997) Lekking, resource defence, and harassment in two subspecies of lechwe antelope.Behav Ecol 8(1):1-9

Nies G, Zachos FE, Hartl GB (2005) The impact of female philopatry on population differentiation in the European roe deer (*Capreoluscapreolus*) as revealed by mitochondrial DNA and allozymes. Mammal Biol 70(2): 130-134

Owen-Smith N (1977) On territoriality in ungulates and an evolutionary model. Quarterly Review of Biology 52(1): 1-38

Peakall R, Smouse PE (2005) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288-295

Perrin R, Mazalov V (2000) Local competition, inbreeding, and the evolution of sex-biased dispersal.AmNat 155: 116-127

Perrin N, Goudet J (2001) Inbreeding, kinship, and the evolution of natal dispersal. In: Clobert, J, Danchin, E, Dhondt, A.A, Nichols, J.D. (Eds) Dispersal. Oxford, UK: Oxford University Press pp. 123-142

Queller DC & Goodnight KF (1989) Estimating relatedness using genetic markers. Evolution 43: 258-275

R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at http://www.R-project.org/

Raymond M &Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86: 248-249

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

Roed KH &Midthjell L (1998) Microsatellites in reindeer, *Rangifertarandus*, and their use in other cervids. Mol Ecol 7: 1773-1776

Rossi L, Lamberti P, Mauri L, Apollonio M (2001) Male and female spatial behaviour of roe deer in a mountainous habitat during pre-rutting and rutting period. J Mt Ecol 6(1): 1-6

Saïd S&Servanty S (2005) The influence of landscape structure on female roe deer home-range size. Landsc Ecol 20: 1003-1012

Saïd S, Gaillard JM, Widmer O, Débias F, Bourgoin G, Delorme D, Roux C (2009) What shapes intra-specific variation in home range size? A case study of female roe deer. Oikos 118: 1299 1306

San José C and Lovari S (1998). Ranging movements of female roe deer: do home-loving does roam tomate? Ethology 104: 721–728

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. SystZool 35: 627-632

Smouse PE &Peakall R (1999) Spatial autocorrelation analysis of multi-allele and multi-locus genetic microstructure. Heredity 82: 561-573

Talbot J, Haigh J, Plante Y (1996) A parentage evaluation test in North American elk (*Wapiti*) using microsatellites of bovine and bovine origin. Anim Genet 27: 117-119

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Note 4: 535-538

Vanpé C, Kjellander P, Galan M, Cosson JF, Aulagnier S, Liberg O, Hewison AJM (2008) Mating system, sexual dimorphism and the opportunity for sexual selection in a territorial ungulate. Behav Ecol 19:309–316

Vanpé C, Morellet N, Kjellander P, Goulard M, Liberg O, Hewison AJM (2009) Access to mates in a territorial ungulate is determined by the size of a male's territory, but not by its habitat quality. J Anim Ecol 78: 42-51.

Vehrencamp SL & Bradbury JW (1984) Mating systems and ecology. -In: Behav Ecol: An Evolutionary Approach (Krebs J.R. & Davies, N.B, eds.). Blackwell Scientific, Oxford, pp 251-278

Vincent JP, Bideau E, Quéré JP, Angibault JM (1983) Occupation de l'espacechez le chevreuil (*Capreoluscapreolus L*.). II. Cas des gemelle.ActaOecol, Oecol Applica 4: 379-389

Vincent JP, Bideau E, Hewison AJM, Angibault JM (1995) The influence of increasing density on body weight, kid production, home range and winter grouping in roe deer (*Capreoluscapreolus*). J Zool 236: 371-382

Wahlström LM, Liberg O (1995) Patterns of dispersal and seasonal migration in roe deer (*Capreoluscapreolus*). J Zool 235: 455-467

White GC, Garrott RA (1990) Analysis of wildlife radio-tracking data. Academic Press Inc, San Diego

Wilson GA, Strobeck C, Wu L, Coffin JW (1997).Characterization of microsatellite loci in caribou *Rangifertarandus*, and their use in other artiodactyls. Mol Ecol 6 : 697–699

Zelano B and Edwards SV (2002). An Mhc Component to Kin Recognition and Mate Choice in Birds: Predictions, Progress, and Prospects. Am. Nat. 160 : S225-S237

Table 1. Body mass of fawns, yearling and adult roe deer in Alpe di Catenaia. Data refer to animals culled in winter (January-March) in hunting districts surrounding the protected area. N = number of culled individuals, SD = standard deviation, Ratio = ratio of average young body mass to average female adult body mass. Fawns: 8-10 months, Yearlings: 20-22 months, Adults: > 2 years.

	Fawns			**Yearlings**			**Adults**		
	Tot	***F***	***M***	***Tot***	***F***	***M***	***Tot***	***F***	***M***
N	96	53	43	25	25	/	222	222	/
Mean	16.48	16.11	16.92	22.91	22.91	/	23.59	23.59	/
SD	2.15	2.03	2.22	2.02	2.02	/	2.34	2.34	/
Ratio		0.68			0.97			/	

Table 2. Correlations (R_{xy}) between genetic relatedness and Euclidean distance among monitored roe deer in Alpe di Catenaia (2002 – 2010). Spatial distances were calculated between summer and winterKernel home range centroids. N = number of monitored individuals, P-value = significance of correlation, * = significant at α = 0.05. Shadowed cells refer to seasons in which the sample size was small (i.e. < 10 individuals).

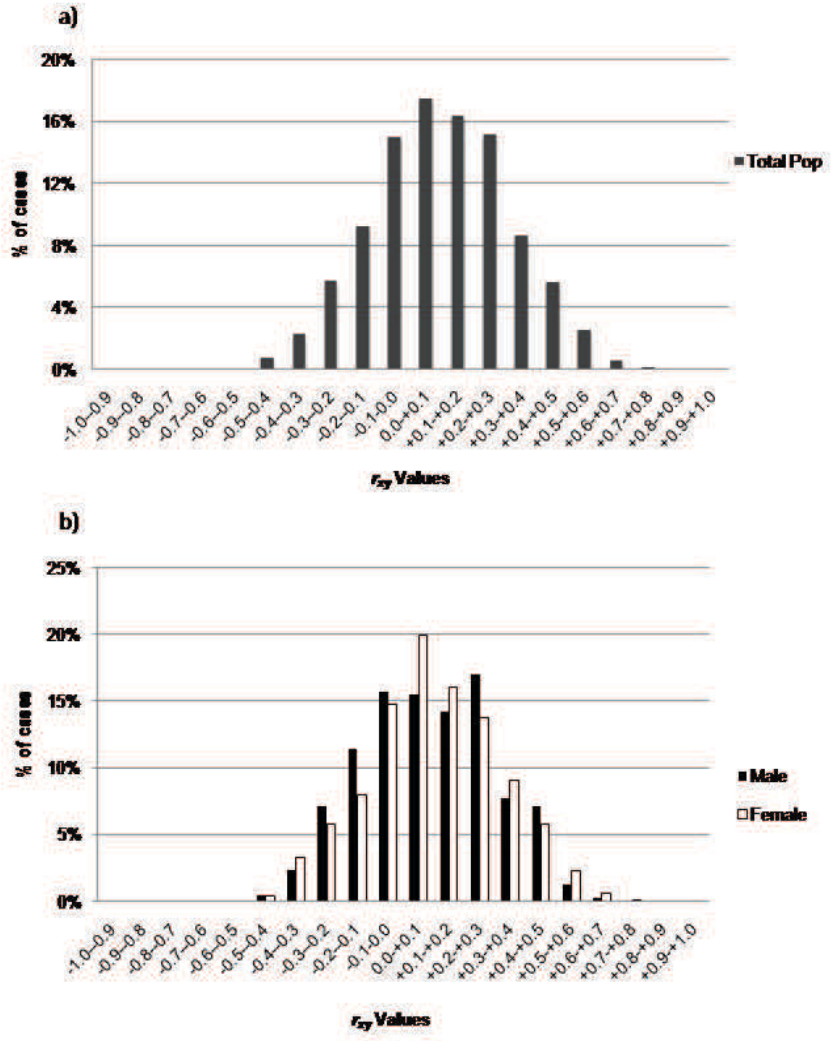
	Summer			**Winter**		
Both sexes						
Year	*R_{xy}*	*P-value*	*N*	*R_{xy}*	*P-value*	*N*
2002	0.033	0.439	8	-0.037	0.365	10
2003	-0.028	0.366	27	-0.196	**0.009**	25
2004	-0.009	0.443	30	0.049	0.220	21
2005	-	-	-	-	-	-
2006	-	-	-	-0.047	0.291	14
2007	-0.121	0.230	12	-	-	-
2008	-0.102	0.276	9	-0.156	0.205	8
2009	-0.036	0.280	28	-	-	-
2010	-0.156	0.207	8	-	-	-
Males						
Year	*R_{xy}*	*P-value*	*N*	*R_{xy}*	*P-value*	*N*
2002	0.466	0.503	3	0.422	0.118	5
2003	0.201	**0.036**	12	-0.062	0.309	9
2004	0.039	0.426	10	0.112	0.402	6
2005	-	-	-	-	-	-
2006	-	-	-	-0.058	0.194	7
2007	-0.366	0.075	6	-	-	-
2008	-0.255	0.502	3	-	-	-
2009	0.105	0.293	10	-	-	-
2010	-			-	-	-
Females						
Year	*R_{xy}*	*P-value*	*N*	*R_{xy}*	*P-value*	*N*
2002	-0.328	0.140	5	-0.324	**0.050**	5
2003	0.180	0.347	15	-0.216	**0.033**	16
2004	-0.085	0.176	20	-0.075	**0.030**	15
2005	-	-	-	-	-	
2006	-	-	-	-0.344	0.094	7
2007	-0.388	0.130	6	-	-	-
2008	0.180	0.347	6	-0.106	0.310	6
2009	-0.114	0.111	18	-	-	-
2010	-0.394	0.075	6	-	-	-

Table 3. Distribution of the relatedness (r_{xy}) in the different overlap classes during summer and during winter. Population was separated in: All, dyads occurred in the complete dataset; FF dyads among females dataset; MM dyads among males dataset; FM dyads among inter-sex pairs.

		SUMMER			WINTER		
		FF	*MM*	*FM*	*FF*	*MM*	*FM*
	N°pairs	410	144	505	176	56	202
O_0	*Mean*	0.075	0.054	0.095	0.053	0.03	0.08
	SE	0.011	0.018	0.010	0.016	0.024	0.015
	N°pairs	66	23	82	53	12	54
O_1	*Mean*	0.123	-0.057	0.149	0.15	0.069	0.18
	SE	0.030	0.040	0.024	0.029	0.057	0.030
	N°pairs	25	15	49	42	15	62
O_2	*Mean*	0.107	0.125	0.072	0.12	0.141	0.154
	SE	0.048	0.040	0.031	0.035	0.047	0.028

Legends figures

Figure 1 Distribution of rxy coefficient of relatedness (Queller& Goodnight, 1989) in 69 roe deer captured from 2002 to 2010 in Alpe di Catenaia, Italy. a) Overall distribution, b) separated distributions of males and females



Conclusions

Conclusions

The environment can shape the genetic variability of a species in a number of ways. The influence of environmental variables on genetic resources and diversity patterns could be detected at multiple geographical scales in relation to different biological aspects of the species.

The present study tried to explain genetic patterns observed in two European ungulate species in relation to multiple environmental factors, with natural or human origin, which the species are inevitably subjected to. Modifications in environmental conditions at a wide geographical scale (e.g. continental) are likely to affect the biology and genetics of natural populations over a long time period and their effects are likely to be visible for thousands of generations. Firstly, climatic variations influence the distribution of species, which have the ability to colonize areas with very different environmental characteristics, producing local adaptations at population level. The phylogeography studies such phylogenetic patterns of species in the context of their geographical distribution (Avise et al. 1987). In other words, it is concerned to geographically contextualize genetic data to assess the relationship between geographical phenomena, the distribution of species'diversity and speciation (Avise et al. 1987, 2000). The current phylogeography of species inhabiting temperate regions seemed to be congruent with an outstanding role of climatic Quaternary fluctuations (Hewitt 2004). Such trend was also identified in the European wild boar (**CHAPTER 1),** despite of the substantial alterations operated by man (translocations, hybridizations, reintroductions). Climatic modifications during the Last Glacial Maximum (LGM) resulted in the disappearance of many species from their former range and/or their retirement into southern refugia (Iberian and Italian peninsula, Balkans and southern of France), from which they naturally recolonized Europe during post-glacial period (Hewitt 2004). Also the genetic diversity of wild boar populations in Europe resulted to have been shaped by past climatic events, and particularly by the LGM, as other ungulates in Europe, such as roe deer (*Capreolus capreolus*, Sommer et al. 2009) and red deer (*Cervus elaphus*, Banks et al., 2008; Sommer et al., 2008). Indeed, genetic patterns showed a south-north gradient of genetic diversity, with high levels of variation in southern populations that substantially decreases towards northern populations, confirming the important role as genetic reservoirs of southern areas during the LGM, accordingly to fossils remains (Sommer & Nadachowski, 2006). A high genetic diversity was found mainly in the Italian peninsula, with the highest values of haplotype and nucleotide diversity, probably originated by a combination of factors: on one hand a strong isolation of the Italian population, due to the Alps, during late Pleistocene glaciations, with maintenance of a high population size; on the other hand, recent translocations from Central and Eastern Europe for hunting purposes increasing the overall variation. Two different haplogroups, E1 and E2, were observed in Italy, the second of which is present only in the Italian peninsula and in Sardinia

and, from the analysis of ancient materials (Larson et al. 2005), seems to have not dispersed to the rest of Europe during the last post-glacial. This evidence would suggest a marginal role of the Italian refugium to the recolonization of Europe. Actually, on the basis of phylogenetic analysis and haplotype sharing, refugia in Italy and south-western France were seemingly the starting point for the Central Europe recolonization, the Balkans for north-eastern Europe, while the Iberian refugium seems not have contributed in a significant way to the post-glacial repopulation. The absence of haplogroup E2 in the rest of Europe can be attributed to a low frequency in populations that promoted the leading-edge post-glacial population expansion, and/or to successive drift events, due to demographic oscillations. In conclusion a latitudinal gradient of genetic diversity was found in the wild boar European populations, as in the majority of mammals (Adams & Hadly, 2013). As proved in the paper, the effect of latitude mostly reflects the climatic suitability to wild boar during the LGM, that was estimated using modelling palaeoclimatic data and species' preference by an ecological niche approach (Phillips et al. 2006). It is important to highlight here how mitochondrial DNA can retain the signature of so ancient events, even if the European populations were perturbed by local extinctions and human manipulation (translocations, hybridization). We argued that the latter events have only modified genetic variation at a small scale, but have not removed the mascoscopic phylogeographic pattern.

Large fluctuations in climatic and environmental conditions are the basis of the large inter and intra-population genetic and phenotypic variability in biological systems (Aspöck 2008). Nevertheless, an appreciable degree of genetic variation can also be found among populations of the same species living under very similar climatic conditions, but with a low degree of gene flow due to habitat modifications or the presence of physical barriers. Such form of isolation leads to genetic discontinuity within the population, breaking up the population into more distinct genetic units that could be subjected to genetic impoverishment till local extinction (Groombridge, 1992; Heywood, 1995). To understand patterns of genetic variation of populations in relation, not only to large climatic modifications, but also in relation to local environmental features, the genetic differentiation within an isolated wild boar population inhabiting Sardinia was analyzed (**CHAPTER 2)**. Such population is considered a distinct subspecies (*Sus scrofa meridionalis* Major 1883), on the basis of both morphological and genetic evidences, has experienced an appreciable genetic introgression from domestic pigs and continental wild boar, and is interested by a relevant genetic structure (Scandura et al. 2011). The previously described genetic structure, found by Scandura et al. 2011, was confirmed and explained as a consequence of the interaction between land use, and specifically the presence of unsuitable environments and recent anthropogenic landscape modifications in Sardinia, that can effectively limit

animal movements across the island. In particular, the unsuitable Campidano plain and the main motorway (SS131) were identified as main causes of substructure, even if a certain degree of individuals' migration between subpopulations was revealed by the number of car impacts along the SS131 (2001-2011, Autonomous Region of Sardinia, unpublished data). Actually genetic abrupt discontinuities that separate the population into three subpopulations (one in the north-west, one in the south-west and one in the east of Sardinia, Scandura et al. 2011) match the previously mentioned environmental elements. Generally, presence of barrier to gene flow and the consequent population separation into smaller groups leads to negative effects on the genetic variability of populations, with its impoverishment by inbreeding and the loss of rare allelic variants (Frankham et al. 2002). Conversely, in the Sardinian wild boar population physical barrier to animal movements has been so far effective in avoiding the spread of introgressed genes from the eastern subpopulation to the rest of the island and probably safeguarded the genetic integrity of the west subpopulations (seemingly the purest ones). For a conservation point of view, anthropogenic barriers are mostly associated to negative effects on the genetic variability of populations, but in such case the presence of the motorway plays a positive role in the maintenance of genetic variability in local Sardinian. As mentioned, detrimental genetic effects due to isolation cannot be excluded and should be investigated.

Hence, human action could create changes in the structure and genetic variability both indirectly, causing environmental changes that affect gene flow between demes, or directly, through translocations and reintroductions to restore or increase the population size, which are very likely to introduce new allelic variants in the populations. After the II World War, many European ungulates had disappeared in various parts of the continent. Reintroductions events were necessary to prevent the complete disappearance of many species caused by overhunting and sudden habitat fragmentation. Mostly, management practices have inevitably led to mix different gene pools coming from different geographical areas, each with their allelic variants adapted to the local environment, producing hybridization and extended loss of genetic variability (Allendorf et al. 2001). In the **CHAPTER 3** the following aspects were evaluated:

i) the degree of genetic admixture between an Italian native roe deer gene pool (*C. c. italicus*) and exotic lineages introduced by man into three areas of central Italy. Indeed, in central regions, the recovery of the species in the last half century has been favoured by a number of releases (Mattioli 1994; Perco & Calò 1995), mostly realized with the northern subspecies (*C. c. capreolus*), thus threatening the genetic integrity of the endemic one (Vernesi et al. 2002; Randi et al. 2004; Gentile et al. 2009; Mucci et al. 2012;

ii) the diffusion and relative success of native versus reintroduced genotypes, evaluated as the proportion of admixture between the two gene pools in the contact zone.

Noticeably, although many generations had passed from reintroductions, genetic admixture was largely incomplete. Both mtDNA and microsatellite results revealed a clear '*italicus*' gradient decreasing from the area of diffusion of the native form to the area of reintroduction of non-indigenous deer. Several factors such as time of the introduction, number of released individuals, habitat and landscape matrix could have played a role in generating this gradient, also producing different situations within the different study areas. Actually, in addition to genetic differences, it was suggested that the two gene pools may also show different ecological requirements, presumably due to the different habitats in which they have adapted (Montanaro et al. 2004). Specifically, the Italian roe deer seems to prefer deciduous woodland, scrubland and maquis vegetation, abundant in Central and Southern Italy. Despite this, habitat does not seem to have been the major determinant of the spread and degree of admixture of the two gene pools, but the main role is likely to have been played by the human management of the species. Nevertheless, a possible role of landscape features as physical barrier (rivers, main roads) was suggested to explain the observed spatial distribution of genetic variants in at least one case and cannot be excluded in the others, in spite of the occurrence of suitable habitats for the species. Thus, landscape features should be taken into account also in assessing the effect of management practices and the introgression risk of native populations. In this frame, the importance of a genetic assessment prior to any management activity and the opportunity of a genetic monitoring in presumed contact zones between different gene pools have to be stressed.

Inevitably, the presence of particular habitats and environmental variables influence individual movements, with consequences that are reflected in the spatial and social behaviour of populations (e.g. Coulon et al. 2004, Jepsen & Topping 2004). Animal populations of the same species seem to react differently in relation to landscape matrices, modifying the social interaction between individuals (Lott 1991) and consequently the resultant genetic structure. In this context, a locally forest roe deer population was used as model species to evaluate the fine-scale genetic structure in relation to environmental features, social behaviour, mating strategy and population parameters (**CHAPTER 4)**. Indeed, environmental matrices could influence the tendency of dispersal in individuals, developing a high variability in the social structure and spatial behaviour among populations (Hewison et al. 1998, 2001; Cargnelutti et al. 2002), which could be reflected by genetic data. In our study area, in Central Italy, we detected a similar pattern of genetic relatedness between sexes and an absence of correlation between genetic and spatial distance both in males and females. Moreover, individuals of both sexes

tended to share their home range with relatives in summer as well as in winter, but in summer (during the mating season) relatedness in close male-female dyads was minimized. Such genetic pattern could be influenced by the environmental context in which the population lives, a continuous forest without barriers that might limit gene flow, and fits the prediction that in high-density and high-quality habitats both sexes should behave similarly and show limited dispersal (Wahlström & Liberg 1995). Thus, habitat constraints could be at the basis of different spatial organization and of population-specific solutions to limit inbreeding and to shuffle genes in the population.

In conclusion, the integration of environmental and climatic features in the description of current genetic patterns of ungulate species is a remarked point to understand past and ongoing processes with implications in ecology, evolution and conservation genetics (Manel et al. 2003). Indeed, often genetic discontinuities seem to overlap with physical barriers or appear to be the result of past events, both of natural and human origin. Therefore a multidisciplinary approach that amalgams genetic data with information on the environment and on the demographic history of populations has to be recommended in evaluating proper management and conservation policies (van Strien et al. In press). Hence, the Landscape Genetics comes to help us, providing conceptual and analytical tools essentially for the prediction of dispersion, migration and gene flow mediated by the environmental characteristics at different geographical scales (Segelbacher et al. 2010). Habitat fragmentation and climatic fluctuations could affect negatively the worldwide biodiversity through modifications in population dynamics (Fischer & Lindenmayer 2007). Environmental discontinuities, associated to interrupted resource distribution and microclimate variation, lead a constriction of suitable habitat and a decreasing in functional connectivity between populations (Kindlmann & Burel 2008). This translates into a reduced probability of individual dispersal and persistence of the species, increasing the risk of local extinction (With & King 1999).

Investigating the genetic composition of a species in relation to external modifications is a complex issue because of the many changes that can occur over time, the interactions between different events and the diversified responses that individuals or populations may have in relation to such changes. Despite the great applicability of the Landscape Genetics in conservation genetics, few studies have addressed the success of the Landscape Genetics approach in conservative management of important species (Epps et al. 2005; Vignieri et al. 2005; Riley et al. 2006; Segelbacher et al. 2008). A multi-factorial and multi-scale approach, going from the study of environmental influence on a wide scale to individual-specific behavioural choices, passing through the evaluation of species' ecological preferences, dispersion patterns and ability in colonize new habitats or recolonize areas previously

occupied, can contribute to increase the knowledge about the best practices to adopt to preserve the genetic component of biodiversity.

References

Adams R & Hadly E (2013) Genetic diversity within vertebrate species is greater at lower latitudes. Evolutionary Ecology, 27 133–143.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001). The problems with hybrids: setting conservation guidelines. Trends Ecol Evol 16: 613–622.

Aspöck H 2008 Postglacial formation and fluctuations of the biodiversity of Central Europe in the light of climate change Parasitology Resources 103:S7-S10.

Avise JC, Reeb CA, Saunders NC (1987) Landscape genetics as a tool for conservation planning: predicting the effects of landscape change on gene flow.

Avise JC (2000) Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA. (447 pp.)

Banks WE, d'Errico F, Peterson AT, Kageyama M & Colombeau G. (2008) Reconstructing ecological niches and geographic distributions of caribou (*Rangifer tarandus*) and red deer (*Cervus elaphus*) during the Last Glacial Maximum. Quaternary Science Reviews, 27, 2568–2575.

Cargnelutti B, Reby D, Desneux L, Angibault JM, Joachim J, Hewison AJM (2002) Space use by roe deer in a fragmented landscape. Some preliminary results. Revue d'Ecologie (Terre Vie) 57: 29-37.

Coulon, A., Cosson, J.F., Angibault, J.M., Cargnelutti, B., Galan, M., Morellet, N., Petit, E., Aulagnier, S., Hewison, A.J.M. (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: An individual-based approach. Molecular Ecology, 13 (9), pp. 2841-2850.

Epps CW, Palsbøll PJ, Wehausen JD, Roderick GD, Ramey RR, McCullough DR (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters 8:1029–1038.

Fischer J & Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. Global Ecology and Biogeography 16:265-280

Frankham R, Ballou , JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge: Cambridge University Press. 640.

Gentile G, Vernesi C, Vicario S, Pecchioli E, Caccone A, Bertorelle G et al (2009). Mitochondrial DNA variation in roe deer (Capreolus capreolus) from Italy: evidence of admixture in one of the last C.c.italicus pure populations from central-southern Italy. Ital J Zool 76:16-27.

Groombridge B. 1992. Global Biodiversity. Chapman & Hall, London

Heywood V. H. 1995. Global Biodiversity Assessment. Cambridge University Press, Cambridge, UK.

Hewison AJM, Vincent JP, Reby D (1998) Social organisation of European roe deer. In: Andersen, R, Duncan, P, Linnell, J.D.C. (Eds). The European roe deer: The biology of success. Oslo, Norway: Scandinavian University Press, pp. 189-219

Hewison AJM, Vincent JP, Joachim J, Angibault JM, Cargnelutti B, Cibien C (2001) The effects of woodlands fragmentation and human activity on roe deer distribution in agricultural landscapes. Can J Zool 79: 679-689

Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B: Biological Sciences, 359, 183–195.

Jepsen, J.U., Topping, C.J. (2004) Modelling roe deer (Capreolus capreolus) in a gradient of forest fragmentation: Behavioural plasticity and choice of cover. Canadian Journal of Zoology, 82 (9), pp. 1528-1541.

Lott DF (1991) Intraspecific variation in the social systems of wild vertebrates. Cambridge University Press, Cambridge.

Kindlmann P, Burel F (2008). Connectivity measures: a review. Landscape Ecology 23:879-890

Larson, G., Dobney, K., Albarella, U., Fang, M. Y., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L. & Cooper, A. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science, 307, 1618-1621.

Mattioli S (1994) Le reintroduzioni di Cervidi in Italia, con particolare riferimento all'Appennino. In: Inglisa M., cur., La fauna appenninica e la sua conservazione. Airone e Corpo Forestale dello Stato, Caramanico.

Montanaro P, Vicari N, Marcone F, Riga F & Focardi S (2004) – Roe deer in a Mediterranean habitat: habitat use, selection and suitability. – In: Arianoutsou M. e V.P., Papanastasis (eds) Ecology, conservation and management of Mediterranean climate ecosystems. Proceedings of the 10th international conference on Mediterranean climate ecosystems, April 25-may 1, 2004, Rhodes, Greece. Millpress (CD-Rom).

Mucci N, Mattucci F, Randi E (2012). Conservation of threatened local gene pools: Landscape genetics of the Italian roe deer (*Capreolus c. italicus)* populations. Evol Ecol Res 14: 897-920.

Olden JD, LeRoy Poff N, Douglas MR, Douglas ME, Fausch KD (2004). Ecological and evolutionary consequences of biotic homogenization. TREE 19: 18-24

Perco F, Calò CM (1995). The status of roe deer in Italy. In: Wotschikowsky U., cur., Proceedings of the 2nd European Roe Deer Meeting. Wildbiologische Gesellschaft, München.

Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190, 231-259

Randi E, Alves PC, Carranza J, Milosevic-Zlatanovic S, Sfougaris A, Mucci N (2004). Phylogeography of roe deer (Capreolus capreolus) populations: the effects of historical genetic subdivisions and recent non-equilibrium dynamics. Mol Ecol 13:3071-3083.

Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne RK (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. 15:1733–1741.

Scandura M, Iacolina L, Cossu A, Apollonio M (2011). Effects of human perturbation on the genetic make-up of an island population: the case of the Sardinian wild boar. Heredity 106: 1012-1020.

Segelbacher, G., Tomiuk, J., & Manel, S. (2008). Temporal and spatial analyses disclose consequences of habitat fragmentation on the genetic diversity in capercaillie (*Tetrao urogallus*).

Molecular Ecology 17: 2356–2367.

Segelbacher, G. et al. 2010. Applications of landscape genetics in conservation biology: Concepts and challenges. Conservation Genetics 11: 375-385.

Sommer RS & Nadachowski A (2006) Glacial refugia of mammals in Europe: evidence from fossil records. Mammal Review 36 4: 251-265.

Sommer RS, Zachos FE, Street M, Jöris O, Skog A & Benecke N (2008) Late Quaternary distribution dynamics and phylogeography of the red deer (*Cervus elaphus*) in Europe. Quaternary Science Reviews, 27: 714–733.

Sommer RS, Fahlke JM, Schmölcke U, Benecke N, Zachos FE (2009) Quaternary history of the European roe deer *Capreolus capreolus*. Mammal Review, 39, 1–16.

van Strien MJ, Keller D, Holderegger R, Ghazoul J, Kienast F, Bolliger J. (In press) Landscape genetics as a tool for conservation planning: predicting the effects of landscape change on gene flow. Ecological Applications, http://dx.doi.org/10.1890/13-0442.1

Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G (2002). The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. Mol. Ecol. 11:1285-1297.

Vignieri SN (2005). Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). Molecular Ecology 14:1925–1937.

Wahlström LM, Liberg O (1995) Patterns of dispersal and seasonal migration in roe deer (Capreoluscapreolus). J Zool 235: 455-467

With KA & King AW (1999) Extinction Thresholds For Species in Fractal Landscapes, Conservation Biology: 13 (2): 314-326.

Aknowledgements

Eccoci qua… è arrivato il momento più divertente di questa tesi: i ringraziamenti.

Ringrazio innanzitutto il Prof. Marco Apollonio per avermi dato la possibilità di lavorare su quello che più mi piace e con persone fantastiche.

Siete tanti da ringraziare e contrariamente a quanto fanno tutti lo vorrei fare uno ad uno con la certezza di non dimenticarmi nessuno.

Ringrazio di cuore Massimo per aver lavorato giorno e notte con me nonostante i numerosi impegni. Grazie per la tua pazienza e soprattutto per essere stato sempre disponibile e aver avuto parole di conforto proprio quando più servivano. Non potevo avere tutor migliore.

Lauretta… tra i miei ringraziamenti non puoi mancare di certo! Grazie per tutte le volte che sei venuta in mio aiuto, in qualsiasi momento della tua giornata e da qualsiasi parte del globo. Grazie per le risate quotidiane nascoste dietro i nostri schermi del pc (più tu che io!!!). Grazie per i numerosi viaggi "di lavoro" che abbiamo fatto insieme, Parigi rimarrà sempre nel mio cuore! Grazie per tutte le volte che hai ascoltato i miei sfoghi in silenzio, senza fare domande. La tua presenza e amicizia ha significato molto per me e sarà così anche in futuro.

Ringrazio tutti quelli dell'"ufficio di là":
Antonio, grazie per aver sempre una parola carina nei miei confronti ogni mattina… "nonostante tutto"…
Antonello, grazie per avermi fatto apprezzare il piacere di condividere la scrivania con qualcuno..
Andrea, grazie per tutte le volte che hai risposto con il sorriso alle mie "tremende ed a volte insensate domande su GIS"
Elena, grazie mille per aver condiviso questa esperienza mistica di fine dottorato con me!E soprattutto grazie di cuore per aver capito quando le cose non andavano e di avermi tirato su nei momenti bui. Ricordati che una missione ci attende…
Paolo Ma, grazie per avermi fatto capire che a volte il silenzio è la salvezza
Paolino, grazie per la tua disponibilità e le tue parole gentili arrivate al momento giusto
Valentina, grazie per avermi portato al livello in cui sono ora..senza la tua ferrea disciplina e senza una fase iniziale di sottrazione punti non penso sarei mai riuscita a superare Lauretta

Roberta, grazie per tutte le volte che sei stata gentile e disponibile

Ovviamente non posso mancare di ringraziare la sezione distaccata dell'ufficio "dei Nordici":

Stefano, grazie per avermi fatto capire che non esiste solo il bianco e il nero, ma che le sfumature di colore sono altrettanto belle e neanche così tanto complicate da capire ed apprezzare

Nico, grazie per le serate passate insieme tra birre, animali totem e promesse non mantenute...ma mi rifarò, PROMESSO!

Francesca, grazie per la tua disponibilità, gentilezza e soprattutto pignoleria nordica

Grazie anche a Mariuccio per il suo sostegno "da lontano" e per aver condiviso con me molte altre cose, al di fuori di questo mondo.

Vorrei inoltre ringraziare gli altri miei compagni di avventura, Antonella per aver condiviso le gioie e i dolori della burocrazia italiana ed essere stata sempre disponibile a risolvere ogni mia domanda e ogni mio dubbio e Fabietto per aver accettato sempre con il sorriso i miei scherzi e le mie battute… e per essere stato un buon "amico di laboratorio"

Un dovuto ringraziamento va anche ai "miei tesisti" Carlo, Antonio, Agostino, Mauro e…… anche tu Mario sì (che ti piaccia o no!!!). Grazie per tutte le volte che mi avete fatto arrabbiare, ma anche sorridere… e soprattutto grazie per la fiducia che avete riposto in me e il bel rapporto di amicizia che abbiamo creato.

Ringrazio inoltre tutte le persone che fanno parte, o che ne hanno fatto parte in passato, del "cast di Zoologia": Daria S, Francesca B, Piero C, Tiziana L, Alberto A, Marco C e tutti i Professori per aver reso piacevole e formativa la mia esperienza e avermi fatto appassionare ancora di più alla biologia.

Ringrazio di cuore anche Stèphane Joost e tutto il laboratorio LASIG dell'EPFL per avermi ospitato due mesi e per avermi insegnato molto. Ringrazio Sylvie, Ivo, Kevin, Timothè e tutti i ragazzi che ho conosciuto durante il mio soggiorno in Svizzera, per avermi fatto sentire a casa.

Un ringraziamento speciale va a Gian Luca per aver capito quando le cose non andavano, per avermi ascoltato, capito, consolato e fatto ridere quando ne avevo più bisogno… Molto probabilmente senza il tuo supporto avrei ceduto… Grazie di cuore, non lo dimenticherò mai.

Grazie a tutti i miei amici che in un modo o nell'altro mi hanno aiutato e anche se non fanno parte di questo strano mondo non se la sono presa se li ho trascurati un po'..

Grazie Katia per essere l'amica più speciale che si può desiderare e grazie Daria, lo sai senza te non sarei mai riuscita a superare quest'ultimo anno di follia.. Grazie a Melania, Ilaria, Sonia, Valentina, Antonella, Claudio che solo con una parola e con un gesto sono capaci di cambiare in meglio le mie giornate e fino ad ora ce l'hanno fatta.

Un ultimo ringraziamento va a una persona speciale, Daniele, che mi è stato sempre affianco, iniziando e finendo con me questa particolare avventura che ci ha portato a vivere nuove esperienze ed affrontare nuove sfide. Finalmente dopo tredici anni insieme sei riuscito a trasferire un po' della tua flemma, del tuo positivismo e della tua solarità che mi hanno permesso di non far esplodere tutto. Grazie di tutto!

Ringrazio infine la mia famiglia che ha capito il periodo impegnativo passato e non se l'è presa se l'ho trascurata un po'.

GRAZIE…