

# Phylogeographic patterns of *Capreolus capreolus* in the centre of the Iberian peninsula

F. Horcajada, L. Alcaraz, I. Barja, A. Machordom

Horcajada, F., Alcaraz, L., Barja, I., Machordom, A., 2018. Phylogeographic patterns of *Capreolus capreolus* in the centre of the Iberian peninsula. *Animal Biodiversity and Conservation*, 41.2: 415–425, Doi: <https://doi.org/10.32800/abc.2018.41.0415>

## Abstract

*Phylogeographic patterns of Capreolus capreolus in the centre of the Iberian peninsula.* One hundred and one samples of muscle tissue were obtained from roe deer in the centre of the Iberian peninsula. We compared the sequences of the control region (D-loop) of the mitochondrial DNA of these samples with those obtained in previous studies. Adding the information from microsatellite markers and derived genetic parameters to study the population structure, we found a philopatric structure, with females maintaining mitochondrial haplotype diversity, while males showed a pattern of genome homogenization. The population can thus be considered panmictic. Different times of palaeohistory of the species may explain these results: glacial–interglacial stages of the Pleistocene and the reduction and recovery of populations throughout the 20th century.

Key words: *Capreolus capreolus*, Central Spain, Microsatellites, Mitochondrial, Phylogenetic

## Resumen

*Patrones filogenéticos de Capreolus capreolus en el centro de la península ibérica.* Se obtuvieron 101 muestras de tejido muscular de ejemplares de corzo en el centro de la península ibérica. Se compararon las secuencias de la región control (D-loop) del ADN mitocondrial con las obtenidas en otros estudios anteriores. Al añadir la información relativa a los marcadores microsatélites y los parámetros genéticos derivados para estudiar la estructura de la población, se constató la existencia de una estructura filopátrida, en la que las hembras mantenían la diversidad de haplotipos mitocondriales, mientras que los machos seguían un patrón de homogeneización del genoma. Por lo tanto, la población puede considerarse panmíctica. Diferentes épocas de la paleohistoria de la especie pueden explicar estos resultados: las etapas glaciales e interglaciales del Pleistoceno, por un lado, y una marcada reducción y recuperación de las poblaciones a lo largo del siglo XX, por otro.

Palabras clave: *Capreolus capreolus*, España central, Microsatélites, Mitocondrial, Filogenética

Received: 20 IV 17; Conditional acceptance: 01 XII 17; Final acceptance: 28 II 18

Fernando Horcajada Centro de Investigación, Seguimiento y Evaluación, Parque Nacional de la Sierra de Guadarrama, ctra. M-604 km 28, 28740 Rascafría, Madrid, Spain.– Lourdes Alcaraz, Annie Machordom, Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (MNCN–CSIC), c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain.– Fernando Horcajada, Isabel Barja, Departamento de Biología, Unidad de Zoología, Facultad de Ciencias, Universidad Autónoma de Madrid, c/ Darwin 2, Campus Universitario de Cantoblanco, 28049 Madrid, Spain.

Corresponding author: F. Horcajada. E-mail: horcajadas@gmail.com

## Introduction

The roe deer (*Capreolus capreolus*) was part of the fauna of the Iberian peninsula during the last glaciations (Gliozzi et al., 1997; Hufthammer and Aaris-Sørensen, 1998). Some authors claim that the populations followed different routes of dispersal and recolonization toward northern Europe after the Ice Age (Hewitt, 1999), while others (Lorenzini and Lovari, 2006) propose that the Iberian peninsula remained an isolated geographical area, not acting as a main source for postglacial recolonization (Sommer and Zachos, 2009).

*Capreolus capreolus* was present in the Middle Pleistocene of Central Spain, where paleontological remains have been found (Buitrago, 1992). In addition to functioning as a shelter in Pleistocene times, the Sierra de Guadarrama area might have played a fundamental role in safeguarding populations of roe deer (Tellería, 1999), although there are references indicating that populations of Central Spain underwent severe reductions in the 19<sup>th</sup> and early 20<sup>th</sup> century (Tellería and Virgós, 1997; Gortázar et al., 2000) that may have left their mark on their genetic parameters.

Among the few genetic analyses that exist on the Iberian populations, only the studies of Lorenzini and Lovari (2006), Lorenzini et al. (2014), Randi et al. (2004) and Royo et al. (2007) included some locations near the Sierra de Guadarrama.

With the aim of contributing to a comprehensive molecular analysis of the populations in the Sierra de Guadarrama we examined two types of markers, mitochondrial and nuclear (microsatellites), with the largest number of samples analysed to date in the Iberian peninsula, in order to increase our knowledge of the phylogeographic pattern of this species in Spain and dispel doubts about its heterogeneity, and to analyse the impact of isolation from recent centuries.

## Material and methods

### Study area

Samples were collected between 2002 and 2007 at 22 locations in the Sierra de Guadarrama, in the mountains north of Madrid, in the centre of the Iberian peninsula (fig. 1, table 1). Currently, four of these localities are within the National Park of the Sierra de Guadarrama (declared a National Park on June 25<sup>th</sup>, 2013) (fig. 1). Sampled locations are within public forests of *Pinus sylvestris* and *Quercus pyrenaica*.

### Sample collection and DNA extraction

We collected muscle tissue from 101 individuals distributed evenly throughout the study area. Samples were collected post-mortem from hunted deer or deer killed in road accidents. A small muscle biopsy was kept in absolute ethanol at -20°C until analysis was undertaken in the laboratory.

DNA was extracted using the standard phenol/chloroform protocol (Sambrook et al., 1989). Agarose

gels were used to test presence, concentration and possible degradation of the DNA extracted from the samples.

### Analysis of the mitochondrial DNA

We amplified the complete mitochondrial D-loop using the primers L-Pro and H-Phe (Jäger et al., 1992). Internal primers used from the monitoring region (D-loop) were L-362 and H-493 (Randi et al., 1998). The PCRs were conducted with 1–3 ml DNA, 5 ml of the corresponding buffer (with 10 x 2 mM MgCl<sub>2</sub>), 1 ml of dNTPs mix (10 mM), 0.8 ml of forward and reverse primers (10 mM), 0.3 ml Taq DNA polymerase (5 U/ml) (Biotools) and double-distilled water (ddH<sub>2</sub>O) for a total volume of 50 ml. The PCR cycles were: 94°C (4 min), 40 cycles of 94°C (45 s), 56°C (1 min), 72°C (1 min) and a final extension at 72°C (10 min). Amplifications were purified for sequencing both forward and reverse using BigDye Terminator and an ABI 3730 genetic analyser (SECUGEN, Madrid).

Chromatograms resulting from each specimen reaction were combined and primers were cleaned with Sequencher (Gene Codes Corporation). To compare the sequences obtained with those from other European populations, we downloaded GenBank available sequences representing the different haplotypes cited mainly by Wiehler and Tiedemann (1998), Versini et al. (2002), Randi et al. (2004) and Royo et al. (2007). We used the complete control region sequence of *Capreolus pygargus* and part of one of its subspecies, *Capreolus pygargus ochracea*, as an outgroup to root the phylogenetic tree.

Two treatments were used to analyse these mitochondrial DNA data, phylogenetic inferences and networks (using Haploviewer, <http://www.cibiv.at/~greg/haploviewer>) based on the haplotypes, their frequencies, and the localities where each sample was collected. For the phylogenetic inferences, we used maximum parsimony (using PAUP; Swofford, 2003). Both the complete matrix of data, and a matrix reduced to 336 characters were prepared, the latter to avoid the problem of missing data, due to the reduced length of some GenBank sequences.

A Mantel test was used to calculate the coefficient of association between genetic distance and altitude.

### Microsatellite analyses

We amplified 12 unlinked microsatellite loci (BM1706, BM757, BM848, CSSM39, HUI1177, NVHRT48, BMC1009, CSSM41, CSSM43, IDVGA29, IDVGA8 and OarFCB304) which had been previously used in other roe deer studies, with the concentrations there referred (Galán et al., 2003; Coulon et al., 2004). Forward primers were labelled with 3 different fluorescent dyes (6-FAM, HEX or NED). Fragment sizes were analysed on an ABI 3730 sequencer, and individuals were genotyped using the program GenMapper 3.1 (Do and Rahm, 2004).

We tested for deviation from Hardy-Weinberg equilibrium at each locus and over all using the

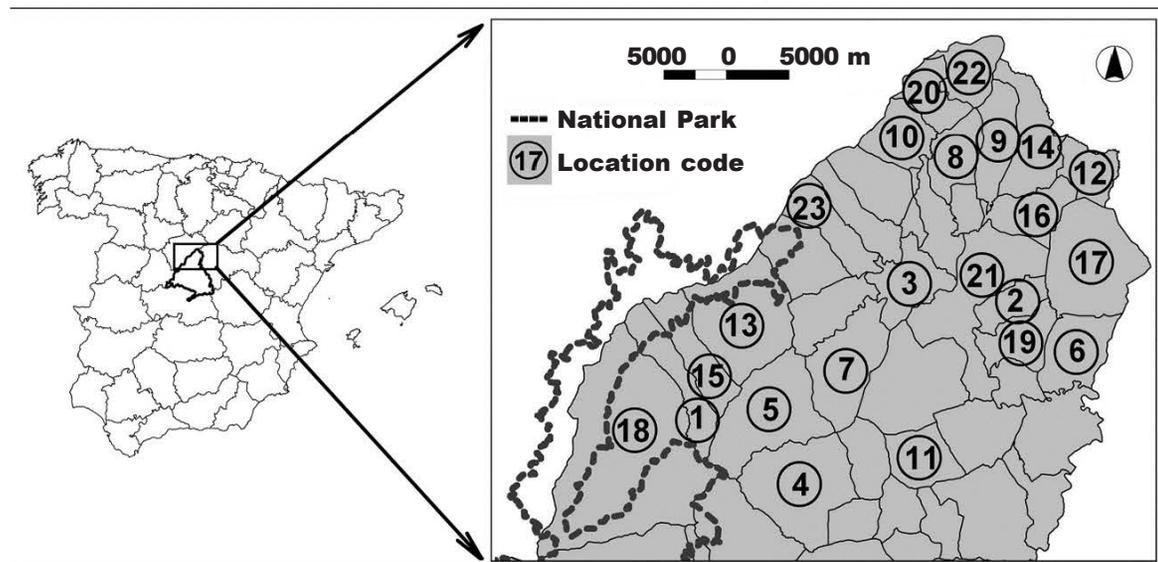


Fig. 1. Location of the study area (Sierra de Guadarrama, Madrid), in the centre of the Iberian peninsula, where the roe deer muscle samples were obtained. Circles contain the location code (table 1).

Fig. 1. Mapa de localización de la zona de estudio (Sierra de Guadarrama, Madrid) en la península ibérica donde se obtuvieron las muestras de músculo de corzo. Cada círculo contiene el código de la localidad (tabla 1).

Fisher's exact test and Markov chain algorithms (1,000 batches, 10,000 iterations), and  $F_{IS}$  (Weir and Cockerham, 1984) was calculated using the program Genepop 3.4 (Raymond and Rousset, 1995). To as-

sess the level of genetic diversity, we calculated the mean expected and observed heterozygosities using Arlequin 3.0 (Schneider et al., 2000). In addition, we calculated the molecular coancestry coefficient and

Table 1. Data on tissue samples collected from roe deer analyzed during the study: L, location code; N, number of samples; and H, haplotypes found.

Tabla 1. Datos sobre las muestras de tejido de corzo recogidas y analizadas durante el estudio: L, código de cada localidad; N, número de muestras; y H, haplotipos encontrados.

L	N	H	L	N	H
1	9	A, E	13	4	C
2	3	C	14	3	A, C
3	3	C	15	5	A, C
4	1	A	16	7	A, B, C
5	1	C	17	1	E
6	8	A, C	18	30	A, B, C, E
7	5	A, B	19	5	C
8	2	C	20	3	A
9	1	D	21	1	C
10	2	A, B	22	1	C
11	2	A, C	23	1	B
12	3	A, C			

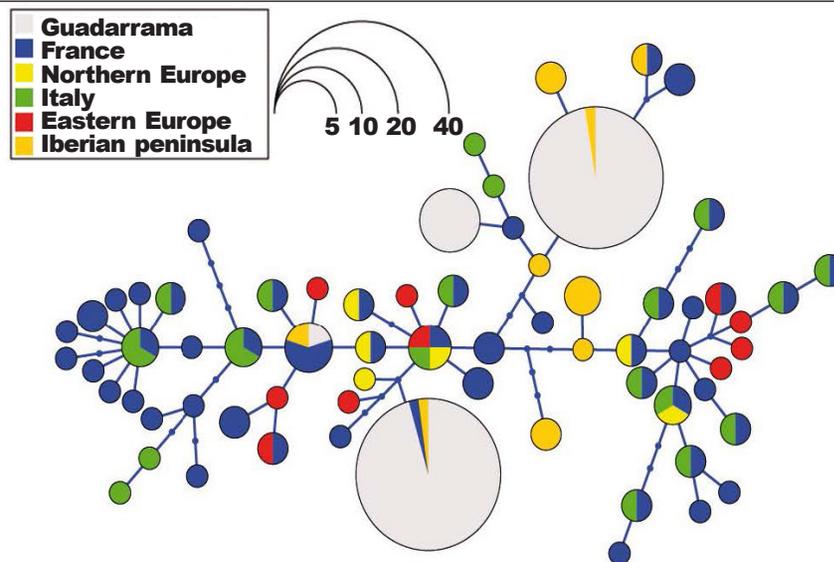


Fig. 2. Network based on 336 base pairs of the mitochondrial D-loop of roe deer analysed (Sierra de Guadarrama) and the available sequences of several European populations. Circles represent the different haplotypes and are proportional to their frequency among the analysed data, as indicated.

*Fig. 2. Red formada por los 336 pares de bases de la región D-loop del ADN mitocondrial del corzo analizados (Sierra de Guadarrama) y las secuencias disponibles de varias poblaciones europeas. Los círculos representan los distintos haplotipos y, como se indica, son proporcionales a su frecuencia relativa entre los datos analizados.*

PIC (Polymorphism Information Content) for each locus using the program Molkin 2.0 (Gutiérrez et al., 2005).

To define the genetic structure of the populations of roe deer in the study area we used the program Structure 2.1 (Pritchard and Wen, 2003). This program uses a Bayesian algorithm to group samples into genetically distinct clusters,  $K$ , based on the similarities between the genotypes of individuals. We tested  $K = 1-15$ , with 10 replicates for each  $K$ -level. The most likely number of clusters was identified using the posterior probability (Pritchard and Wen, 2003).

## Results

### Mitochondrial DNA analysis

The complete D-loop region (928–930 base pairs) was sequenced for the 101 studied samples (GenBank accession numbers MG760247 to MG760343). A maximum of only five gaps was necessary to align the complete data set, rendering a matrix length of 933 positions. We found 13 parsimony-informative characters in this matrix that exclusively contained the newly sequenced samples. Subsequently, sequences from GenBank were incorporated, including the different haplotypes previously found, most of them from partially sequenced D-loop, with lengths from 340–342, 427, 678 to the complete sequence, rea-

ching a matrix of 201 samples and 936 characters. In this last matrix, the number of parsimony informative characters increased to 62.

The 101 individuals analysed comprised five different haplotypes, A, B, C, D and E (table 1), grouped in four clusters. Differences between haplotypes A and E were due to 9 polymorphic positions in haplotype E, probably because of some degree of heteroplasmy. To avoid spurious results and reticulations, haplotype E was eliminated from the matrix to the following data treatments.

The network constructed showed the complex relationships between the samples from different locations (fig. 2). Usually in this kind of figure, all the haplotypes are included to show the relative weight of each haplotype and deduce the potential centre of dispersion. Here we have only represented the different haplotypes or those representatives of groups of haplotypes cited in the principal roe deer D-loop analyses (Wiehler and Tiedemann, 1998; Versini et al., 2002; Randi et al., 2004; Royo et al., 2007) plus the samples analysed here. This rendered a matrix of 68 haplotypes of 201 specimens. The large number of haplotypes in the Italian and French populations stands out, distributed throughout the network. The Iberian samples were situated in the centre of this representation, and their different haplotypes showed more similarity to other groups, than among them.

By analysing the phylogenetic relationship of the total matrix (201 specimens, 940 pb), the unique-

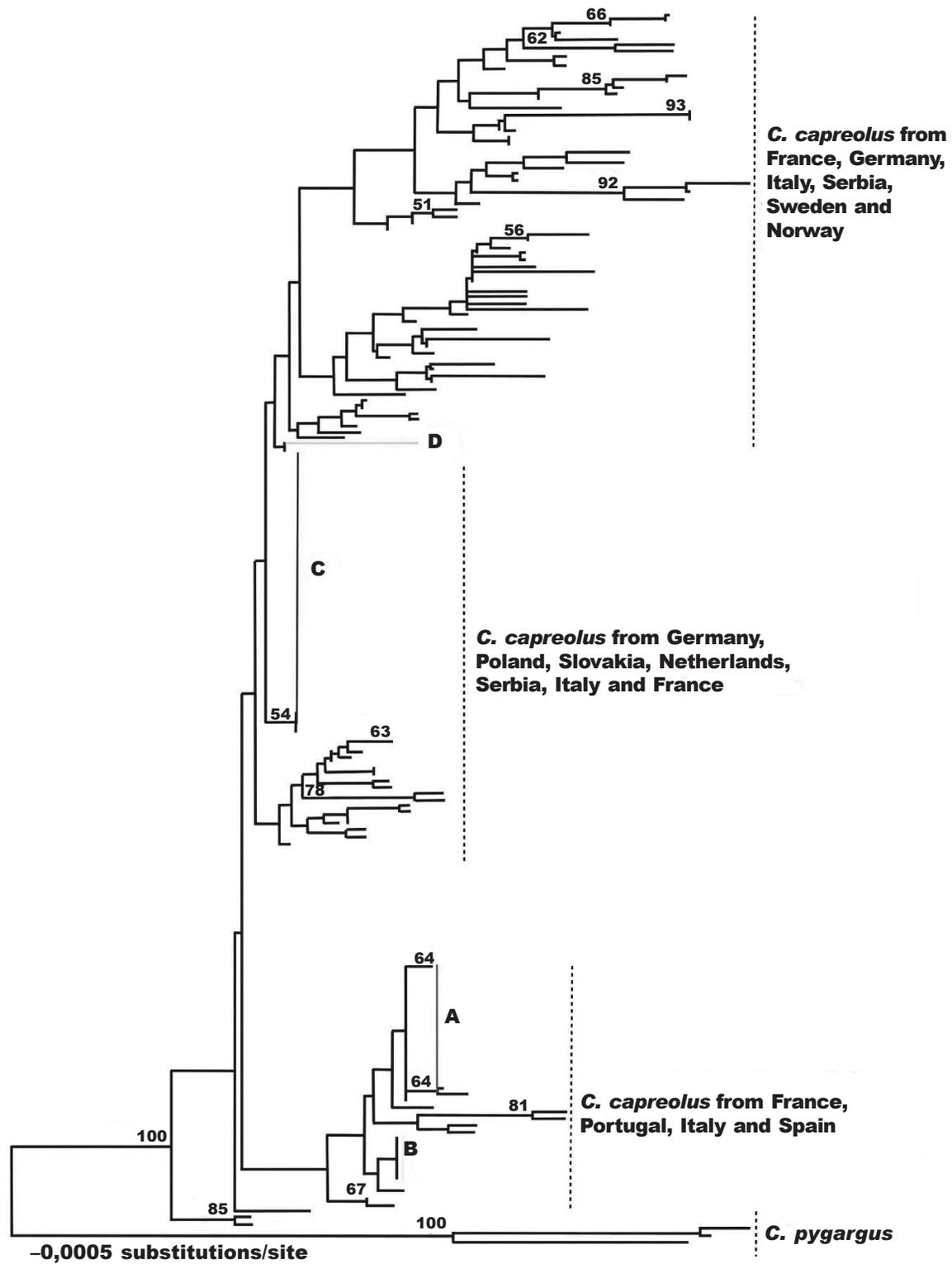


Fig. 3. Maximum parsimony phylogenetic tree for European roe deer. The numbers on the branches indicate bootstrap values above 50%. A, B, C and D are the haplotypes found in this study.

Fig. 3. Árbol filogenético basado en el principio de la máxima parsimonia para el corzo europeo. Los números sobre las ramas indican los valores de bootstrap superiores al 50%. A, B, C y D son los haplotipos encontrados en este estudio.

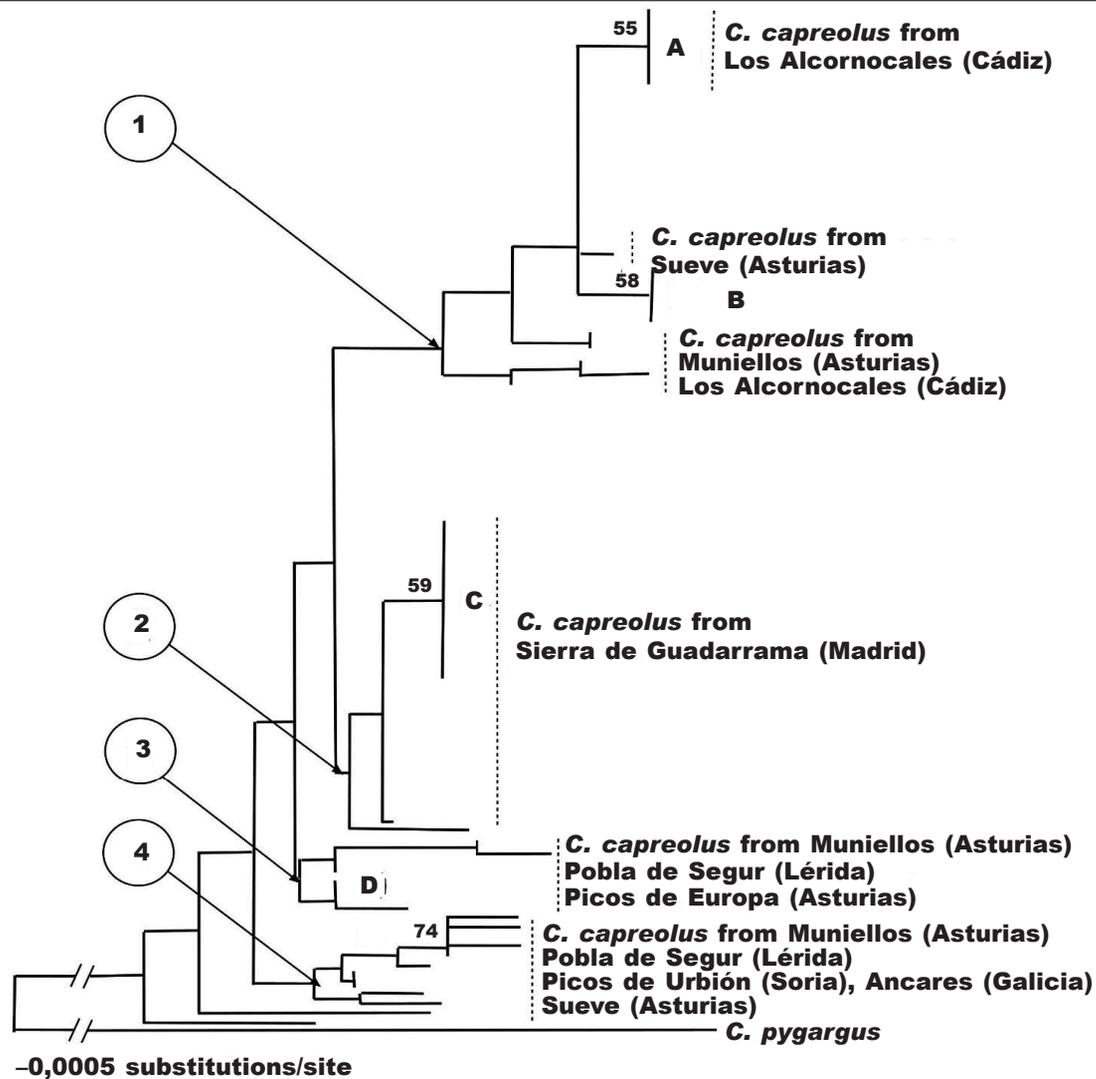


Fig. 4. Maximum parsimony phylogenetic tree with individuals analysed in this study together with the haplotypes found by Royo et al. (2007) for populations in the Iberian peninsula. The numbers in the circles indicate the four main nodes, and those located on the branches indicate bootstrap values of above 50%. A, B, C and D are the haplotypes found in this study.

Fig. 4. Árbol filogenético basado en el principio de la máxima parsimonia con los individuos analizados en este estudio junto con los haplotipos encontrados por Royo et al. (2007) para la península ibérica. Los números dentro de los círculos señalan los cuatro nodos principales y los situados sobre las ramas indican los valores de bootstrap superiores al 50%. A, B, C y D son los haplotipos encontrados en este estudio.

ness of some of the specimens from the Sierra de Guadarrama was confirmed and the relationships with some European haplotypes (fig. 3) was again demonstrated. Furthermore, this study showed the non-monophyly (common source) of Iberian haplotypes, since some appeared more related to other European haplotypes than with others found in the Sierra de Guadarrama. Fundamentally, lineages A and B were related to others analysed from Spain and Portugal, but also to some from France

and Italy. The C group was related to haplotypes from Germany, Poland, Slovakia, the Netherlands, Serbia, Italy and France, while the groups related to the unique roe deer with haplotype D were mostly from central and northern Europe (France, Germany, Italy, Serbia, Sweden and Norway). However, the most part of the groups signalled were hardly supported (moderate to low bootstrap values), since the number of differences between sequences was limited.

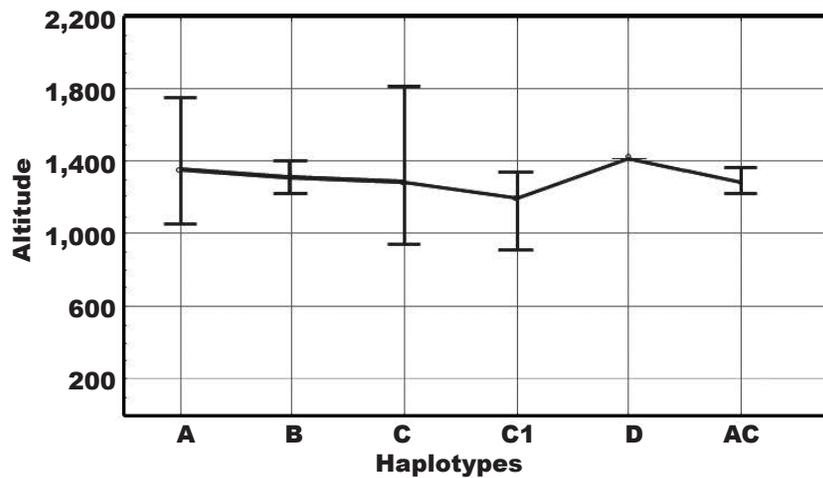


Fig. 5. Relationship between genetic distance and altitude.

*Fig. 5. Relación entre distancia genética y altitud.*

In figure 4, centring attention on the Iberian samples, we divided the clades into four groups. The first of these groups included specimens with A and B haplotypes, where haplotype 'A' coincided exactly with one haplotype found in the study by Royo et al. (2007). This coincident haplotype of Royo et al. (2007) comes from Alcornocales (Cádiz, south of Spain). In this first cluster, there were others Iberian specimens coming from Asturias (Sueve and Picos de Europa), Asturias (Muniellos), and France (Bordeaux), although the latter location is not decisive, as reintroductions from different sources were conducted.

In our haplotype clustering C, in the group numbered 2 (fig. 4), there was also an individual sequenced by Royo et al. (2007), also coming also from the Sierra de Guadarrama. Group 3 included animals presenting the 'D' haplotype, which coincided with one haplotype sequenced by Royo et al. (2007) from northern locations such as the Picos de Europa or Muniellos, related to a lineage composed of specimens from the Picos de Europa, plus others from Poblá de Segur (Lleida) and Bordeaux. The group marked 4 was composed entirely of specimens from the northern Iberian peninsula.

An analysis based on the Mantel test (fig. 5) showed that genetic distance was not significantly correlated with altitude ( $p > 0.05$ ), suggesting that altitude was not the principal factor influencing genetic differentiation in the 101 individuals analysed.

#### Microsatellite analysis

Of the 12 loci analysed, 11 were polymorphic (table 2), rendering a polymorphism equal to 0.91. The average number of alleles for the population was  $4.75 \pm 0.96$ , varying between 2 (CSSM41, CSSM43 and IDVGA29) and 13 (OarFCB304). The

PIC reached its lowest value for one of the less polymorphic loci (PIC IDVGA29 = 0.1035) and the highest value for one of the most polymorphic loci (PIC BM757 = 0.7465). The observed heterozygosity ranged from 0.11 (IDVGA29) to 0.73 (BM1706), with a mean of 0.43. The mean molecular coancestry, which characterizes the degree of genetic similarity among individuals in a population, was 0.5067 regardless of the informational value of different markers; the average value considering the PIC was 0.3701.

The population deviated significantly from Hardy–Weinberg equilibrium ( $p < 0.05$ ) due to 5 markers (BM757, HJJ1177, BMC1009, IDVGA8 and OarFCB304). These 5 markers had positive  $F_{IS}$  (reflecting a deficit of heterozygotes). This was a general trend since most  $F_{IS}$  values were positive. Only the microsatellites IDVGA29 and BM1706 with negative  $F_{IS}$ , of  $-0.031$  and  $-0.056$ , respectively, showed a weak excess of heterozygotes. The population  $F_{IS}$  calculated by Weir and Cockerham (1984) was positive and equal to 0.1115.

The various simulations with the computer program Structure 2.1 determined that the estimated log-likelihood of the data ( $\ln Pr(X/K)$ ) was maximum when  $K = 1$ , i.e., when considering a single population. When a factorial correspondence analysis was performed, where roe deer were in the space with a number of dimensions equal to the markers analysed, the specimens differed slightly, appearing overlapped because of the genotype similarity (fig. 6). The lower right area of figure 6 distinguished a unique individual, which could force the greater overlap of the other specimens that shared alleles other than this animal. This distinction is produced because this individual had a private allele at one of the loci analysed.

Table 2. Summary of basic analysis of genetic variability for the population of roe deer:  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity under the Hardy–Weinberg equilibrium;  $F_{IS}$ , was calculated according to Weir and Cockerham (1984). Values of the average number of alleles, average heterozygosity and  $F_{IS}$  calculated for all loci are shown in the last row.

*Tabla 2. Resumen del análisis básico de variabilidad genética para la población de corzo:  $H_O$ , heterocigosis observada;  $H_E$ , heterocigosis esperada bajo equilibrio de Hardy–Weinberg;  $F_{IS}$ , se calculó según Weir y Cockerham (1984). En la última fila aparecen los valores del número medio de alelos, la heterocigosis media y la  $F_{IS}$  calculada para todos los loci.*

Locus name	Allele range	Nº alleles	$H_O$	$H_E$	$F_{IS}$
BM1706	205–251	6	0.7303	0.7085	–0.031
BM757	193–295	8	0.3529	0.7829	0.551
BM848	355–361	4	0.6235	0.6235	0
CSSM39	177–181	3	0.6046	0.6242	0.032
HUJ1177	198–226	5	0.5662	0.6990	0.191
NVHRT48	86	1	0	0	–
BMC1009	277–281	3	0.1785	0.1851	0.036
CSSM41	121–123	2	0.3222	0.3880	0.170
CSSM43	240–246	2	0.4588	0.4960	0.075
IDVGA29	140–148	2	0.1162	0.1102	–0.056
IDVGA8	207–227	8	0.6363	0.6851	0.072
OarFCB304	151–193	13	0.6140	0.7543	0.187
Total	–	4.75	0.4337	0.5047	0.112

## Discussion

The combined analysis of two types of markers increased knowledge of the roe deer population in the Sierra de Guadarrama. Data from mitochondrial genomes (maternally inherited) and nuclear (from both parents) were complementary.

Analysis of the variation in the mtDNA control region revealed a multiple relationship between roe deer from the Sierra de Guadarrama and those from the Iberian peninsula populations (Randi et al., 2004; Royo et al., 2007) and European populations (Douzery and Randi, 1997; Vernesi et al., 2002; Randi et al., 2004). This relationship could be interpreted as a multiple origin of the population, with migration coming from different lineages of roe deer to the mountains of Madrid, either naturally by range expansion or through reintroductions. If migration had been natural by expansion of the distribution range, the observed variation should be lower than in the centre of origin, as it would represent the limit distribution of these populations; however, it is inconsistent with the results obtained. The variation stands in the range that it was shown in other studies, even in the upper part of this range (Wiehler and Tiedemann, 1998; Vernesi et al., 2002; Randi et al., 2004). The possibility of re-introductions in this area is unlikely because there are no records of such strategies in the Community of Madrid (FIDA, 2008).

The most likely explanation is therefore the acceptance of the theories that argue for the existence of shelters in the Pleistocene in the Iberian peninsula (Gliozzi et al., 1997; Hufthammer and Aaris-Sørensen, 1998; Taberlet et al., 1998), followed by later recolonization routes to northern Europe (Hewitt, 1999), or the area remaining as an isolated geographical region that would not have been influenced as a source for postglacial recolonization (Lorenzini and Lovari, 2006).

It should be noted that various proposals of phylogenetic relationships between haplotypes can be obtained when the characters are taken in whole or in part. The sequences provided by Royo et al. (2007) have different lengths. In the present study, the data were handled by taking all available information (930 pb) for the analyses together with the European available data or by reducing our matrix to the number of characters in common with other information studies (up to 446 pb), thereby obtaining various proposed relationships between the detected lineages. Lineages A to D showed different clustering, with C being more closely related to D in the global analysis, while in the treatment where only the specimens from the Iberian peninsula were considered, C was apparently closer to A and B. In any case, it should be kept in mind that most of the nodes were poorly supported, and thus, if the nodes not fully supported were collapsed, no incongruence occurred. Besides, in these proposals,

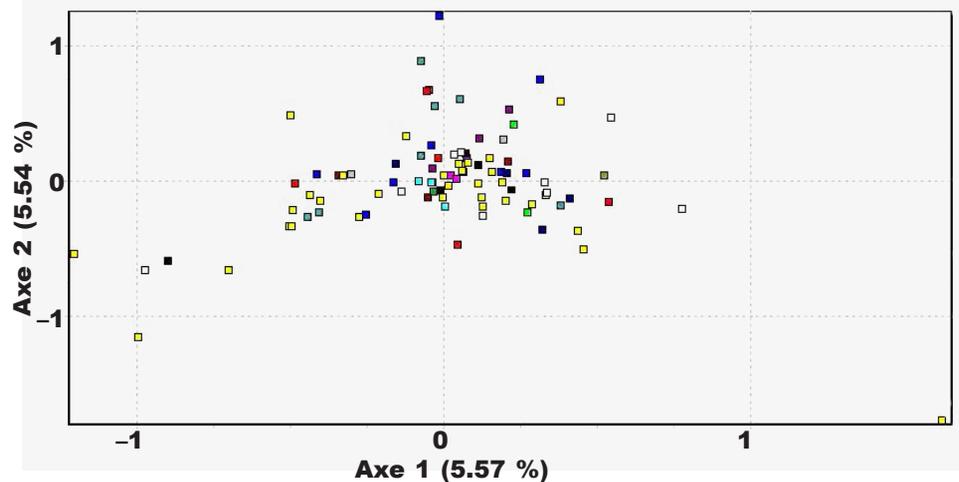


Fig. 6. Factorial correspondence analysis of the roe deer individuals in the Sierra de Guadarrama (Madrid, Spain), determined by the alleles detected in 12 microsatellite loci. The colours of the individuals indicate the different localities.

*Fig. 6. Análisis factorial de correspondencias de los individuos de corzo en la Sierra de Guadarrama (Madrid, España), determinado por los alelos detectados en 12 loci de microsatélites. Los colores de los individuos indican las diferentes localidades.*

not only the relationship between Sierra de Guadarrama samples and other specimens from central or southern Iberian peninsula remained constant in different clusters, but also their relationship with samples from the north, contrary to what is expected according to the findings of Royo et al. (2007). This confirms the point made in the comparative analysis of the study population with other European studies (Douzery and Randi, 1997; Vernesi et al., 2002; Randi et al., 2004). As the history of very recent populations, from an evolutionary point of view, shows this lack of support, since the number of diagnostic characters is limited, in this study we considered other markers such as microsatellites, which could elucidate the genetic structure of the population as they are more variable than mitochondrial markers. In addition, results from the population structure and Mantel test suggest that the relationship between genetic diversity and altitude is not significant, and hence it is possible to hypothesize that the species has not had sufficient time for evolutionary differentiation to occur along an altitude gradient.

Analysis of the genetic variability using microsatellite markers indicated that there is probably a single genetically homogeneous population. The existence of a single population can be expected since the roe deer population has recovered recently in the mountains of Madrid (Tellería and Virgós, 1997). Only if movements of recolonization arose from different populations or if different nuclei had remained isolated for a long time would a clear genetic structure be expected. This also implies that gene flow between groups of roe deer is adequate

and that, in principle, there are no major barriers that prevent adequate dispersion of individuals in the Madrid mountains. While this is a good sign, in order to preserve the population, it should be kept in mind that although the roe deer is a species with a high ecological plasticity (Tellería and Virgós 1997; Acevedo et al., 2005), it moves in small areas and is therefore highly susceptible to fragmentation and the presence of barriers (Coulon et al., 2004).

In-depth analysis of the genetic variability reveals that the population deviates from the Hardy–Weinberg equilibrium, so it is violating some of the assumptions of this principle (infinite population panmictic reproduction, absence of migration, mutation or selection). This was shown in five markers with positive values of  $F_{IS}$ . This parameter has values ranging from  $-1$  to  $1$ , with negative values indicating an excess of heterozygotes and positive values indicating a deficit. Thus, the results point to some inbreeding in the population. Studies in other populations of roe deer (Wang and Schreiber, 2001; Coulon et al., 2004) also found deviations from the Hardy–Weinberg equilibrium due to a deficit of heterozygotes. In a study on the population in south–western France (Coulon et al., 2004), the same set of markers used in the present study (except NVHRT48) was analysed in 1,148 individuals in an area of  $40 \times 55$  km. Two of the four microsatellites showing deficit of heterozygotes in the French population also showed this deficit in the Sierra de Guadarrama population. The average number of alleles and the heterozygosity values were higher than those observed here and the  $F_{IS}$

for this population was lower, but also positive. The results suggest that, in comparison with other studies with similar methodologies in European populations (Coulon et al., 2004), the genetic variability of the Guadarrama population is somewhat impoverished, although respect to the Spanish populations, it shows a genetic variability similar to or even superior to other more southern populations (Lorenzini et al., 2003; Royo et al., 2007).

The study by Royo et al. (2007) shared two microsatellites used in this study, NVHRT48 and BM757. These authors found four and 20 alleles for each marker respectively, while in our case, BM757, being one of the most polymorphic, showed only eight alleles and NVHRT48 was monomorphic. Although nine different populations were analysed in the study by Royo et al. (2007) without detailing the number of alleles in each of them, it should be noted that if in the rest of the Iberian peninsula four alleles were present for NVHRT48, in the population of Sierra de Guadarrama the variation was reduced to only one.

The average coefficient of coancestry helps us to analyse these results in more detail, giving an idea of how individuals are alike. Analyses in other populations of roe deer found mean coancestry values ranging between 0.277 and 0.476 (Royo et al., 2007). If we compare these results with the average molecular coancestry in our study, the value in the Madrid population was higher, indicating lower polymorphism or increased homozygosity, related to some inbreeding, probably due to the bottleneck the population underwent in the recent past (Tellería and Virgós, 1997; Gortázar et al., 2000). However, it should be kept in mind that the set of microsatellites used in the two studies and the population sizes were different, so it is more appropriate to compare the coancestry coefficient corrected by PIC. In this case, our value falls within the range found by Royo et al. (2007) and the value that these authors found for the 'central' population, including individuals from the Sierra de Guadarrama and Toledo (average coancestry coefficient of 0.373).

In conclusion, the combination of both results provided part of the complex story of the roe deer in the Iberian peninsula and especially in the Sierra de Guadarrama. On one hand, there is a philopatric structure with females maintaining the variation in mitochondrial haplotypes, clustered in four different groups without a common origin. These haplotypes are related to those from different populations of the Iberian peninsula and Europe. This complex origin leads to the consideration that no *C. capreolus* subspecies can be identified in the Iberian peninsula, agreeing in this case with Royo et al. (2007). On the other hand, as males contribute to the homogenisation of the genome, the population cannot thus be subdivided into different breeding units, but has to be considered as a single panmictic population. Additionally, heterozygosity was below the expected level in most markers and allelic richness was generally lower than in other populations. This provides evidence of a relative genetic impoverishment of the population.

In view of the above, the different times of the palaeohistory of the species may explain these re-

sults: glacial–interglacial stages of the Pleistocene (Gliozzi et al., 1997; Hufthammer and Aaris–Sørensen, 1998) and the reduction and recovery of populations throughout the 20<sup>th</sup> century (Tellería and Virgós, 1997; Gortázar et al., 2000). The two markers used, mitochondrial DNA of maternal inheritance and nuclear DNA of both parents, have a different heritage and unequal mutation rate, reflected in each of these two main stages in the recent evolution of the roe deer.

Our results have a strong implication for the management and conservation of roe deer in the Sierra de Guadarrama. Fortunately, in recent years, the roe deer population in the Iberian peninsula has benefited from a decline in ranching, abandonment of agriculture in foothill areas, and a significant increase in forest areas (Acevedo et al. 2005), all positive events to avoid population isolation. Even so, to preserve the populations, management must keep in mind that the species is highly susceptible to fragmentation and to barriers (Coulon et al., 2004) such as highways and fencing. Finally, taken together, our results indicate that restocking within the Sierra de Guadarrama should be prevented if we want to preserve the diversity among populations and the patterns of natural gene flow.

Genetic assessment of structure and connectivity of roe deer populations that recently recolonized a fragmented landscape could be a promising approach for future studies.

### Acknowledgements

We are grateful to the Consejería de Medio Ambiente y Ordenación del Territorio de la Comunidad de Madrid, for making this study in the Sierra de Guadarrama possible. We also thank M. J. Ruiz for her full involvement in genetic testing. Thanks too to R. Fernández–Mellado, A. Navarro–Castilla and the Editor and referees for their comments and improvement of the manuscript.

### References

- Acevedo, P., Delibes–Mateos, M., Escudero, M. A., Vicente, J., Marco, J., Gortázar, C., 2005. Environmental constraints in the colonization sequence of roe deer across the Iberian Mountains, Spain. *Journal of Biogeography*, 32: 1671–1680.
- Buitrago, A. M., 1992. Estudio de los Artiodáctilos del yacimiento del Pleistoceno medio de Pinilla del Valle (Madrid). PhD thesis, Universidad Complutense de Madrid, España.
- Coulon, A., Cosson, J. F., Angibault, J. M., Cargnelutti, B., Galan, M., Morellet, N., Petit, E., Aulagnier, S., Hewison, A. J., 2004. Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual–based approach. *Molecular Ecology*, 13: 2841–2850.
- Do, H. H., Rahm, E., 2004. Flexible integration of molecular–biological annotation data: The genmapper approach. In: *Advances in database technology–EDBT 2004*: 811–822 (E. Berlino, S. Christodoulakis, D. Plexousakis, V. Christophi-

- des, M. Koubarakis, K. Böhm, E. Ferrari, Eds.). Springer, Berlin Heidelberg.
- Douzery, E., Randi, E., 1997. The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content. *Molecular Biology and Evolution*, 14: 1154–1166.
- FIDA (Fundación para la Investigación y el Desarrollo Ambiental), 2008. Filogenia, evolución y variabilidad genética. In: *Ecología y caracterización genética de las poblaciones de corzo de la Comunidad de Madrid: 79–95* (Fundación para la Investigación y el Desarrollo Ambiental, Eds.). Consejería de Medio Ambiente, Vivienda y Ordenación del Territorio, Madrid.
- Galán, M., Cosson, J. F., Aulagnier, S., Maillard, J. C., Thévenon, S., Hewison, A. J. M., 2003. Cross-amplification tests of ungulate primers in roe deer (*Capreolus capreolus*) to develop a multiplex panel of 12 microsatellite loci. *Molecular Ecology Notes*, 3: 142–146.
- Giozzi, E., Abbazzi, L., Argenti, P., Azzaroli, A., Caloi, L., Capasso Barbato, L., Torre, D., 1997. Biochronology of selected mammals, molluscs and ostracods from the middle pliocene to the late pleistocene in Italy. *The state of the art. Rivista italiana di paleontologia e stratigrafia*, 103: 369–388.
- Gortázar, C., Herrero, J., Villafuerte, R., Marco, J., 2000. Historical examination of the status of large mammals in Aragón, Spain. *Mammalia*, 64: 411–422.
- Gutiérrez, J. P., Royo, L. J., Álvarez, I., Goyache, F., 2005. MolKin v2.0: A computer program for genetic analysis of populations using molecular coancestry information. *Journal of Heredity*, 96: 718–721.
- Hewitt, G. M., 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68: 87–112.
- Hufthammer, A. K., Aaris-Sørensen, K., 1998. Late- and postglacial European roe deer. In: *The European roe deer: the biology of success: 47–69* (R. Andersen, P. Duncan, J. D. C. Linnell, Eds.). Scandinavian university press, Oslo.
- Jäger, F., Hecht, W., Herzog, A., 1992. Untersuchungen an mitochondrialer DNS (mtDNS) von hessischem Rehwild (*C. capreolus*). *Zeitschrift für Jagdwissenschaft*, 38: 26–33.
- Lorenzini, R., Garofalo, L., Qin, X., Voloshina, I., Lovari, S., 2014. (Artiodactyla: Cervidae), a Palearctic meso-mammal. *Zoological Journal of the Linnean Society*, 170: 209–221.
- Lorenzini, R., Lovari, S., 2006. Genetic diversity and phylogeography of the European roe deer: the refuge area theory revisited. *Biological Journal of the Linnean Society*, 88: 85–100.
- Lorenzini, R., San José, C., Braza, C., Aragón, S., 2003. Genetic differentiation and phylogeography of roe deer in Spain, as suggested by mitochondrial DNA and microsatellite analysis. *Italian Journal of Zoology*, 70: 89–99.
- Pritchard, J. K., Wen, W., 2003. Documentation for STRUCTURE software: version 2. Available from: <http://pritch.bsd.uchicago.edu>.
- Randi, E., Alves, P., Carranza, J., Milosevi-Zlatanova, S., Sfougaris, A., 2004. Phylogeography of roe deer (*Capreolus capreolus*) populations: the effects of historical genetic subdivisions and recent nonequilibrium dynamics. *Molecular Ecology*, 13: 3071–3083.
- Randi, E., Pierpaoli, M., Danilkin, A., 1998. Mitochondrial DNA polymorphism in populations of Siberian and European roe deer (*Capreolus pygargus* and *C. capreolus*). *Heredity*, 80: 429–437.
- Raymond, M., Rousset, F., 1995. Genepop (version 1.2): population genetics software for exact tests and ecumnicism. *Journal of Heredity*, 86: 248–249.
- Royo, L. J., Pajares, G., Álvarez, I., Fernández, I., Goyache, F., 2007. Genetic variability and differentiation in Spanish roe deer (*Capreolus capreolus*): A phylogeographic reassessment within the European framework. *Molecular Phylogenetics and Evolution*, 42: 47–61.
- Sambrook, J., Fritsch, E. F., Maniatis, T., 1989. *Molecular Cloning: a Laboratory Manual*, 2nd Edition. Cold Spring Harbor Laboratory Press, New York.
- Schneider, S., Roessli, D., Excoffier, L., 2000. *Arlequin v.2.0: a software for population genetics data analysis*. University of Geneva, Geneva, Switzerland.
- 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.
- Swofford, D. L., 2003. *PAUP\*: Phylogenetic analysis using parsimony (\*and other methods), version 4.0b 10*. Sinauer associates, Sunderland, Massachusetts.
- 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.
- Tellería, J. L., 1999. La diversidad de vertebrados del Valle del Paular (Madrid). In: *Primeros Encuentros Científicos del Parque Natural de Peñalara: 155–162* (Consejería de Medio Ambiente, Eds.). Comunidad de Madrid, Madrid.
- Tellería, J. L., Virgós, E., 1997. Distribution of an increasing roe deer population in fragmented Mediterranean landscape. *Ecography*, 20: 247–252.
- 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.
- Wang, M., Schreiber, A., 2001. Impact of the habitat fragmentation and social structure on the population genetics of roe deer (*Capreolus capreolus* L.) in Central Europe. *Heredity*, 86: 703–715.
- Weir, B. S., Cockerham, C. C., 1984. Estimating  $F$ -statistics for the analysis of population structure. *Evolution*, 38: 1358–1370.
- Wiehler, J., Tiedemann, R., 1998. Phylogeography of the European roe deer (*Capreolus capreolus*) as revealed by sequence analysis of the mitochondrial control region. *Acta Theriologica*, 5: 187–197.

