Spatial genetic structure across a hybrid zone between European rabbit subspecies

The Iberian Peninsula is the only region in the world where the two existing subspecies of the European rabbit (*Oryctolagus cuniculus*) naturally occur and hybridize. In this study we explore the relative roles of historical and contemporary processes in shaping the spatial genetic structure of the rabbit across its native distribution range, and how they differently affect each subspecies and the hybrid zone. For that purpose multilocus genotypes and mitochondrial DNA data were obtained for 771 rabbits across most of the species' distribution range in Spain. Nuclear markers defined a hierarchical genetic structure firstly comprised by two genetic groups, largely congruent with the mitochondrial lineages and subspecies distributions (*O. c. algirus* and *O. c. cuniculus*), which were subsequently subdivided into seven genetic groups probably shaped by environmental or ecological factors. Geographic distance alone emerged as an important factor explaining genetic differentiation across the whole range, without the need to invoke for the effect for geographical barriers. Thus, when considering the overall genetic structure, differences at a local level seem to be of greater importance. The significantly positive spatial correlation up to a distance of only 100 km supported this hypothesis. However, northern populations of *O. c. cuniculus* showed more spatial genetic structure and differentiation than *O. c. algirus*, which could be due to local geographic barriers, limited resources, soil type and/or social behaviours limiting dispersal. The hybrid zone showed similar genetic structure to the southern populations but a larger introgression from the northern lineage genome. These differences have been attributed to selection against the hybrids rather than to behavioural differences between subspecies. Ultimately, the genetic structure of the rabbit in its native distribution range is the result of an ensemble of factors, from geographical and ecological, to behavioural and molecular, that

hierarchically interact in time and space.

Fernando ALDA[1, 2*], Ignacio DOADRIO[1]

[1]Dpto. Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), Madrid, 28006, Spain

[2]Current address: Tulane/Xavier Center for Bioenvironmental Research (CBR), Tulane University, New Orleans, Louisiana, 70118, USA

Corresponding author:

Fernando ALDA

Tulane/Xavier Center for Bioenvironmental Research (CBR)

Tulane University

New Orleans, LA 70118, USA

Tel: (+1) 504 862 8441

Fax: (+1) 504 862 8455

e-mail: alda.fernando@gmail.com

INTRODUCTION

In most species populations are genetically structured. This genetic structure may be a consequence of many factors. Foremost among these, geographical factors may lead to vicariant events and divide populations (Knowles & Carstens, 2007), or ecological may determine habitat suitability across space, and consequently population connectivity (Pilot et al., 2006). Population structure can also be shaped by species behavioral traits, such as family groups in primates, or colonies of social insects (Shoemaker & Ross, 1996; Bradley et al., 2002). At last, genetic structure will result from the balance between gene flow, drift, and the time necessary to reach a balance between both forces (Hutchison & Templeton, 1999).

The processes leading to population structure can act at different temporal and spatial scales. Temporarily, historical processes such as isolation in glacial refugia and subsequent expansions can leave detectable signals in the current populations (Avise, 2000), as well as contemporary dispersal (Palsboll, 1999). Spatially, gene flow may vary among individuals within a geographical region, between adjacent regions, or at larger scales, between populations that presumably have little genetic exchange but share a more ancient genetic history. Thus, investigating population relationships and their spatial patterns of genetic variation is useful in order to infer these hierarchical and interacting processes (Hedrick, 2005).

The European rabbit (*Oryctolagus cuniculus* Linnaeus 1758) is a species with worldwide biological and economic importance that has long attracted scientific interest (reviewed in Ferreira, 2012). This, along with a well-documented history, has allowed the development of multiple studies on the evolutionary history of this lagomorph from a wide range of molecular

(Ferrand, 1989; Biju-Duval et al., 1991; Branco, Ferrand & Monnerot, 2000; Geraldes et al., 2008; Carneiro et al., 2012), temporal (Hardy et al., 1994; Monnerot et al., 1994) and geographical perspectives (Webb et al., 1995; Fuller, Wilson & Mather, 1997; Surridge et al., 1999b; Queney et al., 2000; Queney et al., 2001; Branco et al., 2002).

In the Iberian Peninsula, two divergent evolutionary lineages occur and contact each other in the middle of their distribution range (Branco, Ferrand & Monnerot, 2000; Branco et al., 2002). Studies conducted on uniparentally inherited molecular markers support the existence of two highly differentiated groups: so-called mitochondrial lineage A, predominant in the subspecies *O. c. algirus* (Loche 1867) inhabiting the southwest of the Iberian Peninsula and lineage B, which predominates *in O. c. cuniculus* in the northeast of the Peninsula (Biju-Duval et al., 1991). It is proposed that mitochondrial lineages A and B diverged following isolation in two glacial refugia in the southwestern and northeastern extremes of the Iberian Peninsula, likely during the Quaternary paleoclimatic oscillations. After climatic amelioration, they expanded their ranges and came again into contact to form a secondary contact zone where they hybridize along a northwest-southeast axis (Branco, Ferrand & Monnerot, 2000; Branco et al., 2002; Geraldes, Rogel-Gaillard & Ferrand, 2005; Geraldes, Ferrand & Nachman, 2006; Ferrand & Branco, 2007).

Hybrid zones like this one are usually interpreted as zones where genetically distinct populations meet and interbreed because, despite genetic differences, they have not reached the status of species and are to some extent interfertile. Therefore, they can be considered as intermediate stages in the process of speciation (Barton & Hewitt, 1989; Harrison, 1993). Hybrid zones may be ephemeral, resulting from the recent meeting and blending of two divergent lineages, or may

have arisen from an ancient contact and last indefinitely. In the former case, many individuals in the centre of the zone will resemble the parental forms, leading to high genetic variance and high linkage disequilibrium between loci. In the latter case, the hybrid zone will consists of individuals that are product of many generations of hybridization, leading to lower genetic variance and lower linkage disequilibrium (Brelsford & Irwin, 2009). Therefore, the comparative study of the genetic structure in a hybrid zone and in the parental populations can provide insights into the evolutionary processes that contribute to its origin and maintenance (Harrison, 1993).

As a matter of fact, hybrid zones are areas of particular interest for evolutionary studies enabling insight into the initial stages of speciation and reproductive isolation, adaptation and selection, and even behavioural processes (Hewitt, 1988; Barton & Gale, 1993; Arnold, 1997; Futuyma, 1998). In the rabbit, extensive studies in the hybrid zone have evidenced highly contrasting degrees of introgression among loci, or even a complete absence of genetic structure (Branco, Machado & Ferrand, 1999; Queney et al., 2001; Geraldes, Ferrand & Nachman, 2006; Ferrand & Branco, 2007; Campos, Storz & Ferrand, 2008; Carneiro, Ferrand & Nachman, 2009; Carneiro et al., 2010; Carneiro et al., 2013). The variation in the introgression of autosomal and sexual chromosomes has also revealed different selective pressures across genes and its importance in the reproductive isolation between the two rabbit subspecies (Geraldes, Ferrand & Nachman, 2006; Campos, Storz & Ferrand, 2008; Carneiro, Ferrand & Nachman, 2009; Carneiro et al., 2010; Carneiro et al., 2013).

However, differences in the spatial genetic structure between the parental lineages, either in allopatry or interacting within the hybrid zone, are still largely unknown, because most

population genetics studies on the rabbit in the Iberian Peninsula have either examined genetic variation between lineages, or the area where they come into contact (Monnerot et al., 1994; Queney et al., 2001; Branco & Ferrand, 2003; Ferrand & Branco, 2007; Carneiro, Ferrand & Nachman, 2009; Carneiro et al., 2010), but rarely all together or in a comparative context. Furthermore, so far, studies relating the observed genetic structure to behavioural traits or habitat have only been carried out in regions where the rabbit is a non-native species (Fuller, Wilson & Mather, 1997; Surridge et al., 1999a; Surridge et al., 1999b). In this context, the Iberian Peninsula is unique, since it is the only region in the world where the two subspecies are native and co-occur. Therefore, due to their long evolutionary history in this region, it is expected that a complex ensemble of interacting factors affect the genetic structure of the rabbit at different hierarchical scales.

Thus the main objective of this study was to describe and study the relative roles that historical and contemporary processes have in (1) shaping the spatial genetic structure of the rabbit in Spain and (2) conditioning differences in the distribution of genetic variability and structure between populations of *O. c. cuniculus*, *O. c. algirus* and the hybrid zone.

MATERIALS AND METHODS

Sample collection

Samples of 771 rabbits were obtained from 30 localities covering most of the species' range in Spain (Supplementary Table 1 and Fig. 1). Sampling was performed mainly on hunting estates by

licensed hunters during legal hunting seasons, or during management or restocking activities carried out by local administrations. Since these activities did not involve any experimental or scientific purpose, no approval was requested from the Ethics Committee at the Spanish Superior Research Council (CSIC).

The number of samples per locality ranged from 2 to 56 individuals (Supplementary Table 1). According to their geographic location (Branco et al., 2002; Geraldes et al., 2008), 6 localities (n = 144) were *a priori* assigned to subspecies *O. c. cuniculus*, 14 localities (n = 410) to subspecies *O. c. algirus*, and 10 localities (n = 217) to the hybrid zone (Fig. 1, Supplementary Table 1).

DNA extraction and amplification of molecular markers

Samples obtained from live rabbits consisted of blood drawn from the femoral vein or a small piece of ear tissue. Ear tissue or muscle samples were taken from dead rabbits, depending on the preservation state of the animal. DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN) following the manufacturer's instructions.

Mitochondrial lineages, A or B, were identified in all samples by amplifying the complete cytochrome *b* gene using the primers OcunCB_F: 5'-ATGACCAACATTCGCAAAACC-3 'and OcunCB_R: 5'-TGTCTCAGGGAGAACTATCTCC-3'. The PCR reaction was performed in a final volume of 25 µL containing: 200 µM of dNTPs, 0.2 µM of each primer, 1 U *Taq* polymerase (Eppendorf), 1x PCR buffer (500 mM KCl, 100 mM Tris-HCl pH8.3, 15 mM Mg^{2+}), and 1 µL of DNA extract. The PCR program consisted of 4 min of denaturation at 94 °C, followed by 40

cycles of 1 min at 94 ºC, 1 min at 55 °C and 1 min 30 s at 72 ºC, plus a final extension of 10 min at 72 ºC. PCR products were digested separately with *HaeIII* and *AluI* restriction enzymes (Promega) and migrated in 3% agarose gels stained with ethidium bromide for subsequent visualization under UV light. RFLP patterns were identified as described previously (Branco, Ferrand & Monnerot, 2000) and the samples were assigned to one of two mitochondrial lineages.

All individuals were genotyped according to 10 microsatellite markers: Sat3, Sat4, Sat5, Sat7, Sat8, Sat12, Sat13, Sat16, Sol33 and Sol44 (Mougel, Mounolou & Monnerot, 1997; Surridge et al., 1997). PCR reactions were performed in a final volume of 13 µL that contained 200 µM dNTPs, 0.2-0.4 µM of each primer, 2-2.5 mM $MgCl_2$, 0.325 U *Taq* polymerase (Eppendorf), 1x PCR buffer (500 mM KCl, 100 mM Tris-HCl pH 8.3), and 0.5 µL of DNA extract. PCR programs involved 2 min of initial denaturation at 95 °C, followed by 35 cycles of 30 sec at 95 ºC, 30 sec between 55 ºC and 65 °C, 30 sec at 72 °C, followed by a final extension step of 7 min at 72 ºC. The amplified fragments were analysed in an ABI3730 automatic sequencer (Applied Biosystems) and allele sizes were assigned using the program GeneMapper v3.7 (Applied Biosystems). The complete data file of microsatellite genotypes and mitochondrial haplotypes was deposited as a Supplementary Dataset in PeerJ.

Statistical analyses

Our first objective was to undertake a formal analysis on the genetic structure of the two rabbit lineages and across the hybrid zone between them. For this purpose, we used Bayesian model-based assignment methods to determine admixture proportions in our rabbit sample. Although we were primarily interested in a model with two clusters for assessing admixture between lineages,

we also analysed whether models with more than two clusters were supported by our data. Therefore, firstly we used the algorithm of STRUCTURE 2.2 (Pritchard, Stephens & Donnelly, 2000) implementing the admixture model with correlated allele frequencies (Falush, Stephens & Pritchard, 2003), since this model is more appropriate for individuals with admixed ancestries and for populations with similar expected frequencies. No information on the localities of origin of individuals was included. Ten independent analyses were run for each value of K, from $K = 1$ to $K = 30$. Each analysis consisted of 1×10^6 Markov chains with a prior burn-in of 1×10^5 chains. Mean posterior probability values were used to calculate ΔK, a measure of the rate of change of the posterior probabilities between successive K values. Thus, it is possible to detect when the increase in $\ln P(X|D)$ is not significant anymore and find the true value of K (Evanno, Regnaut & Goudet, 2005).

To visually explore the distribution of the inferred genetic groups across the hybrid zone, the proportion of genetic admixture and the frequency of each mitochondrial lineage were plotted along a one-dimensional transect perpendicular to the proposed rabbit hybrid zone (Branco, Ferrand & Monnerot, 2000; Branco et al., 2002; Geraldes et al., 2008), with the exception of the localities of Galicia and Mallorca which were excluded (Fig. 1), Geographical locations along this linear transect were measured considering km 0 at the approximate intersection between the transect and the hybrid zone (Fig. 1), and fitted to a sigmoid curve (3 parameters) as expected by hybrid zone tension zone models (Barton & Hewitt, 1985; Barton & Gale, 1993) in SigmaPlot 10.0 (Systat Software, Inc.). Additionally, congruence between the assignment probabilities and the frequency of mitochondrial lineages was evaluated by performing χ^2 tests in both parental lineages and in the hybrid zone.

158 Secondly, we used the Bayesian method implemented in the program BAPS 5.1 (Corander &
159 Marttinen, 2006). In addition to the genetic data, we included the geographical coordinates of
160 each individual and used the spatial model in BAPS (Corander, Sirén & Arjas, 2008). This model
161 estimates genetic structure assuming that the structure within a particular area depends on the
162 neighbouring areas, thereby increasing the statistical power to detect the true genetic structure
163 (Corander, Sirén & Arjas, 2008). We undertook 10 independent replicates from 1 to a maximum
164 of 30 genetic clusters. The average admixture values obtained for each individual were plotted
165 using the *maps* package in R (R Core Team, 2014).

166 For the inferred genetic clusters and the localities analysed with more than 10 sampled
167 individuals, we tested significant deviations from Hardy-Weinberg equilibrium through a Fisher
168 exact test (Guo & Thompson, 1992) after applying the Bonferroni correction (Rice, 1989) in
169 GENEPOP 3.4 (Raymond & Rousset, 1995). Parameters of genetic diversity such as number of
170 alleles (N_A), allelic richness (A_R), observed and expected heterozygosity (H_o and H_e) and
171 inbreeding coefficient (F_{IS}) were calculated for each locus and genetic group using the programs
172 GENETIX 4.02 (Belkhir et al., 2004) and FSTAT 2.9.3 (Goudet, 1995).

173 The distribution of genetic variation among the sampled localities, as well as within and among
174 the inferred genetic groups was assessed by an analysis of molecular variance (AMOVA)
175 (Excoffier, Smouse & Quattro, 1992). AMOVA was performed using GenoDive 2.0b11
176 (Meirmans & Van Tienderen, 2004) which allows the calculation of a F_{ST} analogue coefficient of
177 differentiation, standardized according to the level of intra-population variation, so that the

results obtained can be compared between markers showing different polymorphism (Meirmans, 2006). To avoid confusion, this ratio is hereafter referred to as F_{ST}'.

The effect of geographical distance on the genetic differentiation between individuals was also tested for all the samples in peninsular Spain (i.e. excluding the population from the island of Mallorca) and for each lineage and the hybrid zone separately. Regression was performed for the kinship coefficient between pairs of individuals (f_{ij}) (Loiselle et al., 1995) and their geographical distance (d_{ij}), to give a regression slope $\ln b_d$ and its statistical significance (Vekemans & Hardy, 2004). Also, spatial autocorrelation methods were applied to examine spatial genetic structure (Smouse & Peakall, 1999). Geographic locations of individuals were permuted 10,000 times among 50 distance intervals with an equal number of comparisons between individuals (5935) to test the null hypothesis that d_{ij} and f_{ij} were not correlated. Positive spatial autocorrelations are expected when gene flow is restricted to short distances. These tests are dependent on the type and scale of sampling (Vekemans & Hardy, 2004), so to compare the extent of spatial genetic structure in each subspecies of rabbit and the hybrid zone, we used the statistical *Sp* (Vekemans & Hardy, 2004), $Sp = -\ln b_d/(1 - F_{ij})$, where F_{ij} is the average kinship coefficient between individuals closer together (the first distance interval, ~ 5km), and $\ln b_d$ is the slope of the linear regression of the correlation coefficients and the logarithm of the geographical distance. All these tests were conducted in SPAGeDi 1.2 (Hardy & Vekemans, 2002). Additionally, a Mantel test (Mantel, 1967) was employed to determine if there was significant correlation between the genetic (F_{ST} /1-F_{ST}) and geographical distances of the localities studied. Also, the presence or absence of putative barriers, such as large rivers and/or mountain ranges, between localities was coded as 1 and 0 in a third data matrix. Using these three distance matrices, a partial Mantel test was performed to

200 determine whether, besides geographical distance, these landforms represented a barrier to gene
201 flow for the rabbit. Both analyses were performed in ARLEQUIN 3.1.

202 RESULTS

203 *Distribution of genetic diversity*

204 Overall, similar proportions of rabbits carried mitochondrial haplotypes from lineages A and B
205 (53.5% and 46.5% respectively). In only five localities with over 10 individuals analysed, all
206 individuals belonged to one lineage. Rabbits from Mallorca, Lérida and Galicia belonged to
207 lineage B, while those from Jaén3 and Sevilla1 belonged to lineage A. Although all the other
208 localities showed a mixture of both lineages, there was a clear predominance of B haplotypes in
209 the northeast of the Iberian Peninsula and Balearic islands, A haplotypes in the southwest, and a
210 mixture of both in the centre of the Peninsula (Table 1).

211 In general, nuclear genetic diversity was high, with a total of 264 alleles at the 10 microsatellites
212 analysed (average N_A per locus = 26.29 ± 9.07). In all localities heterozygosity values were larger
213 than H_o = 0.6 (average H_o = 0.7 ± 0.15). The localities showing greatest diversity, measured as
214 allelic richness, were Madrid1 and Sevilla2 (A_R = 7.70 and 7.62), whereas the least diverse were
215 Mallorca (A_R = 5.15) and La Rioja (A_R = 6.20) (Table 1 and Supplementary Table 2).

Deviations from Hardy-Weinberg equilibrium were detected in 17 of the 26 locations analysed (Table 1 and Supplementary Table 2). In all cases, these deviations were due to heterozygote deficits. Locations in the centre and southwest of the Iberian Peninsula showed larger deviations from equilibrium, mainly attributed to locus Sat16 and to a lesser extent to Sat3 and Sol33. Interestingly, none of these loci appeared to be in disequilibrium in the northeastern localities. A comparison of the observed genotypes with a random distribution of genotypes generated by MICRO-CHECKER (van Oosterhout et al., 2004) suggested the presence of null alleles at locus Sat16, as had been proposed earlier for this microsatellite (Queney et al., 2001).

Structure and assignment of rabbit genetic clusters

The Bayesian clustering analysis performed in STRUCTURE revealed that $\ln P(X|D)$ increased substantially from $K = 1$ to $K = 2$ and then was attenuated as the number of K increased but without reaching a clear asymptote. Notwithstanding, calculation of ΔK clearly revealed the existence of 2 genetic populations or groups (K1 and K2). Taken together, these results suggest that the sampled rabbits belong to two large distinct genetic groups, but do not completely exclude the possibility that more gene pools exist. The distribution of the two genetic groups exhibited high geographical correlation such that the localities to the south of the Iberian Peninsula were assigned with a greater likelihood to group K1 and the northern localities to K2 group (Table 2 and Fig. 2).

Both the assignment probabilities to the genetic groups based on the nuclear DNA and the frequencies of mitochondrial lineages closely conformed to sigmoidal functions ($R^2 = 0.878$, $F = 57.478$, $P < 0.0001$ for the microsatellite data; $R^2 = 0.628$, $F = 13.501$, $P < 0.0001$ for the mtDNA

data; Fig. 3). In addition, the inferred genetic groups were in agreement with the mitochondrial haplotypes of each individual. Of the 224 individuals assigned to K1with a posterior probability greater than 0.9, 91% carried mitochondrial haplotypes belonging to lineage A, while 82% of the 239 rabbits assigned with equal probability to K2 showed haplotypes from lineage B (Fig. 2). Thus, the frequency of mitochondrial haplotypes in the parental lineages did not differ significantly from the individuals' assignment frequencies to the inferred genetic groups (χ^2 = 0.781, df = 1, P = 0.377).

The analysis in BAPS of the genetic data together with the geographical information for individuals, detected a maximum marginal likelihood (corresponding to the maximum posterior probability) for 7 genetic clusters. Most of the inferred genetic clusters showed a well-defined geographical distribution (Table 2 and Fig. 1). The first cluster corresponded to all individuals from the Balearic Islands (K-BAL). In the north of the Iberian Peninsula, a second cluster appeared consisting mainly of individuals from Lérida (K-NE) and a third cluster comprising individuals from the localities of La Rioja, Madrid1 and Madrid2 (KN). In the south of the Iberian Peninsula, one large cluster was inferred, that included most of the individuals from Badajoz1, Badajoz2, Jaén2, Sevilla1, Sevilla2 and Cádiz (K-S1), and a small group of individuals from Sevilla2 (K-S2). At the heart of the Iberian Peninsula we found a large cluster fully or partially covering the localities of Galicia, Valladolid, Madrid1-3, Cuenca, Toledo1-6, Albacete, Ciudad Real1-3 and Cáceres1 and Cáceres2 (K-ZH) and another small group of individuals from the locality of Cuenca (K-CU) (Table 2 and Fig. 1). Also, these clusters were congruent with the genetic groups inferred in STRUCTURE and with the mitochondrial haplotypes of the individuals, so that the northern and central clusters (K-BAL, K-NE, K-N and K-CU) had

assignment probabilities greater than 90% to K2, and in the same way, the southern clusters (K-S1 and K-S2) to K1.

Genetic diversity parameters estimated for the inferred genetic clusters in BAPS, indicated a greater diversity for K-ZH and K-S1, which were also the only clusters in Hardy-Weinberg disequilibrium, due to a significant deficit of heterozygotes (Table 2 and Supplementary Table 3). All genetic clusters displayed unique alleles, which were usually found at low frequencies. In those genetic clusters mostly including lineage A rabbits (K-S1 and K-S2), 37 unique alleles were found, while clusters with mostly lineage B rabbits (K-BAL, K-NE, KN and K-CU) showed 21 unique alleles. Forty-nine unique alleles were detected in cluster K-ZH.

The percentage of genetic variation explained by the 7 genetic clusters was similar, but not greater than that obtained among all localities (F_{ST}' = 0.325, $P < 0.001$, F_{ST}'= 0.370, $P < 0.001$, respectively). Conversely, R_{ST} was much higher ($R_{ST} = 0.627$, $P < 0.001$ and $R_{ST} = 0.110$, $P < 0.001$, respectively), indicating that the effect of mutation is of greater importance than drift in the differentiation of rabbit genetic clusters.

The Mantel test revealed significant correlation between geographical distances and genetic distances for all pairs of populations ($r = 0.538$, $P < 0.001$, Fig. 4). However, this correlation was not improved by including the effect of geographical barriers, such as rivers or mountain ranges, in the partial Mantel test. Across the whole distribution of rabbits in peninsular Spain, the regression slope between kinship and geographical distance was negative and significant ($\ln b_d =$

-0.011, $P < 0.001$). In effect, relationships between individuals decreased rapidly as geographical distances increased and this autocorrelation was significantly positive up to a distance of approximately 100 km (Fig. 5A). By comparing the spatial genetic structure of the two subspecies of rabbit and those of the hybrid zone, spatial autocorrelation analyses indicated much higher f_{ij} values and a steeper decline of kinship with distance in *O. c. cuniculus* for the first distance intervals. Similarly, we obtained a higher *Sp* value for *O. c. cuniculus* (*Sp* = 0.0137) than for *O. c. algirus* and the hybrid zone, which showed similar values (*Sp* = 0.0062 and *Sp* = 0.0063, respectively), indicating a greater genetic structure of rabbit populations in the northeast of the Iberian Peninsula (Fig. 5B).

DISCUSSION

Variation in genetic diversity

Overall, the microsatellites analysed were highly polymorphic, and showed a similar variability to that reported previously for 9 of the 10 loci studied (Queney et al., 2001). The general trend was greater genetic variability in populations from the central Iberian Peninsula and lower diversity in the northeast mainland populations (Table 1 and Supplementary Table 2). This reduced genetic diversity of northeast populations has been attributed to a lower effective size during their isolation in Quaternary glacial refugia, compared to the southern populations (Branco, Ferrand & Monnerot, 2000). Further, the lowest diversity and high genetic differentiation found in the Balearic Islands is most likely due to the founder effect caused by the introduction of the rabbit in these Mediterranean islands following the first human arrival to Mallorca 4300-4100 years ago (Flux, 1994; Alcover, 2008). Interestingly, a much older estimate has been proposed for the most recent common ancestor between island and mainland rabbits

between 170,000 years and present, according to mitochondrial sequence data (Seixas et al., 2014). While the number of alleles and allelic richness detected for this locality were lower than in most of the other samples analysed, this was not the case for its heterozygosity values (Table 1 and Supplementary Table 2). Loss of heterozygosity depends on the time it takes a population to recover a large size (Nei, Maruyama & Chakraborty, 1975). Thus, because of its rapid expansion capability, the rabbit may have managed to retain more diversity during different colonizations or bottlenecks suffered (Queney et al., 2000). Similarly, this could explain why in La Rioja, where demographic explosions are frequent, a low number of alleles are detected, but not a low heterozygosity. In contrast, in Galicia, both allelic richness and heterozygosity are low and show a significant excess of homozygotes, since rabbit populations in this region have continuously declined in recent years with the consequent loss of genetic diversity (Table 1 and Supplementary Table 2).

Most localities, particularly the southern ones, revealed loci in Hardy-Weinberg disequilibrium, because of a deficiency in heterozygotes. In this study, as in earlier ones (Queney et al., 2001), putative null alleles were detected at Sat16, although the exclusion of this locus and others showing large deviations from equilibrium (e.g. Sat3 and Sol33), did not significantly alter the results. The absence of disequilibrium in the northeastern localities should not be strange considering that the microsatellites analysed were originally developed for the domestic rabbit (i.e. subspecies *O. c. cuniculus*) (Mougel, Mounolou & Monnerot, 1997; Surridge et al., 1997), therefore a higher chance for null alleles could occur in the southern *O. c. algirus*. Another non-exclusive explanation for the significant deficit of heterozygotes could be due to a Wahlund effect. This is also quite likely in view of the territorial behaviour and social structure of the

rabbit (Surridge et al., 1999a; Surridge et al., 1999b), which could lead to an underlying genetic structure, at a small geographical scale, not detected by Bayesian clustering methods.

Hierarchical genetic structure dependent on geographic distance

Overall, the rabbit in Spain showed considerable genetic structure, which was similar to that described for rabbit populations in the northeast and southwest of the Iberian Peninsula (Queney et al., 2001) and slightly lower than that reported for Britain (Surridge et al., 1999a). The fact that the largest percentage of genetic variation was explained separately by each locality indicates that genetic structure exists at a very local scale and reaffirms the importance of rabbit social behaviour in shaping its genetic structure (Surridge et al., 1999a).

Bayesian methods and AMOVA, as well as comparisons of F_{ST} and R_{ST} statistics, indicate that the rabbit has a hierarchical genetic structure. First, the oldest and largest differences are mainly reflected by the two genetic groups, based on nuclear markers, and their high R_{ST} values. Within these, there are other genetic groups identified in BAPS and determined by other factors that could be either environmental or ecological. In turn, these inferred populations will consist of even smaller groups conditioned by the social behaviour of the rabbit, and are reflected by the significant values of F_{ST} between localities and the significantly positive spatial autocorrelation (Lugon-Moulin et al., 1999; Balloux & Lugon-Moulin, 2002)

Besides this hierarchical genetic structure of the rabbit in Spain, geographic distance emerged as an important factor explaining genetic differentiation (Figs. 4 and 5). This contradicts the situation in Britain, where significant differences observed between locations could not be correlated with geographical distance (Surridge et al., 1999a). Similarly, it seems logical that main rivers, or other geographical features, constitute a barrier to gene flow in rabbits given their low dispersal capacity (Webb et al., 1995; Richardson et al., 2002). However, these barriers did not determine an increase in genetic differentiation explained solely by geographic distance. This is probably because, when considering the overall genetic structure and distribution of the rabbit, differences at a more local level have greater importance (Surridge et al., 1999a; Surridge et al., 1999b; Branco, Ferrand & Monnerot, 2000). This hypothesis is further supported by the results of our spatial autocorrelation analyses indicating significantly positive correlation up to a distance of about 100 km (Fig. 5A). However, the influence of geographic distance on genetic differentiation was not the same for all rabbit populations. Northern populations of *O. c. cuniculus* (K2) showed greater relatedness among close individuals and more spatial genetic structure and differentiation than the southern populations or those in the hybrid zone (Fig. 5B). This contrasting pattern could be due, in the first place, to the existence of genetic barriers among populations within each region. For example, it is well known that the Ebro River, running across northeastern Spain, has historically acted both as a physical and an ecological barrier for mammal species (O'Regan, 2008). As a matter of fact, the inferred genetics groups of K-NE and K-N are located to the north and south of the Ebro Valley, thus suggesting its role as a current barrier to gene flow (Fig. 1). Conversely, other large rivers in southern Spain (e.g. Guadiana River) do not seem to hinder gene flow among southern rabbit populations. Secondly, at a smaller geographic scale low dispersal might also be due both to resource availability and soil type, which at last

largely influence the distribution and social relationships of rabbits (Baker & Dunning, 1975; Cowan & Garson, 1985; Blanco & Villafuerte, 1993; Richardson et al., 2002; Lombardi et al., 2003) and other fossorial mammals (Lovegrove, 1989; Ebensperger & Cofré, 2001). It has been shown for other burrowing mammals, such as the wombat (*Lasiorhinus latifrons*), that where soft soils occur the construction of burrows is facilitated so animals do not need to share their shelter with other groups of individuals (Walker, Taylor & Sunnucks, 2007). Thus, the social structure of wombats in soft soils is characterized by closely related social groups and positive spatial correlation within a short distance, as observed in *O. c. cuniculus* in the northern Iberian Peninsula where softer soils also exist (Blanco & Villafuerte, 1993). In contrast, in hard soils, wombats share burrows with other individuals, and therefore are less related and spatial correlation is observed at a greater distance (Walker, Taylor & Sunnucks, 2007), as observed for the southern *O. c. algirus* populations (Fig. 5B).

Genetic variation within the hybrid zone

The large differences between the two rabbit lineages were evidenced by the maximum R_{ST} value obtained when considering the genetic variation among lineages A and B, which represent a divergence of 1,800,000 years – 2,000,000 (Branco, Ferrand & Monnerot, 2000; Carneiro, Ferrand & Nachman, 2009). The transition between these two genetic groups and mitochondrial lineages is well explained by a sigmoid curve. This was consistent with the Bayesian clustering of STRUCTURE, which indicates that the hybrid zone is not formed by individuals with a bimodal distribution of genotypes from the parental lineages, but instead they form a gradual cline of assignment probabilities to each group (Figs. 2 and 3). On the other hand, when geographic information was incorporated in the Bayesian clustering analysis of BAPS, which

usually helps to increase the power of analysis in cases where hierarchical structure might hinder the delineation of discrete groups on a smaller scale (Corander, Sirén & Arjas, 2008), the hybrid zone was shown as a large genetic cluster itself (Fig. 1). However, this result should be taken cautiously, since it could represent an artefact of the method. Firstly, Bayesian clustering methods can overestimate genetic structure when analysing scenarios under a pattern of isolation by distance (Frantz et al., 2009), or under strong linkage disequilibrium or departures from Hardy-Weinberg equilibrium (Falush, Stephens & Pritchard, 2003). Secondly, a kind of mixture linkage disequilibrium can occur even between physically unlinked loci, due to the correlation of allelic frequencies within populations. As a consequence, highly contrasting parental genotypes can lead to differences in this pattern of linkage disequilibrium and intermediate allele frequencies between these populations be interpreted as a distinct genetic cluster (Falush, Stephens & Pritchard, 2003; Kaeuffer et al., 2007).

Notwithstanding, hybrid zones can also be characterized by new genotypic combinations, resulting from crossings between genetically divergent individuals (Arnold et al., 1999). As expected from a region comprising the gene pools from both lineages of rabbit, the genetic diversity found in the hybrid zone was higher than in the parental populations, as reflected by the total number of alleles, allelic richness and expected heterozygosity (Table 1). Interestingly, the higher number of alleles was mainly due to. 49 alleles exclusively observed in this region, as opposed to the 37 and 21 exclusive alleles found in the parental populations. Although, this could be a mere consequence of the higher number of individuals found in this inferred cluster (n = 457), it is surprising that the hybrid zone shows so many exclusive alleles, when we would initially expect to represent only the sum of the parental alleles. Unique alleles have been previously described in the rabbit hybrid zone for the HBA haemoglobin alpha chain gene,

probably originated by recombination of alleles from the parental lineages (Campos, Storz & Ferrand, 2008). However, in the case of microsatellite loci, further evidence could suggest that these new alleles might be the result of an increased mutation rate caused by higher heterozygosity of the hybrids (Bradley et al., 1993; Hoffman & Brown, 1995; Amos & Harwood, 1998).

In the hybrid zone cluster, considered as the region with intermediate frequencies not belonging to any of the parental groups, the genetic contribution of each rabbit lineage was not balanced. In this area, the frequency of the two mitochondrial lineages is virtually the same (A = 0.485 and B = 0.515), but significantly greater proportions of individuals had been assigned to K2 (K1 = 0.37 and K2 = 0.63) (c^2 = 25.187, df = 1, $P < 0.0001$), showing a greater genetic introgression of lineage B, characteristic of the northern *O. c. cuniculus,* into lineage A, *O. c. algirus,* than vice versa. This is consistent with recent findings related to autosomal loci (Carneiro, Ferrand & Nachman, 2009; Carneiro et al., 2013), yet contrasts with that described for the X chromosome, suggesting a slightly higher number of migrants from the southwest to the northeast of the Iberian Peninsula (Geraldes, Ferrand & Nachman, 2006). The fact that the greatest contribution of the northern rabbit lineage is only reflected in the frequencies of nuclear markers and not in those of maternal inheritance could suggest that males are primarily responsible for this bias. Hence, if this would be the case, it would be expected that the Y chromosome was more introgressed than autosomal loci. On the contrary, it has been evidenced that the Y chromosome cline is highly stepped, as well as the mtDNA, which suggests some kind of selection acting against introgression (Geraldes, Ferrand & Nachman, 2006; Geraldes et al., 2008; Carneiro et al., 2013). In this regard, preliminary behavioural work discarded the existence of pre-mating reproductive selection between lineages, and found instead lower fertility in F1 males, thus following the

expectations of Haldane's rule (Haldane, 1922; Blanco-Aguiar et al., 2010). In this context it seems that the relative role of selection leading to postzygotic barriers has a stronger importance in shaping the genetic structure in the rabbit hybrid zone than behavioural and prezygotic barriers. Similarly, different types of selection have been detected at several autosomal loci, suggesting a wide range of evolutionary pressures across the rabbit's genome as well as across distribution range in the Iberian Peninsula (Campos, Storz & Ferrand, 2008; Carneiro et al., 2012; Carneiro et al., 2013).

Ultimately, multiple factors ranging from geographical and ecological, to behavioural and molecular, are interacting and shaping the overall genetic structure of the rabbit subspecies and their hybrid zone. Future studies based on genomic data coupled with behavioural and ecological information could potentially clarify these issues related to differences in genetic variation and to the structure of rabbit subspecies.

ACKNOWLEDGEMENTS

The authors wish to thank all those who provided samples for this study: the Consellería de Territori i Habitatge from Generalitat Valenciana, the Hunting Federation of Lérida, Hunting Society of Ajalvir (Madrid), M. Sanmartín from the University of Santiago de Compostela, S. Agudín, J. Inogés, J. Layna, F. Leiva, F. Silvestre and specially to F. Guil from Fundación CBD-Habitat. We also thank N. Ferrand and S. Lopes for data reviewing and helpful discussion of the results, as well as two anonymous referees for their thorough revisions. A. Burton reviewed the

449 English text, L. Alcaraz assisted in the laboratory and A. Benítez-López provided assistance with
450 statistical analyses.

451 REFERENCES

452 Alcover JA. 2008. The first Mallorcans: prehistoric colonization in the Western Mediterranean.
453 *Journal of Prehistory* 21:19-84.

454 Amos W, and Harwood J. 1998. Factors affecting levels of genetic diversity in natural
455 populations. *Philosophical Transactions of the Royal Society of London Series B-Biological*
456 *Sciences* 353:177-186.

457 Arnold ML. 1997. *Natural hybridisation and evolution*. New York: Oxford University Press.

458 Arnold ML, Bulger MR, Burke JM, Hempel AL, and Williams JH. 1999. Natural hybridization:
459 how long can you go and still be important? *Ecology* 80:371-381.

460 Avise JC. 2000. *Phylogeography: The history and formation of species*. Cambridge, MA:
461 Harvard University Press.

462 Baker AN, and Dunning RA. 1975. Effects of soil type and crop density on the activity and
463 abundance of the epigeic fauna, particularly Carabidae, in sugar-beet fields. *The Journal of*
464 *Applied Ecology* 12:809-818.

465 Balloux F, and Lugon-Moulin N. 2002. The estimation of population differentiation with
466 microsatellite markers. *Molecular Ecology* 11:155-165.

467 Barton N, and Gale KS. 1993. Genetic analysis of hybrid zones. In: Harrison RG, ed. *Hybrid*
468 *Zones and the Evolutionary Process*. New York: Oxford University Press.

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

471 Barton NH, and Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature* 341:497-503.

472 Belkhir K, Borsa P, Chikhi L, Raufaste N, and Bonhomme F. 2004. *GENETIX 4.05, logiciel sous*
473 *Windows*^{TM} *pour la génétique des populations.* Montpellier: Laboratoire Génome, Populations,
474 Interactions, CNRS UMR 5000, Université de Montpellier II.

475 Biju-Duval C, Ennafaa H, Dennebouy N, Monnerot M, Mignotte F, Soriguer RC, El Gaaïed A, El
476 Hili A, and Mounolou JC. 1991. Mitochondrial DNA evolution in lagomorphs: origin of

systematic heteroplasmy and organization of diversity in European rabbits. *Journal of Molecular Evolution* 33:92-102.

Blanco JC, and Villafuerte R. 1993. *Factores ecológicos que influyen sobre las poblaciones de conejos*. Madrid: Instituto Nacional para la Conservación de la Naturaleza.

Blanco-Aguiar JA, Carneiro M, Villafuerte R, Ríos-Saldaña CA, Ferreira C, and Ferrand N. 2010. Preliminary results assessing pre and postzygotic isolation between two subspecies, *Oryctogalus cuniculus algirus* and *O. c. cuniculus*, of the European rabbit. Trends in Biodiversity and Evolution (TiBE). Vairao, Portugal: Research Centre in Biodiversity and Genetic Resources, University of Porto.

Bradley BJ, Bull JJ, Johnson AD, and Hillis DM. 1993. Origin of a novel allele in a mammalian hybrid zone. *Proceedings of the National Academy of Sciences of the United States of America* 90:8939-8941.

Bradley BJ, Doran D, Robbins MM, Williamson E, Boesch C, and Vigilant L. 2002. Comparative analyses of genetic social structure in wild gorillas (*Gorilla gorilla*) using DNA from feces and hair. *American Journal of Physical Anthropology*:47-48.

Branco M, and Ferrand N. 2003. Biochemical and population genetics of the rabbit, *Oryctolagus cuniculus*, carbonic anhydrases I and II, from the Iberian Peninsula and France. *Biochemical Genetics* 41:391-404.

Branco M, Ferrand N, and Monnerot M. 2000. Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome b gene. *Heredity* 85:307-317.

Branco M, Machado JC, and Ferrand N. 1999. Extensive genetic polymorphism of peptidases A, B, C, and D, in wild rabbit (*Oryctolagus cuniculus*) populations from the Iberian Peninsula. *Biochemical Genetics* 37:237-249.

Branco M, Monnerot M, Ferrand N, and Templeton AR. 2002. Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution* 56:792-803.

Brelsford A, and Irwin DE. 2009. Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. *Evolution* 63:3050-3060.

Campos R, Storz JF, and Ferrand N. 2008. Evidence for contrasting modes of selection at interacting globin genes in the European rabbit (*Oryctolagus cuniculus*). *Heredity* 100:602-609.

Carneiro M, Albert FW, Melo-Ferreira J, Galtier N, Gayral P, Blanco-Aguiar JA, Villafuerte R, Nachman MW, and Ferrand N. 2012. Evidence for widespread positive and purifying selection across the European rabbit (*Oryctolagus cuniculus*) genome. *Molecular Biology and Evolution* 29:1837-1849.

Carneiro M, Baird SJE, Afonso S, Ramirez E, Tarroso P, Teotónio H, Villafuerte R, Nachman MW, and Ferrand N. 2013. Steep clines within a highly permeable genome across a hybrid zone between two subspecies of the European rabbit. *Molecular Ecology* 22:2511-2525.

Carneiro M, Blanco-Aguiar JA, Villafuerte R, Ferrand N, and Nachman MW. 2010. Speciation in the European rabbit (*Oryctolagus cuniculus*): islands of differentiation on the X chromosome and autosomes. *Evolution* 64:3443-3460.

Carneiro M, Ferrand N, and Nachman MW. 2009. Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* 181:593-606.

Corander J, and Marttinen P. 2006. Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* 15:2833-2843.

Corander J, Sirén J, and Arjas E. 2008. Bayesian spatial modeling of genetic population structure. *Computational Statistics* 23:111-129.

Cowan DP, and Garson PJ. 1985. Variations in the social structure of rabbit populations: causes and demographic consequences. In: Smith RM, and Sibly RH, eds. *Behavioural Ecology: Ecological Consequences of Adaptive Behaviour The 25th Symposium of the British Ecological Society*. Oxford, Reading: Blackwell Scientific Publications, 537-555.

Ebensperger LA, and Cofré H. 2001. On the evolution of group-living in the new world cursorial hystricognath rodents. *Behavioural Ecology* 12:227-236.

Evanno G, Regnaut S, and Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.

Excoffier L, Smouse PE, and Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. *Genetics* 131:479-491.

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

Ferrand N. 1989. Biochemical and genetic studies on rabbit hemoglobin. I. Electrophoretic poymorphism of the b-chain. *Biochemical Genetics* 27:673-678.

Ferrand N, and Branco M. 2007. The evolutionary history of the European rabbit (*Oryctolagus cuniculus*): Major patterns of population differentiation and geographic expansion inferred from protein polymorphism. In: Weiss S, and Ferrand N, eds. *Phylogeography of European Refugia*. Netherlands: Springer, 207-235.

Ferreira C. 2012. European rabbit research in the Iberian Peninsula: state of the art and future perspectives. *European Journal of Wildlife Research* 58:885-895.

Flux JEC. 1994. World distribution. In: Thompson HV, and King CM, eds. *The European Rabbit The History and Biology of a Successful Colonizer*. Oxford: Oxford University Press, 8-21.

Frantz AC, Cellina S, Krier A, Schley L, and 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.

Fuller SJ, Wilson JC, and Mather PB. 1997. Patterns of differentiation among wild rabbit populations *Oryctolagus cuniculus* L. in arid and semiarid ecosystems of north-eastern Australia. *Molecular Ecology* 6:145-153.

Futuyma DJ. 1998. *Evolutionary Biology*. Sunderland: Sinauer Associates.

Geraldes A, Carneiro M, Delibes-Mateos M, Villafuerte R, Nachman MW, and Ferrand N. 2008. Reduced introgression of the Y chromosome between subspecies of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula. *Molecular Ecology* 17:4489-4499.

Geraldes A, Ferrand N, and Nachman MW. 2006. Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* 173:919-933.

Geraldes A, Rogel-Gaillard C, and Ferrand N. 2005. High levels of nucleotide diversity in the European rabbit (*Oryctolagus cuniculus*) SRY gene. *Animal Genetics* 36:349-351.

Goudet J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86:485-486.

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

Haldane JBS. 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics* 12:101-109.

Hardy C, Vigne J-D, Casane D, Dennebouy N, Mounolou J-C, and Monnerot M. 1994. Origin of European rabbit (*Oryctolagus cuniculus*) in a Mediterranean island: Zooarchaeology and ancient DNA examination. *Journal of Evolutionary Biology* 7:217-226.

Hardy OJ, and Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2:618-620.

Harrison RG. 1993. Hybrids and hybrid zones: historical perspective. In: Harrison RG, ed. *Hybrid zones and the evolutionary process*. New York: Oxford University, 3-12.

Hedrick PW. 2005. *Genetics of populations*. Boston, MA: Jones and Bartlett.

Hewitt G. 1988. Hybrid zones-Natural laboratories for evolutionary studies. *Trends in Ecology & Evolution* 3:158-167.

Hoffman SMG, and Brown WM. 1995. The molecular mechanism underlying the "rare allele phenomenon" in a subspecific hybrid zone of the california field mouse, *Peromyscus californicus*. *Journal of Molecular Evolution* 41:1165-1169.

Hutchison DW, and Templeton AR. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* 53:1898-1914.

585 Kaeuffer R, Réale D, Coltman DW, and Pontier D. 2007. Detecting population structure using
586 STRUCTURE software: effect of background linkage disequilibrium. *Heredity* 99:374-380.

587 Knowles LL, and Carstens BC. 2007. Estimating a geographically explicit model of population
588 divergence. *Evolution* 61:477-493.

589 Loiselle BA, Sork VL, Nason J, and Graham C. 1995. Spatial genetic structure of a tropical
590 understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82:1420-1425.

591 Lombardi L, Fernández N, Moreno S, and Villafuerte R. 2003. Habitat-related differences in
592 rabbit (*Oryctolagus cuniculus*) abundance, distribution, and activity. *Journal of Mammalogy*
593 84:26-36.

594 Lovegrove BG. 1989. The cost of burrowing by the social mole rats (Bathyergidae) *Cryptomys*
595 *damarensis* and *Heterocephalus glaber*: the role of soil moisture. *Physiological Zoology* 62:449-
596 469.

597 Lugon-Moulin N, Brunner H, Wyttenbach A, Hausser J, and Goudet J. 1999. Hierarchical
598 analyses of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora : Soricidae).
599 *Molecular Ecology* 8:419-431.

600 Mantel NA. 1967. The detection of disease clustering and a generalized regression approach.
601 *Cancer Research* 27:209-220.

602 Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic
603 differentiation measure. *Evolution* 60:2399-2402.

604 Meirmans PG, and Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for
605 the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792-794.

606 Monnerot M, Vigne JD, Biju-Duval C, Casane D, Callou C, Hardy C, Mougel F, Soriguer R,
607 Dennebouy N, and Mounolou J. 1994. Rabbit and man: genetic and historic approach. *Genetics*
608 *Selection Evolution* 26:167-182.

609 Mougel F, Mounolou J, and Monnerot M. 1997. Nine polymorphic microsatellite loci in the
610 rabbit, *Oryctolagus cuniculus*. *Animal Genetics* 28:58-71.

611 Nei M, Maruyama T, and Chakraborty R. 1975. The bottleneck effect and genetic variability in
612 populations. *Evolution* 29:1-10.

613 O'Regan HJ. 2008. The Iberian Peninsula - corridor or cul-de-sac? Mammalian faunal exchange
614 and possible routes of dispersal in the last 2 million years. *Quaternary Science Reviews* 27:2136-
615 2144.

616 Palsboll PJ. 1999. Genetic tagging: contemporary molecular ecology. *Biological Journal of the*
617 *Linnean Society* 68:3-22.

618 Pilot M, Jedrzejewski W, Branicki W, Sidorovich VE, Jedrzejewska B, Stachura K, and Funk SM.
619 2006. Ecological factors influence population genetic structure of European grey wolves.
620 *Molecular Ecology* 15:4533-4553.

621 Pritchard JK, Stephens M, and Donnelly P. 2000. Inference of population structure using
622 multilocus genotype data. *Genetics* 155:945-959.

623 Queney G, Ferrand N, Marchandeau S, Azevedo M, Mougel F, Branco M, and Monnerot M.
624 2000. Absence of a genetic bottleneck in a wild rabbit (*Oryctolagus cuniculus*) population
625 exposed to a severe viral epizootic. *Molecular Ecology* 9:1253-1264.

626 Queney G, Ferrand N, Weiss S, Mougel F, and Monnerot M. 2001. Stationary distributions of
627 microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus*
628 *cuniculus*). *Molecular Biology and Evolution* 18:2169-2178.

629 R Core Team. 2014. R: A language and environment for statistical computing. Vienna, Austria: R
630 Foundation for Statistical Computing.

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

633 Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.

634 Richardson BJ, Hayes RA, Wheeler CM, and Yardin MR. 2002. Social structures, genetic
635 structures and dispersal strategies in Australian rabbit (*Oryctolagus cuniculus*) populations.
636 *Behavioural Ecology and Sociobiology* 51:113-121.

637 Seixas FA, Juste J, Campos PF, Carneiro M, Ferrand N, Alves PC, and Melo-Ferreira J. 2014.
638 Colonization history of Mallorca Island by the European rabbit, *Oryctolagus cuniculus*, and the
639 Iberian hare, *Lepus granatensis* (Lagomorpha: Leporidae). *Biological Journal of the Linnean*
640 *Society* 111:748-760.

641 Shoemaker DD, and Ross KG. 1996. Effects of social organization on gene flow in the fire ant
642 *Solenopsis invicta*. *Nature* 383:613-616.

643 Smouse PE, and Peakall R. 1999. Spatial autocorrelation analysis of individual multiallele and
644 multilocus genetic structure. *Heredity* 82:561-573.

645 Surridge AK, Bell DJ, Ibrahim KM, and Hewitt G. 1999a. Population structure and genetic
646 variation of European wild rabbits (*Oryctolagus cuniculus*) in East Anglia. *Heredity* 82:479-487.

647 Surridge AK, Bell DJ, Rico C, and Hewitt GM. 1997. Polymorphic microsatellite loci in the
648 European rabbit (*Oryctolagus cuniculus*) are also amplified in other lagomorph species. *Animal*
649 *Genetics* 28:302-305.

650 Surridge AK, Ibrahim KM, Bell DJ, Webb NJ, Rico C, and Hewitt GM. 1999b. Fine-scale genetic
651 structuring in a natural population of European wild rabbits (*Oryctolagus cuniculus*). *Molecular*
652 *Ecology* 8:299-307.

653 van Oosterhout C, Hutchinson WF, Wills DPM, and Shipley P. 2004. MICRO-CHECKER:
654 software for identifying and correcting genotyping errors in microsatellite data. *Molecular*
655 *Ecology Notes* 4:535-538.

656 Vekemans X, and Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses
657 in plant populations. *Molecular Ecology* 13:921-935.

658 Walker FM, Taylor AC, and Sunnucks P. 2007. Does soil type drive social organization in
659 southern hairy-nosed wombats? *Molecular Ecology* 16:199-208.

660 Webb NJ, Ibrahim KM, Bell DJ, and Hewitt GM. 1995. Natal dispersal and genetic structure in a
661 population of the European wild rabbit (*Oryctolagus cuniculus*). *Molecular Ecology* 4:239-248.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Rabbit localities sampled in this work, number of individuals analysed, geographical coordinates and subspecies occurring in each locality according to it its natural distribution range. Numbers correspond to those indicated in Fig. 1.

Supplementary Table 2. Genetic diversity statistics for all the rabbit localities analysed. N = number of samples, N_A = number of alleles, H_o = observed heterozygosity, H_e = expected heteroygosity, F_{IS} = inbreeding coefficient.

Supplementary Table 3. Genetic diversity statistics for the rabbit populations inferred in BAPS. N = number of samples, N_A = number of alleles, H_o = observed heterozygosity, H_e = expected heteroygosity, F_{IS} = inbreeding coefficient.

Table 1 (on next page)

Genetic diversity statistics based on 10 microsatellite loci genotypes for the rabbit localities analysed

Genetic diversity statistics based on 10 microsatellite loci genotypes for the rabbit localities (≥10 individuals) analysed and the genetic clusters inferred in BAPS. n = number of samples, N_A = mean number of alleles per locus, A_R = allelic richnes, H_o = observed heteroyzygosity, H_e = expected heterozygosity, F_{IS} = inbreeding coefficient. The proportion of haplotypes from rabbit A and B lineages is shown for each locality and cluster. Numbers of each locality correspond to those in Fig. 1.

Table 1. Genetic diversity statistics based on 10 microsatellite loci genotypes for the rabbit localities (≥10 individuals) analysed and the genetic clusters inferred in BAPS. n = number of samples, N_A = mean number of alleles per locus, A_R = allelic richnes, H_o = observed heteroyzygosity, H_e = expected heterozygosity, F_{IS} = inbreeding coefficient. The proportion of haplotypes from rabbit A and B lineages is shown for each locality and cluster. Numbers of each locality correspond to those in Fig. 1.

	n	N_A	A_R	H_o	H_e	F_{IS}	Hap A	Hap B	Subspecies
Locality									
1. Mallorca	14	6	5.15	0.63	0.71	0.15	0.08	0.92	*O. c. cuniculus*
2. Lérida	50	9.64	6.20	0.76	0.79	0.06	0.00	1.00	*O. c. cuniculus*
3. Valencia	18	9.18	7.13	0.68	0.80	**0.18**	0.06	0.94	*O. c. cuniculus*
4. La Rioja	19	7.00	5.62	0.72	0.74	0.06	0.05	0.95	*O. c. cuniculus*
5. Galicia	27	8.73	6.42	0.61	0.80	**0.26**	0.00	1.00	*O. c. cuniculus*
6. Valladolid	16	7.91	6.41	0.67	0.73	0.11	0.07	0.93	*O. c. cuniculus*
7. Madrid1	51	12.45	7.70	0.68	0.86	**0.22**	0.40	0.60	hybrid zone
10. Cuenca	42	12.18	7.40	0.75	0.82	**0.11**	0.02	0.98	hybrid zone
11. Toledo1	26	8.91	6.61	0.72	0.81	**0.13**	0.33	0.67	hybrid zone
12. Toledo2	33	10.91	7.09	0.68	0.83	**0.19**	0.40	0.60	hybrid zone
13. Toledo3	24	10.45	7.08	0.77	0.83	**0.09**	0.60	0.40	hybrid zone
15. Toledo5	19	8.73	6.99	0.68	0.81	**0.19**	0.37	0.63	hybrid zone
16. Toledo6	11	7.82	7.08	0.78	0.82	0.09	0.36	0.64	hybrid zone
17. Ciudad Real1	51	12.36	7.26	0.69	0.84	**0.19**	0.84	0.16	*O. c. algirus*
18. Ciudad Real2	27	9.73	7.09	0.64	0.83	**0.25**	0.89	0.11	*O. c. algirus*
19. Ciudad Real3	50	12.27	7.34	0.70	0.82	**0.16**	0.82	0.18	*O. c. algirus*
20. Albacete	25	9.64	6.86	0.75	0.81	0.10	0.48	0.52	*O. c. algirus*
21. Cáceres1	10	6.64	6.22	0.70	0.72	**0.10**	0.90	0.10	*O. c. algirus*
22. Cáceres2	28	9.00	6.51	0.63	0.78	0.21	0.89	0.11	*O. c. algirus*
23. Badajoz1	20	9.73	6.82	0.71	0.79	0.13	0.95	0.05	*O. c. algirus*
24. Badajoz2	29	9.45	6.99	0.79	0.82	0.05	0.90	0.10	*O. c. algirus*
25. Jaén1	15	8.64	6.89	0.63	0.78	**0.23**	0.00	1.00	*O. c. algirus*
27. Jaén3	22	9.64	7.13	0.70	0.82	**0.16**	1.00	0.00	*O. c. algirus*
28. Sevilla1	43	11.36	6.89	0.72	0.80	**0.10**	1.00	0.00	*O. c. algirus*
29. Sevilla2	32	11.82	7.62	0.72	0.83	**0.14**	0.63	0.38	*O. c. algirus*
30. Cádiz	56	11.64	6.92	0.69	0.80	**0.14**	0.96	0.04	*O. c. algirus*
Cluster									
K-BAL	14	6.00	4.35	0.63	0.71	0.15	0	1.00	
K-NE	52	9.55	5.09	0.74	0.80	0.07	0	1.00	
K-N	21	7.00	4.61	0.69	0.74	0.10	0	1.00	
K-CU	14	6.73	4.99	0.69	0.74	0.13	0.07	0.93	
K-ZH	457	21.00	6.20	0.70	0.87	**0.20**	0.49	0.52	
K-S1	206	16.36	5.72	0.71	0.83	**0.14**	1.00	0.00	

F_{IS} values in bold represent significant deviations form Hardy-Weinberg equilibrium, after Bonferroni correction.

Table 2(on next page)

Proportion of mitochondrial lineage and average assignment probabilities for each locality to the genetic populations inferred.

Proportion of lineage A and B rabbits and average assignment probability of each locality to the populations inferred in STRUCTURE and BAPS. Numbers of each locality correspond to those in Fig. 1.

Table 2. Proportion of lineage A and B rabbits and average assignment probability of each locality to the populations inferred in STRUCTURE and BAPS. Numbers of each locality correspond to those in Fig. 1.

		mtDNA		STRUCTURE		BAPS					
	n	Hap A	Hap B	K1	K2	K-BAL	K-NE	K-N	K-CU	K-ZH	K-S1	K-S2
Locality												
1. Mallorca	14	0.08	0.92	0.09	0.91	1.00						
2. Lérida	50		1.00	0.04	0.96		0.98			0.02		
3. Valencia	18	0.06	0.94	0.26	0.74		0.11			0.89		
4. La Rioja	19	0.05	0.95	0.06	0.94			0.95		0.05		
5. Galicia	27		1.00	0.44	0.56					0.89	0.11	
6. Valladolid	16	0.07	0.93	0.19	0.82		0.06			0.88	0.06	
7. Madrid1	51	0.40	0.60	0.38	0.62			0.02		0.75	0.24	
8. Madrid2	7	0.29	0.71	0.06	0.94			0.29		0.71		
9. Madrid3	2	1.00		0.15	0.85					1.00		
10. Cuenca	42	0.02	0.98	0.12	0.88				0.33	0.67		
11. Toledo1	26	0.33	0.67	0.23	0.77					1.00		
12. Toledo2	33	0.40	0.60	0.19	0.81					1.00		
13. Toledo3	24	0.60	0.40	0.16	0.84					1.00		
14. Toledo4	2		1.00	0.08	0.92					1.00		
15. Toledo5	19	0.37	0.63	0.16	0.84					1.00		
16. Toledo6	11	0.36	0.64	0.21	0.79					1.00		
17. Ciudad Real1	51	0.84	0.16	0.67	0.33					0.92	0.08	
18. Ciudad Real2	27	0.89	0.11	0.47	0.53					0.96	0.04	
19. Ciudad Real3	50	0.82	0.18	0.66	0.35					0.96	0.04	
20. Albacete	25	0.48	0.52	0.16	0.84					1.00		
21. Cáceres1	10	0.90	0.10	0.70	0.30					0.70	0.30	
22. Cáceres2	28	0.89	0.11	0.60	0.40					0.89	0.11	
23. Badajoz1	20	0.95	0.05	0.92	0.09					0.05	0.95	
24. Badajoz2	29	0.90	0.10	0.94	0.06					0.03	0.97	
25. Jaén1	15		1.00	0.52	0.48					0.50	0.50	
26. Jaén2	2	0.93	0.07	0.87	0.14					0.47	0.53	
27. Jaén3	22	1.00		0.82	0.18					0.82	0.18	
28. Sevilla1	43	1.00		0.91	0.09					0.02	0.98	
29. Sevilla2	32	0.63	0.38	0.86	0.14					0.03	0.75	0.22
30. Cádiz	56	0.96	0.04	0.91	0.09					0.09	0.91	
Population												
K-BAL	14		1.00	0.09	0.91							
K-NE	52		1.00	0.04	0.96							
K-N	21		1.00	0.05	0.95							
K-CU	14	0.07	0.93	0.06	0.94							
K-ZH	457	0.49	0.52	0.37	0.63							
K-S1	206	0.90	0.10	0.93	0.06							
K-S2	7	1.00		0.89	0.11							

Table 3(on next page)

AMOVA analyses

AMOVA analysis performed for different levels of genetic structure among the rabbit localities analysed and the inferred clusters.

Table 3. AMOVA analysis performed for different levels of genetic structure among the rabbit localities analysed and the inferred clusters.

Structure	% variation		*F'*	*p*		*R*	*p*
Localities							
All localities	0.287	F_{ST}'	0.37	0.001	R_{ST}	0.11	0.001
A Haplotypes Vs. B Haplotypes	0.111	F_{ST}'	0.173	0.001	R_{ST}	0.637	0.001
Clusters							
STRUCTURE K1 Vs. K2	0.192	F_{ST}'	0.26	0.001	R_{ST}	0.635	0.001
All clusters BAPS	0.253	F_{ST}'	0.325	0.001	R_{ST}	0.627	0.001

Figure 1

Maps of the Iberian Peninsula indicating the rabbit localities analysed

Maps of the Iberian Peninsula indicating the localities analysed (numbers correspond to localities in Supplementary Table 1), the hybrid zone (white dotted line), the perpendicular transect (black dotted line) and the average individual assignment probabilities for the 7 clusters inferred in BAPS. Colour gradient from grey (or green) to white denotes assignment probabilities for each population from 0 to 1.

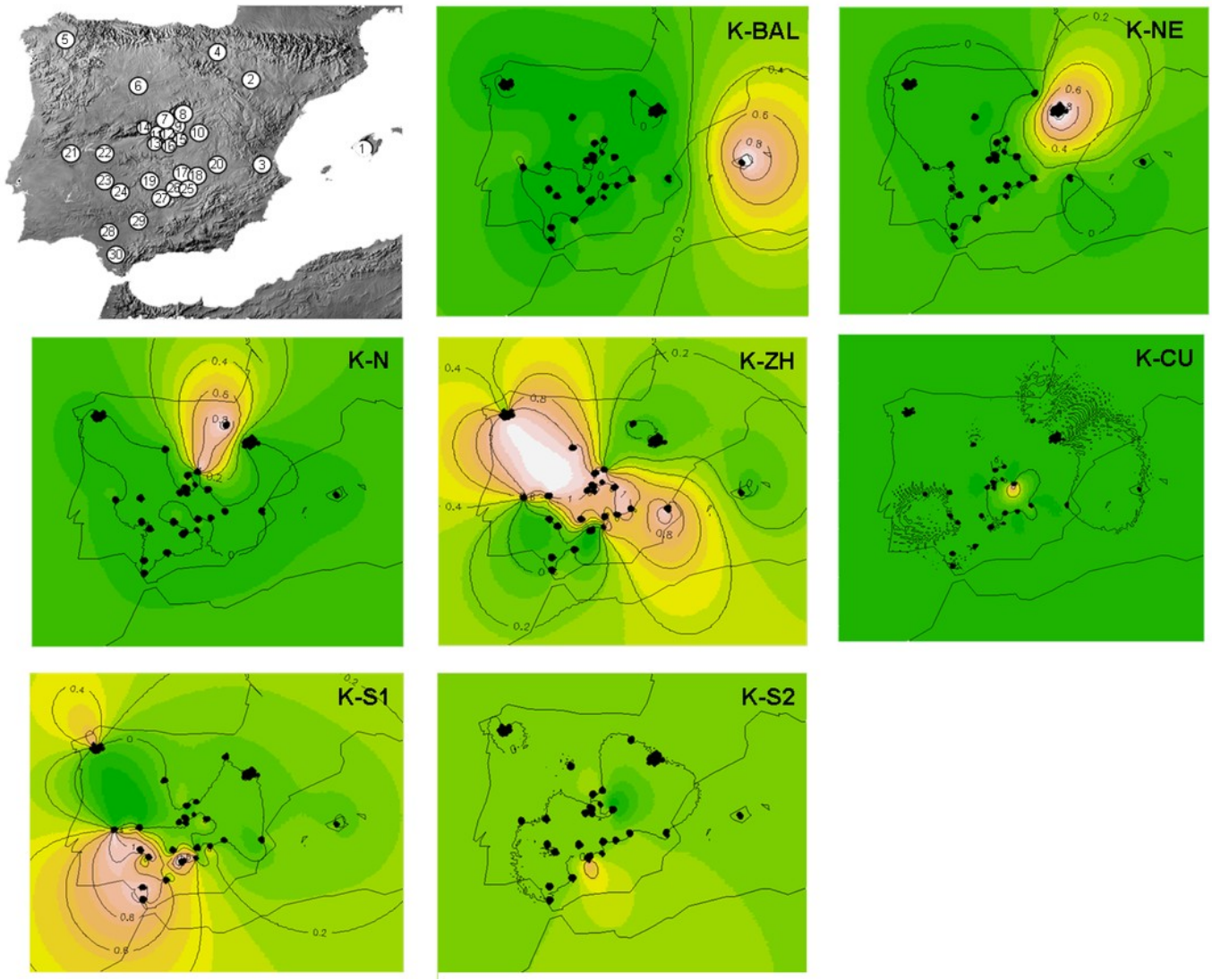

Figure 2

Individual assignment probabilities to the genetic groups inferred in STRUCTURE.

Individual assignment probabilities (*q*) to genetic group K1 (A). Each dot represents an individual, and grey bars are the confidence intervals obtained for their assignment probabilities. Dotted lines indicate assignment probabilities to group K1 larger than 0.9 and lower than 0.1. The number of individuals assigned within these intervals and the proportion of their mitochondrial lineages are indicated.

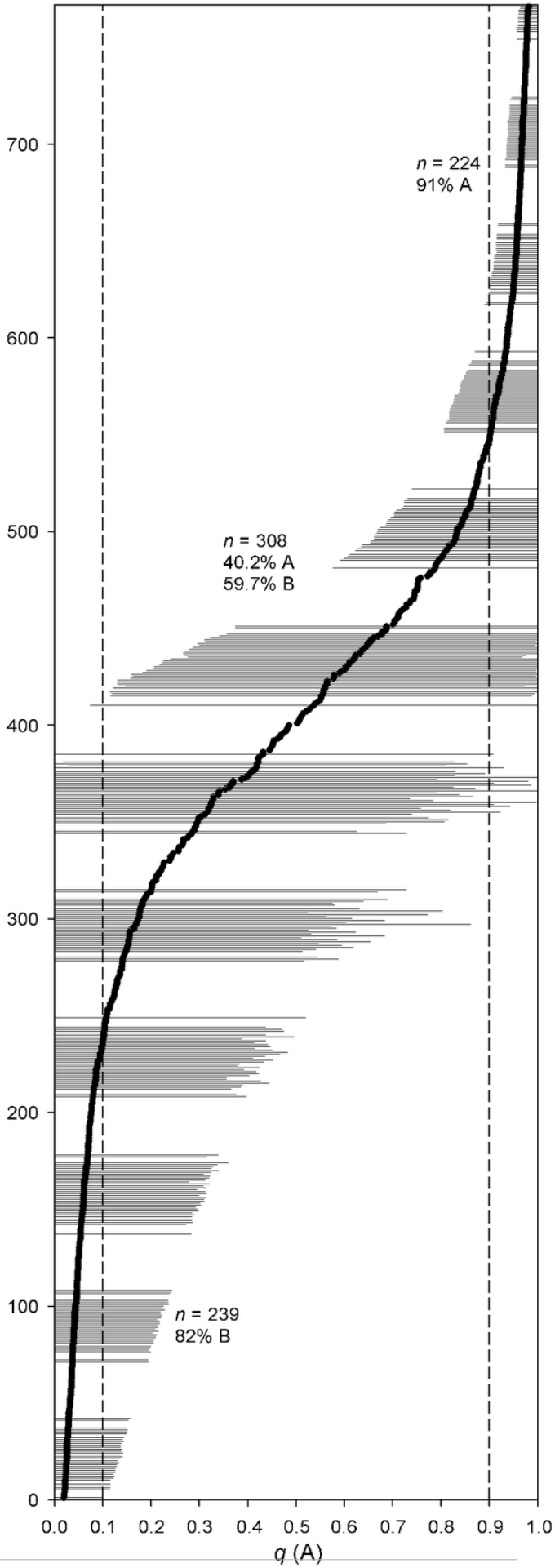

Figure 3

Clinal patterns for the mitochondrial and nuclear markers along the rabbit hybrid zone.

Clinal patterns for the mitochondrial (grey lines) and nuclear markers (black lines) along the hybrid zone transect from southwestern to northeastern Spain (see Fig. 1). Dots represent the frequency of lineage A mitochondrial haplotypes (grey) and mean assignment probabilities (*q*) to genetic group K1 (A) in each locality. Distances are in km, starting (km 0) at the intersection between the transect and the hybrid zone. Negative distance values indicate km to the south and positive values km to the north.

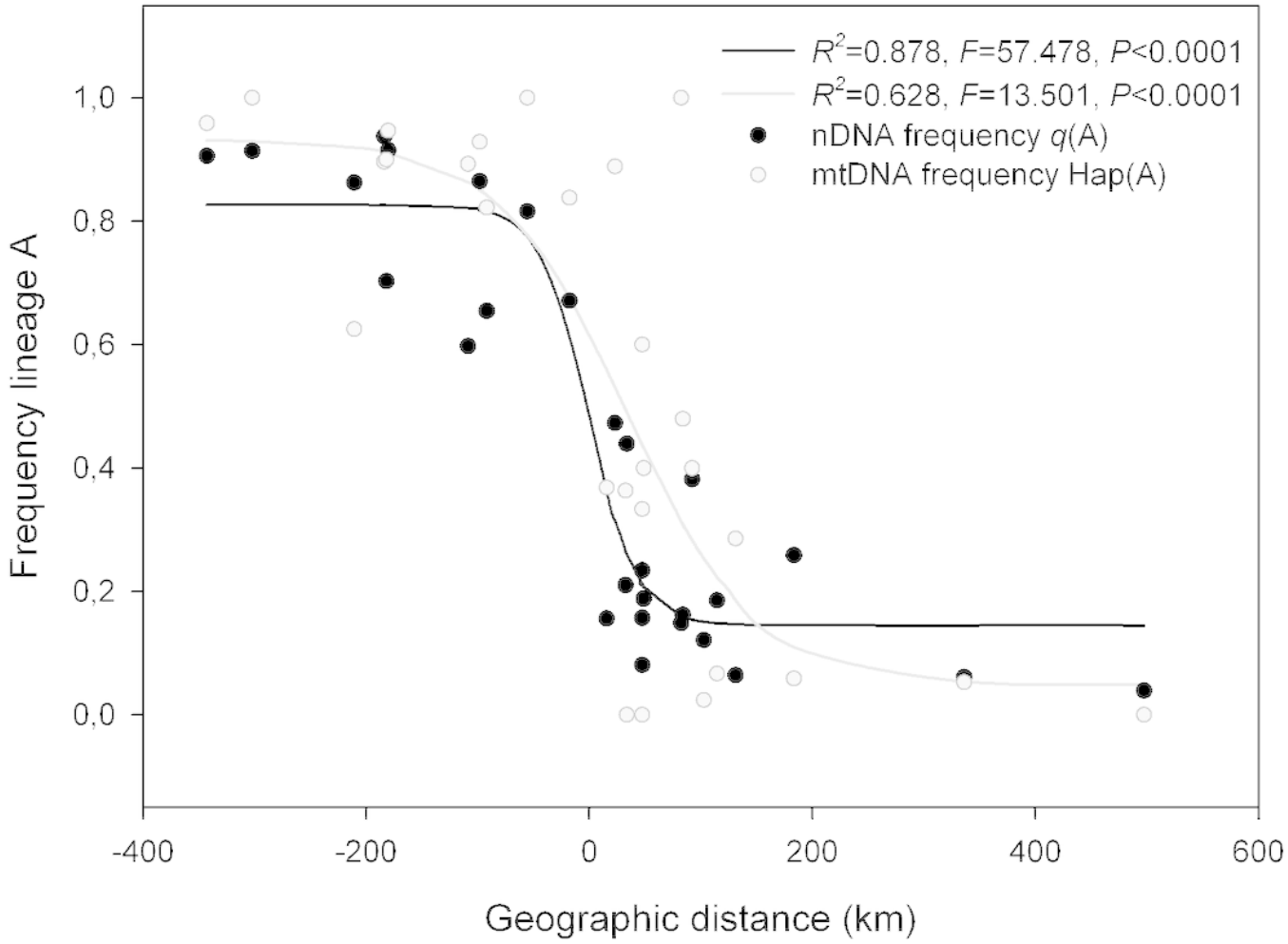

Figure 4

Isolation by distance among rabbit localities

Isolation by distance for all the localities of *O. cuniculus* analysed, as shown by the correlation of genetic distances (F_{ST}/1-F_{ST}) and geographic distances (Mantel test).

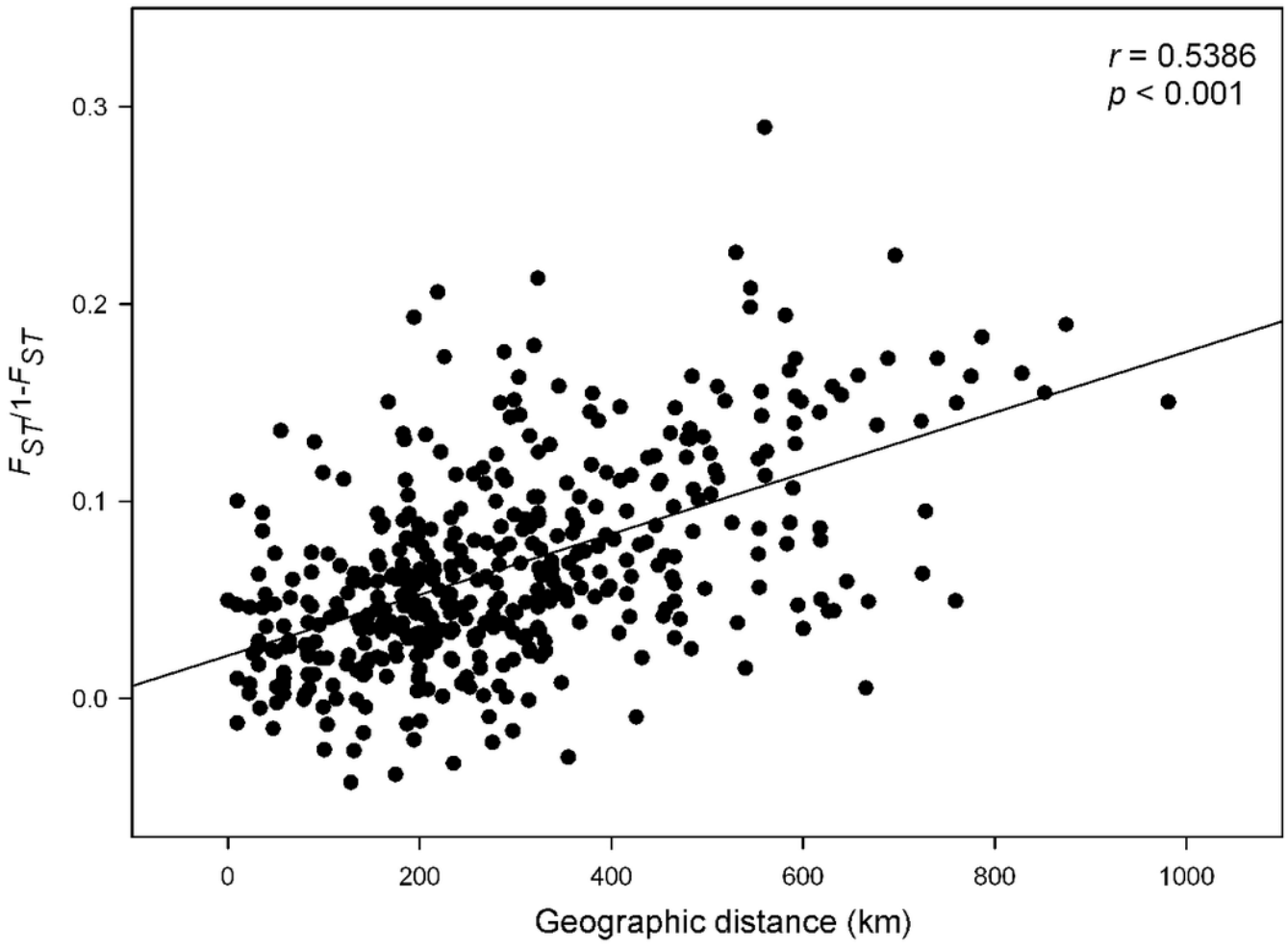

Figure 5

Spatial autocorrelation analyses

Spatial autocorrelation analyses showing the average inbreeding coefficient (f_{ij}) for each distance interval among individuals (d_{ij}), for the complete dataset (A) and for each of the subspecies analysed and the hybrid zone separately (B). Black symbols represent significant correlations between f_{ij} and d_{ij}.