

A peer-reviewed version of this preprint was published in PeerJ on 30 September 2014.

[View the peer-reviewed version](https://doi.org/10.7717/peerj.582) (peerj.com/articles/582), which is the preferred citable publication unless you specifically need to cite this preprint.

Alda F, Doadrio I. 2014. Spatial genetic structure across a hybrid zone between European rabbit subspecies. PeerJ 2:e582
<https://doi.org/10.7717/peerj.582>

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.

1 Fernando ALDA^{1,2*}, Ignacio DOADRIO¹

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

4 ²Current address: Center for Bioenvironmental Research, Tulane University, New Orleans,
5 Louisiana, 70118, USA

6 Corresponding author:

7 Fernando ALDA

8 Center for Bioenvironmental Research (CBR)

9 Tulane University

10 New Orleans, LA 70118, USA

11 Tel: (+1) 504 862 8441

12 Fax: (+1) 504 862 8455

13 e-mail: alda.fernando@gmail.com

14 INTRODUCTION

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

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

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

35 (Ferrand, 1989; Biju-Duval et al., 1991; Branco, Ferrand & Monnerot, 2000; Geraldès et al.,
36 2008; Carneiro et al., 2012), temporal (Hardy et al., 1994; Monnerot et al., 1994) and
37 geographical perspectives (Webb et al., 1995; Fuller, Wilson & Mather, 1997; Surridge et al.,
38 1999b; Queney et al., 2000; Queney et al., 2001; Branco et al., 2002).

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

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

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

63 Hybrid zones are areas of particular interest for evolutionary studies enabling insight into the
64 initial stages of speciation and reproductive isolation, adaptation and selection, and even
65 behavioral processes (Hewitt, 1988; Barton & Gale, 1993; Arnold, 1997; Futuyma, 1998). In the
66 European rabbit, extensive studies in the hybrid zone have evidenced highly contrasting degrees
67 of introgression among loci, or even a complete absence of genetic structure (Branco, Machado
68 & Ferrand, 1999; Queney et al., 2001; Geraldes, Ferrand & Nachman, 2006; Ferrand & Branco,
69 2007; Campos, Storz & Ferrand, 2008; Carneiro, Ferrand & Nachman, 2009; Carneiro et al.,
70 2010; Carneiro et al., 2013). The variation in the introgression of autosomal and sexual
71 chromosomes has also revealed different selective pressures across genes and its importance in
72 the reproductive isolation between the two rabbit subspecies (Geraldes, Ferrand & Nachman,
73 2006; Campos, Storz & Ferrand, 2008; Carneiro, Ferrand & Nachman, 2009; Carneiro et al.,
74 2010; Carneiro et al., 2013).

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

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

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

92 MATERIALS AND METHODS

93 *Sample collection*

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

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

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

104 *DNA extraction and amplification of molecular markers*

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

109 Mitochondrial lineages, A or B, were identified in all samples by amplifying the complete
110 cytochrome *b* gene using the primers OcunCB_F: 5'-ATGACCAACATTCGCAAAACC-3' and
111 OcunCB_R: 5'-TGTCTCAGGGAGAACTATCTCC-3'. The PCR reaction was performed in a
112 final volume of 25 μ L containing: 200 μ M of dNTPs, 0.2 μ M of each primer, 1 U *Taq* polymerase
113 (Eppendorf), 1x PCR buffer (500 mM KCl, 100 mM Tris-HCl pH8.3, 15 mM Mg²⁺), and 1 μ L of
114 DNA extract. The PCR program consisted of 4 min of denaturation at 94 °C, followed by 40

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

120 All individuals were genotyped according to 10 microsatellite markers: Sat3, Sat4, Sat5, Sat7,
121 Sat8, Sat12, Sat13, Sat16, Sol33 and Sol44 (Mougel, Mounolou & Monnerot, 1997; Surridge et
122 al., 1997). PCR reactions were performed in a final volume of 13 µL that contained 200 µM
123 dNTPs, 0.2-0.4 µM of each primer, 2-2.5 mM MgCl₂, 0.325 U *Taq* polymerase (Eppendorf), 1x
124 PCR buffer (500 mM KCl, 100 mM Tris-HCl pH 8.3), and 0.5 µL of DNA extract. PCR programs
125 involved 2 min of initial denaturation at 95 °C, followed by 35 cycles of 30 sec at 95 °C, 30 sec
126 between 55 °C and 65 °C, 30 sec at 72 °C, followed by a final extension step of 7 min at 72 °C.
127 The amplified fragments were analyzed in an ABI3730 automatic sequencer (Applied
128 Biosystems) and allele sizes were assigned using the program GeneMapper v3.7 (Applied
129 Biosystems). The complete data file of microsatellite genotypes and mitochondrial haplotypes
130 was deposited as a Supplementary Dataset in PeerJ.

131 *Statistical analyses*

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

136 between lineages, we also analyzed whether models with more than two clusters were supported
 137 by our data. Therefore, firstly we used the algorithm of STRUCTURE 2.2 (Pritchard, Stephens &
 138 Donnelly, 2000) implementing the admixture model with correlated allele frequencies (Falush,
 139 Stephens & Pritchard, 2003), since this model is more appropriate for individuals with admixed
 140 ancestries and for populations with similar expected frequencies. No information on the localities
 141 of origin of individuals was included. Ten independent analyses were run for each value of K ,
 142 from $K = 1$ to $K = 30$. Each analysis consisted of 1×10^6 Markov chains with a prior burn-in of 1
 143 $\times 10^5$ chains. Mean posterior probability values were used to calculate ΔK , a measure of the rate
 144 of change of the posterior probabilities between successive K values. Thus, it is possible to detect
 145 when the increase in $\ln P(X|D)$ is not significant anymore and find the true value of K (Evanno,
 146 Regnaut & Goudet, 2005).

147 To visually explore the distribution of the inferred genetic groups across the hybrid zone, the
 148 proportion of genetic admixture and the frequency of each mitochondrial lineage were plotted
 149 along a one-dimensional transect perpendicular to the proposed rabbit hybrid zone (Branco,
 150 Ferrand & Monnerot, 2000; Branco et al., 2002; Geraldès et al., 2008), with the exception of the
 151 localities of Galicia and Mallorca which were excluded (Fig. 1). We considered km 0 of the
 152 linear transect to be the approximate intersection between the transect that the hybrid zone (Fig.
 153 1). Geographical locations along this linear transect were fitted to a sigmoid curve (3 parameters)
 154 as expected by hybrid zone tension zone models (Barton & Hewitt, 1985; Barton & Gale, 1993)
 155 in SigmaPlot 10.0 (Systat Software, Inc.). Additionally, congruence between the assignment
 156 probabilities and the frequency of mitochondrial lineages was evaluated by performing χ^2 tests in
 157 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 neighboring 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 analyzed 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). We calculated parameters of genetic diversity such
170 as number of alleles (N_A), allelic richness (A_R), observed and expected heterozygosity (H_o and H_e)
171 and inbreeding coefficient (F_{IS}) for each locus and genetic group using the programs GENETIX
172 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

178 results obtained can be compared between markers showing different polymorphism (Meirmans,
179 2006). To avoid confusion, this ratio is hereafter referred to as F_{ST} .

180 The effect of geographical distance on the genetic differentiation between individuals was also
181 tested for all the samples in peninsular Spain (i.e. excluding the population from the island of
182 Mallorca) and for each lineage and the hybrid zone separately. Regression was performed for the
183 kinship coefficient between pairs of individuals (f_{ij}) (Loiselle et al., 1995) and their geographical
184 distance (d_{ij}), to give a regression slope $\ln b_d$ and its statistical significance (Vekemans & Hardy,
185 2004). Also, spatial autocorrelation methods were applied to examine spatial genetic structure
186 (Smouse & Peakall, 1999). Geographic locations of individuals were permuted 10,000 times
187 among 50 distance intervals with an equal number of comparisons between individuals (5935) to
188 test the null hypothesis that d_{ij} and f_{ij} were not correlated. Positive spatial autocorrelations are
189 expected when gene flow is restricted to short distances. These tests are dependent on the type
190 and scale of sampling (Vekemans & Hardy, 2004), so to compare the extent of spatial genetic
191 structure in each subspecies of rabbit and the hybrid zone, we used the statistical Sp (Vekemans &
192 Hardy, 2004), $Sp = -\ln b_d / (1 - F_{ij})$, where F_{ij} is the average kinship coefficient between individuals
193 closer together (the first distance interval, $\sim 5\text{km}$), and $\ln b_d$ is the slope of the linear regression of
194 the correlation coefficients and the logarithm of the geographical distance. All these tests were
195 conducted in SPAGeDi 1.2 (Hardy & Vekemans, 2002). Additionally, a Mantel test (Mantel,
196 1967) was employed to determine if there was significant correlation between the genetic ($F_{ST} / 1 -$
197 F_{ST}) and geographical distances of the localities studied. Also, the presence or absence of putative
198 barriers, such as large rivers and/or mountain ranges, between localities was coded as 1 and 0 in a
199 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 analyzed, 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 center of the Peninsula (Table 1).

211 In general, nuclear genetic diversity was high, with a total of 264 alleles at the 10 microsatellites
212 analyzed (average N_A per locus = 26.29 ± 9.07). In all localities heterozygosity values were larger
213 than $H_o = 0.6$ (average $H_o = 0.7 \pm 0.15$). 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).

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

224 *Structure and assignment of rabbit genetic clusters*

225 The Bayesian clustering analysis performed in STRUCTURE revealed that $\ln P(X|D)$ increased
 226 substantially from $K = 1$ to $K = 2$ and then was attenuated as the number of K increased but
 227 without reaching a clear asymptote. Notwithstanding, calculation of ΔK clearly revealed the
 228 existence of 2 genetic populations or groups (K1 and K2). Taken together, these results suggest
 229 that the sampled rabbits belong to two large and distinct genetic groups, but do not completely
 230 exclude the possibility that more gene pools exist. The distribution of the two genetic groups
 231 exhibited high geographical correlation such that the localities to the south of the Iberian
 232 Peninsula were assigned with a greater likelihood to group K1 and the northern localities to K2
 233 group (Table 2 and Figs. 2 and 3).

234 Both the assignment probabilities to the genetic groups based on the nuclear DNA and the
 235 frequencies of mitochondrial lineages closely conformed to sigmoidal functions ($R^2 = 0.878$, $F =$
 236 57.478 , $P < 0.0001$ for the microsatellite data; $R^2 = 0.628$, $F = 13.501$, $P < 0.0001$ for the mtDNA

237 data; Fig. 3). In addition, the inferred genetic groups were in agreement with the mitochondrial
 238 haplotypes of each individual. Of the 224 individuals assigned to K1 with a posterior probability
 239 greater than 0.9, 91% carried mitochondrial haplotypes belonging to lineage A, while 82% of the
 240 239 rabbits assigned with equal probability to K2 showed haplotypes from lineage B (Fig. 2).
 241 Thus, the frequency of mitochondrial haplotypes in the parental lineages did not differ
 242 significantly from the individuals' assignment frequencies to the inferred genetic groups ($\chi^2 =$
 243 0.781, $df = 1$, $P = 0.377$).

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

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

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

268 The percentage of genetic variation explained by the 7 genetic clusters was similar, but not
269 greater than that obtained among all localities ($F_{ST}' = 0.325$, $P < 0.001$, $F_{ST}' = 0.370$, $P < 0.001$,
270 respectively). Conversely, R_{ST} was much higher ($R_{ST} = 0.627$, $P < 0.001$ and $R_{ST} = 0.110$, $P <$
271 0.001 , respectively), indicating that the effect of mutation is of greater importance than drift in
272 the differentiation of rabbit genetic clusters.

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

significant ($\ln b_d = -0.011$, $P < 0.001$). 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 analyzed were highly polymorphic, and showed similar variability to that reported 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 northeastern 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). The lowest diversity and high genetic differentiation found in the island of Mallorca 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,

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

312 Most localities, particularly the southern ones, revealed loci in Hardy-Weinberg
 313 disequilibrium, because of a deficiency in heterozygotes. In this study, as in earlier ones (Queney
 314 et al., 2001), putative null alleles were detected at Sat16, although the exclusion of this locus and
 315 others showing large deviations from equilibrium (e.g. Sat3 and Sol33), did not significantly alter
 316 the results. The absence of disequilibrium in the northeastern localities could be explained by the
 317 fact that the microsatellites analyzed were originally developed for the domestic rabbit (i.e.
 318 subspecies *O. c. cuniculus*) (Mougel, Mounolou & Monnerot, 1997; Surridge et al., 1997),
 319 therefore a higher chance for null alleles could occur in the southern *O. c. algirus*. Another non-
 320 exclusive explanation for the significant deficit of heterozygotes could be a Wahlund effect. The
 321 territorial behavior and social structure of the rabbit (Surridge et al., 1999a; Surridge et al.,

1999b) could lead to an underlying genetic structure at a small geographical scale that is not detected by Bayesian clustering methods.

Hierarchical genetic structure dependent on geographic distance

Overall, the European 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 sampling locality indicates that genetic structure exists at a very local scale and reaffirms the importance of rabbit social behavior 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 that are determined by other factors that could be either environmental or ecological. In turn, these inferred populations consist of even smaller groups conditioned by the social behavior 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)

339 In addition to the hierarchical genetic structure of the rabbit in Spain, geographic distance
 340 emerged as an important factor explaining genetic differentiation (Figs. 4 and 5). This contradicts
 341 the situation in Britain, where significant differences observed between locations could not be
 342 correlated with geographical distance (Surridge et al., 1999a). Similarly, it seems logical that
 343 main rivers, or other geographical features, constitute a barrier to gene flow in rabbits given their
 344 low dispersal capacity (Webb et al., 1995; Richardson et al., 2002). However, these barriers did
 345 not determine an increase in genetic differentiation explained solely by geographic distance. This
 346 is probably because, when considering the overall genetic structure and distribution of the rabbit,
 347 differences at a more local level have greater importance (Surridge et al., 1999a; Surridge et al.,
 348 1999b; Branco, Ferrand & Monnerot, 2000). This hypothesis is further supported by the results of
 349 our spatial autocorrelation analyses indicating significantly positive correlation up to a distance
 350 of about 100 km (Fig. 5A). However, the influence of geographic distance on genetic
 351 differentiation was not the same for all rabbit populations. Northern populations of *O. c.*
 352 *cuniculus* (K2) showed greater relatedness among close individuals and more spatial genetic
 353 structure and differentiation than the southern populations or those in the hybrid zone (Fig. 5B).
 354 This contrasting pattern could be due to the existence of genetic barriers among populations
 355 within each region. For example, it is well known that the Ebro River, running across
 356 northeastern Spain, has historically acted both as a physical and an ecological barrier for mammal
 357 species (O'Regan, 2008). The inferred genetics groups of K-NE and K-N are located to the north
 358 and south of the Ebro Valley, thus suggesting its role as a current barrier to gene flow (Fig. 1).
 359 Conversely, other large rivers in southern Spain (e.g. Guadiana River) do not seem to hinder gene
 360 flow among southern rabbit populations. At a smaller geographic scale, low dispersal might also
 361 be due both to resource availability and soil type, which 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 of other fossorial mammals (Lovegrove, 1989; Ebensperger & Cofré, 2001). It has been shown for other burrowing species, such as the wombat (*Lasiorninus 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 artifact of the method. Firstly, Bayesian clustering methods can overestimate genetic structure when analyzing 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 the crossing of genetically divergent individuals (Arnold et al., 1999). As expected from a region comprising the gene pool 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. Though this could simply be the 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 it to only hold the sum of the parental alleles. Unique alleles have been previously described in the rabbit hybrid zone for the HBA haemoglobin alpha chain gene, which probably originated by

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

411 In the hybrid zone cluster, considered as the region with intermediate frequencies not belonging
412 to any of the parental groups, the genetic contribution of each rabbit lineage was not balanced. In
413 this area, the frequency of the two mitochondrial lineages is virtually the same ($A = 0.485$ and B
414 $= 0.515$), but significantly greater proportions of individuals had been assigned to K2 ($K1 = 0.37$
415 and $K2 = 0.63$) ($\chi^2 = 25.187$, $df = 1$, $P < 0.0001$), showing a greater genetic introgression of
416 lineage B, characteristic of the northern *O. c. cuniculus*, into lineage A, *O. c. algirus*, than vice
417 versa. This is consistent with recent findings related to autosomal loci (Carneiro, Ferrand &
418 Nachman, 2009; Carneiro et al., 2013), yet contrasts with that described for the X chromosome,
419 suggesting slight introgression from the southwest to the northeastern lineages of the Iberian
420 Peninsula (Geraldes, Ferrand & Nachman, 2006). The fact that the greatest contribution of the
421 northern rabbit lineage is only reflected in the frequencies of nuclear markers and not in those of
422 maternal inheritance could suggest that males are primarily responsible for this bias. If this is the
423 case, it would be expected that the Y chromosome would be more introgressed than autosomal
424 loci. Conversely, it has been evidenced that the Y chromosome cline is highly stepped, as well as
425 the mtDNA, which suggests some kind of selection acting against introgression (Geraldes,
426 Ferrand & Nachman, 2006; Geraldes et al., 2008; Carneiro et al., 2013). In this regard,
427 preliminary behavioral work discarded the existence of pre-mating reproductive selection
428 between lineages, and found instead lower fertility in F1 males, thus following the expectations
429 of Haldane's rule (Haldane, 1922; Blanco-Aguilar et al., 2010). In this context it seems that the

430 relative role of selection leading to postzygotic barriers has a stronger importance in shaping the
 431 genetic structure in the rabbit hybrid zone than behavioral and prezygotic barriers. Similarly,
 432 different types of selection have been detected at several autosomal loci, suggesting a wide range
 433 of evolutionary pressures across the rabbit's genome as well as across distribution range in the
 434 Iberian Peninsula (Campos, Storz & Ferrand, 2008; Carneiro et al., 2012; Carneiro et al., 2013).

435 Ultimately, multiple factors ranging from geographical and ecological, to behavioral and
 436 molecular, are interacting and shaping the overall genetic structure of the rabbit subspecies and
 437 their hybrid zone. Future studies using genomic data coupled with behavioral and ecological
 438 information could further clarify how these issues are related to the differences in genetic
 439 variation and structure of the rabbit subspecies.

440 ACKNOWLEDGEMENTS

441 The authors wish to thank all those who provided samples for this study: the Consellería de
 442 Territori i Habitatge from Generalitat Valenciana, the Hunting Federation of Lérida, Hunting
 443 Society of Ajalvir (Madrid), M. Sanmartín from the University of Santiago de Compostela, S.
 444 Agudín, J. Inogés, J. Layna, F. Leiva, F. Silvestre and specially to F. Guil from Fundación CBD-
 445 Habitat. We also thank N. Ferrand and S. Lopes for data reviewing and helpful discussion of the
 446 results, as well as two anonymous referees for their thorough revisions. A. Peterson reviewed the
 447 English text, L. Alcaraz assisted in the laboratory and A. Benítez-López provided assistance with
 448 statistical analyses.

449 REFERENCES

- 450 Alcover JA. 2008. The first Mallorcans: prehistoric colonization in the Western Mediterranean.
451 *Journal of Prehistory* 21:19-84.
- 452 Amos W, and Harwood J. 1998. Factors affecting levels of genetic diversity in natural
453 populations. *Philosophical Transactions of the Royal Society of London Series B-Biological*
454 *Sciences* 353:177-186.
- 455 Arnold ML. 1997. *Natural hybridisation and evolution*. New York: Oxford University Press.
- 456 Arnold ML, Bulger MR, Burke JM, Hempel AL, and Williams JH. 1999. Natural hybridization:
457 how long can you go and still be important? *Ecology* 80:371-381.
- 458 Avise JC. 2000. *Phylogeography: The history and formation of species*. Cambridge, MA:
459 Harvard University Press.
- 460 Baker AN, and Dunning RA. 1975. Effects of soil type and crop density on the activity and
461 abundance of the epigeic fauna, particularly Carabidae, in sugar-beet fields. *The Journal of*
462 *Applied Ecology* 12:809-818.
- 463 Balloux F, and Lugon-Moulin N. 2002. The estimation of population differentiation with
464 microsatellite markers. *Molecular Ecology* 11:155-165.
- 465 Barton N, and Gale KS. 1993. Genetic analysis of hybrid zones. In: Harrison RG, ed. *Hybrid*
466 *Zones and the Evolutionary Process*. New York: Oxford University Press.
- 467 Barton NH, and Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and*
468 *Systematics* 16:113-148.
- 469 Barton NH, and Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature* 341:497-503.
- 470 Belkhir K, Borsa P, Chikhi L, Raufaste N, and Bonhomme F. 2004. *GENETIX 4.05, logiciel sous*
471 *WindowsTM pour la génétique des populations*. Montpellier: Laboratoire Génome, Populations,
472 Interactions, CNRS UMR 5000, Université de Montpellier II.
- 473 Biju-Duval C, Ennafaa H, Dennebouy N, Monnerot M, Mignotte F, Soriguer RC, El Gaaïed A, El
474 Hili A, and Mounolou JC. 1991. Mitochondrial DNA evolution in lagomorphs: origin of

- 475 systematic heteroplasmy and organization of diversity in European rabbits. *Journal of Molecular*
476 *Evolution* 33:92-102.
- 477 Blanco JC, and Villafuerte R. 1993. *Factores ecológicos que influyen sobre las poblaciones de*
478 *conejos*. Madrid: Instituto Nacional para la Conservación de la Naturaleza.
- 479 Blanco-Aguiar JA, Carneiro M, Villafuerte R, Ríos-Saldaña CA, Ferreira C, and Ferrand N.
480 2010. Preliminary results assessing pre and postzygotic isolation between two subspecies,
481 *Oryctolagus cuniculus algirus* and *O. c. cuniculus*, of the European rabbit. Trends in Biodiversity
482 and Evolution (TiBE). Vairao, Portugal: Research Centre in Biodiversity and Genetic Resources,
483 University of Porto.
- 484 Bradley BJ, Bull JJ, Johnson AD, and Hillis DM. 1993. Origin of a novel allele in a mammalian
485 hybrid zone. *Proceedings of the National Academy of Sciences of the United States of America*
486 90:8939-8941.
- 487 Bradley BJ, Doran D, Robbins MM, Williamson E, Boesch C, and Vigilant L. 2002. Comparative
488 analyses of genetic social structure in wild gorillas (*Gorilla gorilla*) using DNA from feces and
489 hair. *American Journal of Physical Anthropology*:47-48.
- 490 Branco M, and Ferrand N. 2003. Biochemical and population genetics of the rabbit, *Oryctolagus*
491 *cuniculus*, carbonic anhydrases I and II, from the Iberian Peninsula and France. *Biochemical*
492 *Genetics* 41:391-404.
- 493 Branco M, Ferrand N, and Monnerot M. 2000. Phylogeography of the European rabbit
494 (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome
495 b gene. *Heredity* 85:307-317.
- 496 Branco M, Machado JC, and Ferrand N. 1999. Extensive genetic polymorphism of peptidases A,
497 B, C, and D, in wild rabbit (*Oryctolagus cuniculus*) populations from the Iberian Peninsula.
498 *Biochemical Genetics* 37:237-249.
- 499 Branco M, Monnerot M, Ferrand N, and Templeton AR. 2002. Postglacial dispersal of the
500 European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested
501 clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution* 56:792-803.

- 502 Brelsford A, and Irwin DE. 2009. Incipient speciation despite little assortative mating: the
503 yellow-rumped warbler hybrid zone. *Evolution* 63:3050-3060.
- 504 Campos R, Storz JF, and Ferrand N. 2008. Evidence for contrasting modes of selection at
505 interacting globin genes in the European rabbit (*Oryctolagus cuniculus*). *Heredity* 100:602-609.
- 506 Carneiro M, Albert FW, Melo-Ferreira J, Galtier N, Gayral P, Blanco-Aguilar JA, Villafuerte R,
507 Nachman MW, and Ferrand N. 2012. Evidence for widespread positive and purifying selection
508 across the European rabbit (*Oryctolagus cuniculus*) genome. *Molecular Biology and Evolution*
509 29:1837-1849.
- 510 Carneiro M, Baird SJE, Afonso S, Ramirez E, Tarroso P, Teotónio H, Villafuerte R, Nachman
511 MW, and Ferrand N. 2013. Steep clines within a highly permeable genome across a hybrid zone
512 between two subspecies of the European rabbit. *Molecular Ecology* 22:2511-2525.
- 513 Carneiro M, Blanco-Aguilar JA, Villafuerte R, Ferrand N, and Nachman MW. 2010. Speciation in
514 the European rabbit (*Oryctolagus cuniculus*): islands of differentiation on the X chromosome and
515 autosomes. *Evolution* 64:3443-3460.
- 516 Carneiro M, Ferrand N, and Nachman MW. 2009. Recombination and speciation: loci near
517 centromeres are more differentiated than loci near telomeres between subspecies of the European
518 rabbit (*Oryctolagus cuniculus*). *Genetics* 181:593-606.
- 519 Corander J, and Marttinen P. 2006. Bayesian identification of admixture events using multilocus
520 molecular markers. *Molecular Ecology* 15:2833-2843.
- 521 Corander J, Sirén J, and Arjas E. 2008. Bayesian spatial modeling of genetic population structure.
522 *Computational Statistics* 23:111-129.
- 523 Cowan DP, and Garson PJ. 1985. Variations in the social structure of rabbit populations: causes
524 and demographic consequences. In: Smith RM, and Sibly RH, eds. *Behavioural Ecology:
525 Ecological Consequences of Adaptive Behaviour The 25th Symposium of the British Ecological
526 Society*. Oxford, Reading: Blackwell Scientific Publications, 537-555.
- 527 Ebensperger LA, and Cofré H. 2001. On the evolution of group-living in the new world cursorial
528 hystricognath rodents. *Behavioural Ecology* 12:227-236.

- 529 Evanno G, Regnaut S, and Goudet J. 2005. Detecting the number of clusters of individuals using
530 the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.
- 531 Excoffier L, Smouse PE, and Quattro JM. 1992. Analysis of molecular variance inferred from
532 metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction
533 data. *Genetics* 131:479-491.
- 534 Falush D, Stephens M, and Pritchard JK. 2003. Inference of population structure using multilocus
535 genotype data: Linked loci and correlated allele frequencies. *Genetics* 164:1567-1587.
- 536 Ferrand N. 1989. Biochemical and genetic studies on rabbit hemoglobin. I. Electrophoretic
537 polymorphism of the b-chain. *Biochemical Genetics* 27:673-678.
- 538 Ferrand N, and Branco M. 2007. The evolutionary history of the European rabbit (*Oryctolagus*
539 *cuniculus*): Major patterns of population differentiation and geographic expansion inferred from
540 protein polymorphism. In: Weiss S, and Ferrand N, eds. *Phylogeography of European Refugia*.
541 Netherlands: Springer, 207-235.
- 542 Ferreira C. 2012. European rabbit research in the Iberian Peninsula: state of the art and future
543 perspectives. *European Journal of Wildlife Research* 58:885-895.
- 544 Flux JEC. 1994. World distribution. In: Thompson HV, and King CM, eds. *The European Rabbit*
545 *The History and Biology of a Successful Colonizer*. Oxford: Oxford University Press, 8-21.
- 546 Frantz AC, Cellina S, Krier A, Schley L, and Burke T. 2009. Using spatial Bayesian methods to
547 determine the genetic structure of a continuously distributed population: clusters or isolation by
548 distance? *Journal of Applied Ecology* 46:493-505.
- 549 Fuller SJ, Wilson JC, and Mather PB. 1997. Patterns of differentiation among wild rabbit
550 populations *Oryctolagus cuniculus* L. in arid and semiarid ecosystems of north-eastern Australia.
551 *Molecular Ecology* 6:145-153.
- 552 Futuyma DJ. 1998. *Evolutionary Biology*. Sunderland: Sinauer Associates.
- 553 Geraldès A, Carneiro M, Delibes-Mateos M, Villafuerte R, Nachman MW, and Ferrand N. 2008.
554 Reduced introgression of the Y chromosome between subspecies of the European rabbit
555 (*Oryctolagus cuniculus*) in the Iberian Peninsula. *Molecular Ecology* 17:4489-4499.

- 556 Geraldès A, Ferrand N, and Nachman MW. 2006. Contrasting patterns of introgression at X-
557 linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus*
558 *cuniculus*). *Genetics* 173:919-933.
- 559 Geraldès A, Rogel-Gaillard C, and Ferrand N. 2005. High levels of nucleotide diversity in the
560 European rabbit (*Oryctolagus cuniculus*) SRY gene. *Animal Genetics* 36:349-351.
- 561 Goudet J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of*
562 *Heredity* 86:485-486.
- 563 Guo SW, and Thompson EA. 1992. Performing the exact test of Hardy-Weinberg proportion for
564 multiple alleles. *Biometrics* 48:361-372.
- 565 Haldane JBS. 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*
566 12:101-109.
- 567 Hardy C, Vigne J-D, Casane D, Dennebouy N, Mounolou J-C, and Monnerot M. 1994. Origin of
568 European rabbit (*Oryctolagus cuniculus*) in a Mediterranean island: Zooarchaeology and ancient
569 DNA examination. *Journal of Evolutionary Biology* 7:217-226.
- 570 Hardy OJ, and Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial
571 genetic structure at the individual or population levels. *Molecular Ecology Notes* 2:618-620.
- 572 Harrison RG. 1993. Hybrids and hybrid zones: historical perspective. In: Harrison RG, ed.
573 *Hybrid zones and the evolutionary process*. New York: Oxford University, 3-12.
- 574 Hedrick PW. 2005. *Genetics of populations*. Boston, MA: Jones and Bartlett.
- 575 Hewitt G. 1988. Hybrid zones-Natural laboratories for evolutionary studies. *Trends in Ecology &*
576 *Evolution* 3:158-167.
- 577 Hoffman SMG, and Brown WM. 1995. The molecular mechanism underlying the "rare allele
578 phenomenon" in a subspecific hybrid zone of the California field mouse, *Peromyscus*
579 *californicus*. *Journal of Molecular Evolution* 41:1165-1169.
- 580 Hutchison DW, and Templeton AR. 1999. Correlation of pairwise genetic and geographic
581 distance measures: inferring the relative influences of gene flow and drift on the distribution of
582 genetic variability. *Evolution* 53:1898-1914.

- 583 Kaeuffer R, Réale D, Coltman DW, and Pontier D. 2007. Detecting population structure using
584 STRUCTURE software: effect of background linkage disequilibrium. *Heredity* 99:374-380.
- 585 Knowles LL, and Carstens BC. 2007. Estimating a geographically explicit model of population
586 divergence. *Evolution* 61:477-493.
- 587 Loiselle BA, Sork VL, Nason J, and Graham C. 1995. Spatial genetic structure of a tropical
588 understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* 82:1420-1425.
- 589 Lombardi L, Fernández N, Moreno S, and Villafuerte R. 2003. Habitat-related differences in
590 rabbit (*Oryctolagus cuniculus*) abundance, distribution, and activity. *Journal of Mammalogy*
591 84:26-36.
- 592 Lovegrove BG. 1989. The cost of burrowing by the social mole rats (Bathyergidae) *Cryptomys*
593 *damarensis* and *Heterocephalus glaber*: the role of soil moisture. *Physiological Zoology* 62:449-
594 469.
- 595 Lugon-Moulin N, Brunner H, Wytenbach A, Hausser J, and Goudet J. 1999. Hierarchical
596 analyses of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora: Soricidae).
597 *Molecular Ecology* 8:419-431.
- 598 Mantel NA. 1967. The detection of disease clustering and a generalized regression approach.
599 *Cancer Research* 27:209-220.
- 600 Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic
601 differentiation measure. *Evolution* 60:2399-2402.
- 602 Meirmans PG, and Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for
603 the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792-794.
- 604 Monnerot M, Vigne JD, Biju-Duval C, Casane D, Callou C, Hardy C, Mougél F, Soriguer R,
605 Dennebouy N, and Mounolou J. 1994. Rabbit and man: genetic and historic approach. *Genetics*
606 *Selection Evolution* 26:167-182.
- 607 Mougél F, Mounolou J, and Monnerot M. 1997. Nine polymorphic microsatellite loci in the
608 rabbit, *Oryctolagus cuniculus*. *Animal Genetics* 28:58-71.

- 609 Nei M, Maruyama T, and Chakraborty R. 1975. The bottleneck effect and genetic variability in
610 populations. *Evolution* 29:1-10.
- 611 O'Regan HJ. 2008. The Iberian Peninsula - corridor or cul-de-sac? Mammalian faunal exchange
612 and possible routes of dispersal in the last 2 million years. *Quaternary Science Reviews* 27:2136-
613 2144.
- 614 Palsboll PJ. 1999. Genetic tagging: contemporary molecular ecology. *Biological Journal of the*
615 *Linnean Society* 68:3-22.
- 616 Pilot M, Jedrzejewski W, Branicki W, Sidorovich VE, Jedrzejewska B, Stachura K, and Funk SM.
617 2006. Ecological factors influence population genetic structure of European grey wolves.
618 *Molecular Ecology* 15:4533-4553.
- 619 Pritchard JK, Stephens M, and Donnelly P. 2000. Inference of population structure using
620 multilocus genotype data. *Genetics* 155:945-959.
- 621 Queney G, Ferrand N, Marchandeu S, Azevedo M, Mougél F, Branco M, and Monnerot M.
622 2000. Absence of a genetic bottleneck in a wild rabbit (*Oryctolagus cuniculus*) population
623 exposed to a severe viral epizootic. *Molecular Ecology* 9:1253-1264.
- 624 Queney G, Ferrand N, Weiss S, Mougél F, and Monnerot M. 2001. Stationary distributions of
625 microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus*
626 *cuniculus*). *Molecular Biology and Evolution* 18:2169-2178.
- 627 R Core Team. 2014. R: A language and environment for statistical computing. Vienna, Austria: R
628 Foundation for Statistical Computing.
- 629 Raymond M, and Rousset F. 1995. GENEPOP (version 1.2) Population genetics software for
630 exact tests and ecumenicism. *Journal of Heredity* 86:248-249.
- 631 Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43:223-225.
- 632 Richardson BJ, Hayes RA, Wheeler CM, and Yardin MR. 2002. Social structures, genetic
633 structures and dispersal strategies in Australian rabbit (*Oryctolagus cuniculus*) populations.
634 *Behavioural Ecology and Sociobiology* 51:113-121.

- 635 Seixas FA, Juste J, Campos PF, Carneiro M, Ferrand N, Alves PC, and Melo-Ferreira J. 2014.
- 636 Colonization history of Mallorca Island by the European rabbit, *Oryctolagus cuniculus*, and the
- 637 Iberian hare, *Lepus granatensis* (Lagomorpha: Leporidae). *Biological Journal of the Linnean*
- 638 *Society* 111:748-760.
- 639 Shoemaker DD, and Ross KG. 1996. Effects of social organization on gene flow in the fire ant
- 640 *Solenopsis invicta*. *Nature* 383:613-616.
- 641 Smouse PE, and Peakall R. 1999. Spatial autocorrelation analysis of individual multiallele and
- 642 multilocus genetic structure. *Heredity* 82:561-573.
- 643 Surridge AK, Bell DJ, Ibrahim KM, and Hewitt G. 1999a. Population structure and genetic
- 644 variation of European wild rabbits (*Oryctolagus cuniculus*) in East Anglia. *Heredity* 82:479-487.
- 645 Surridge AK, Bell DJ, Rico C, and Hewitt GM. 1997. Polymorphic microsatellite loci in the
- 646 European rabbit (*Oryctolagus cuniculus*) are also amplified in other lagomorph species. *Animal*
- 647 *Genetics* 28:302-305.
- 648 Surridge AK, Ibrahim KM, Bell DJ, Webb NJ, Rico C, and Hewitt GM. 1999b. Fine-scale genetic
- 649 structuring in a natural population of European wild rabbits (*Oryctolagus cuniculus*). *Molecular*
- 650 *Ecology* 8:299-307.
- 651 van Oosterhout C, Hutchinson WF, Wills DPM, and Shipley P. 2004. MICRO-CHECKER:
- 652 software for identifying and correcting genotyping errors in microsatellite data. *Molecular*
- 653 *Ecology Notes* 4:535-538.
- 654 Vekemans X, and Hardy OJ. 2004. New insights from fine-scale spatial genetic structure analyses
- 655 in plant populations. *Molecular Ecology* 13:921-935.
- 656 Walker FM, Taylor AC, and Sunnucks P. 2007. Does soil type drive social organization in
- 657 southern hairy-nosed wombats? *Molecular Ecology* 16:199-208.
- 658 Webb NJ, Ibrahim KM, Bell DJ, and Hewitt GM. 1995. Natal dispersal and genetic structure in a
- 659 population of the European wild rabbit (*Oryctolagus cuniculus*). *Molecular Ecology* 4:239-248.

660 SUPPLEMENTARY MATERIAL

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

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

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

Table 1 (on next page)

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

Genetic diversity statistics based on 10 microsatellite loci genotypes for the rabbit localities (≥ 10 individuals) analyzed and the genetic clusters inferred in BAPS. n = number of samples, N_A = mean number of alleles per locus, A_R = allelic richness, H_o = observed heterozygosity, 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) analyzed and the genetic clusters inferred in BAPS. n = number of samples, N_A = mean number of alleles per locus, A_R = allelic richness, H_o = observed heterozygosity, 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	<i>O. c. cuniculus</i>
2. Lérida	50	9.64	6.20	0.76	0.79	0.06	0.00	1.00	<i>O. c. cuniculus</i>
3. Valencia	18	9.18	7.13	0.68	0.80	0.18	0.06	0.94	<i>O. c. cuniculus</i>
4. La Rioja	19	7.00	5.62	0.72	0.74	0.06	0.05	0.95	<i>O. c. cuniculus</i>
5. Galicia	27	8.73	6.42	0.61	0.80	0.26	0.00	1.00	<i>O. c. cuniculus</i>
6. Valladolid	16	7.91	6.41	0.67	0.73	0.11	0.07	0.93	<i>O. c. cuniculus</i>
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	<i>O. c. algirus</i>
18. Ciudad Real2	27	9.73	7.09	0.64	0.83	0.25	0.89	0.11	<i>O. c. algirus</i>
19. Ciudad Real3	50	12.27	7.34	0.70	0.82	0.16	0.82	0.18	<i>O. c. algirus</i>
20. Albacete	25	9.64	6.86	0.75	0.81	0.10	0.48	0.52	<i>O. c. algirus</i>
21. Cáceres1	10	6.64	6.22	0.70	0.72	0.10	0.90	0.10	<i>O. c. algirus</i>
22. Cáceres2	28	9.00	6.51	0.63	0.78	0.21	0.89	0.11	<i>O. c. algirus</i>
23. Badajoz1	20	9.73	6.82	0.71	0.79	0.13	0.95	0.05	<i>O. c. algirus</i>
24. Badajoz2	29	9.45	6.99	0.79	0.82	0.05	0.90	0.10	<i>O. c. algirus</i>
25. Jaén1	15	8.64	6.89	0.63	0.78	0.23	0.00	1.00	<i>O. c. algirus</i>
27. Jaén3	22	9.64	7.13	0.70	0.82	0.16	1.00	0.00	<i>O. c. algirus</i>
28. Sevilla1	43	11.36	6.89	0.72	0.80	0.10	1.00	0.00	<i>O. c. algirus</i>
29. Sevilla2	32	11.82	7.62	0.72	0.83	0.14	0.63	0.38	<i>O. c. algirus</i>
30. Cádiz	56	11.64	6.92	0.69	0.80	0.14	0.96	0.04	<i>O. c. algirus</i>
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 from 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.

Locality	n	mtDNA		STRUCTURE		BAPS						
		Hap A	Hap B	K1	K2	K-BAL	K-NE	K-N	K-CU	K-ZH	K-S1	K-S2
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 analyzed and the inferred clusters.

Table 3. AMOVA analysis performed for different levels of genetic structure among the rabbit localities analyzed 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 analyzed

Maps of the Iberian Peninsula indicating the localities analyzed (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.

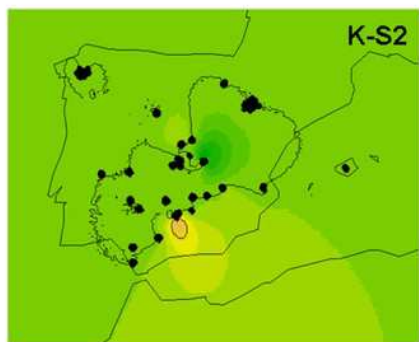
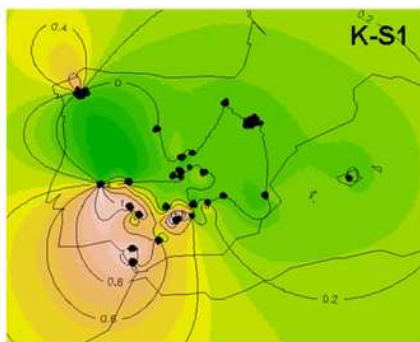
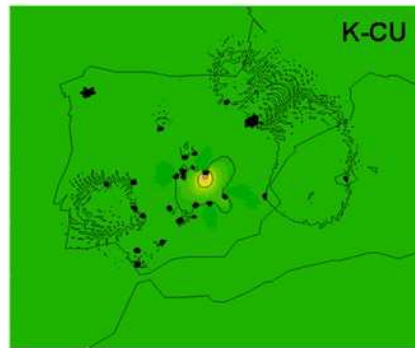
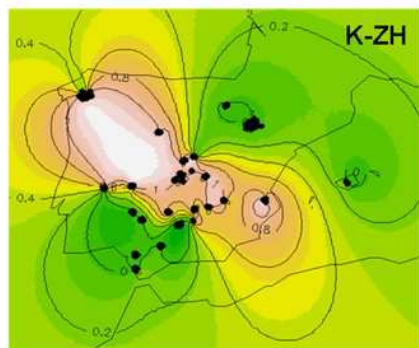
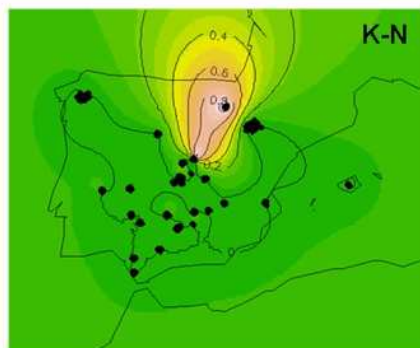
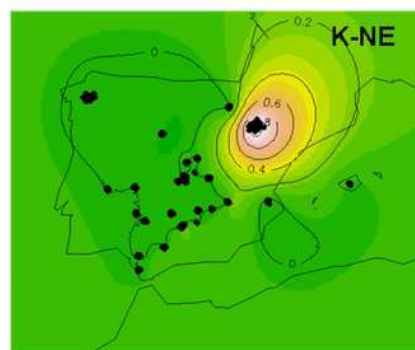
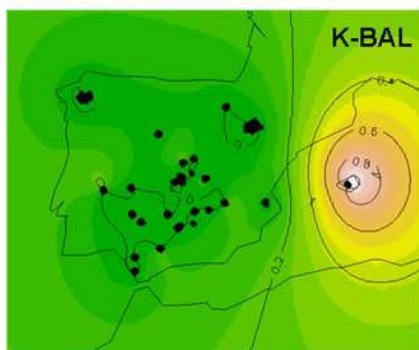
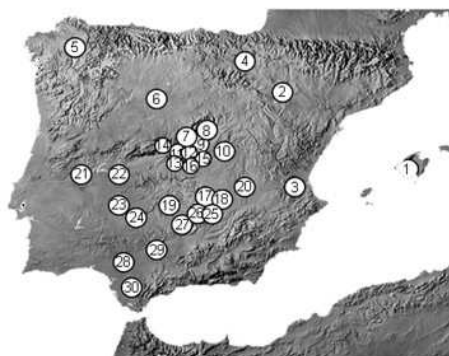


Figure 2

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.

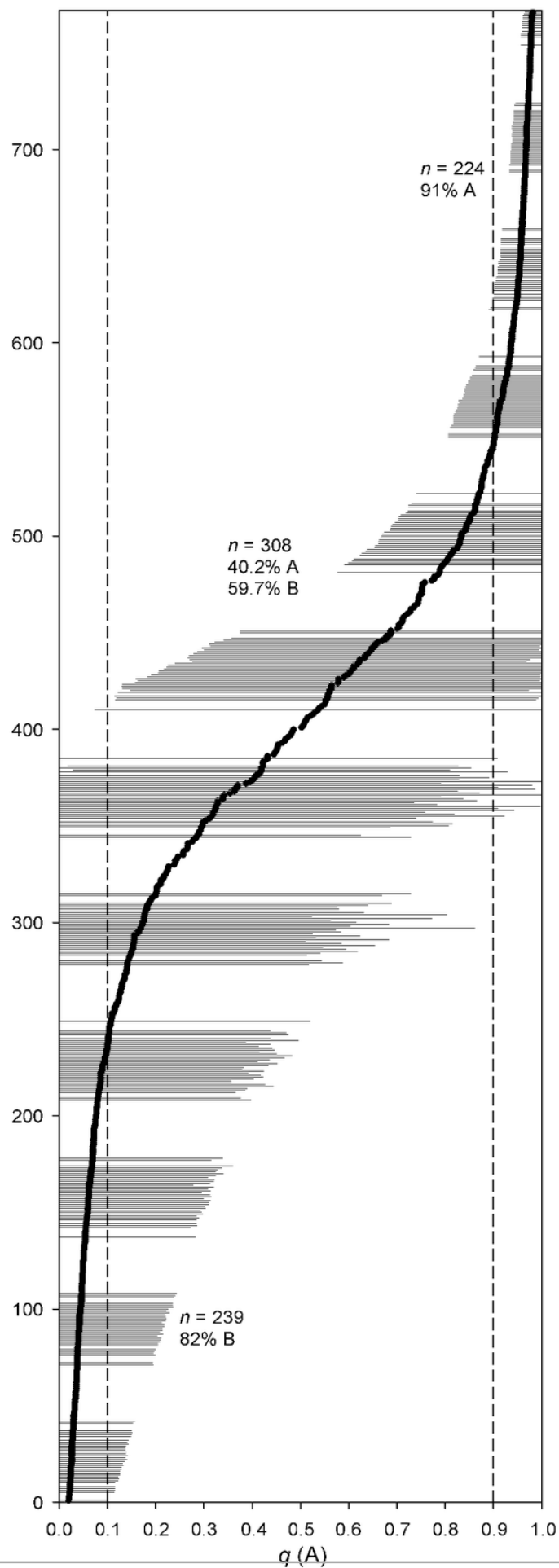


Figure 3

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.

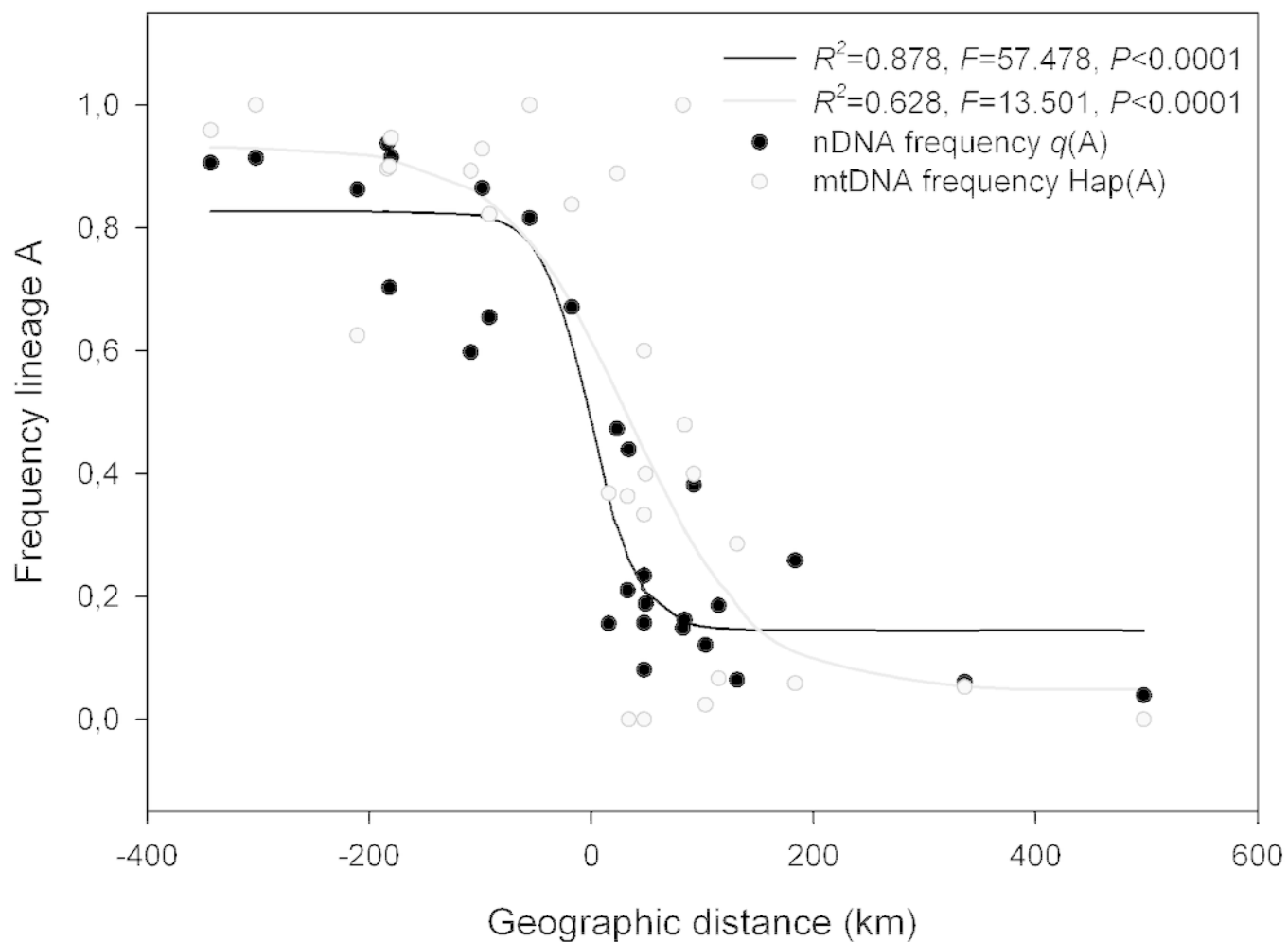


Figure 4

Isolation by distance among rabbit localities

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

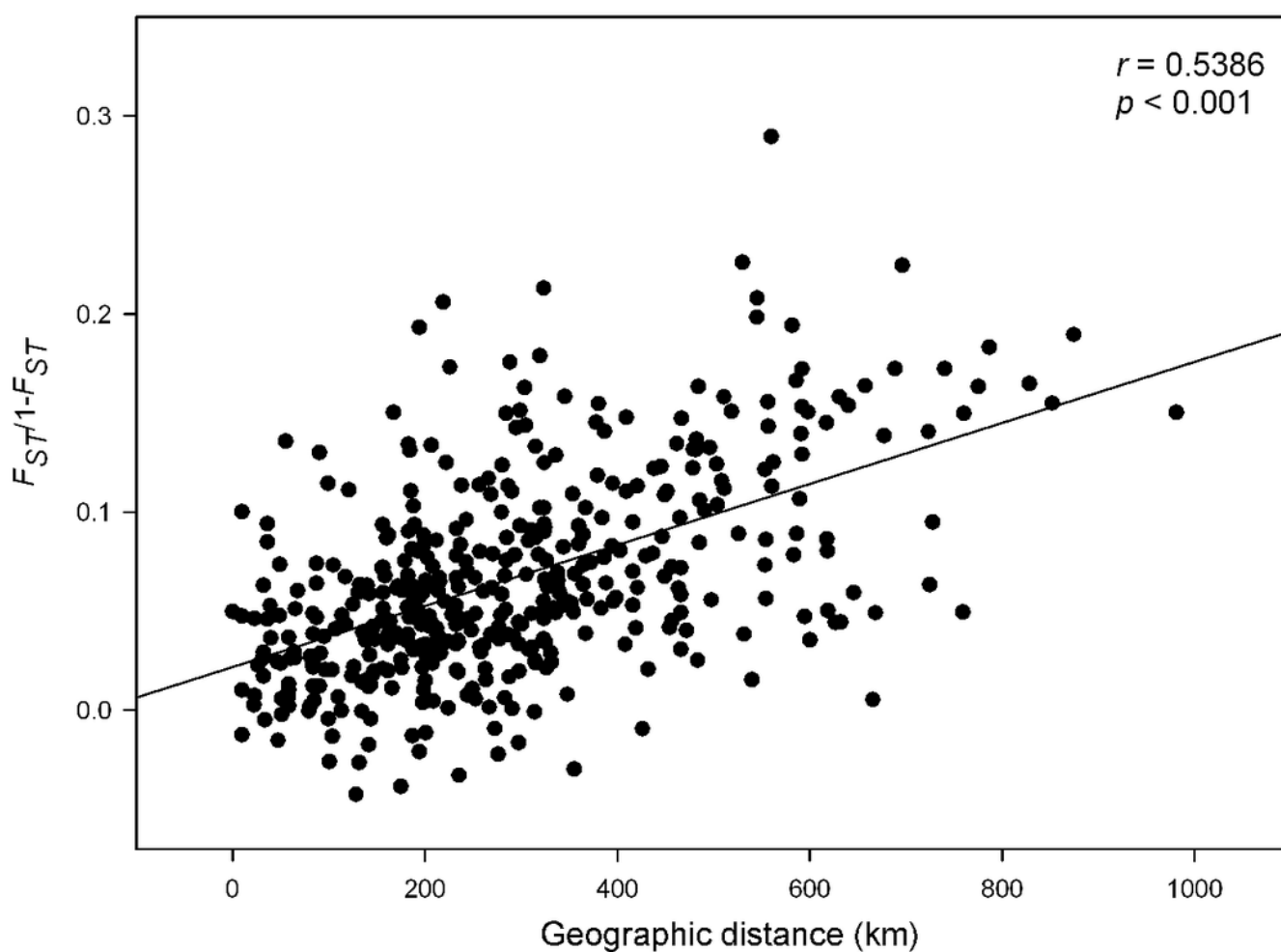


Figure 5

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 analyzed and the hybrid zone separately (B). Black symbols represent significant correlations between f_{ij} and d_{ij} .

