Reduced introgression of the Y chromosome between subspecies of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula

A. GERALDES,* † M. CARNEIRO,* † M. DELIBES-MATEOS,‡ R. VILLAFUERTE,‡ M. W. NACHMAN† and N. FERRAND*

*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal, and Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal, †Department of Ecology & Evolutionary Biology, Biosciences West Building, The University of Arizona, PO Box 210088, Tucson, AZ 85721, USA, ‡IREG, Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCLM), Ronda de Toledo s/n, 13005, Ciudad Real, Spain

Abstract

The role of the Y chromosome in speciation is unclear. Hybrid zones provide natural arenas for studying speciation, as differential introgression of markers may reveal selection acting against incompatibilities. Two subspecies of the European rabbit (*Oryctolagus cuniculus*) form a hybrid zone in the Iberian Peninsula. Previous work on mitochondrial DNA (mtDNA), Y- and X-linked loci revealed the existence of two divergent lineages in the rabbit genome and that these lineages are largely subspecies-specific for mtDNA and two X-linked loci. Here we investigated the geographic distribution of the two Y chromosome lineages by genotyping two diagnostic single nucleotide polymorphisms in a sample of 353 male rabbits representing both subspecies, and found that Y chromosome lineages are also largely subspecies-specific. We then sequenced three autosomal loci and discovered considerable variation in levels of differentiation at these loci. Finally, we compared estimates of population differentiation between rabbit subspecies at 26 markers and found a surprising bimodal distribution of $F_{ST}$ values. The vast majority of loci showed little or no differentiation between rabbit subspecies while a few loci, including the SRY gene, showed little or no introgression across the hybrid zone. Estimates of population differentiation for the Y chromosome were surprisingly high given that there is male-biased dispersal in rabbits. Taken together, these data indicate that there is a clear dichotomy in the rabbit genome and that some loci remain highly differentiated despite extensive gene flow following secondary contact.

Keywords: European rabbit, gene flow, hybrid zone, speciation, Y chromosome

Received 25 May 2008; revision received 30 July 2008; accepted 20 August 2008

The genetic basis of reproductive isolation is a key problem in evolutionary biology, and it has been studied both with laboratory crosses and with natural hybrid populations. Laboratory crosses have the advantage of providing a controlled setting, and wild populations have the advantage of providing more realistic biological conditions for detecting fitness differences. Both kinds of studies have provided abundant evidence for the importance of the X chromosome in reproductive isolation in taxa in which the male is the heterogametic sex, such as *Drosophila* and mammals (Coyne & Orr 2004). For example, Haldane’s rule appears to be largely a consequence of epistatic interactions involving recessive X-linked mutations (Turelli & Orr 1995).

The role of the Y chromosome in speciation is far less clear (Coyne & Orr 2004). Several studies have looked at the role of the Y in *Drosophila*. For example, in crosses involving species in the *Drosophila virilis* group, some Y-chromosome introgression experiments reveal a strong phenotypic effect, while others do not (Orr & Coyne 1989). In crosses involving *D. arizonae* and *D. mojavensis*, the Y chromosome from *D. arizonae* interacts with *D. mojavensis*...
alleles at autosomal loci to cause male hybrid sterility (Vigneault & Zouros 1986). In crosses between *D. sechellia* and *D. simulans*, the role of the Y is asymmetrical. While the Y from *D. sechellia* on a *D. simulans* background produces fully fertile males, the reciprocal cross yields sterile males (Johnson *et al.* 1993; Zeng & Singh 1993). Finally, in crosses between *D. yakuba* and *D. santomea*, one Y chromosome locus has been found to cause hybrid male sterility (Coyne *et al.* 2004) and the Y chromosome shows reduced introgression across the hybrid zone between them (Llopart *et al.* 2005).

Less attention has been devoted to the role of the Y chromosome in speciation in mammals. In house mice, primary sex determination is disrupted in consomic strains in which the *Mus domesticus* (also referred to as *Mus musculus domesticus*) Y chromosome is introduced onto the genetic background of some laboratory strains containing *M. musculus* (also referred to as *M. m. musculus*) alleles (Eicher & Washburn 1986; Washburn *et al.* 2001). The phenotypic effect depends on the genetic composition of both the *M. domesticus* Y chromosome and of the laboratory strains; in some cases, males show complete sterility, while in other cases, males are fully fertile. In two different transects of the *M. musculus–M. domesticus* hybrid zone, the Y chromosome introgresses less than other chromosomes (Vanlerbergh et al. 1986; Tucker *et al.* 1992). However, in another transect, the Y chromosome shows a fair amount of introgression (Munclinger *et al.* 2002). Patterns of Y chromosome introgression between species of shrews have also been documented. In particular, two studies found little or no introgression of the Y chromosome between different races of *Sorex araneus* and *S. antinorii*, suggesting that the Y chromosome may harbour genes involved in isolation (Balloux *et al.* 2000; Yannic *et al.* 2008). In Sweden, a contact zone with noncoincident, but steep, clines for mitochondrial DNA (mtDNA) and Y chromosome has been detected between populations of the field vole, *Microtus agrestis*. The pattern has been suggested to be due to selection against Y chromosome introgression (Jaarola *et al.* 1997). Finally, Rocca *et al.* (2005) detected reduced introgression of the Y chromosome relative to mtDNA and X-linked markers between African forest (*Loxodonta cyclotis*) and African savannah elephants (*L. africana*), despite male-biased dispersal.

The European rabbit (*Oryctolagus cuniculus*) provides another opportunity to study the Y chromosome in the context of hybridizing taxa. This species is native to the Iberian Peninsula and has two recognized subspecies, *O. c. algirus* in the southwest and *O. c. cuniculus* in the northeast. The two subspecies form a contact zone that runs in a northwest–southeast direction. Data from multiple loci are consistent in suggesting that these subspecies diverged at or before the beginning of the Pleistocene. For example, these taxa are characterized by two divergent mtDNA lineages (11.9% uncorrected nucleotide divergence based on cytochrome *b* (Cytb) restriction fragment length polymorphism (RFLP) data; Branco *et al.* 2000), suggesting that they diverged approximately 2 million years ago. Branco *et al.* (2002) inferred that the subspecies have recently come into secondary contact following Pleistocene climatic changes. Analyses of protein (Ferrand & Branco 2007) and immunoglobulin (van der Loo *et al.* 1991; van der Loo *et al.* 1999; Esteves *et al.* 2004) variability in the Iberian Peninsula are consistent with the mtDNA data in showing two major groups, but the level of differentiation between subspecies is often low for these markers, suggesting introgression following secondary contact. Recently, Geraldes *et al.* (2006) sequenced four X-linked loci (two centromeric and two telomeric) from the range of both subspecies and from the contact zone. Two divergent lineages were observed at each of the four loci. The estimated time of divergence for these lineages conformed well to the pre-Pleistocene isolation scenario inferred from mtDNA. The two centromeric loci showed low levels of nucleotide variability, high levels of linkage disequilibrium (LD) and low levels of introgression. In contrast, the two telomeric loci showed high levels of nucleotide variability, low levels of LD and high levels of introgression. Thus, all genes revealed an old divergence between these subspecies, but there was considerable variation among genes in the amount of introgression following secondary contact.

Here, we investigate geographic variation in the Y chromosome in a large sample of rabbits representing both subspecies. First, we genotyped two diagnostic Y-specific single nucleotide polymorphisms (SNPs) in a sample of 353 male rabbits from the Iberian Peninsula and south of France. By doing this, we documented levels of introgression on the Y chromosome in a species with well-known male-biased dispersal (Webb *et al.* 1995; Kunkele & von Holst 1996; Richardson *et al.* 2002). Second, we compared levels of introgression of the Y chromosome with other molecular markers. Namely, we (i) sequenced intronic fragments of three autosomal genes in a small sample of rabbits belonging to both subspecies, and (ii) compiled data from previous publications on mtDNA (Branco *et al.* 2000), X chromosome (Geraldes *et al.* 2006) and protein polymorphism (Ferrand & Branco 2007) to compare estimates of genetic differentiation between subspecies at these markers with estimates from the Y chromosome. Our results show that levels of introgression for the Y chromosome are very low. In addition, we show that the *F*<sub>ST</sub> values of the 26 loci studied follow a striking bimodal distribution, with a majority of loci with low *F*<sub>ST</sub> values, and very few with high *F*<sub>ST</sub> values.

**Materials and methods**

**Sampling**

We sampled 353 male rabbits from 30 natural populations. Populations were divided into three groups (Fig. 1):
southwest Iberian Peninsula (SW), corresponding to the
distribution of the subspecies *O. c. algirus*; northeast Iberian
Peninsula and France (NE), where *O. c. cuniculus* occurs; and
the hybrid zone (HZ). The centre of the hybrid zone was
estimated from a combination of mtDNA, X chromosome
and polymorphic protein loci following the procedures
described in Ferrand (2008). Briefly, data for each genomic
compartment were interpolated in a geographical information
system (GIS) environment, using the ordinary kriging
algorithm. This yields a continuous surface with interpolated
values for each genomic compartment which were then
averaged to a single synthetic map. The latter was used to
extract the contour lines of 40%, 50% and 60%, indicating the
geographic distribution of rabbit lineages. All the analyses
were carried in ArcGIS 9.2 software with Spatial Analyst and
Geostatistical Analyst extensions (ESRI 2006). The approximate
geographic location of the populations is shown in Fig. 1.

**Molecular methods**

Genomic DNA was extracted from blood, liver, kidney or
muscle following standard protocols (Sambrook & Russell
2001). Two different fragments of the SRY gene region
(Geraldes et al. 2005) were amplified through polymerase
chain reaction (PCR). Genotyping was performed using RFLP
to distinguish the two SRY lineages described by Geraldes
et al. (2005) who sequenced approximately 2 kb of the Y
chromosome in four wild and eight domestic rabbits. They
found two divergent lineages separated by seven nucleotide
differences (0.40% average divergence), corresponding
to the two subspecies. We used the primers TSPYF267–
GCAAAGCTGTGATTTTCAAAGGC and TSPYR716–
GTATTGCACTGGTGGTTTGTGC to amplify a 450-bp
fragment, with 35 cycles of 25 s at 94 °C, 25 s at 59 °C and 25 s
at 72 °C, preceded by an initial denaturation step at 94 °C for
2 min, and followed by a final extension of 5 min at 72 °C. The
primers MAEYF1086–GCAGCTAATCTGCTCACAGCC and
MAEYR1376–AACAATCATACCCATTGGTCGAG were
used to amplify a 291-bp fragment using the PCR conditions
above but with an annealing temperature of 58 °C. For both
assays, primer names indicate their location in the SRY
sequence (GenBank Accession no. AY785433). The first PCR
fragment was digested with the restriction enzyme *Tsp509I*
(New England Biolabs), and the second PCR fragment
was digested with the restriction enzyme *MaeIII* (Roche
Diagnostics). For each fragment, restriction maps for the

---

Fig. 1 Geographic distribution of Y chromosome lineages A and B. Pie charts indicate the frequency of each lineage in the population. Numbers in the figure correspond to populations as follows (for each population the sample size and frequency of the A lineage are given in parentheses): 1, Vila Real (n = 11, 1.00); 2, Idanha (n = 18, 1.00); 3, Vila Viçosa (n = 12, 1.00); 4, Elvas (n = 7, 1.00); 5, Sevilla (n = 27, 1.00); 6, Huelva (n = 11, 1.00); 7, Doñana (n = 8, 1.00); 8, Las Lomas (n = 10, 1.00); 9, Fuente Piedra (n = 14, 1.00); 10, Córdoba (n = 10, 1.00); 11, Verin (n = 7, 1.00); 12, Bragança (n = 12, 0.86); 13, Toledo (n = 32, 0.94); 14, Ciudad Real (n = 18, 1.00); 15, Albacete SW (n = 8, 0.87); 16, Las Amoladeras (n = 4, 0.00); 17, Benavente (n = 21, 0.00); 18, Zamora (n = 14, 0.00); 19, Madrid (n = 17, 0.00); 20, Tudela (n = 8, 0.00); 21, La Rioja (n = 4, 0.00); 22, Lleida (n = 7, 0.00); 23, Zaragoza (n = 9, 0.00); 24, Rosell (n = 8, 0.25); 25, Cuenca (n = 5, 0.20); 26, Albacete N (n = 8, 0.00); 27, Valencia (n = 9, 0.44); 28, Alicante (n = 18, 0.00); 29, Cartagena (n = 10, 0.00); and 30, Perpignan (n = 11, 0.00). The estimated centre of the
hybrid zone is indicated by a full line, and the dashed lines provide approximate confidence boundaries assuming a 10% variation to each side of the hybrid zone.
four wild rabbits sequenced in Geraldes et al. (2005) were generated using the BioEdit software (Hall 1999). Tsp509I either cuts the 450 bp fragment of the SRY gene three times, at nucleotides 319, 516 and 596, producing four different fragments of 53, 197, 80 and 120 bp (profile Tsp509I-A), or two times, at nucleotides 319 and 516, producing three different fragments of 53, 197, 200 bp (profile Tsp509I-B). MaeIII was used to digest the other fragment of SRY (291 bp), producing either one cut at nucleotide 1177 with two fragments of 92 and 199 bp (profile MaeIII-A), or two cuts at nucleotides 1177 and 1233 with fragments of 92, 56 and 143 bp (profile MaeIII-B). Restriction products were visualized by silver staining after non-denaturing electrophoresis separation in 9% polyacrylamide gels.

Short intronic fragments of three loci, EDNRA (endothelin receptor type A), PROC (protein C), and NNT (nicotinamide nucleotide transhydrogenase), located on rabbit chromosomes 15q11dist, 7q14 and 11q13–q14 (Chantry-Darmon et al. 2003, 2005), respectively, were PCR-amplified using primers EDNRAF–TGCTGGTCTCCTCTCATTT, EDNRAAR–GAATTCATGGTCGCCAAGTT, PROC–AATCGAGAATTCACTATTCG, PROCF–AATCGAGAAATCGAGAATTCACTATTCG. PCR conditions were as follows: 35 cycles of 30 s at 92 °C, 30 s at 55 °C (EDNRA and NNT) and 63 °C (PROC), and 90 s at 72 °C, preceded by an initial denaturation step at 92 °C for 2 min, and followed by a final extension of 5 min at 72 °C. PCR products were purified using the QIAquick PCR purification kit (QIAGEN) prior to sequencing. Sequencing was carried out using the amplification primers in an ABI 3700 automated sequencer. For PROC only four, instead of five, individuals from the sample Lepus granatensis were sequenced. We also PCR-amplified and sequenced one wild rabbit sequenced in Geraldes et al. (2005) was deposited in GenBank under Accession nos EU862090–EU862121. Alignments were checked and manually edited with BioEdit (Hall 1999). All indel polymorphisms were excluded from subsequent analyses. Haplotypes were inferred with Phase 2.1.1 (Stephens et al. 2001; Stephens & Donnelly 2003) after checking for convergence of three independent runs for each locus. The program sites (Hey & Wakeley 1997) was used to calculate a number of summary statistics, including π (Nei & Li 1979) and θ (Wattoner 1975), two estimators of the population mutation parameter 4μN (where μ is the neutral mutation rate and N is the effective population size), γ (Hey & Wakeley 1997), an estimator of the population recombination parameter 4Nc (where c is the recombination rate), Rm, the minimum number of recombination events (Hudson & Kaplan 1985) and Dxy (Nei 1987), the average pairwise divergence (between all rabbit alleles and Lepus granatensis). To test for departures from a neutral model of molecular evolution, we performed two tests based on the frequency spectrum of polymorphisms (within each population group and also in the entire rabbit sample), Tajima’s D (Tajima 1989) and Fu and Li’s D (Fu & Li 1993). The Hudson–Kreitman–Aguade (HKA) test (Hudson et al. 1987) was used to calculate the ratio of polymorphism (within each population group, and also within the entire rabbit sample) to divergence (to L. granatensis) among loci. Frequency spectrum tests were performed using sites (Hey & Wakeley 1997). Statistical significance for all neutrality tests was obtained by performing 1000 coalescent simulations conditioned on the parameters estimated from our data using the program HKA (http://lifesci.rutgers.edu/~heylab/HeylabSoftware.htm#HKA). Median-joining networks (Bandelt et al. 1999) depicting the evolutionary relationships among alleles were inferred with the software Network 4.2.0.1 (http://www.fluxus-technology.com).

Levels of genetic differentiation on the Y chromosome between NE and SW population groups were compared to the sequence data from the three autosomal loci sequenced here and also to previously published data: four X-linked loci (Geraldes et al. 2006), mtDNA (Branco et al. 2000) and 17 protein loci (Ferrand & Branco 2005), HBB (hemoglobin, beta) is located on chromosome 15q23 (Chantry-Darmon et al. 2005), GC (group-specific component) is located on chromosome 15q23dist (Chantry-Darmon et al. 2005), HBB (haemoglobin, beta) is located on chromosome 1q14–q21 (Xu & Hardison 1989) and SOD1 (superoxide dismutase 1, soluble) is located on chromosome 6p12 (Lemieux & Dutrillaux 1992). The remaining loci are not currently mapped in the rabbit genome. While both microsatellites (Queney et al. 2001) and immunoglobulin allotypes (Esteves et al. 2004) are also available for the same set of populations, they have not been included in this analysis due to the effects of homoplasy, in the first case, and natural selection, in the second. To account for differences in sample size between SW and NE groups, we resampled randomly from the group with the highest sample size. Sample sizes for each locus are shown.

Data analysis

Haplotype frequencies for the SRY data were calculated for each population. Population average heterozygosity was calculated as 1 – Σ x^2, where x is the frequency of each of the alleles. Population differentiation among the SW and NE population groups was assessed with FST (Wright 1951) and Nei’s genetic distance, D (Nei 1972).

Sequence data were trimmed, assembled and edited in phred/phrap/consed/polypyred (Nickerson et al. 1997; Ewing & Green 1998; Ewing et al. 1998; Gordon et al. 1998) coupled with automated shell scripts and Perl programs kindly provided by August Woerner. The resulting contigs were deposited in GenBank under Accession nos EU862090–EU862121. Alignments were checked and manually edited
in Table 1. For the protein data, $F_{ST}$ and $D$ were calculated based on electrophoretic allele frequencies for each locus. For the mtDNA data, differentiation was calculated from the two major haplogroups (Branco et al. 2000). To make the sequence data comparable to the other data sets, $F_{ST}$ and $D$ were calculated for a diagnostic SNP between the two observed lineages at each X-linked (Geraldes et al. 2006) and the two autosomal centromeric loci (EDNRA and PROC). Since each of these genes contained two divergent lineages, there were multiple SNPs at each gene that showed equivalent patterns of differentiation. We used the following SNPs: PHKA2 (site 1388), SMCX (site 252), MSN (site 64) and HPRT1 (site 951) (Geraldes et al. 2006), EDNRA (site 335) and PROC (site 58). For NNT, we failed to detect a deep split in the genealogy of alleles (Fig. 2), and therefore, $F_{ST}$ and $D$ were calculated using a random SNP with intermediate frequency (site 167). The use of other SNPs at this locus produced equally low estimates of population differentiation.

We used the program IMa (Hey & Nielsen 2007) to conduct likelihood ratio tests comparing models of divergence between rabbit subspecies with and without gene flow. Non-recombining data sets were obtained for the four X-linked and the three autosomal loci for which we had nucleotide sequences using the program IMgc (Woerner et al. 2007). We ran IMa under metropolis coupled Markov chain Monte Carlo, using six chains with a two-step heating scheme and parameters that allowed for proper chain swapping. We ran the program for 17 million steps after a burn-in period of 2 million steps. We checked for convergence between the three replicates, and used the trees generated in the longest replicate to perform the likelihood ratio tests comparing models with and without gene flow.

### Results

#### Y chromosome data

PCR and restriction digests were performed on all 353 samples and no novel restriction profiles were observed. The restriction profile Tsp509I-A was always associated...
with the restriction profile MaeIII-A, indicating complete linkage disequilibrium between the mutations at SRY nucleotides 596 and 1233. In the total sample, 203 individuals had the A profile and 150 individuals had the B profile, referred to below as lineage A and lineage B. The frequency of each lineage in the different populations is shown in Fig. 1. All 10 populations in the SW group were fixed for lineage A. Of the 13 populations in the NE group, 10 were fixed for lineage B and three were polymorphic, with frequencies of lineage A of 0.20, 0.25 and 0.44, respectively. In the hybrid zone, two populations were fixed for lineage A, one for lineage B, and three were polymorphic (frequency of lineage A was 0.68, 0.94 and 0.97 in these populations). These results are in general agreement with the distribution of a 7-bp insertion in the 3′ untranslated region of the SRY gene (Geraldes & Ferrand 2006). Haplotype diversity was significantly partitioned between subspecies (\(F_{ST} = 0.93\) and \(D = 3.20\) between SW and NE, Table 1). This reveals a high level of genetic differentiation between the two major groups. Thus, the two divergent Y chromosome lineages (Geraldes et al. 2005) are highly structured geographically.

**Autosomal sequence data**

For EDNRA, PROC and NNT, estimates of nucleotide polymorphism (π and θ), recombination (γ) and divergence (Dxy) to *Lepus granatensis* are shown in Table 2. Levels of nucleotide polymorphism in the entire sample [average π for the three loci = 1.108% (SE = 0.154%)], and in each population group [average π for the three loci in NE = 1.010% (SE = 0.043%), and average π for the three loci in SW = 0.880% (SE = 0.090%)] were high and similar to each other. Recombination was detected at every locus (Rm in the entire sample was higher than one at each locus), but estimates of the recombination parameter (γ) were generally lower than estimates of the mutation parameter (π). Divergence (Dxy) to *L. granatensis* averaged over the three loci was 4.43% (SE = 0.18%). No deviations from neutral expectations were detected either using tests based on the frequency spectrum of polymorphisms (Tajima's D and Fu and Li's D), or the ratio of polymorphism to divergence (HKA test). The visual inspection of the tables of polymorphism for each locus and the resulting networks (Fig. 2) revealed the existence of two evolutionary lineages at EDNRA (defined by four sites) and PROC (defined by five sites), the two loci located near centromeres. The same was not true for NNT, where no sites consistently defined two distinct evolutionary lineages.

**Comparison of estimates of population differentiation**

Estimates of heterozygosity, \(F_{ST}\) and \(D\) between the NE and SW population groups for all loci are presented in Table 1. \(F_{ST}\) for MSN (X-linked) and SRY (Y-linked) were above 0.93, indicating extremely high levels of population differentiation at these two loci. \(F_{ST}\) values above 0.74 were detected at five loci (SRY, Cytb, SMCX, MSN and PROC) out of a total of 26 loci under comparison. \(F_{ST}\) for the remaining 21 loci indicated low to moderate levels of genetic differentiation (\(F_{ST}\) values below 0.32, Table 1 and Fig. 3). Because the data sets differ slightly in terms of sample localities, we performed two additional analyses of the data. First, \(F_{ST}\) values were also calculated using only the populations that are common to the four different studies, and similar results were obtained (data not shown). Second, we calculated Nei’s genetic distance, \(D\), because it is less sensitive to different
sampling regimes, and similar results were obtained (Table 1). The data sets also differ in terms of sample size, with some loci having as few as 16 to 20 chromosomes and others having over 200 chromosomes. We note simply that there is no association between sample size and level of differentiation (Table 1).

Finally, a likelihood ratio test (Hey & Nielsen 2007) comparing models with and without gene flow revealed a significantly better fit to a model with gene flow ($P < 0.001$).

**Discussion**

**Reduced Y-chromosome introgression between rabbit subspecies**

We observed substantial differentiation between subspecies of rabbit for the Y chromosome. All *O. c. algirus* populations studied are fixed for the Y chromosome lineage A, and most *O. c. cuniculus* populations are fixed for lineage B ($F_{ST} = 0.93$). In fact, only three populations in the hybrid zone (Fig. 1) exhibited the two divergent Y chromosome lineages, supporting the view of a recent secondary contact of the two subspecies after population expansion in postglacial times (e.g. Ferrand 2008). Additionally, three populations from NE Spain are polymorphic for the Y chromosome (Fig. 1). While a more patchy structure of the hybrid zone is a possibility that must be examined by further studies, a more plausible explanation is related to the common practice of rabbit restocking in regions of low rabbit density or following outbreaks of rabbit epizootic diseases (Moreno & Villafuerte 1995; Moreno _et al._ 2004). Branco _et al._ (2000) invoked a similar process to explain the occurrence of mtDNA typically found in *O. c. algirus* in populations of *O. c. cuniculus* from the Ebro valley, in NE Spain. While a detailed genetic characterization of these populations will be needed to resolve this issue, two decades of rabbit population genetic studies in Iberia indicate that the effects of restocking are marginal and localized (Ferrand 2008). In general, however, the degree of Y chromosome introgression between subspecies is very low.

This observation is in stark contrast with the analysis of population differentiation for three autosomal DNA fragments that were also included in this study, which varied from very low (0.02 and 0.08 for NNT and EDNRA, respectively) to very high $F_{ST}$ values (0.77 for PROC). To put our results in perspective, we compared them with estimates of population differentiation for mtDNA, X chromosome and protein loci (Table 1). When we plot all $F_{ST}$ values obtained

---

**Table 2** Levels of polymorphism, allele-frequency spectrum tests of neutrality and recombination, for all rabbit samples and NE and SW population groups, and divergence between all rabbit samples and *Lepus granatensis*

<table>
<thead>
<tr>
<th>Length (bp)</th>
<th>Polymorphism</th>
<th>Divergence</th>
<th>Frequency spectrum</th>
<th>Recombination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S$</td>
<td>$\pi$ (%)</td>
<td>$\theta$ (%)</td>
<td>$D_{xy}$ (%)</td>
</tr>
<tr>
<td>EDNRA All</td>
<td>697</td>
<td>20</td>
<td>25</td>
<td>1.068</td>
</tr>
<tr>
<td>NE 697</td>
<td>10</td>
<td>18</td>
<td>1.081</td>
<td>0.913</td>
</tr>
<tr>
<td>SW 697</td>
<td>10</td>
<td>22</td>
<td>1.011</td>
<td>1.116</td>
</tr>
<tr>
<td>PROC All</td>
<td>627</td>
<td>18</td>
<td>27</td>
<td>1.393</td>
</tr>
<tr>
<td>NE 627</td>
<td>10</td>
<td>18</td>
<td>1.015</td>
<td>0.872</td>
</tr>
<tr>
<td>SW 627</td>
<td>8</td>
<td>15</td>
<td>0.923</td>
<td>0.946</td>
</tr>
<tr>
<td>NNT All</td>
<td>950</td>
<td>20</td>
<td>33</td>
<td>0.864</td>
</tr>
<tr>
<td>NE 950</td>
<td>10</td>
<td>26</td>
<td>0.933</td>
<td>0.967</td>
</tr>
<tr>
<td>SW 950</td>
<td>10</td>
<td>19</td>
<td>0.707</td>
<td>0.662</td>
</tr>
</tbody>
</table>

$s$, the number of segregating sites; $\pi$, the average number of nucleotide differences per site (Nei & Li 1979); $\theta$, is the proportion of polymorphic sites (Watterson 1975); $D_{xy}$, the average pairwise divergence per site (Nei 1987); $TD$, Tajima’s $D$ (Tajima 1989); $FLD$, Fu and Li’s $D$ (Fu & Li 1993); $\gamma$, a maximum-likelihood estimator of the population recombination parameter between adjacent sites (Hey & Wakeley 1997); $Rm$, the minimum number of recombination events (Hudson & Kaplan 1985).
for a total of 26 markers distributed in the four rabbit genomic compartments, a clear bimodal distribution is observed, with a majority (81%) of $F_{ST}$ values below 0.4, and a minority (19%) above 0.7. To our knowledge, this is the first time that such a striking bimodality is apparent in the study of subspecies divergence in a mammalian species, and thus deserves further explanation. Of particular interest is the class of markers that exhibits high levels of differentiation. In this class, we find not only genomic regions which are recombination free (SRY and mtDNA Cytb), but also X chromosome (SMCX and MSN) and autosomal markers (PROC) that are located close to centromeres, which are expected to experience low levels of recombination. Within this class, results for the Y chromosome show a near-complete absence of introgression between subspecies. This observation is remarkable given that male rabbits tend to disperse before reaching sexual maturity, while females are much more phylopatric (Webb et al., 1995; Kunkele & vonHolst 1996; Richardson et al., 2002). This would have led us to predict male-mediated gene flow between rabbit subspecies. Notably, we observe precisely the opposite and the Y chromosome shows in fact very high levels of differentiation between rabbit subspecies.

**Ancestral polymorphism vs. recent gene flow between rabbit subspecies**

In principle, differences in levels of differentiation among loci could be due to differences in the degree of sorting of ancestral polymorphism or to differences in levels of gene flow following secondary contact. We discuss each of these in turn.

As two populations diverge, gene genealogies will typically proceed from polyphyly to paraphyly to reciprocal monophyly, and the rate at which this occurs will depend on population size (Avisé, 1994). Assuming a sex ratio of one, the effective population size of mtDNA and the Y chromosome is one-fourth that of the autosomes, and the effective population size of the X chromosome is three-fourths that of the autosomes. Thus, the predicted order of lineage sorting is Y chromosome and mtDNA first, then the X chromosome, and finally the autosomes. This is precisely the pattern we observe. In the case of the nine loci for which genealogical information was obtained, patterns consistent with long-term evolution in allopatry (a deep split in the genealogy) are observed for the Y-linked locus (Geraldes et al., 2005), the mtDNA locus (Branco et al., 2000), all X-linked loci (Geraldes et al., 2006) and two of the three autosomal loci (this study, Fig. 2). Although the sample size for X-linked and autosomal loci is rather small, this suggests that considerable time has elapsed since the divergence of these subspecies and that most X-linked and many autosomal genes will have attained reciprocal monophyly. Thus, it seems unlikely that unsorted ancestral polymorphism is a sufficient explanation by itself for the low $F_{ST}$ values observed at most loci. Coalescent models have been developed in a likelihood framework to distinguish between unsorted ancestral polymorphism and gene flow as alternative explanations for shared polymorphism (Hey & Nielsen, 2004, 2007). When applied to the X chromosome and autosomal data here, models with substantial gene flow are a significantly better fit to the data than models without gene flow (likelihood ratio test, $P < 0.001$). In the face of this evidence for gene flow, one has to explain why some genomic regions, such as the Y chromosome, show little introgression.

We suggest that some of the variation in $F_{ST}$ among markers in Table 1 is likely due to differential introgression of genes following secondary contact (Geraldes et al., 2006). The comparisons presented here are consistent with the notion that different portions of the genome show different levels of genetic isolation (e.g., Rieseberg et al., 1999; Machado et al., 2002; Wu & Ting, 2004). Interestingly all genes showing low introgression between rabbit subspecies are located in regions either devoid of recombination (SRY and Cytb), or located near centromeres (MSN, SMCX and PROC) where recombination is known to be reduced in other species (e.g., Mahtani & Willard, 1998; Jensen-Seaman et al., 2004). Currently, recombination rates in the rabbit genome are not available, but estimates of recombination based on LD at four X-linked genes support the notion that, at least on the X, recombination might be reduced near the centromere. The low values of $F_{ST}$ observed at all 17 protein loci presented in Table 1 are noteworthy in two respects. First, none of these protein loci that have so far been mapped in the rabbit genome are located close to a centromere. Second, the absence of differentiation at a locus in the present does not necessarily mean that differentiation was not high in the past. For example, in a series of detailed electrophoretic, molecular and simulation studies, Campos et al., (2007, 2008) suggested that rabbit HBB consists of two highly divergent haplotypes that may have been fixed in the past in the two rabbit subspecies (corresponding then to a $F_{ST} = 1$) and that, after secondary contact, extensive admixture led to the homogenization of allelic frequencies (corresponding now to a $F_{ST} = 0$). The rabbit genome project will enable us to study genes throughout the genome. We will then be able to ask how many genes show two divergent lineages, suggesting a long period of isolation. We will also be able to obtain a picture of the bimodality of the distribution of $F_{ST}$ values throughout the rabbit genome.

**A role for the Y chromosome in reproductive isolation?**

Several authors have proposed a model of speciation where regions with suppressed recombination are more likely to promote speciation by extending the effects of isolation genes to linked sites (Noor et al., 2001; Rieseberg, 2001).
Evidence from many studies indicates that reproductive isolation between divergent taxa is often due to negative epistatic interactions (i.e. ‘Dobzhansky–Muller incompatibilities’, Coyne & Orr 2004). For example, mitochondrial function requires the expression of nuclear and mitochondrial encoded genes that form enzyme complexes. This requires the co-evolution of the nuclear and mitochondrial genomes to ensure proper function. In several instances, cytonuclear incompatibilities are known to arise following hybridization (e.g. Sackton et al. 2003; Ellison & Burton 2006; Fishman & Willis 2006).

Genomic regions that introgress less are strong candidates for containing genes involved in such incompatibilities. In the rabbit, five such regions have now been described: one autosomal locus (PROC), two X-linked centromeric loci (MSN and SMCX), the mitochondria and the Y chromosome. There has been little attention devoted to the role of the Y chromosome in speciation despite the fact that the male-specific region of the Y is devoid of heterologous recombination, is highly enriched for testis-specific genes (Skatesky et al. 2003) and, has variants in human populations which are the most common cause of spermatogenic failure (Kuroda-Kawaguchi et al. 2001). Our finding of reduced introgression for the Y chromosome between rabbit subspecies is consistent with hybrid zone studies in other species (e.g. house mice, Vanlengerbe et al. 1986; Tucker et al. 1992; shrews, Balloux et al. 2000; elephants, Roca et al. 2005; field voles, Jaarola et al. 1997; fruit flies Llopard et al. 2005), suggesting that incompatibilities involving the Y chromosome might be common and its role in speciation underappreciated.

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia (SFRH/BD/4621/2001 and SFRH/BPD/24743/2005 grants to A.G. and Research Projects POCI/BSE/40280/2001 and PTDC/BIA-BDE/72304/2006), JCCM PA106-0170 and INIA FAU 2006-00014-C02-02, and by National Science Foundation grant to M.W.N. We would like to thank P. Tarroso for invaluable help with the map in Fig. 1. M. Branco, R. Campos, M. Dean, J. Good, J. Hausser, J. Pialek, C. Pinho, R. Storchova, G. Wlasiuk, E. Wood and one anonymous reviewer who made valuable comments to the manuscript.

References


© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd


This paper is part of the PhD thesis of A. Geraldes, who is currently a postdoctoral fellow in the laboratory of M. W. Nachman where he is working on patterns of gene flow between three species of house mice. M. Carneiro is a PhD student working on speciation in wild rabbits. R. Villafuerte and M. Delibes-Mateos work on conservation biology of several Iberian mammals including the European rabbit. M.W. Nachman studies population, evolutionary and ecological genetics and genomics. N. Ferrand heads the CIBIO and is interested in a variety of questions in evolutionary and conservation genetics.