Molecular Ecology (2008) **17**, 4489–4499 doi: 10.1111/j.1365-294X.2008.03943.x

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*

Correspondence: Armando Geraldes, Fax: (520) 621-9190; E-mail: geraldes@email.arizona.edu

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 (Vanlerberghe *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* (Cyt*b*) 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 & vonHolst 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_{ST} values of the 26 loci studied follow a striking bimodal distribution, with a majority of loci with low F_{ST} values, and very few with high F_{ST} values.

Materials and methods

Sampling

We sampled 353 male rabbits from 30 natural populations. Populations were divided into three groups (Fig. 1):

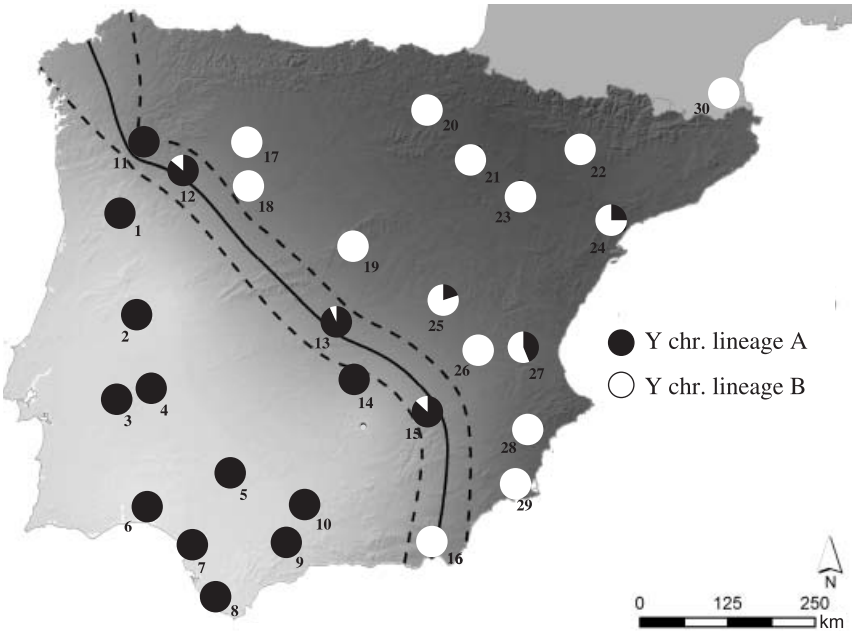

Fig. 1 Geographic distribution of Y chromosome lineages A and B. Pie charts indicate the frequency of each lineage in the population. Numbers in the figure correspond to populations as follows (for each population the sample size and frequency of the A lineage are given in parentheses): 1, Vila Real (*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.

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 *Tsp*509I (New England Biolabs), and the second PCR fragment was digested with the restriction enzyme *Mae*III (Roche Diagnostics). For each fragment, restriction maps for the

four wild rabbits sequenced in Geraldes *et al*. (2005) were generated using the BioEdit software (Hall 1999). *Tsp*509I 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 *Tsp*509I-A), or two times, at nucleotides 319 and 516, producing three different fragments of 53, 197, 200 bp (profile *Tsp*509I-B). *Mae*III 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 *Mae*III-A), or two cuts at nucleotides 1177 and 1233 with fragments of 92, 56 and 143 bp (profile *Mae*III-B). Restriction products were visualized by silver staining after nondenaturing 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– TGCTGGTTCCCTCTTCATTT, EDNRAR–GAATTCATGGTCGCCAAGTT, PROCF–AATCGAGAAGAAACGCGGTA, PROCR–GGCCAGCTTCTTCTTGGAGT, NNTF–ATCCACATTTTGCCATGGAT and NNTR–CAGCAAGCCTGCATTGAGTA. 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 each locus, five individuals from NE (population 23, $n = 3$ and population 28, $n = 2$) and SW (population 5, $n = 2$ and population 6, $n = 3$) population groups were sequenced. For PROC only four, instead of five, individuals from the SW population group were sequenced. We also PCR-amplified and sequenced one *Lepus granatensis* sample from each of these loci and used it as an outgroup.

Data analysis

Haplotype frequencies for the SRY data were calculated for each population. Population average heterozygosity was calculated as $1 - \sum x_i^2$, where x_i is the frequency of each of the alleles. Population differentiation among the SW and NE population groups was assessed with F_{ST} (Wright 1951) and Nei's genetic distance, *D* (Nei 1972).

Sequence data were trimmed, assembled and edited in phred/phrap/consed/polyphred (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 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 (Watterson 1975), two estimators of the population mutation parameter $4N\mu$ (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 compare 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 2007). ALB (albumin) 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

Table 1 Estimates of population average heterozygosity for NE and SW population groups and for the entire sample, and estimates of differentiation between NE and SW

			Heterozygosity					
	*n**	Genome location†	NE	SW	Total	F_{ST}	Nei's *D*	Reference‡
SRY	256	Y-linked	0.075	0	0.499	0.925	3.204	This study
Cytb	290	Cytoplasmic	0.079	0.092	0.500	0.829	2.368	1
PHKA2	28	X-linked (Xd)	0	0.254	0.135	0.060	0.014	2
SMCX	28	X-linked (Xc)	0.254	0	0.499	0.745	1.805	2
MSN	28	X-linked (Xc)	0	0	0.509	1	∞	2
HPRT1	28	X-linked (Xd)	0.254	0.519	0.444	0.130	0.206	2
EDNRA	20	Autosomal (15c)	0.525	0.335	0.456	0.075	0.154	This study
PROC	16	Autosomal (7c)	0.233	0	0.508	0.770	1.955	This study
NNT	20	Autosomal (11d)	0.505	0.505	0.513	0.015	0.080	This study
ADA	256	Autosomal	0.469	0.435	0.552	0.181	0.454	3
ALB	256	Autosomal (15d)	0.519	0.522	0.521	0.001	0.002	3
CAI	256	Autosomal	0.157	0.016	0.090	0.038	0.004	3
CAII	256	Autosomal	0.540	0.400	0.527	0.109	0.235	3
GALT	204	Autosomal	0.251	0.199	0.228	0.015	0.008	3
GC	248	Autosomal (15d)	0	0.288	0.153	0.057	0.006	3
HBA	256	Autosomal	0.609	0.402	0.628	0.195	0.660	3
HBB	256	Autosomal (1d)	0.219	0.144	0.182	0.006	0.002	3
HPX	252	Autosomal	0.643	0.676	0.685	0.037	0.158	3
NP	252	Autosomal	0	0.198	0.104	0.050	0.006	3
PEPA	248	Autosomal	0	0.508	0.373	0.319	0.323	3
PEPB	256	Autosomal	0.119	0.228	0.175	0.008	0.001	3
PEPC	256	Autosomal	0.146	0.626	0.442	0.127	0.119	3
PEPD	256	Autosomal	0	0.514	0.351	0.268	0.229	3
PGD	256	Autosomal	0.134	0.504	0.466	0.317	0.530	3
SOD	180	Autosomal (6d)	0	0.044	0.022	0.011	0	3
TF	240	Autosomal	0.033	0.446	0.277	0.137	0.067	3

*Total sample size for each locus. The sample size for each population group is half the value presented; †Approximate chromosome location; c indicates that the locus is located near a centromere and d that it has a more distal position on the chromosome. See Materials and methods for details; ‡References are: 1, Branco *et al*. 2000; 2, Geraldes *et al*. 2006; 3, Ferrand & Branco 2007.

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 *Tsp*509I-A was always associated

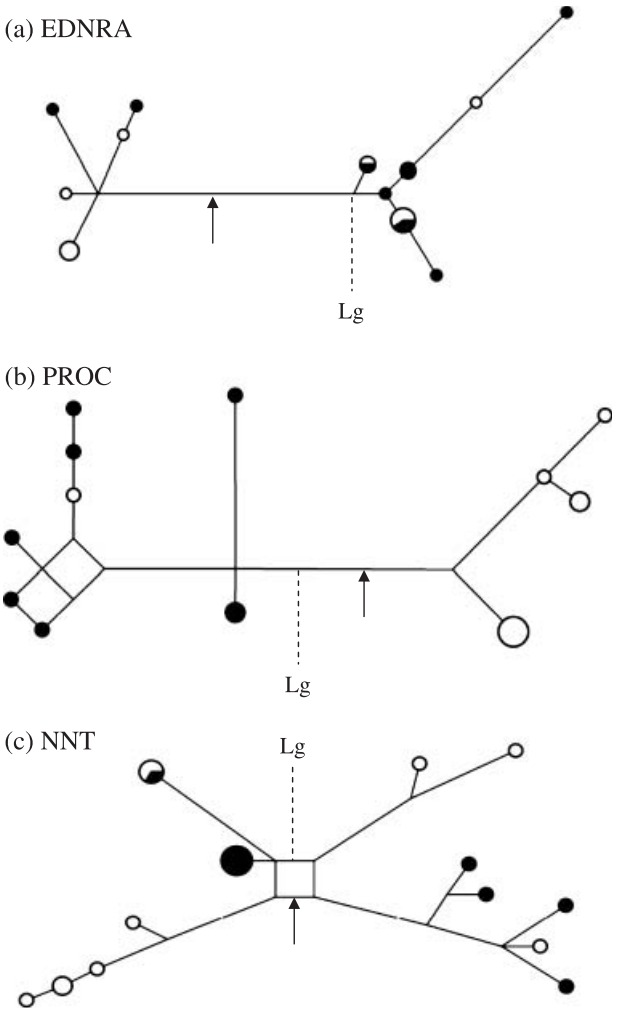

Fig. 2 Haplotype networks depicting the relationship among the alleles found at each autosomal locus sequenced here. (a) EDNRA, (b) PROC and (c) NNT. White circles represent haplotypes found in NE samples, black circles haplotypes found in SW. The length of the branches is proportional to the number of mutational steps separating the haplotypes, and the size of the circles is proportional to the haplotype frequencies. The root is indicated by a dashed line and the letters Lg (*Lepus granatenis*). The arrow indicates the position of the SNP used to calculate F_{ST} and *D*.

with the restriction profile *Mae*III-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, Cyt*b*, 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

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*

				Polymorphism			Divergence	Frequency spectrum		Recombination	
		Length (bp)	*n*	S	π (%)	θ (%)	Dxy (%)	TD	FLD	γ (%)	Rm
EDNRA	All	697	20	25	1.068	1.011	4.183	0.217	0.012	1.272	2
	NE	697	10	18	1.081	0.913		0.859	0.103	0.749	1
	SW	697	10	22	1.011	1.116		–0.445	–0.293	1.499	2
PROC	All	627	18	27	1.393	1.252	4.799	0.448	0.446	0.554	3
	NE	627	10	18	1.015	0.872		–0.658	–0.751	0	0
	SW	627	8	15	0.923	0.946		–0.751	–0.067	3.154	1
NNT	All	950	20	33	0.864	0.979	4.348	–0.465	–0.567	0.478	1
	NE	950	10	26	0.933	0.967		–0.168	–0.167	0.239	1
	SW	950	10	19	0.707	0.662		–0.298	–0.893	0	0

S, the number of segregating sites; π, the average number of nucleotide differences per site (Nei & Li 1979); θ, is the proportion of polymorphic sites (Watterson 1975); Dxy, the average pairwise divergence per site (Nei 1987); TD, Tajima's *D* (Tajima 1989); FLD, Fu and Li's *D* (Fu & Li 1993); γ, 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).

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.

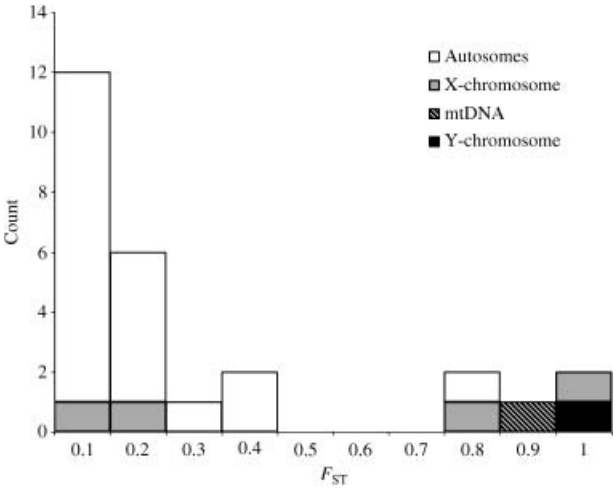

Fig. 3 Histogram of F_{ST} values among subspecies of rabbits (data from Table 1). Autosomal loci are shown in white, X-linked loci in grey, Cytoplasmic loci dashed and Y-linked loci in black.

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

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 Cyt*b*), 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 (Avise 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 Cyt*b*), 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 $F_{ST} \approx 1$) and that, after secondary contact, extensive admixture led to the homogenization of allelic frequencies (corresponding now to a $F_{ST} \approx 0$). 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 (Skaletsky *et al*. 2003) and, has variants in human populations which are the most common cause of spermatogenetic 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, Vanlerberghe *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 Llopart *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 POCTI/BSE/40280/2001 and PTDC/BIA-BDE/72304/2006), JCCM PA106-0170 and INIA FAU 2006-00014-C02-02, and by a 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

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution*, **54**, 1414–1422.

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

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

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

Campos R, Branco M, Weiss S, Ferrand N (2007) Patterns of hemoglobin polymorphism [α-globin (HBA) and β-globin (HBB)] across the contact zone of two distinct phylogeographic lineages of the European rabbit (*Oryctolagus cuniculus*). In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds Weiss S, Ferrand N). Springer, Amsterdam, The Netherlands.

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

Chantry-Darmon C, Rogel-Gaillard C, Bertaud M *et al*. (2003) 133 new gene localizations on the rabbit cytogenetic map. *Cytogenetic and Genome Research*, **103**, 192–201.

Chantry-Darmon U, Urien C, Hayes H *et al*. (2005) Construction of a cytogenetically anchored microsatellite map in rabbit. *Mammalian Genome*, **16**, 442–459.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Coyne JA, Elwyn S, Kim SY, Llopart A (2004) Genetic studies of two sister species in the *Drosophila melanogaster* subgroup, *D. yakuba* and *D. santomea*. *Genetical Research*, **84**, 11–26.

Eicher EM, Washburn LL (1986) Genetic control of primary sex determination in mice. *Annual Review of Genetics*, **20**, 327–360.

Ellison CK, Burton RS (2006) Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. *Evolution*, **60**, 1382–1391.

ESRI (2006) *ArcGis 9.2*. Environmental Systems Research Institute, Inc., Redlands, California.

Esteves PJ, Lanning D, Ferrand N *et al*. (2004) Allelic variation at the V(H)a locus in natural populations of rabbit (*Oryctolagus cuniculus*, L.). *Journal of Immunology*, **172**, 1044–1053.

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, **8**, 186–194.

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research*, **8**, 175–185.

Ferrand N (2008) Inferring the evolutionary history of the European rabbit (*Oryctolagus cuniculus*) from molecular markers. In: *Lagomorph Biology: Evolution, Ecology and Conservation* (eds Alves PC, Ferrand N, Hacklander K). Springer, Amsterdam, The Netherlands.

Ferrand N, Branco M (2007) The evolutionary history of the European rabbit (*Oryctolagus cuniculus*): major patterns of population differentiation and geographic expansion inferred from protein polymorphism. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity* (eds Weiss S, Ferrand N). Springer, Amsterdam, The Netherlands.

Fishman L, Willis JH (2006) A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, **60**, 1372–1381.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Geraldes A, Ferrand N (2006) A 7-bp insertion in the 3′ untranslated region suggests the duplication and concerted evolution of the rabbit SRY gene. *Genetics Selection and Evolution*, **38**, 313–320.

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

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

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. *Genome Research*, **8**, 195–202.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analyses program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration with the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785–2790.

Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics*, **145**, 833–846.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.

Jaarola M, Tegelstrom H, Fredga K (1997) A contact zone with noncoincident clines for sex-specific markers in the field vole (*Microtus agrestis*). *Evolution*, **51**, 241–249.

Jensen-Seaman MI, Furey TS, Payseur BA *et al*. (2004) Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research*, **14**, 528–538.

Johnson NA, Hollocher H, Noonburg E, Wu CI (1993) The effects of interspecific Y chromosome replacements on hybrid sterility within the *Drosophila simulans* clade. *Genetics*, **135**, 443–453.

Kunkele J, vonHolst D (1996) Natal dispersal in the European wild rabbit. *Animal Behaviour*, **51**, 1047–1059.

Kuroda-Kawaguchi T, Skaletsky H, Brown LG *et al*. (2001) The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nature Genetics*, **29**, 279–286.

Lemieux N, Dutrillaux B (1992) New gene assignments to rabbit chromosomes; implications for chromosome evolution. *Cytogenetics and Cell Genetics*, **61**, 132–134.

Llopart A, Lachaise D, Coyne JA (2005) Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics*, **171**, 197–210.

van der Loo W, Ferrand N, Soriguer RC (1991) Estimation of gene diversity at the b-locus of the constant region of the immunoglobulin light chain in natural populations of the European rabbit (*Oryctolagus cuniculus*) in Portugal, Andalusia and the Azorean islands. *Genetics*, **127**, 789–799.

van der Loo W, Mougel F, Sanchez MS *et al*. (1999) Cytonuclear disequilibria in wild populations of rabbit (*Oryctolagus cuniculus* L.) suggest unequal allele turnover rates at the *b* locus (IGKC1). *Immunogenetics*, **49**, 629–643.

Machado CA, Kliman RM, Markert JA, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution*, **19**, 472–488.

Mahtani MM, Willard HF (1998) Physical and genetic mapping of the human X chromosome centromere: repression of recombination. *Genome Research*, **8**, 100–110.

Moreno S, Villafuerte R (1995) Traditional management of scrubland for the conservation of rabbits *Oryctolagus cuniculus* and their predators in Doñana National Park, Spain. *Biological Conservation*, **73**, 81–85.

Moreno S, Villafuerte R, Cabezas S, Lombardi L (2004) Wild rabbit restocking for predator conservation in Spain. *Biological Conservation*, **118**, 183–193.

Munclinger P, Bozikova E, Sugerkova M, Pialek J, Macholan M (2002) Genetic variation in house mice (*Mus*, Muridae, Rodentia) from the Czech and Slovak republics. *Folia Zoologica*, **51**, 81–92.

Nei M (1972) Genetic distance between populations. *The American Naturalist*, **106**, 283–292.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.

Nickerson DA, Tobe VO, Taylor SL (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Research*, **25**, 2745–2751.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences, USA*, **98**, 12084–12088.

Orr HA, Coyne JA (1989) The genetics of Postzygotic isolation in the *Drosophila virilis* group. *Genetics*, **121**, 527–537.

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

Richardson BJ, Hayes RA, Wheeler SH, Yardin MR (2002) Social structures, genetic structures and dispersal strategies in Australian rabbit (*Oryctolagus cuniculus*) populations. *Behavioral Ecology and Sociobiology*, **51**, 113–121.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351–358.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Rocca AL, Georgiadis N, O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. *Nature Genetics*, **37**, 96–100.

Sackton TB, Haney RA, Rand DM (2003) Cytonuclear coadaptation in *Drosophila*: disruption of cytochrome *c* oxidase activity in backcross genotypes. *Evolution International Journal of Organic Evolution*, **57**, 2315–2325.

Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ *et al*. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, **423**, 825–837.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

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

Tucker PK, Sage RD, Warner J, Wilson AC, Eicher EM (1992) Abrupt cline for sex-chromosomes in a hybrid zone between 2 species of mice. *Evolution*, **46**, 1146–1163.

Turelli M, Orr HA (1995) The dominance theory of Haldane's rule. *Genetics*, **140**, 389–402.

Vanlerberghe F, Dod B, Boursot P, Bellis M, Bonhomme F (1986) Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. *Genetical Research*, **48**, 191–197.

Vigneault G, Zouros E (1986) The genetics of asymmetrical male-sterility in *Drosophila mojavensis* and *Drosophila arizonensis* hybrids – interactions between the Y-chromosome and autosomes. *Evolution*, **40**, 1160–1170.

Washburn LL, Albrecht KH, Eicher EM (2001) C57BL/6J-T-associated sex reversal in mice is caused by reduced expression of a *Mus domesticus* Sry allele. *Genetics*, **158**, 1675–1681.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Webb NJ, Ibrahim KM, Bell DJ, Hewitt GM (1995) Natal dispersal and genetic structure in a population of the European wild rabbit (*Oryctolagus cuniculus*). *Molecular Ecology*, **4**, 239–247.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

Wu CI, Ting CT (2004) Genes and speciation. *Nature Reviews Genetics*, **5**, 114–122.

Xu J, Hardison RC (1989) Localization of the beta-like globin gene cluster and the genes for parathyroid hormone and c-Harvey-ras 1 to region q14–q21 of rabbit chromosome 1 by *in situ* hybridization. *Cytogenetics and Cell Genetics*, **52**, 157–161.

Yannic G, Basset P, Hausser J (2008) A hybrid zone with coincident clines for autosomal and sex-specific markers in the *Sorex araneus* group. *Journal of Evolutionary Biology*, **21**, 658–667.

Zeng LW, Singh RS (1993) The genetic basis of Haldane's rule and the nature of asymmetric hybrid male sterility among *Drosophila simulans*, *Drosophila mauritiana* and *Drosophila sechellia*. *Genetics*, **134**, 251–260.

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.
