Contrasting Patterns of Introgression at X-Linked Loci Across the Hybrid Zone Between Subspecies of the European Rabbit (*Oryctolagus cuniculus*)

Armando Geraldes,[*,†,‡,1] **Nuno Ferrand**[*,†] **and Michael W. Nachman**[‡]

CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal, †Departamento de Zoologia e Botânica, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal and ‡Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721

ABSTRACT

Hybrid zones provide an excellent opportunity for studying the consequences of genetic changes between closely related taxa. Here we investigate patterns of genetic variability and gene flow at four X-linked loci within and between the two subspecies of European rabbit (*Oryctolagus cuniculus cuniculus* and *O. c. algirus*). Two of these genes are located near the centromere and two are located near the telomeres. We observed a deep split in the genealogy of each gene with the root located along the deepest branch in each case, consistent with the evolution of these subspecies in allopatry. The two centromeric loci showed low levels of variability, high levels of linkage disequilibrium, and little introgression between subspecies. In contrast, the two telomeric loci showed high levels of variability, low levels of linkage disequilibrium, and considerable introgression between subspecies. These data are consistent with suppression of recombination near the centromere of the rabbit X chromosome. These observations support a view of speciation where genomic incompatibilities at different loci in the genome create localized differences in levels of gene flow between nascent species.

A key problem in evolutionary genetics concerns the origin of reproductive isolation between incipient species (COYNE and ORR 2004). Two important conclusions come from previous studies of the genetics of reproductive isolation. First, barriers to gene flow often derive from incompatibilities between allelic variants at two or more loci, *i.e.*, epistasis (BATESON 1909; DOBZHANSKY 1936; MULLER 1940, 1942). Empirical support for epistasis comes from a large body of work in Drosophila, beginning with DOBZHANSKY (1936). More recently, specific genes underlying reproductive isolation have been identified, and all involve epistatic interactions (MALITSCHEK *et al.* 1995; TING *et al.* 1998; BARBASH *et al.* 2003; PRESGRAVES *et al.* 2003). Second, loci contributing to reproductive isolation tend to be overrepresented on the X chromosome in groups in which males are heterogametic (COYNE and ORR 1989). Evidence for the "large X effect" comes from mapping studies of hybrid sterility and hybrid inviability (*e.g.*, DOBZHANSKY 1936; GRULA and TAYLOR 1980; TRUE *et al.* 1996; PRESGRAVES 2003; PRESGRAVES *et al.* 2003; TAO *et al.* 2003). Moreover, HALDANE's (1922) rule (the sterility

or inviability of heterogametic hybrids) seems to be due largely to incompatibilities involving recessive X-linked mutations (TURELLI and ORR 1995, 2000). Finally, in a number of cases where sister species hybridize in nature, X-linked loci introgress less than autosomal loci (*e.g.*, HAGEN 1990; SPERLING and SPENCE 1991; TUCKER *et al.* 1992).

The genetic basis of reproductive isolation has been studied both with laboratory crosses and in natural hybrid zones, and both approaches have advantages and disadvantages. For example, laboratory crosses make it possible to control the genetic background as well as the environment, and they are repeatable. Hybrid zones offer the advantage of many generations of recombination, making fine-scale mapping more feasible. In hybrid zones, it is possible to identify genes contributing to isolation simply from patterns of gene flow without prior knowledge of the phenotype. Hybrid zones also allow us to study species that cannot be crossed in the laboratory. Finally, hybrid zones provide a picture of the fitness of hybrid genotypes under natural conditions.

The European rabbit (*Oryctolagus cuniculus*) provides an opportunity to study the genetic basis of reproductive isolation between recently evolved taxa. This species consists of two subspecies, *O. cuniculus algirus* in the southwestern portion of the Iberian Peninsula and *O. cuniculus cuniculus* in the northeast of the Iberian Peninsula and France. These two groups diverged in allopatry
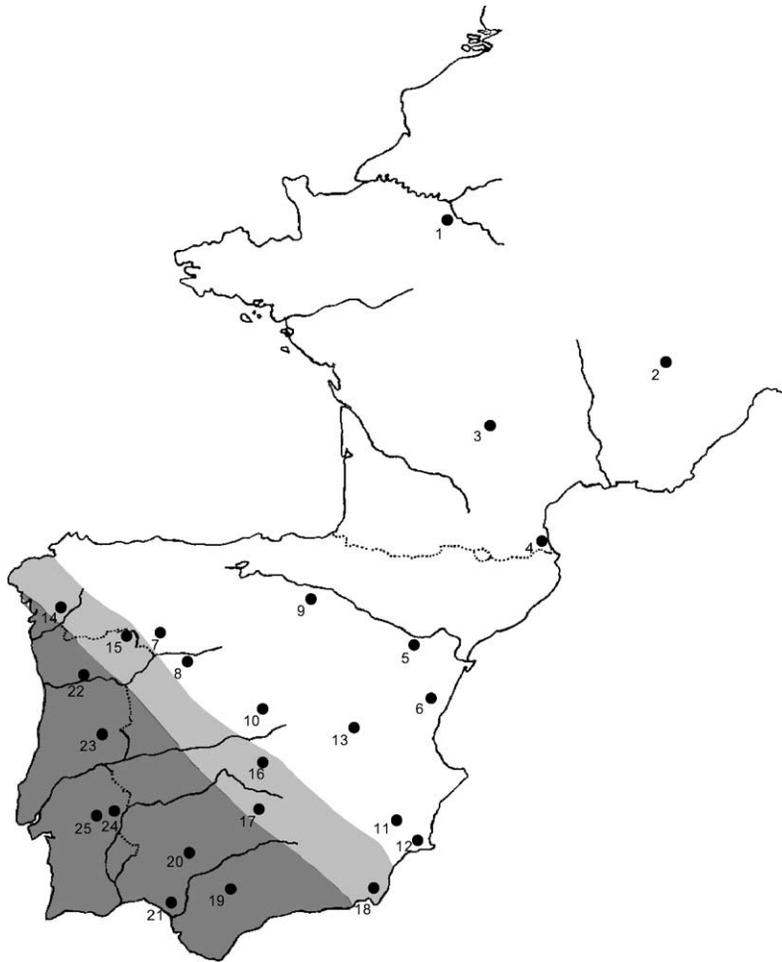
FIGURE 1.—Populations of European rabbit sampled and their geographic locations. Dark shading indicates SW populations, light shading indicates CZ populations, and no shading indicates NE populations. The name and number of samples from each population and population names are specified in Table 1.

during the early Pleistocene and have subsequently come into secondary contact in central Iberia, forming a contact zone that runs in a NW–SE direction (Figure 1) (Branco et al. 2000, 2002). The two subspecies are well differentiated with respect to mtDNA (Branco et al. 2000), the Y chromosome (Geraldes et al. 2005), and some allozyme loci (Ferrand and Branco 2006).

Motivated by the large X-effect documented in other species, here we focus on four X-linked loci to understand the nature of reproductive isolation in rabbits. Two of these loci are near the centromere and two are near the telomeres. We address three main questions. First, what are the levels and patterns of genetic variation at genes on the rabbit X chromosome? Second, are patterns of variation and introgression heterogeneous among loci, and if so, do the differences correlate with the physical location of genes on the X chromosome? Third, are the data compatible with a model of divergence without gene flow? We surveyed nucleotide variability at four X-linked loci in a sample of 43 male rabbits representing both subspecies and the area of contact. All four loci showed two divergent lineages. Despite this deep divergence, there is still evidence of gene flow between subspecies. Patterns of gene flow and nucleotide variability were heterogeneous among loci,

being low at the centromeric loci and high at the telomeric loci. We hypothesize that the centromeric region of the X chromosome of the European rabbit may be involved in reproductive isolation between these two subspecies.

MATERIALS AND METHODS

Samples: Forty-three male European rabbits were sampled (Table 1). The samples were divided into three groups: 20 from the northeastern region of the Iberian Peninsula and from France, corresponding to *O. c. cuniculus* (NE), 14 from the southwestern region of the Iberian Peninsula, corresponding to *O. c. algirus* (SW), and 9 from the contact zone (CZ) as defined by mtDNA variation (Branco et al. 2000). The geographic locations of the populations sampled are shown in Figure 1, and collecting localities are given in Table 1. Additionally, one male *Lepus granatensis* was used as an outgroup.

PCR amplification and sequencing: Genomic DNA was extracted from blood, muscle, or liver following Sambrook and Russell (2001). Introns of four X-linked loci were PCR amplified; two are located near the centromere and two are near the telomeres (Figure 2). Amplification and sequencing primers are listed in supplemental Table 1 at http://www.genetics.org/supplemental/. For each locus, two pairs of amplification primers were designed. The first was based either on published rabbit sequences (*Phka2*) (Davidson et al. 1992) or on conserved exonic regions among human, mouse, and

TABLE 1

Individuals sampled and their geographic locations

Population	Sample size	Population no.[a]	Group	Individual ID
Versailles	1	1	NE	Ver1827
Vaulx-en-Velin	1	2	NE	Vau1
Carlucet	1	3	NE	Cau19
Perpignan	1	4	NE	Pep18
Zaragoza	2	5	NE	Zrg16, Zrg20
Castelló	2	6	NE	Rsl4, Rsl10
Benavente	1	7	NE	Bnv3
Zamora	2	8	NE	Zam1, Zam20
La Rioja	2	9	NE	Lrj3, Lrj6
Madrid	1	10	NE	Mdr7
Alicante	3	11	NE	Alic1, Alt107, Alt120
Cartagena	1	12	NE	Cat12
Cuenca	2	13	NE	Cue1, Cue3
Galicia	1	14	CZ	Gal25c3
Bragança	2	15	CZ	Bra1, Bra13
Toledo	3	16	CZ	Tol25, Tol50, Tol64
Ciudad Real	2	17	CZ	Cre1, Vdm12
Las Amoladeras	1	18	CZ	Amo2
Córdoba	3	19	SW	Luc4, Luc9, Luc17
Sevilla	3	20	SW	Pfr1, Pfr5, Pfr7
Doñana	1	21	SW	Don6
Vila Real	3	22	SW	Vrl1, Vrl4, Vrl7
Idanha-a-nova	1	23	SW	Id85
Elvas	2	24	SW	Elv3, Elv6
Vila Viçosa	1	25	SW	Vv1_1/94

[a] Population numbers are from Figure 1.

rat. Nested primers were then designed specifically for the rabbit on the basis of the first sequences obtained. Amplifications were carried out in 50-µl volumes using Platinum *Taq* High Fidelity DNA Polymerase (Invitrogen, San Diego) following manufacturer recommendations. Cycling temperatures were as follows: an initial denaturation step at 94° for 1 min and 20 sec followed by 35 cycles of 94° for 20 sec, annealing for 20 sec, and extension at 68° for 4 min. Annealing temperatures for each PCR are specified in supplemental Table 1 at http://www.genetics.org/supplemental/. PCR products were purified using the QIAquick PCR purification kit (QIAGEN, Chatsworth, CA) prior to sequencing. Sequencing was carried out using an ABI 3700 automated sequencer. All sequences have been deposited in GenBank under accession nos. DQ306315–DQ306490.

Data analyses: Sequences were inspected and concatenated using the computer program Sequencher (Gene Codes, Ann Arbor, MI) and then aligned manually using the BioEdit software (HALL 1999). By sequencing the X chromosome in males we were able to recover haplotypes directly. The analyses below were based on single nucleotide polymorphisms in introns only.

Basic population genetic parameters, including the number of segregating sites, number of haplotypes, levels of nucleotide diversity, π (NEI and LI 1979), and the proportion of segregating sites, θ (WATTERSON 1975), were estimated using the program DnaSP 4.00 (ROZAS *et al.* 2003) for the entire data set and also for the NE, CZ, and SW groups (Figure 1). Phylogenetic relationships among alleles were estimated using the median-joining algorithm (BANDELT *et al.* 1999) as imple-

mented in Network v4.1.0.8 (http://www.fluxus-technology. com/).

We estimated divergence three ways. First, divergence between all *O. cuniculus* alleles and *L. granatensis* was calculated as the average pairwise distance per nucleotide site, D_{xy} (NEI 1987), and as the number of net nucleotide substitutions per site, D_a (NEI 1987). D_a is defined as $D_{xy} - 0.5 (D_x + D_y)$, where D_{xy} is the average pairwise distance between groups and D_x and D_y are the average pairwise distances within groups. Second, D_{xy} and D_a were calculated between the NE and SW groups of *O. cuniculus*. Finally, to estimate the divergence time of the two subspecies of *O. cuniculus*, maximum-likelihood net nucleotide distances between *L. granatensis* and *O. cuniculus*, and between the two main lineages found in *O. cuniculus* (see RESULTS), were calculated using PAUP v 4.0 (SWOFFORD 2002). Divergence time between subspecies of *O. cuniculus* was calculated assuming a divergence time between *L. granatensis* and *O. cuniculus* of 11.8 million years (MY) (MATTHEE *et al.* 2004).

The population recombination parameter, R ($R = 3Nc$ for X-linked loci, where c is the recombination rate per generation and N is the population size) between adjacent sites (HUDSON 1987), the minimum number of recombination events, R_m (HUDSON and KAPLAN 1985), and the number of pairs of sites showing four gametic types were calculated using DnaSP 4.00 (ROZAS *et al.* 2003). Another estimator of the population recombination parameter, γ (HEY and WAKELEY 1997), was calculated using the software SITES. While Hudson's R is based on the variance of the number of base-pair differences between DNA sequences, γ is a maximum-likelihood estimator developed using a coalescent model for a sample of four DNA
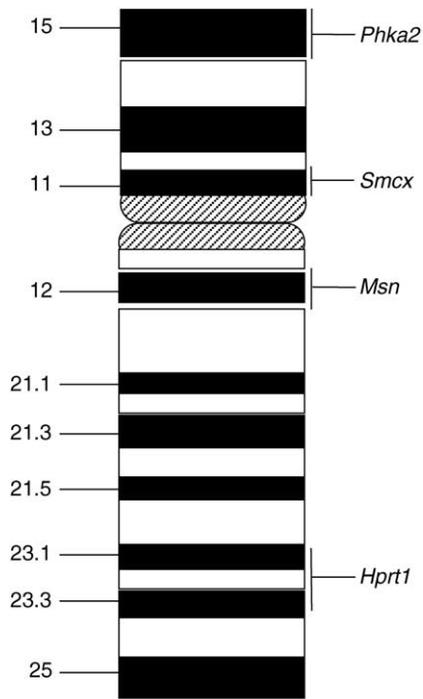
FIGURE 2.—Chromosomal location of the four X-linked loci used in this study. Modified from CHANTRY-DARMON *et al.* (2003) and HAYES *et al.* (2002).

sequences with recombination. Linkage disequilibrium (LD) between pairs of polymorphic sites present at a frequency of at least 10% was calculated within and between all loci, using the statistics D' (LEWONTIN 1964) and r^2 (HILL and ROBERTSON 1968) as implemented in DnaSP 4.00 (ROZAS *et al.* 2003).

Tajima's D (TAJIMA 1989) and Fu and Li's D (FU and LI 1993) were calculated to test for deviations from a neutral equilibrium frequency distribution using DnaSP 4.00 (ROZAS *et al.* 2003). Ratios of polymorphism within *O. cuniculus* to divergence between *O. cuniculus* and *L. granatensis* were compared with the expectations under a neutral model using the Hudson–Kreitman–Aguadé (HKA) test (HUDSON *et al.* 1987). We performed one four-locus test and six pairwise comparisons between loci using the HKA software (HEY and KLIMAN 1993).

At each of the four loci we detected a deep split in the genealogy (see RESULTS). We asked if the observed pattern of nucleotide polymorphism is compatible with a single panmitic population, as opposed to some form of population subdivision. If two populations have evolved in allopatry, the basal branch of a gene genealogy may be longer than in a panmitic population. Furthermore, mutations arising in an isolated subpopulation are unable to recombine with mutations in a different subpopulation, resulting in higher levels of LD. WALL (2000) suggested two measures based on LD that could be powerful indicators of population subdivision. The first, l_b, is the number of congruent sites, defined as the number of mutations that, on a pairwise basis, result in only two haplotypes. The second, g_d, is the maximum physical distance between congruent sites. Coalescent simulations of panmixia were performed with the computer program ms (HUDSON 2002). For each locus, 50,000 genealogies of 43 individuals were simulated conditioned on the estimated values of θ and γ. Additionally, for each locus, coalescent simulations were performed using two different values of the population recombination parameter ($3Nc = 0.0015$ and $3Nc = 0.015$ per site),

chosen to reflect the range of recombination rates known for other mammals (*e.g.*, DIETRICH *et al.* 1994; KONG *et al.* 2002; JENSEN-SEAMAN *et al.* 2004). A computer program (GARRIGAN *et al.* 2005) was used to calculate l_b and g_d from the simulated data sets, and the distributions of the two statistics for each set of conditions were plotted against each other. The probability of obtaining the observed values of l_b and g_d was calculated as the proportion of simulated genealogies for which the values of l_d and g_d were greater than the observed values.

F_{ST} and Nm were calculated using the method of HUDSON *et al.* (1992a) implemented in DnaSP 4.00 (ROZAS *et al.* 2003). Genetic differentiation was also calculated using the test statistic K_S^* (HUDSON *et al.* 1992b), and significance was assessed by performing 1000 permutations. To test for significant population structure among populations and among groups of populations, analyses of molecular variance (AMOVA; EXCOFFIER *et al.* 1992) between the SW and NE groups were performed using ARLEQUIN (SCHNEIDER *et al.* 2000).

One simple model of divergence is an isolation model in which two populations become separated with no subsequent gene exchange. The HKA model (HUDSON *et al.* 1987) takes this form and further assumes that the ancestral species has a population size that is the average of the two descendant species. More recent models relax this assumption. For example, WAKELEY and HEY (1997) proposed a model that is similar to the HKA model but includes an additional parameter, θ_A, the population mutation parameter for the ancestral species. While the HKA test uses only the number of polymorphic sites and divergence, this model also incorporates the total number of polymorphic positions in the two groups (S), the number of polymorphisms exclusive to one group ($S_{x_{NE}}$ and $S_{x_{SW}}$), the number of shared polymorphisms (S_s), and the number of fixed differences (S_f). We tested the fit of our data to these two models in two different ways. First, we performed pairwise comparisons among all loci, and second, we performed tests with all four loci together. The fit of our data to the Wakeley and Hey model of divergence without gene flow was tested using the program WH (WANG *et al.* 1997).

These models assume that there has been no gene flow between the two populations since the initial split. In many cases this is an unrealistic assumption. HEY and NIELSEN (2004) developed a model of population divergence that allows for genetic drift (increasing population divergence) and gene flow (preventing population divergence) to act together, which they call the isolation with migration model. The computer program IM is an implementation of the Markov chain Monte Carlo method for the analyses of genetic data under this model. We used IM to estimate the effective population size of *O. c. cuniculus* and of *O. c. algirus* and to estimate migration rates for each locus between subspecies in each direction. IM assumes that there is no recombination within loci. For *Phka2* and *Hprt1*, we used the largest region showing no evidence of recombination, following WON and HEY (2005). For *Phka2*, a portion of 687 bp containing 23 polymorphic sites was used, and for *Hprt1*, a region of 501 bp with 20 polymorphic sites was used. For *Smcx*, all 20 NE and 14 SW individuals were used since the data are free of recombination. For *Msn*, we removed three recombinant individuals (Vau1, Rsl4, and Rsl10) from the NE group. We assigned wide prior distributions of the parameters on the basis of preliminary trial runs. We ran the program under Metropolis Coupled Monte Carlo Markov Chains, using 10 chains with linear heating. We used a burn-in period of 1,000,000 steps and recorded results every 40 steps. To test whether the chains were mixing well, we ran the program with different random seed numbers and the results were similar. We ran the program for 25,625,601 steps after the burn-in period and recorded the results of 625,014 steps.

a

```
                                                              1111111111111111111111111111111111111111222222222222222222222222222222222222222222222222222222222222222233333
                                  1111222223444444555666777777788888999999990001111222333444445555556666777788888888889999900011111112222233333344444445555555566666777777788889999900011
                              127779022501178801556836639977789992357703335677811723552445682240348122723412347888133486990126689901148024589347788812224469146772470781145027790 2704
                              6179990864198338017247063366736730796824672349114307668296788782496533312784846205902361472203857068188023953826419170479289579376635920479590 72707934732
                                                                         *                                           *           *
L. granatensis  ATCAGGCTGTGCAAGTGCATTCAGAGAATGCGGTAGGCTGTCCAGCGGGGTACAAGCGTCCTCCGGGGGCGACGAGGTTCCGCGACTGTGCGTTAAGGGGGCGGCGCTCGGGACTACCCGTGCAGAAGCCCACAGCTCAAGCGGCGGCGGG

Ver1827  (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Vau1     (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Cau19    (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Pep18    (P02)  ....C.....A....C.......A..C....TA....TGG.....G..C..T..T......T........T..T..AA.......T......C......C.......C..........T..CA.C.G..........GGA......CA.
Zrg16    (P03)  ..T.CA....A.......G..T...A..CA....TA....TGG.....A.G.C..G.T......AT.T..AA.......T...C...C.....C..CAA........C..C.G..........A......C..
Zrg20    (P04)  ....CA....A.......G..T...A..CA....TA....TGG.T.A.A.G.C..G.T......AT.T..AA.......T...C.G....C..A.............C.TC.G.....GA.......C..
Rs14     (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Rs110    (P05)  .C..C.....A.......T...A..CA....TA....TGG...A.A.G.C.A..G.T...A...AT.T..AA.......T...A..C.G....C..A............C.TC.G.....C..A....G.......C..
Bnv3     (P06)  ..T.CA....A....C...T...A..CA....TA....TGG...A.A.G.C..G.T......AT.T..AA...T.T....CC....C..A............A.....C..C.G..........T.A.......CA.
Zam1     (P07)  .......C............GGC....C...G.....A...A.....G.............TTAG.........TC........GA..AC........G.T.C..T.G....T.......GA.....A..C..
Zam20    (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Lrj3     (P08)  ...GC....A.......A..C....A..TA....TGG.....A.G.C......T........T..TA.AA.......T......C.....C.....C.......G....C..C.G.......GA..A....TC.
Lrj6     (P09)  ..T.C.....AT...C...T...A..CA....TA....TGG...A.A.GGC.....G.T.T....AT.T..AA.......T......C......C...C.........C..C.G.......GA..A....C..
Mdr7     (P10)  ...GC....A.......A..C....A..TA....TGG.....A.G.C......T........T..TA.AA.......T...C....AA.C.......C.....G....C..C.G....T....GA..A....TC.
Alt1     (P11)  ..T.CAT...A.......T...A..CA..A.TA....ACTGG....A.ACG.C..G.T......AT.T..AA.......T...C...C....CGAA..T..........C..C.G.......GA......C..
Alt107   (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Alt120   (P12)  GC..C.....A..G........A..C....TA....TGG.....G..C..T..T......T........T..T..AA....A..T....C......C.......G........T..CA.C.G.....GGA......CA.
Cat12    (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Cue1     (P13)  .C..C....CA.G.......T...A..CA....TA....TGG...A.A.G.C..G.T......AT.T..AA.......T...C...C.....T..AT.T............C..C.G.......GA...T...C.A
Cue3     (P14)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...A.A..C....A..C.....C....A..G....C..C.G..T.....GA..A....C..

Gal25C3  (P15)  .......C............GGC.TT..C...G.....A...A.....G....T.......A...TTAG.........TC........GA...C........T..G.T.C..T.G.........GA........C..
Bra1     (P16)  ....C.....A....C.......A..C....TA....TGG.....A.G.C..G.T......AT.T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
Bra13    (P01)  ....C.....A....C.......A..C....TA....TGG.....A.G.C......T........T..T..AA.......T......C...A..C.....C.......G....C..CAG......GA..A....C..
To125    (P06)  ..T.CA....A....C...T...A..CA....TA....TGG...A.A.G.C..G.T......AT.T..AA...T.T....CC....C..A............A.....C..C.G..........T.A.......CA.
To150    (P17)  ..T.CA....A..AC...T...A..CA....TA....TGG...A.A.G.C..G.T......AT.T..AA....AT....CC....C..A.A.........T.AC..C.G..........GA........C..
To164    (P18)  .......C............GGC....C...G.....A...A.....G.............TTAG.........TC........GA..AC........G.T.C..T.G....T.......GA.....C..
Cre1     (P13)  .C..C....CA.G.......T...A..CA....TA....TGG...A.A.G.C..G.T......AT.T..AA.......T...C...C.....T..AT.T............C..C.G.......GA...T...C.A
Vdm12    (P19)  ..T.CA....A....C...T...A..C....TAA....TGG...A.A.G.C..G.T.T....AT.T..AA.......T...C....CC....C..A............C..C.G.......C...GA....C..
Amo2     (P20)  ..T.CA....A....T...A..CA..A.TA....ACTGG....A.ACG.C..G.T......AT.T..AA...T..T....C....C....C..AA..T............C..C.G.......GA......C..

Luc4     (P21)  ....C.....A.......C....A........TA....TGG.....A.G.C.......TC......AT.T..AA.......T......C......C.......A..T.......C..C.G.......GA.A....C..
Luc9     (P22)  .C..C.....A.....T...TA.A..C....C..T...TGG.....G.C..............T.AG.........TC........C....T...A....G.T.C..T.G.T..T.......G.T......T..
Luc17    (P23)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...C...A..C.....C.......G....C..C.G.......GA..A....C..
Pfr1     (P24)  ....CA....A....C...T...A..CA....TA....TGG..AA.A.G.C..G.T......AT.T..AA.......T......C......C.......G....CA.C.G.......C...T..GGAT......C..
Pfr5     (P25)  .C..C.....A.......T...A..CA....T......TGG...A.A...C.A.AG.T..T.A...T.TA.AA.C.......T......C......C.............C..C.GT.T.....C..GA......C..
Pfr7     (P26)  ..T.CA....AT...C...T...A..CA....TA....TGG...A.A.G.C..G.T.T....AT.T..AA.......T...C....CC....C..A............C..C.G.......C...GA....C..
Don6     (P25)  .C..C.....A.......T...A..CA....T......TGG...A.A...C.A.AG.T..T.A...T.TA.AA.C.......T......C......C.............C..C.GT.T.....C..GA......C..
Vrl1     (P27)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...T.C...A..C.....C.......G....C..C.G.......GA..A....C..
Vrl4     (P27)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...T.C...A..C.....C.......G....C..C.G.......GA..A....C..
Vrl7     (P18)  .......C............GGC....C...G.....A...A.....G.............TTAG.........TC........GA..AC........G.T.C..T.G....T.......GA........C..
Id85     (P28)  .C..C.....A.......A..C....GTA....TGG.....A.G.C......T......AT.TA..A..T....T...C...A..C.....C.......G....C..C.G.......GA..A....C..
Elv3     (P27)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...T.C...A..C.....C.......G....C..C.G.......GA..A....C..
Elv6     (P27)  ....C.....A.......A..C....TAA....TGG.....A.G.C......T........T..TA.AA.......T...T.C...A..C.....C.......G....C..C.G.......GA..A....C..
VV1_1/94 (P29)  ....C..A.A.........GA..C....C..T...TGG.......C.....T....A..TA.T...AC......T....AC.........C..........T..........C.G.......GA.....A.C..
```

FIGURE 3.—Polymorphisms for the four X-linked loci. (a) *Phka2*, (b) *Smcx*, (c) *Msn*, and (d) *Hprt1*. The nucleotide found in the *L. granatensis* sample was used to infer the ancestral state for each position. Numbers in the columns indicate the position of the polymorphism in the alignment. Each row represents one individual. A dot represents the ancestral state, and letters (A, C, G, and T) represent the derived nucleotide. The name of the haplotype defined by each sequence is shown in parentheses. Positions polymorphic for bases other than the one found in *L. granatensis* are indicated with an asterisk (*) and positions segregating for three different bases in *O. cuniculus* are indicated by "#" in d. Whenever a polymorphic site in *O. cuniculus* corresponded to a deletion in *L. granatensis*, we treated the most frequent base in our sample of *O. cuniculus* as the ancestral state.

RESULTS

Levels and patterns of variation: We observed considerable variation at all four genes. Polymorphic sites for each gene are shown in Figure 3, and summaries of variation are given in Table 2. The number of polymorphic sites varied from 56 at *Msn* to 151 at *Phka2*, but the number of haplotypes observed at each gene was much more constant (from 23 at *Smcx* to 29 at *Phka2*). Levels of polymorphism were high, both in the total sample and within each subspecies. However, nucleotide diversity (π) at the two centromeric loci was considerably lower (0.52% at *Smcx* and 0.55% at *Msn*) than at the two telomeric loci (0.70% at *Phka2* and 1.26% at *Hprt1*). This contrast was even more striking in the proportion of segregating sites, where θ at the centromeric loci was roughly half the value seen at the telomeric loci.

We assessed the amount of LD in our sample in several ways, and all were consistent in revealing more recombination (less LD) at the telomeric loci than at the centromeric loci (Table 2), consistent with suppression of recombination near the centromere. The num-

ber of pairs of sites showing all four gametic types was zero at *Smcx*, 77 at *Msn*, 390 at *Phka2*, and 180 at *Hprt1*. R_m, the minimum number of recombination events in the history of the sample (HUDSON and KAPLAN 1985), was zero at *Smcx*, intermediate at *Msn* and *Hprt1* (7 and 6, respectively), and highest at *Phka2* (17). Similarly, R (HUDSON 1987) between adjacent sites was low at the two centromeric loci (*Smcx* and *Msn*) and much higher at the telomeric loci (*Phka2* and *Hprt1*). The values for γ (HEY and WAKELEY 1997), a maximum-likelihood estimator of the population recombination parameter, are concordant with R in showing more recombination at the telomeric loci than at the centromeric loci. Thus, although there are small differences among estimators, levels of recombination are higher at the telomeric loci than at the centromeric loci. We assessed the significance of LD through pairwise comparisons of polymorphic sites present at a frequency of at least 10% using a Fisher's exact test. From the 8515 pairwise comparisons performed, 2497 were significant in a two-tailed test, and 980 remained significant after a Bonferroni correction for multiple tests. Of these, 197 were between polymorphic sites at *Smcx*, 219 between sites at *Msn*, 93

b

```
                                    1111111111111111111111111111222222222222
                    12222444456666669990000001111223333444555666678901224455666
                    8725780236303579926790444781111234789458067479926560745175777
                    791227657304237563981314743136867116226267558886109309012445
                                                  *
L. granatensis  GTAGTGAGGAAGTGGTATCTAGGCCCGCGAGAGAGGCGCGGTACAGCAGACCGCCTTTCA

Ver1827  (S01) ...A..G.A...C..G.C.CG.AT..A.......C..C.A.CG.....A..........
Vau1     (S02) ...A..G.A...C..G.C.CG.AT..A.......C.....CG.....A..........
Cau19    (S03) ...A..G.A...C..G.C.CG.AT..A.......C..C...CG.....A..........
Pep18    (S02) ...A..G.A...C..G.C.CG.AT..A.......C.....CG.....A..........
Zrg16    (S04) ...A..G....C..G.C.CG.AT..A.C...C................G..
Zrg20    (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....
Rs14     (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Rsl10    (S07) ...A..GA....C..G.C.CG.AT..A.......C........................
Bnv3     (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Zam1     (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Zam20    (S08) ...A..G..G..C..G.C.CG.AT..A.......C........................
Lrj3     (S09) ...A..G....C..G.CTCG.AT..A.......C.............G..........
Lrj6     (S10) ...A..GA....C..G.C.CG.AT..A..TA...C...A...................
Mdr7     (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Alic1    (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Alt107   (S02) ...A..G.A...C..G.C.CG.AT..A.......C.....CG.....A..........
Alt120   (S11) ...A..G....C..G.C.CGAAT..A.......C........................
Cat12    (S03) ...A..G.A...C..G.C.CG.AT..A.......C..C...CG.....A..........
Cue1     (S12) ..C..AG.....TCT.A........T..T....T..............AG.G....TCC....
Cue3     (S13) ..C..A.....TCTAA.........T...T.........T.A...G....TC....

Gal25c3  (S14) A.C..A.....TCT.A.........T...T..........A...G.....TC....
Bra1     (S06) ...A..G....C..G.C.CG.AT..A.......C........................
Bra13    (S06) ...A..G....C..G.C.CG.AT..A.......C........................
To125    (S15) ..C..A.....T.T.AG.........T...T..........A...G....TC....
To150    (S16) ..C..A.....TCTAA.........T...T..........A...G....TC....
To164    (S17) ..C..AG.....TCT.A.........T...T.........C..G.G....TCC....
Cre1     (S15) ..C..A.....T.T.AG.........T...T..........A...G....TC....
Vdm12    (S15) ..C..A.....T.T.AG.........T...T..........A...G....TC....
Amo2     (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....

Luc4     (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....
Luc9     (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....
Luc17    (S18) ..C..AG.....TCT.A.........T...T..........A...G....TC....
Pfr1     (S18) ..C..AG.....TCT.A.........T...T..........A...G....TC....
Pfr5     (S16) ..C..A.....TCTAA.........T...T..........A...G....TC....
Pfr7     (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....
Don6     (S19) ..C..A.....TCT.A.........T...GT...T.......A...G.G..TC....
Vrl1     (S20) ..C..A.....TCTAA.........T...T..........A...G....TTC....
Vrl4     (S16) ..C..A.....TCTAA.........T...T..........A...G....TC....
Vrl7     (S21) .AC..AG.....TCT.A.........T...T..........A...G..A.TC....
Id85     (S18) ..C..AG.....TCT.A.........T...T..........A...G....TC....
Elv3     (S22) ..C..A....GTCT.A.........T...GTT..........A...G.G..TC..AG
Elv6     (S05) ..C..A.....TCT.A.........T...GT..........A...G.G..TC....
VV1_1/94 (S23) ..C.AAG....TCT.A.......TT.T...T..A.........AG.G....TCC...
```

c

```
                                    1111111111111111111111111122222222222222
                    11234456667788990001223444556677777888991233345556 7788
                    6722806700256846345671487269395656677899093334450577 0400
                    42461247745510585032319375099132613195582964905593460557
L. granatensis  TCCGCACCCGGCAGTTCCGGCTCTTGCTAGTCGCATTTTTACTAGGACGGTACGGC

Ver1827  (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Vau1     (M02) C...G..T.A.TTA...T.A..T....TC.AC...GC.C.C....CT....C..A..
Cau19    (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Pep18    (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Zrg16    (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Zrg20    (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Rs14     (M03) C...G..T.A...A...T....T...C..C..TGC.C.C..CCT....C..A.A
Rsl10    (M04) C...G..A...A...T....T...C..C..TGC.C.C..CCT....C..A.A
Bnv3     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Zam1     (M05) C...G..A...A...T....C..C...C.C.C....CT.T..C..A..
Zam20    (M06) C...G..T.A...A...T....T...TC.AC...GC.CCC....CT....C..A..
Lrj3     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Lrj6     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Mdr7     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Alic1    (M07) C...G..TGA...A...T....T...C..C...GC.C.C....CT....C..A..
Alt107   (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Alt120   (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Cat12    (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Cue1     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Cue3     (M08) C...G..T.AC..A...T....T...TC.AC...GC.C.C....CT....C..A..

Gal25c3  (M09) .T.A.G.........CA....G........C...G.C.....CC.....T.....
Bra1     (M01) C...G..T.A...A...T....T...TC.AC...GC.C.C....CT....C..A..
Bra13    (M10) .T.A.G.........CA....G...G.C..G.C.....CC...T.....
To125    (M11) .T.A.G.........CA...G.......C...G.C.....C.......
To150    (M12) .....G.........CA...G.A....C...G.C.....C.......
To164    (M09) .T.A.G.........CA....G........C...G.C.....CC.....T.....
Cre1     (M13) .....G....T...CA..A.G........G..G.C.........
Vdm12    (M14) .....G.........CA...G........C...G.....T........
Amo2     (M15) C.T.G..T.A...A...T....T...C..C..TGC.C.C....CT....C..A..

Luc4     (M16) .....G....T...CA..A.G..C.....G..G.C.........
Luc9     (M16) .....G....T...CA..A.G..C.....G..G.C.........
Luc17    (M12) .....G.........CA...G.A....C...G.C.....C.......
Pfr1     (M17) .....GG...T...CA..A.G........G..G.C.........
Pfr5     (M18) ...A.G.........CA...G.....C.A.G.C....T.C....C.......
Pfr7     (M19) .....G.........CA...G..A....C...G.C........G....G..A.
Don6     (M20) .....G.......ACA...G........C...G.C.....C.........
Vrl1     (M21) .T.A.G.........CA...G........C...G.C.....CC..G..T.....
Vrl4     (M22) .....G....T...CA.AA.G........G..G.C.........G......T...
Vrl7     (M21) .T.A.G.........CA...G........C...G.C.....CC..G..T.....
Id85     (M23) .....G....T...CA..A.G........G..G.C..............T...
Elv3     (M21) .T.A.G.........CA...G........C...G.C.....CC..G..T.....
Elv6     (M24) .....G.........CA...G........C...G.C.....CC..G..T.....
VV1_1/94 (M25) .T.A.G.........CA...TG........C...G.C.....CC..G..T.....
```

d

```
                                              111111111111111
                    11122222223333444455567777777777777788888888888888999990000012223333344
                    15734703488904490224374001122334458000133345889902355146712557445601
                    75321245157570474459160928372819071458189011906181124062905891606 21
                                                  #  *    #   #
L. granatensis  GAACTCGGATAGCCTTCATCCCTCCAGGCGCTTGGAACCGAGCCCTGACGGTTCAAACAAGGAAAGGG

Ver1827  (H01) ......A.........T.G...AA......CCAT..A.A..TA....T.....GG....A...G...
Vau1     (H01) ......A.........T.G...AA......CCAT..A.A..TA....T.....GG....A...G...
Cau19    (H02) ......A.........T.....AA......CCAT..A.A..TA....T.....GG....A...G...
Pep18    (H03) ........G............A......CCAT..A.A..TA....GT....G..GT.....A..
Zrg16    (H04) ......A..G....T.GT......A.T..CG..C.......TC.G..TGG.GG..GT.....AAA
Zrg20    (H05) ......A.........G........A......CCAT..A.A..GTA....G...GT....A..
Rs14     (H06) ........G.........A..A..TCCAT..A.A..TA....GT.....G..GT.....A..
Rsl10    (H06) ........G.........A..A..TCCAT..A.A..TA....GT.....G..GT.....A..
Bnv3     (H07) ........G............A......CCAT..A.A..TA....GT....G..GT....G.A.
Zam1     (H08) ......A..G......T..T......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Zam20    (H04) ......A..G....T.GT......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Lrj3     (H09) ......A.C.A..........AA......CCAT..A.A..TA....GT....TGG....A..G...
Lrj6     (H04) ......A..G....T.GT......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Mdr7     (H10) ........G..........C.A......CCAT..A.A..TA....GT.....G..GT.....A..
Alic1    (H11) ......A.C..A..GT......AA......A.CCAT..A.A..TA....GT....G......A..G...
Alt107   (H12) ......A.C.............AA......A.CCAT..A.A..TA....GT..G..T..A...G...
Alt120   (H01) ......A.........T.G...AA......CCAT..A.A..TA....T.....GG....A...G...
Cat12    (H13) ......A.C............A......CCAT..A.A..TA....GT....G..GT.....A..
Cue1     (H04) ......A..G....T.GT......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Cue3     (H09) ......A.C.A..........AA......CCAT..A.A..TA....GT....TGG....A..G...

Gal25c3  (H14) ......A.C............AA......CCAT..A.A..TA....GTA...G......A..G...
Bra1     (H04) ......A..G....T.GT......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Bra13    (H15) ...T..A.......T..TTG......A.T..CG.T....T.....TC.G...GG.GG......A..G...
To125    (H16) ...T..A.......T.G..T..GA.T...G.T...T......TCAG...GG.GG..GT.....A..
To150    (H17) ........G............A......CCAT..A.A..TA....GT......G..T.A..A..
To164    (H18) ........GC.....T.G....A......CCAT..A.AG.TA..GT......G..GT.....A..
Cre1     (H19) ......A.......T.G.T...GA.T...G.T.......TC.G..GG.GG..GT.....A..
Vdm12    (H19) ......A.......T.G.T...GA.T...G.T.......TC.G..GG.GG..GT.....A..
Amo2     (H20) ......C.....G....T.G.......A.G.TC.......TC.G..GG.GG..GT.....A.A

Luc4     (H21) A.C.............T.G....A.A.T..CG.T....T.A..TC.G..GG.GG......A...
Luc9     (H22) ........G..........AA....A.CCAT..A..TA....GT....G....A..G...
Luc17    (H04) ......A..G....T.GT......A.T..CG..C.......TC.G.TGG.GG..GT.....AAA
Pfr1     (H23) .G.....G.............A......CCAT..A.A..TA....GT....G..GT..T..A..
Pfr5     (H05) ......A.........G........A......CCAT..A.A..TA....GTA....G.....A..
Pfr7     (H24) ...T..A.......T..TTG......A.T..CG.T.T..T....TC.G..GG.GG....A..G...
Don6     (H25) ...T..A.C.A.............A......CCAT..A.A..TA....GT....GGG...A..G...
Vrl1     (H14) ......A.C............AA....C.CCAT..A.A..TA....GTA...G......A..G...
Vrl4     (H15) ...T..A.......T..TTG......A.T..CG.T....T.....TC.G...GG.GG......A..G...
Vrl7     (H14) ......A.C............AA....C.CCAT..A.A..TA....GTA...G......A..G...
Id85     (H26) ............T..TTG......A.T..CG.T....T.....TC.G..GG.GG......A..
Elv3     (H27) ........G....A.......A......CCAT..A.A..TA....GT......G..GT.....A..
Elv6     (H28) .....T.......T.TTG......A.T..CG.T.T..T....AC.G..GG.GG......A..G...
VV1_1/94 (H04) ......A..G....T.GT......A.TCGCC..GCA.CGAG...CGGC.GG.GG..GT.....AAA
```

FIGURE 3.—Continued.

between sites at *Phka2*, and 130 between sites at *Hprt1*. Interlocus LD was detected only between *Smcx* and *Msn* where there were 341 pairs of sites that showed significant LD.

The distribution of allele frequencies as measured by Tajima's D and Fu and Li's D generally conformed to expectations under a neutral model of molecular evolution (Table 2). For example, in the total sample, Tajima's D was positive (*Msn* and *Hprt1*), very close to zero (*Smcx*), or negative (*Phka2*), but not significantly different from zero. When the population groups were analyzed separately, Tajima's D was negative (except for *Hprt1* in the NE and SW groups), but not significantly so ($P > 0.05$ for all tests). We also tested a neutral model of molecular evolution by comparing ratios of polymorphism within *O. cuniculus* (in the total sample) to divergence between *O. cuniculus* and *L. granatensis*, using the HKA test. We performed one four-locus comparison as well as six pairwise comparisons among loci, and none of these were significant ($P > 0.05$ for each).

Divergence and gene flow between subspecies: F_{ST} estimates between *O. c. cuniculus* and *O. c. algirus* are shown in Table 3. F_{ST} was very high at the centromeric loci, *Smcx* (0.680) and *Msn* (0.829), and one order of magnitude lower at the telomeric loci, *Phka2* (0.027) and *Hprt1* (0.022). The two subspecies were significantly differentiated, using the K_s^* test statistic, at all loci but *Hprt1* (Table 3). Similarly, the AMOVA analyses revealed

TABLE 2

Levels of polymorphism, allele-frequency spectrum tests of neutrality, divergence, and recombination

Locus	Sample	n	L^a	K^b	S^c	π (%)	θ (%)	Tajima's D	Fu and Li's D	D_a^d	D_{xy}^e	D_a^f	D_{xy}^g	γ^h	R^i	R_m^j	$4gt^k$
				Polymorphism				Frequency spectrum tests of neutrality		Between NE and SW		Between *O. cuniculus* and *L. granatensis*		Recombination			
Phka2	All	43	3168	29	151	0.699	1.102	−1.337	−0.995	0.008	0.659	5.719	6.059	0.0051	0.0025	17	390
	NE	20	3168	14	105	0.574	0.934	−1.584	−2.106					0.0019	0.0018	10	95
	CZ	9	3168	9	86	0.924	0.999	−0.385	0.121					0.0052	0.0028	3	25
	SW	14	3168	10	95	0.708	0.945	−1.109	−1.109					0.0074	0.0032	10	137
Smcx	All	43	2709	23	60	0.517	0.512	0.035	−1.499	0.531	0.782	1.475	1.735	0	0.0012	0	0
	NE	20	2709	13	44	0.341	0.458	−1.021	−0.647					0	0	0	0
	CZ	9	2709	6	29	0.390	0.394	−0.054	0.667					0	0.0003	0	0
	SW	14	2709	8	21	0.160	0.244	−1.438	−2.172					0	0.0020	0	0
Msn	All	43	2825	25	56	0.553	0.458	0.733	−0.849	0.786	0.949	3.803	4.083	0.0023	0.0011	7	77
	NE	20	2825	8	14	0.079	0.140	−1.591	−1.390					0.0003	0.0002	1	3
	CZ	9	2825	8	36	0.462	0.469	−0.073	0.423					0	0.0005	0	0
	SW	14	2825	11	23	0.246	0.256	−0.161	−0.649					0.0016	0.0117	2	7
Hprt1	All	43	1473	28	68	1.256	1.115	0.454	−0.381	0.027	1.256	3.848	4.428	0.0088	0.0086	6	180
	NE	20	1473	13	47	1.129	0.899	1.025	0.978					0.0071	0.0034	2	58
	CZ	9	1473	8	49	1.228	1.249	−0.087	0.134					0.0094	0.0109	4	104
	SW	14	1473	12	50	1.328	1.132	0.760	0.642					0.0079	0.0120	4	111

[a] Length of the sequence in base pairs.

[b] No. of haplotypes.

[c] No. of polymorphic sites.

[d] Net nucleotide divergence per site (NEI 1987) between NE and SW (CZ was excluded from this analysis).

[e] Average pairwise nucleotide substitutions per site (NEI 1987) between NE and SW (CZ was excluded from this analysis).

[f] Net nucleotide divergence (NEI 1987) between the haplotype found in *L. granatensis* and all *O. cuniculus* sequences.

[g] Average pairwise nucleotide substitutions per site (NEI 1987) between the haplotype found in *L. granatensis* and all *O. cuniculus* sequences.

[h] HUDSON's (1987) estimator of the population recombination parameter between adjacent sites (HEY and WAKELEY 1997).

[i] HUDSON's (1987) estimator of the population recombination parameter between adjacent sites.

[j] Minimum number of recombination events in the history of the sample (HUDSON and KAPLAN 1985).

[k] No. of the pairs of sites that show all four gametic types.

TABLE 3

Genetic differentiation between NE and SW groups at four X-linked loci

	F_{ST}[a]	Nm[b]	ϕ_{ct}[c]	ϕ_{st}[d]	ϕ_{sc}[e]	D_a (%)[f]
Phka2	0.0266*	12.22	0.02	0.07	0.05	0.008
Smcx	0.6796***	0.16	0.64	0.80	0.43	0.531
Msn	0.8286***	0.07	0.84	0.86	0.11	0.786
Hprt1	0.0218	14.99	0.02	0.07	0.05	0.027

*$P < 0.05$, **$P < 0.01$, ***$P < 0.005$.

[a] F_{ST} was calculated using the method proposed by Hudson *et al.* (1992). Statistical significance for the estimation of F_{ST} between the two groups was obtained with the K_{st}^* statistic (Hudson *et al.* 1992).

[b] Nm was calculated according to Wright's (1951) island model of population structure, using the expression $F_{ST} = 1/(1 + 3Nm)$ for X-linked loci.

[c] ϕ_{ct} is the fixation index for the amount of variation segregating between NE and SW groups, calculated using the AMOVA framework (Excoffier *et al.* 1992). For the NE group, the 13 populations studied were pooled into seven subgroups (NE1: populations 1, 2, 3, and 4; NE2: population 5; NE3: population 6; NE4: populations 7 and 8; NE5: populations 9 and 10; NE6: populations 11 and 12; and NE7: population 13). For the SW group, the 7 populations studied were pooled into four subgroups (SW1: population 19; SW2: populations 20 and 21; SW3: population 22; and SW4: populations 23, 24, and 25).

[d] ϕ_{st} is the fixation index for the amount of variation segregating within each subgroup, calculated using the AMOVA framework (Excoffier *et al.* 1992).

[e] ϕ_{sc} is the fixation index for the amount of variation segregating among subgroups within each group, calculated using the AMOVA framework (Excoffier *et al.* 1992).

[f] D_a is the net nucleotide distance per base pair between populations (Nei 1987).

that at the two centromeric loci, *Smcx* and *Msn*, most of the genetic variation is partitioned among the two subspecies (64 and 84%, respectively) while at the telomeric loci, *Phka2* and *Hprt1*, most of the observed variation was partitioned among populations within each subspecies (93% at both loci), and only a marginal proportion (2%) of the variation segregated between subspecies.

This differentiation can also be seen in the phylogeny of alleles for each gene (Figure 4). At each locus there were two divergent groups of haplotypes, and in each case the root fell along the deep branch separating these two groups. In this analysis, only *Smcx* was free of homoplasy. The locus with the most homoplasy was *Phka2*. This homoplasy may be due to recombination or recurrent mutation. Evidence for recurrent mutation comes from the observation that at *Hprt1* three different positions have three nucleotides segregating (Figure 3d). Other evidence of recurrent mutation is the fact that the amount of homoplasy is slightly reduced if CpG sites, which are known to be hypermutable, are excluded. For example, for *Phka2*, the consistency index (CI) increased from 0.778 to 0.803 when CpG sites were removed. However, much of the homoplasy is probably

due to recombination, as evidenced by the fact that the CI was 1.0 for *Smcx*, 0.889 at *Msn*, but was 0.778 and 0.717 at *Phka2* and *Hprt1*, respectively. Moreover, at *Phka2* and *Hprt1*, 1 and 16 individuals, respectively, were identified as recombinants between the two divergent lineages on the basis of their position on the haplotype network and by visual inspection of the table of polymorphism (Figure 3, a and d).

The degree of introgression between subspecies can also be seen by the concordance (or lack thereof) between geography and phylogeny. For the two centromeric loci (*Smcx* and *Msn*), there was good concordance between phylogeny and geography; *i.e.*, the two major lineages correspond well with each subspecies (Figure 4). At *Msn* we did not detect any introgressed haplotypes and at *Smcx* we observed only three NE individuals with haplotypes from the lineage that is otherwise restricted to the SW and CZ groups. At the two telomeric genes (*Phka2* and *Hprt1*), in contrast, there seems to be little or no concordance between phylogeny and geography. At all four genes, individuals from the CZ group are scattered throughout the haplotype networks.

The proportion of congruent sites, l_b, is greater at the two centromeric loci (representing 30 and 32% of all polymorphic sites at *Smcx* and at *Msn*, respectively) than at the telomeric loci (11% at *Phka2* and 16% at *Hprt1*). Similarly, the maximum distance between congruent sites, g_d, is greater at *Smcx* (85% of the total locus length) and at *Msn* (95%) than at *Phka2* (65%) and *Hprt1* (16%). We calculated the probability of observing these values of l_b and g_d using coalescent simulations of 50,000 genealogies of 43 individuals evolving neutrally under panmixia with mutation (θ) and recombination (γ) parameters estimated from the data. Results are shown in Table 4. Under these conditions, the null model was rejected for *Msn* ($P = 0.00214$). This test is quite conservative using γ estimated from the data since population subdivision will increase LD and thus underestimate the true value of recombination. Therefore, we also conducted simulations with a population size of 10^5 and per-site recombination rates of 0.5×10^{-8} and 5×10^{-8}, reflecting the range of recombination rates seen in other mammals (*e.g.*, Dietrich *et al.* 1994; Kong *et al.* 2002; Jensen-Seaman *et al.* 2004). At the two centromeric loci, *Msn* and *Smcx*, the null model was rejected using either value of recombination. For *Phka2*, the null hypothesis was rejected only with the higher recombination rate, and *Hprt1* was marginally significant ($P = 0.066$) only for the higher recombination rate (Table 4).

Another way of looking at divergence is to quantify the amount of shared and fixed variation between the two groups (Table 5). The number of shared polymorphisms was low at the centromeric loci (16 and 6% of all polymorphisms at *Smcx* and *Msn*, respectively) and high at the telomeric loci (42 and 64% at *Phka2* and *Hprt1*, respectively). Only *Msn* showed fixed differences
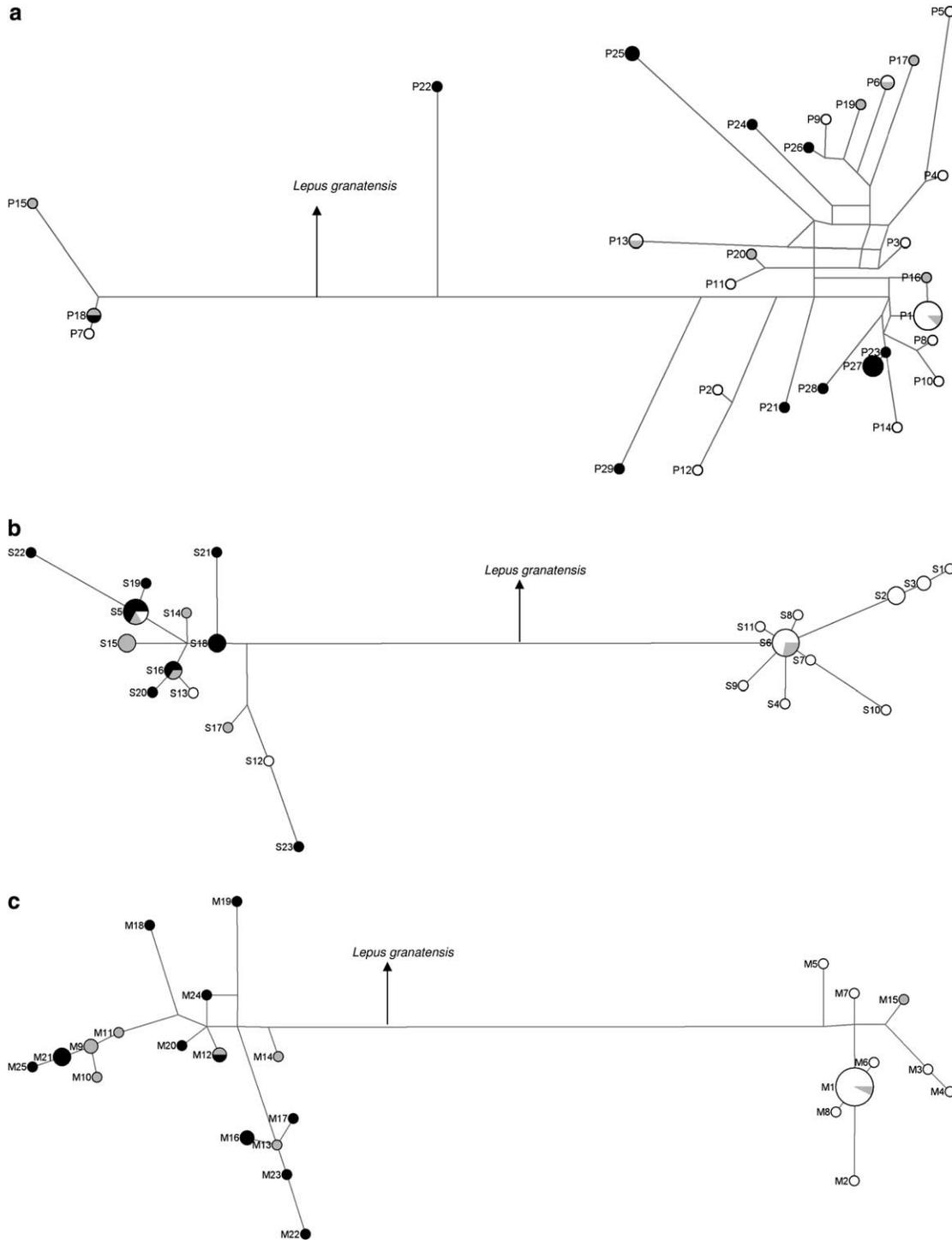
FIGURE 4.—Median-joining haplotype networks representing the phylogenetic relationships among all the alleles found in the European rabbit. (a) *Phka2*, (b) *Smcx*, (c) *Msn*, and (d) *Hprt1*. The size of the circles is proportional to the frequency of each haplotype. The population group of the individuals represented in each haplotype is denoted by solid (SW), shaded (CZ), and open (NE) treatment. The point in the network from which the outgroup sequence of *L. granatensis* stems is indicated by an arrow. Haplotype IDs correspond to Figure 3.

between the two groups. These patterns of variation suggest that there has been gene flow between *O. c. cuniculus* and *O. c. algirus* at some, but not all, loci. To further test this, we performed an HKA test between NE and SW population groups. A multilocus test between

all four loci failed to reject the null model, but in one pairwise comparison (between *Phka2* and *Msn*) the model was rejected ($P = 0.035$). This result seems to be mainly driven by the fact that divergence at *Msn* was much higher than expected (observed $D = 26.80$; expected
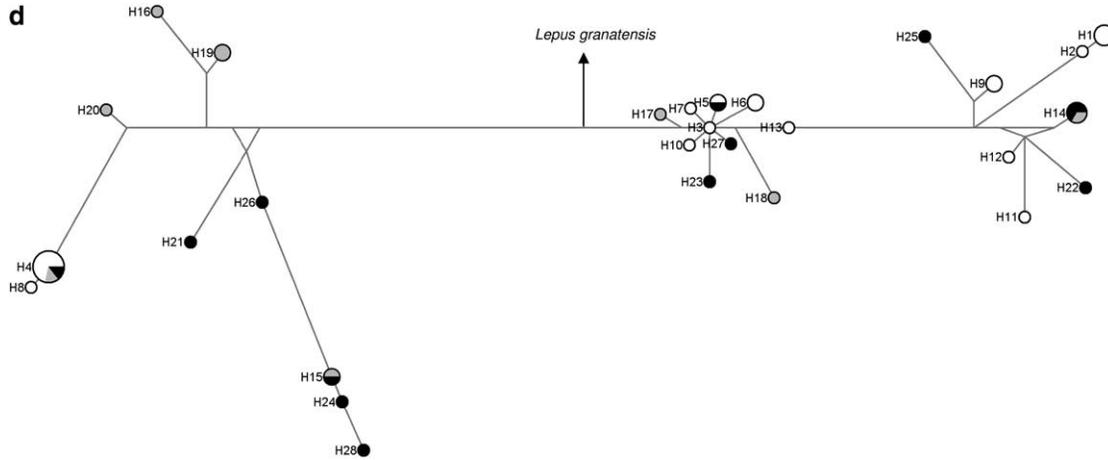
Figure 4.—*Continued.*

$D = 10.68$). The isolation without migration model (Wakeley and Hey, 1997) shares most of the assumptions with the HKA model, but estimates the ancestral population size instead of assuming that it is the average of the population size of the extant populations. A four-locus test using this model also failed to reject the null hypothesis. We also tested the fit of the data using all pairwise comparisons to *Msn*. Only the comparisons to *Msn* were performed because in the other comparisons there are no fixed differences and the program is unable to simulate the distribution of the expected values. The comparison between *Smcx* and *Msn* failed to reject the null model while the other two comparisons did reject the null model (*Phka2/Msn*: $P_{(\chi2)} = 0.032$ and $P_{(WH)} = 0.030$; *Msn/Hprt1*: $P_{(\chi2)} = 0.024$ and $P_{(WH)} = 0.009$).

We used IM (Hey and Nielsen 2004) to obtain maximum likelihood estimates (MLE) of the effective population size for each subspecies. We also estimated migration rates for each locus in each direction. The average estimate of the effective population size was

~882,000 for *O. c. algirus* and 422,000 for *O. c. cuniculus* (Table 6). The probability distribution of the ancestral population parameter was flat (not shown), as expected if the ancestral population existed long ago (Won and Hey 2005). Similarly, the probability distribution of t, the time since divergence, was flat, but nonzero (not shown). This suggests that the two subspecies were isolated in the past, but this analysis does not provide a reliable estimate of the time of isolation. Gene flow at the telomeric loci was higher from NE to SW than from SW to NE. For the centromeric loci, introgression of *Msn* is quite low in both directions, while *Smcx* shows some unidirectional introgression from SW to NE. Thus it seems that levels and patterns of gene flow are very different between centromeric and telomeric loci.

We estimated divergence time between the two subspecies of *O. cuniculus* using a phylogenetic approach. Assuming a divergence time of 11.8 MY (Matthee *et al.* 2004) between *O. cuniculus* and *L. granatensis*, divergence time between *O. c. cuniculus* and *O. c. algirus* was estimated to be on the order of 2–5 MYA (Table 7).

TABLE 4

Probabilities of observing the number of congruent sites, l_b, and maximum distance between congruent sites, g_d, under a single panmitic population

	γ estimated from data[a]			γ = 0.0015[b]			γ = 0.015[c]		
	l_b	g_d	l_b and g_d	l_b	g_d	l_b and g_d	l_b	g_d	l_b and g_d
Phka 2	0.35364	0.96288	0.34002	0.76470	0.99352	0.76296	0.04484	0.89190	0.04032
Smcx	0.39566	0.79952	0.37298	0.09558	0.33292	0.04248	0.00010	0.10748	0.00004
Msn	0.03470	0.03466	0.00214	0.06722	0.04908	0.00638	0.00006	0.01082	0.00000
Hprt 1	0.17324	0.99916	0.17302	0.62354	0.99992	0.62354	0.06626	0.99742	0.06626

Probabilities were calculated as the proportion of simulated genealogies with values of l_b, g_d, or both, equal to or greater than those observed in our data.

[a] For *Phka2*, γ = 0.0051 per site; for *Smcx*, γ = 0 per site; for *Msn*, γ = 0.0023 per site; and for *Hprt1*, γ = 0.0088 per site.

[b] γ = 3*Nc*, where $N = 1 \times 10^5$ and $c = 0.5 \times 10^{-8}$ per site.

[c] γ = 3*Nc*, where $N = 1 \times 10^5$ and $c = 5 \times 10^{-8}$ per site.

TABLE 5

Shared and fixed variation between NE and SW groups at four X-linked loci

	S^a	$S_{x_{NE}}{}^b$	$S_{x_{SW}}{}^c$	$S_s{}^d$	$S_f{}^e$
Phka2	140	45 (64.6)	35 (56.5)	60 (41.1)	0 (10.9)
Smcx	56	35 (16.9)	12 (14.7)	9 (10.7)	0 (2.8)
Msn	53	11 (15.2)	20 (13.3)	3 (9.7)	19 (2.6)
Hprt1	59	9 (20.3)	15 (17.7)	38 (12.9)	0 (3.4)

The expected values under the population parameters estimated with WH software (WAKELEY and HEY 1997) are shown in parentheses.

[a] S, no. of polymorphic positions.
[b] $S_{x_{NE}}$, no. of exclusive polymorphisms in the NE group.
[c] $S_{x_{SW}}$, no. of exclusive polymorphisms in the SW group.
[d] S_s, no. of shared polymorphisms.
[e] S_f, no. of fixed differences.

DISCUSSION

We documented genetic variation at four X-linked loci in natural populations of the European rabbit, *O. cuniculus*. At each locus, we observed a deep split in the phylogeny with the root lying along the long internal branch. This pattern is consistent with the evolution of each subspecies in allopatry and subsequent secondary contact. Despite this broad similarity among loci, we detected heterogeneity among loci in terms of levels of nucleotide polymorphism, recombination, and introgression between the two subspecies. This heterogeneity corresponds well with the physical location of the loci on the rabbit X chromosome. The two centromeric loci had lower levels of nucleotide polymorphism, higher levels of LD, and reduced introgression in comparison with the two telomeric loci. Although we do not have direct estimates of the frequency of crossing over in rabbits, these observations are consistent with suppression of recombination near the centromere, as has been observed in other species (*e.g.*, KONG *et al.* 2002).

Levels and patterns of variation: Across the entire sample (*i.e.*, including both subspecies), the average heterozygosity among all loci ($\pi = 0.76\%$) was high and roughly one order of magnitude higher than heterozygosity at X-linked loci in humans ($\pi = 0.081\%$; HAMMER *et al.* 2004) and mice ($\pi = 0.078\%$; NACHMAN 1997). Clearly, this high level of nucleotide variability reflects not only nucleotide polymorphism within each subspecies but also the divergence between subspecies. One gene (*Msn*) showed no introgression between subspecies. Levels of nucleotide polymorphism at this gene were 0.14% for *O. c. cuniculus* and 0.26% for *O. c. algirus*, closer to values observed in humans (HAMMER *et al.* 2004) and mice (NACHMAN 1997).

We also observed variation in levels of polymorphism among loci. Interestingly, the two centromeric loci had lower levels of π and θ than observed at the two

TABLE 6

MLE and the 90% highest posterior density intervals of demographic parameters

	$N_{e_{SW}}$	$N_{e_{NE}}$	Nm from NE to SW				Nm from SW to NE			
			Phka2	*Smcx*	*Msn*	*Hprt1*	*Phka2*	*Smcx*	*Msn*	*Hprt1*
MLE	882,675	422,149	10.215	0.008[a]	0.008[a]	3.470	0.281	0.404	0.012	0.951
HPD	520,833–1,737,939	202,851–685,307	2.600–16.092	0.008–6.480	0.008–0.411	0.733–14.695	0.004–7.696[b]	0.004–3.569	0.004–0.358	0.004–6.687

The 90% HPD intervals contain 90% of the probability density for each estimate. $N_{e_{SW}}$, the estimated effective population size of SW population group; $N_{e_{NE}}$, the estimated effective population size of NE population group; HPD, highest posterior density.

[a] The estimated value of Nm is at the lower limit of resolution; *i.e.*, 0.008 corresponds to the first bin of the parameter space surveyed for the migration parameter from SW to NE.

[b] The actual interval was larger than this and could not be estimated reliably, because the likelihood surface was relatively flat.

<div align="center">TABLE 7</div>

Uncorrected and corrected net nucleotide (D_a) divergences between *O. cuniculus* and *L. granatensis* and between subspecies of *O. cuniculus*, and estimates of divergence time (MY) between subspecies of *O. cuniculus*

| | *O. cuniculus/L. granatensis* | | *O. c. cuniculus/O. c. algirus* | | | |
| | Uncorrected | Corrected | Uncorrected | | Corrected | |
	D_a (%)	D_a (%)	D_a (%)	Divergence time (MY)	D_a (%)	Divergence time (MY)
Phka2	5.719	8.483[a]	1.153	2.38	1.263[a]	1.76
Smcx	1.475	1.529[b]	0.735	5.88	0.750[b]	5.79
Msn	3.803	4.634[c]	0.711	2.21	0.819[c]	2.08
Hprt1	3.848	5.313[d]	1.633	5.01	1.911[d]	4.24

Appropriate models of nucleotide substitution to correct for multiple hits were selected for each gene using MODELTEST 3.06 (Posada and Crandall 1998) with the Akaike Information Criterion (Posada and Buckley 2004). Pairwise distances (D_{xy}) per site were calculated using PAUP v 4.0 (Swofford 2002) with locus-specific estimated models of substitution. Net nucleotide divergences (D_a) per site were calculated as $D_{xy} − 0.5 (D_x + D_y)$. Recombinant haplotypes were excluded from this analysis.

[a] Tamura–Nei (1993) model with proportion of invariable sites of 0.6578 and estimated α-parameter describing the gamma distribution of 1.0022.

[b] Tamura–Nei (1993) model with proportion of invariable sites of 0.5536.

[c] Transversion model with estimated α-parameter describing the gamma distribution of 0.2204.

[d] Tamura–Nei (1993) model with proportion of invariable sites of 0.8524.

telomeric loci, both for the entire data set and for each subspecies considered separately. Within each subspecies, this difference may be explained by different levels of introgression. In other words, *Smcx* and *Msn* may be less variable within each subspecies because they contain relatively few introgressed haplotypes, compared to *Phka2* and *Hprt1*. However, we also observe less variation at *Smcx* and *Msn* in the total sample. This may be due in part to lower mutation rates at these genes. For example, divergence between Oryctolagus and Lepus is lower at *Smcx* ($D_{xy} = 1.74\%$) and *Msn* ($D_{xy} = 4.08\%$) than at *Phka2* ($D_{xy} = 6.06\%$) or *Hprt1* ($D_{xy} = 4.43\%$) (Table 2). If recombination is suppressed near the centromere, these differences in mutation rate may reflect an association between mutation and recombination (*e.g.,* Hellmann *et al.* 2003). It is also possible that reduced variation at *Smcx* and *Msn* may be due partly to the effect of either positive or negative selection at linked sites (Maynard-Smith and Haigh 1974; Charlesworth *et al.* 1993).

Indirect evidence that recombination is suppressed near the centromere comes from our observation of increased LD at *Smcx* and *Msn* compared to *Phka2* and *Hprt1*. Patterns of LD are affected by many factors, including selection, mutation, recombination, and changes in population size (*e.g.,* Ardlie *et al.* 2002). However, in humans, there is good evidence that levels of LD are inversely correlated with recombination rate over much of the genome (*e.g.,* Reich *et al.* 2001; McVean *et al.* 2004; Myers *et al.* 2005). Moreover, in many organisms, recombination is suppressed near the centromeres, particularly in metacentric chromosomes (*e.g.,* Kong *et al.* 2002). Thus, our observation of increased LD at *Smcx* and *Msn* relative to *Phka2* and *Hprt1* is consistent with, but not proof of, reduced recombination near the rabbit X centromere.

Divergence and gene flow between subspecies: RFLP surveys of mtDNA polymorphism in the Iberian Peninsula and France have shown that *O. cuniculus* is composed of two deeply divergent mtDNA lineages that are thought to have diverged ∼2 MYA (Biju-Duval *et al.* 1991; Branco *et al.* 2000). A survey of nucleotide variability at *Sry* also found evidence for the existence of two divergent lineages in the Y chromosome (Geraldes *et al.* 2005). These two lineages are associated with *O. c. algirus* and *O. c. cuniculus* (Branco *et al.* 2000) and are thought to have evolved in allopatry. Our X chromosome data confirm the existence of two divergent evolutionary units in *O. cuniculus,* and we show that in general the data reject the evolution of the two lineages under panmixia. The divergence time estimated from these loci is in good agreement with divergence time estimated from mitochondrial genes and places the origin of these two subspecies at the Pliocene/Pleistocene boundary. We observed high levels of population differentiation at the two centromeric loci, but not at the telomeric loci. At the centromeric loci the two divergent lineages correspond well with the described subspecies and are broadly concordant with the patterns of differentiation seen at the Y chromosome and at the mtDNA. The same was not observed at the two telomeric loci, where geography and phylogeny are largely decoupled.

If two populations evolve in allopatry for a sufficiently long time and then come into secondary contact with little or no gene flow, a high percentage of fixed differences and a small number of shared polymorphisms are expected. In our data this is seen only at *Msn* where 36% of all polymorphisms correspond to fixed differences between groups, and 6% correspond to shared polymorphisms. At all other loci, there are no fixed

differences between subspecies and the percentage of shared polymorphisms varies from 16% at *Smcx* (centromeric) to 64% at *Hprt1* (telomeric). This heterogeneity among loci is also reflected in the rejection of an isolation without gene flow model using the HKA test between *Phka2* and *Msn* and the rejection of the null model using the WH test between *Phka2* and *Msn* and between *Hprt1* and *Msn*.

It is noteworthy that the patterns of reduced introgression seen at *Smcx* and *Msn*, which may experience reduced recombination, are similar to the patterns seen previously at the mtDNA (BRANCO *et al.* 2000) and the Y chromosome (A. GERALDES, unpublished data), genomic regions with no recombination. In contrast, a survey of 14 allozyme loci revealed higher, but variable, levels of introgression (ϕ_{ct} between subspecies ranged from 0 to 0.46), comparable to the patterns observed at X chromosome loci. Differences among loci in levels of introgression have also been documented in other organisms. For example, genomic regions with suppressed recombination as a result of chromosomal rearrangements introgress less than colinear regions in comparisons between *Drosophila pseudoobscura* and *D. persimilis* (NOOR *et al.* 2001; MACHADO *et al.* 2002) and between hybridizing sunflowers of the genus Helianthus (RIESEBERG *et al.* 1999). On the basis of such observations, NOOR *et al.* (2001) and RIESEBERG (2001) have argued that chromosomal rearrangements may promote speciation, not through underdominance directly as in traditional models (*e.g.*, WHITE 1978), but by suppressing recombination and thereby extending the effects of isolation genes to linked sites. Our finding of low levels of introgression in an area of high LD near the X chromosome centromere of the rabbit is consistent with similar observations in fruit flies and sunflowers. Similarly, in Anopheles mosquitoes, two (of three) areas of reduced introgression map to centromeres (TURNER *et al.* 2005).

Our observations also have some interesting parallels with studies of hybridization in the house mice, *Mus musculus* and *M. domesticus*. In the house mouse hybrid zone in Western Europe, the Y chromosome shows reduced introgression (VANLERBERGHE *et al.* 1986; TUCKER *et al.* 1992; DOD *et al.* 1993), and the X chromosome shows lower levels of introgression than do the autosomes (TUCKER *et al.* 1992; DOD *et al.* 1993; MUNCLINGER *et al.* 2002), although there is also considerable variability in levels of introgression among loci on the X chromosome (PAYSEUR *et al.* 2004). In a similar fashion, we observe some X-linked loci with much reduced introgression in rabbits, providing further support for the importance of the X chromosome in reproductive isolation. Interestingly, the differences among loci in migration estimates (Table 6) may provide some clues to the nature of incompatibilities underlying reproductive isolation. In particular we note that estimates of the number of migrants from NE to SW for the centromeric loci are slightly lower than in the opposite direction. This is in agreement with the expected asymmetric behavior of young Dobzhansky–Muller interactions (ORR 1995) and suggests that incompatibilities may derive from interactions between the *cuniculus* X chromosome and an *algirus* genetic background.

One ultimate goal of speciation studies is to determine the identity of genes involved in reproductive isolation between nascent species. With the completion of the sequence of the rabbit genome expected in the next few years, it may soon be possible to identify candidate genes for reproductive isolation in this species. The results presented here suggest that some of these genes may lie near the centromere of the X chromosome.

LITERATURE CITED

ARDLIE, K. G., L. KRUGLYAK and M. SEIELSTAD, 2002 Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. **3:** 299–309.

BANDELT, H. J., P. FORSTER and A. ROHL, 1999 Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16:** 37–48.

BARBASH, D. A., D. F. SIINO, A. M. TARONE and J. ROOTE, 2003 A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA **100:** 5302–5307.

BATESON, W., 1909 Heredity and variation in modern lights, pp. 85–101 in *Darwin and Modern Science*, edited by A. C. SEWARD. Cambridge University Press, Cambridge, UK.

BIJU-DUVAL, C., H. ENNAFAA, N. DENNEBOUY, M. MONNEROT, F. MIGNOTTE *et al.*, 1991 Mitochondrial DNA evolution in Lagomorphs: origin of systematic heteroplasmy and organization of diversity in European rabbits. J. Mol. Evol. **33:** 92–102.

BRANCO, M., N. FERRAND and M. MONNEROT, 2000 Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analyses of the cytochrome *b* gene. Heredity **85:** 307–317.

BRANCO, M., M. MONNEROT, N. FERRAND and A. R. TEMPLETON, 2002 Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA variation. Evolution **56:** 792–803.

CHANTRY-DARMON, C., C. ROGEL-GAILLARD, M. BERTAUD, C. URIEN, M. PERROCHEAU *et al.*, 2003 133 new gene localizations on the rabbit cytogenetic map. Cytogenet. Genome Res. **103:** 192–201.

CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.

COYNE, J. A., and H. A. ORR, 1989 Two rules of speciation, pp. 180–207 in *Speciation and Its Consequences*, edited by D. OTTE and J. ENDLER. Sinauer Associates, Sunderland, MA.

COYNE, J. A., and H. A. ORR, 2004 *Speciation*. Sinauer Associates, Sunderland, MA.

DAVIDSON, J. J., T. OZCELIK, C. HAMACHER, P. J. WILLEMS, U. FRANCKE *et al.*, 1992 cDNA cloning of a liver isoform of the phosphorylase-kinase alpha-subunit and mapping of the gene to Xp22.2-p22.1, the region of human X-linked liver glycogenosis. Proc. Natl. Acad. Sci. USA **89:** 2096–2100.

DIETRICH, W. F., J. C. MILLER, R. G. STEEN, M. MERCHANT, D. DAMRON *et al.*, 1994 A genetic map of the mouse with 4,006 single sequence length polymorphisms. Nat. Genet. **7:** 220–245.

Dobzhansky, T. H., 1936 Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics **21:** 113–135.

Dod, B., L. S. Jermiin, P. Boursot, V. H. Chapman, J. Tonnes-Nielsen et al., 1993 Counterselection on sex-chromosomes in the *Mus musculus* European hybrid zone. J. Evol. Biol. **6:** 529–546.

Excoffier, L., P. E. Smouse and J. M. Quattro, 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131:** 479–491.

Ferrand, N., and M. Branco, 2006 The evolutionary history of the European rabbit (*Oryctolagus cuniculus*): major patterns of population differentiation and geographic expansion inferred from protein polymorphism, pp. 207–235 in *Phylogeography of European Refugia*, edited by S. Weiss and N. Ferrand. Springer, The Netherlands (in press).

Fu, Y. X., and W. H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 696–709.

Garrigan, D., Z. Mobasher, S. B. Kingan, J. A. Wilder and M. F. Hammer, 2005 Deep haplotype divergence and long-range linkage disequilibrium at Xp21.1 provide evidence that humans descended from a structured ancestral population. Genetics **170:** 1849–1856.

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

Grula, J. W., and O. R. Taylor, 1980 Some characteristics of hybrids derived from the sulphur butterflies, *Colias eurytheme* and *Colias philodice*: phenotypic effects of the X chromosome. Evolution **34:** 673–687.

Hagen, R. H., 1990 Population structure and host use in hybridizing subspecies of *Papilio glaucus* (Lepidoptera: Papilionidae). Evolution **44:** 1914–1930.

Haldane, J. B. S., 1922 Sex-ratio and unisexual sterility in hybrid animals. J. Genet. **12:** 101–109.

Hall, T. A., 1999 BioEdit: a user friendly biological sequence alignment editor and analyses program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41:** 95–98.

Hammer, M. F., D. Garrigan, E. Wood, J. A. Wilder, Z. Mobasher et al., 2004 Heterogeneous patterns of variation among multiple X-linked loci: the possible role of diversity-reducing selection in non-Africans. Genetics **167:** 1841–1853.

Hayes, H., C. Rogel-Gaillard, C. Zijlstra, N. A. de Haan, C. Urien et al., 2002 Establishment of an R-banded rabbit karyotype nomenclature by FISH localization of 23 chromosome-specific genes on both G- and R-banded chromosomes. Cytogenet. Genome Res. **98:** 199–205.

Hellmann, I., I. Ebersberger, S. E. Ptak, S. Paabo and M. Przeworski, 2003 A neutral explanation for the correlation of diversity with recombination rates in humans. Am. J. Hum. Genet. **72:** 1527–1535.

Hey, J., and R. M. Kliman, 1993 Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol. Biol. Evol. **10:** 804–822.

Hey, J., and R. Nielsen, 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., and J. Wakeley, 1997 A coalescent estimator of the population recombination rate. Genetics **145:** 833–846.

Hill, W. G., and A. Robertson, 1968 Linkage disequilibrium in finite populations. Theor. Appl. Genet. **38:** 226–231.

Hudson, R. R., 1987 Estimating the recombination parameter of a finite population model without selection. Genet. Res. **50:** 245–250.

Hudson, R. R., 2002 Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics **18:** 337–338.

Hudson, R. R., and N. L. Kaplan, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111:** 147–164.

Hudson, R. R., M. Kreitman and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Hudson, R. R., D. D. Boos and N. L. Kaplan, 1992 A statistical test for detecting geographic subdivision. Mol. Biol. Evol. **9:** 138–151.

Hudson, R. R., M. Slatkin and W. P. Maddison, 1992 Estimation of levels of gene flow from DNA sequence data. Genetics **132:** 583–589.

Jensen-Seaman, M. I., T. S. Furey, B. A. Payseur, Y. Lu, K. M. Roskin et al., 2004 Comparative recombination rates in the rat, mouse and human genomes. Genome Res. **14:** 528–538.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson et al., 2002 A high-resolution recombination map of the human genome. Nat. Genet. **31:** 241–247.

Lewontin, R. C., 1964 Interaction of selection and linkage. I. General considerations: heterotic models. Genetics **49:** 49–67.

Machado, C. A., R. M. Kliman, J. A. Markert and J. Hey, 2002 Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. Mol. Biol. Evol. **19:** 472–488.

Malitschek, B., D. Fornzler and M. Schartl, 1995 Melanoma formation in *Xiphophorus*: a model system for the role of receptor tyrosine kinases in tumorigenesis. BioEssays **17:** 1017–1023.

Matthee, C. A., B. J. van Vuuren, D. Bell and T. J. Robinson, 2004 A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. Syst. Biol. **53:** 433–447.

Maynard-Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favorable gene. Genet. Res. **23:** 23–35.

McVean, G. A. T., S. R. Myers, S. Hunt, P. Deloukas, D. R. Bentley et al., 2004 The fine-scale structure of recombination rate variation in the human genome. Science **304:** 581–584.

Muller, H. J., 1940 Bearing of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by J. S. Huxley. Clarendon Press, Oxford.

Muller, H. J., 1942 Isolating mechanisms, evolution and temperature. Biol. Symp. **6:** 71–125.

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

Myers, S., L. Bottolo, C. Freeman, G. McVean and P. Donnelly, 2005 A fine-scale map of recombination rates and hotspots across the human genome. Science **310:** 321–324.

Nachman, M. W., 1997 Patterns of DNA variability at X-linked loci in *Mus domesticus*. Genetics **147:** 1303–1316.

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

Nei, M., and W. H. Li, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:** 5269–5273.

Noor, M. A. F., K. L. Grams, L. A. Bertucci and J. Reiland, 2001 Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. USA **98:** 12084–12088.

Orr, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139:** 1805–1813.

Payseur, B. A., J. G. Krenz and M. W. Nachman, 2004 Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. Evolution **58:** 2064–2078.

Posada, D., and T. R. Buckley, 2004 Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. **53:** 793–808.

Posada, D., and K. A. Crandall, 1998 MODELTEST: testing the model of DNA substitution. Bioinformatics **14:** 817–818.

Presgraves, D. C., 2003 A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics **163:** 955–972.

Presgraves, D. C., L. Balagopalan, S. M. Abymayr and H. A. Orr, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature **423:** 715–719.

Reich, D. E., M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti et al., 2001 Linkage disequilibrium in the human genome. Nature **411:** 199–204.

Rieseberg, L. H., 2001 Chromosomal rearrangements and speciation. Trends Ecol. Evol. **16:** 351–358.

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

Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer and R. Rozas, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:** 2496–2497.

Sambrook, J., and D. W. Russell, 2001 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schneider, S., D. Roessli and L. Excoffier, 2000 *Arlequin: A Software Program for Population Genetics Data Analysis*. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva.

Sperling, F. A. H., and J. R. Spence, 1991 Structure of an asymmetric hybrid zone between two water strider species (Hemiptera: Gerridae: Limnoporus). Evolution **45:** 1370–1383.

Swofford, D. L., 2002 *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods*, Version 2.0.b10. Sinauer Associates, Sunderland, MA.

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

Tamura, K., and M. Nei, 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10:** 512–526.

Tao, Y., S. Chen, D. L. Hartl and C. C. Laurie, 2003 Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics **164:** 1383–1397.

Ting, C. T., S. C. Tsaur, M. L. Wu and C. I. Wu, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. Science **282:** 1501–1504.

True, J. R., B. S. Weir and C. C. Laurie, 1996 A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. Genetics **142:** 819–837.

Tucker, P. K., R. D. Sage, J. Warner, A. C. Wilson and E. M. Eicher, 1992 Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. Evolution **46:** 1146–1163.

Turelli, M., and H. A. Orr, 1995 The dominance theory of Haldane's rule. Genetics **140:** 389–302.

Turelli, M., and H. A. Orr, 2000 Dominance, epistasis and the genetics of postzygotic isolation. Genetics **154:** 1663–1679.

Turner, T. L., M. W. Hahn and S. V. Nuzhdin, 2005 Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. **3:** e285.

Vanlerberghe, F., B. Dod, P. Boursot, M. Bellis and F. Bonhomme, 1986 Absence of Y chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. Genet. Res. **48:** 191–197.

Wakeley, J., and J. Hey, 1997 Estimating ancestral population parameters. Genetics **145:** 847–855.

Wall, J. D., 2000 Detecting ancient admixture in humans using sequence polymorphism data. Genetics **154:** 1271–1279.

Wang, R. L., J. Wakeley and J. Hey, 1997 Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics **147:** 1091–1106.

Watterson, G. A., 1975 On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7:** 256–276.

White, M. J. D., 1978 *Modes of Speciation*. W. H. Freeman, San Francisco.

Won, Y. J., and J. Hey, 2005 Divergence population genetics of chimpanzees. Mol. Biol. Evol. **22:** 297–307.

Wright, S., 1951 The genetical structure of populations. Ann. Eugen. **15:** 323–354.