DOI: 10.1534/genetics.108.096826

Recombination and Speciation: Loci Near Centromeres Are More Differentiated Than Loci Near Telomeres Between Subspecies of the European Rabbit (*Oryctolagus cuniculus*)

Miguel Carneiro,*,†,‡,[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 Antropologia, Faculdade de Ciências, Universidade do Porto, 4099-002 Porto, Portugal and*
‡*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721*

Manuscript received September 27, 2008
Accepted for publication November 10, 2008

ABSTRACT

Recent empirical and theoretical studies suggest that regions of restricted recombination play an important role in the formation of new species. To test this idea, we studied nucleotide variation in two parapatric subspecies of the European rabbit (*Oryctolagus cuniculus*). We surveyed five loci near centromeres, where recombination is expected to be suppressed, and five loci near telomeres, where recombination is expected to be higher. We analyzed this multilocus data set using a divergence-with-gene flow framework and we report three main findings. First, we estimated that these subspecies diverged ~1.8 MYA and maintained large effective population sizes (*O. c. algirus* $N_e \approx 1{,}600{,}000$ and *O. c. cuniculus* $N_e \approx 780{,}000$). Second, we rejected a strict allopatric model of divergence without gene flow; instead, high rates of gene flow were inferred in both directions. Third, we found different patterns between loci near centromeres and loci near telomeres. Loci near centromeres exhibited higher levels of linkage disequilibrium than loci near telomeres. In addition, while all loci near telomeres showed little differentiation between subspecies, three of five loci near centromeres showed strong differentiation. These results support a view of speciation in which regions of low recombination can facilitate species divergence in the presence of gene flow.

UNDERSTANDING the conditions that give rise to new species is an important topic in evolutionary biology. In the classic allopatric model of speciation, divergence occurs in geographic isolation, and, after secondary contact, gene flow is impeded (Mayr 1963). However, in the last few years a large body of literature focusing on patterns of gene flow via interspecific hybridization has demonstrated that introgression between young species can be quite common (Rieseberg *et al.* 1999; Machado *et al.* 2002; Saetre *et al.* 2003; Stump *et al.* 2005; Putnam *et al.* 2007; Yatabe *et al.* 2007; Teeter *et al.* 2008). One clear result to emerge from this work is that levels of differentiation vary across the genome. While some regions of the genome can cross species boundaries, others, despite hybridization, are preserved from introgression. The consequence is semipermeable species boundaries (Endler 1977; Caisse and Antonovics 1978; Barton 1979; Bengtsson 1979; Wu 2001), in which the movement of genomic regions between species depends on their fitness effects or the fitness effects of linked regions (Barton and Hewitt 1989; Harrison 1990).

Recent studies have demonstrated that genomic regions harboring chromosomal rearrangements show less admixture between species than colinear regions (*e.g.*, Rieseberg *et al.* 1999; Machado *et al.* 2002, 2007; Feder *et al.* 2003; Noor *et al.* 2007). Moreover, genes involved in reproductive isolation sometimes map preferentially to rearranged regions (Noor *et al.* 2001). These observations led several authors to propose models of speciation where regions of low recombination represent islands of differentiation that facilitate or maintain species identity in the face of gene flow (Noor *et al.* 2001; Rieseberg 2001; Navarro and Barton 2003). Rieseberg's (2001) model is based on the combined effects of multiple isolation genes that are held together on a single haplotype, while the model of Noor *et al.* (2001) is based on the asymmetric nature of Dobzhansky–Muller incompatibilities and their potential to impede introgression when two loci with opposite asymmetries are located within a single nonrecombining region. The model of Navarro and Barton (2003), in contrast, is based on coadapted gene complexes held together by the absence of recombination (see also Kirkpatrick and Barton 2006). These models differ in many ways but share the common feature that the effects of genes involved in isolation are extended in regions with chromosomal rearrangements due to decreased recombination (reviewed in Butlin 2005). Besides structural rearrangements, the role of other regions of

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. FJ000560–FJ001233.

[1]*Corresponding author:* Department of Ecology and Evolutionary Biology, Biosciences West Bldg., University of Arizona, P.O. Box 210088, Tucson, AZ 85721. E-mail: miguel.carneiro@mail.icav.up.pt

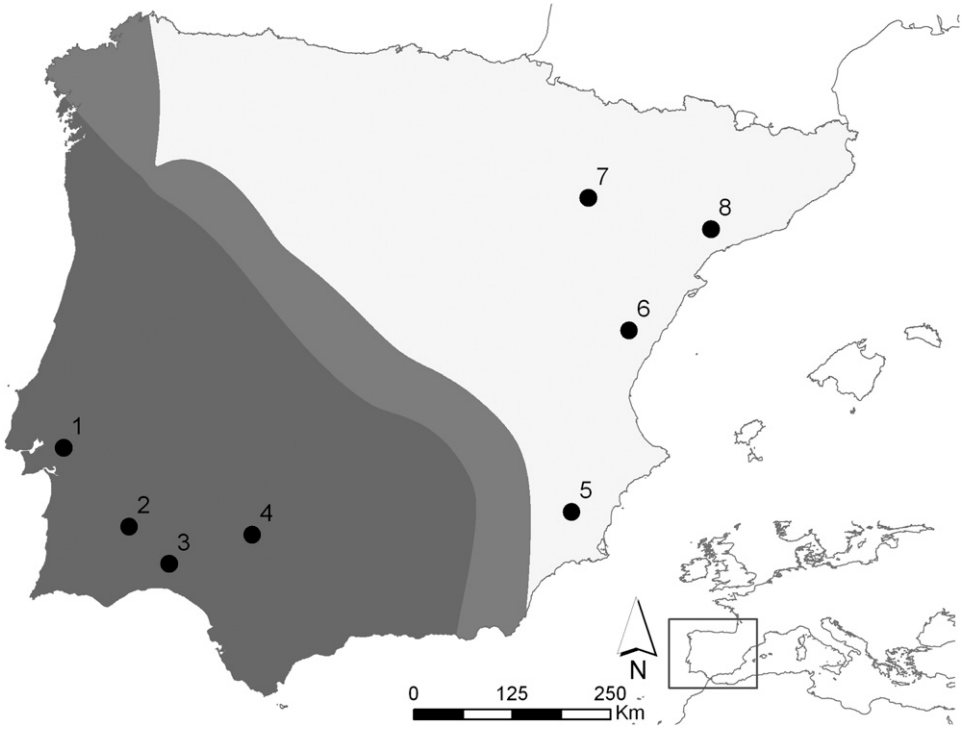

Figure 1.—Geographical location of the populations used in this study. Dark and light shading indicates the range of *O. c. algirus* and *O. c. cuniculus*, respectively. The putative hybrid zone separates both ranges (adapted from Geraldes *et al.* 2008). Numbers correspond to populations as follows: 1, Pancas; 2, Mértola; 3, Huelva; 4, Sevilla; 5, Alicante; 6, Rosell; 7, Zaragoza; 8, Tarragona.

restricted recombination in speciation is far less clear. However, nonstructural chromosomal features, such as proximity to the centromere, where recombination is often suppressed, have also been associated with lower rates of gene flow (Stump *et al.* 2005). While both theoretical and empirical studies suggest that the higher differentiation in low-recombination regions might be caused by a higher density of genes contributing to reproductive isolation in such regions, greater differentiation in regions of low recombination could also arise because genes involved in reproductive isolation influence longer neutral regions of the genome in regions of reduced crossing over.

Surprisingly, with the exception of the house mouse (*e.g.*, Boursot *et al.* 1993; Britton-Davidian *et al.* 2005) and to a lesser extent primates (Won and Hey 2005; Patterson *et al.* 2006), the genetics of speciation have received relatively little attention in mammals. The European rabbit (*Oryctolagus cuniculus*) is a good mammalian system for studying the genetic architecture of speciation under a model of divergence with gene flow. It is composed of two parapatric subspecies: *O. cuniculus algirus*, distributed in the southwestern part of the Iberian Peninsula, and *O. cuniculus cuniculus* in the northeastern part of the Iberian Peninsula (Figure 1). Although there are slight morphological differences between the two forms in craniometric measurements (Sharples *et al.* 1996) and body size (Villafuerte 2002), this intraspecific taxonomy is mainly based on patterns of genetic differentiation (*e.g.*, Branco *et al.* 2000; Geraldes *et al.* 2006). Crosses between these subspecies produce viable hybrids, and the fertility of these hybrids is under investigation. Previous work showed contrasting patterns of differentiation across multiple markers, and the extensive sharing of genetic variation at some loci has been interpreted as evidence of gene flow following secondary contact (Geraldes *et al.* 2006, 2008; Ferrand and Branco 2007). In particular, Geraldes *et al.* (2006) described a striking contrast between two centromeric and two telomeric loci on the X chromosome. Levels of gene flow were low at the two loci near the centromere and high at the two loci near the telomere. These authors suggested that the centromeric region of the X chromosome may be involved in reproductive isolation between the two subspecies.

Here, we report the first large-scale survey of patterns of polymorphism and divergence at multiple autosomal regions in the European rabbit. To test the role of recombination on levels of differentiation we surveyed five loci near centromeres (hereafter "centromeric loci"), where recombination is expected to be lower (Choo 1998; Kong *et al.* 2002), and five loci near telomeres (hereafter "telomeric loci"), where recombination is expected to be higher. Additionally, using a divergence–population-genetics framework (Hey and Nielsen 2004, 2007), we estimated relevant demographic parameters that provide insight into the genetic architecture of speciation in this group. We address three main questions: (1) What are the levels and patterns of genetic variation on the autosomes in wild rabbit subspecies and what are the inferred effective population sizes?, (2) When did these subspecies start to diverge and can we reject a strict allopatric model without subsequent gene flow?, and (3) Does the physical location of loci (near centromeres *vs.* near telomeres)

TABLE 1

Chromosome location, position in the gene, and length of the regions sequenced

Gene	Chromosome	Chromosome location	Position in the gene	Length
ATP12A	8	Centromere	Intron 2	1488
CYTC	4	Centromere	Intron 1	1120
MGST3	13	Centromere	Intron 1	925
PRL	12	Centromere	5′	1270
STAG1	14	Centromere	Intron 6	1425
EXT1	3	Telomere	Intron 4	1227
LUM	4	Telomere	Intron 1	1223
T	12	Telomere	5′	957
TIAM1	14	Telomere	Intron 2	1262
UD14	7	Telomere	Intron 4	1260

Chromosome location is based on the rabbit cytogenetic map (Chantry-Darmon *et al.* 2005). Position in the gene was obtained in Ensembl Build 48.1e.

correlate with patterns of genetic differentiation, as predicted by recent speciation models?

MATERIALS AND METHODS

Samples and sampling strategy: We surveyed 10 individuals from each of eight populations, for a total of 80 individuals (Figure 1). Populations were divided into two major groups on the basis of the mtDNA phylogeographical division described by Branco *et al.* (2000): (1) the southwestern part of the Iberian Peninsula, corresponding to the subspecies *O. c. algirus* ($n = 40$), and (2) the northeastern part of the Iberian Peninsula, corresponding to the subspecies *O. c. cuniculus* ($n = 40$). One individual of *Lepus granatensis* was sampled to provide an outgroup. Genomic DNA was isolated from tissues, using a standard phenol-chloroform DNA extraction (Sambrook and Russell 2001).

Loci, DNA amplification, and sequencing: We investigated patterns of genetic diversity in wild rabbits by resequencing 10 effectively unlinked autosomal loci. The loci consisted entirely of noncoding DNA, either intronic or 5′ flanking (Table 1). Five loci were selected near centromeres and five near telomeres (Figure 2) from genes mapped on the rabbit cytogenetic map (Chantry-Darmon *et al.* 2005).

Given the high levels of nucleotide diversity previously described for the European rabbit (Geraldes *et al.* 2006), there is a strong possibility of mutations occurring in priming sites, creating allelic dropout. To detect such situations we amplified two overlapping amplicons, one contained within the other, for each segment. The products were sequenced in both directions. We used additional primers for individuals heterozygous for multiple-indel polymorphisms and when instances of allele-specific PCR or allele-specific sequencing were detected. All PCR and sequencing primers were designed from the rabbit genomic sequence and are provided in supplemental Table 1. Sequences were edited, aligned, and assembled using phred/phrap/consed/polyphred (Nickerson *et al.* 1997; Ewing and Green 1998; Ewing *et al.* 1998; Gordon *et al.* 1998) together with auxiliary shell scripts and Perl programs kindly provided by August Woerner. Manual adjustments were further performed in some contigs, using Bioedit (Hall 1999). Individuals or positions with >15% of missing data were eliminated from the data set. Sequences have been deposited in GenBank under accession nos. FJ000560–FJ001233.

Nucleotide variation, recombination, and neutrality tests: Haplotypes were assigned by the computer program PHASE (Stephens *et al.* 2001; Stephens and Donnelly 2003). For each data set, we applied the algorithm five times with different seeds for the random number generator, and we checked for consistency of the results across independent runs.

Most standard population-genetic analyses were carried out using the program SITES (Hey and Wakeley 1997). All summary statistics were obtained for the global sample as well as for each subspecies independently. Indels were not included in the analyses. We estimated the neutral mutation parameter $4N_e\mu$ (where N_e is the effective population size and μ is the mutation rate per site per generation) by calculating Watterson's (1975) θ_w, the proportion of segregating sites in a sample, and π (Nei 1987), the average number of pairwise differences per sequence in a sample. Several descriptors of the amount of linkage disequilibrium (LD) were obtained. The population recombination parameter ($R = 4N_ec$, where c is the recombination rate per generation) between adjacent sites was estimated using γ (Hey and Wakeley 1997) and ρ (Hudson 2001). γ is a maximum-likelihood estimator developed using a coalescent model for a sample of four DNA sequences with recombination. ρ is adapted to a finite-sites model and is estimated by means of a composite-likelihood method, as implemented in the LDhat 2.0 package (McVean *et al.* 2002). We also estimated the minimum number of recombination events (R_m) in the sample (Hudson and Kaplan 1985). Finally, we assessed the significance of LD through pairwise comparisons of polymorphic sites, using a chi-square test excluding individuals with missing data, using DnaSP 4.50.3 (Rozas *et al.* 2003).

We tested for deviations from neutrality in allele-frequency distributions, using Tajima's (1989) D and Fu and Li's (1993) D^*. Significance of Tajima's D and Fu and Li's D^* values was assessed by coalescent simulations, using DnaSP 4.50.3 (Rozas *et al.* 2003). We ran 10^4 coalescent simulations of the standard neutral model with the same number of chromosomes and base pairs as in the actual data and with the θ_W and population recombination rate equal to the estimated values. We also used the HKA test (Hudson *et al.* 1987) to evaluate heterogeneity in the ratio of polymorphism to divergence across loci under a neutral model. We compared ratios of polymorphism within rabbit to divergence between rabbit and Lepus. We performed three 10-locus tests, using polymorphism for both the two subspecies together and separately by means of the HKA software (http://lifesci.rutgers.edu/~heylab).

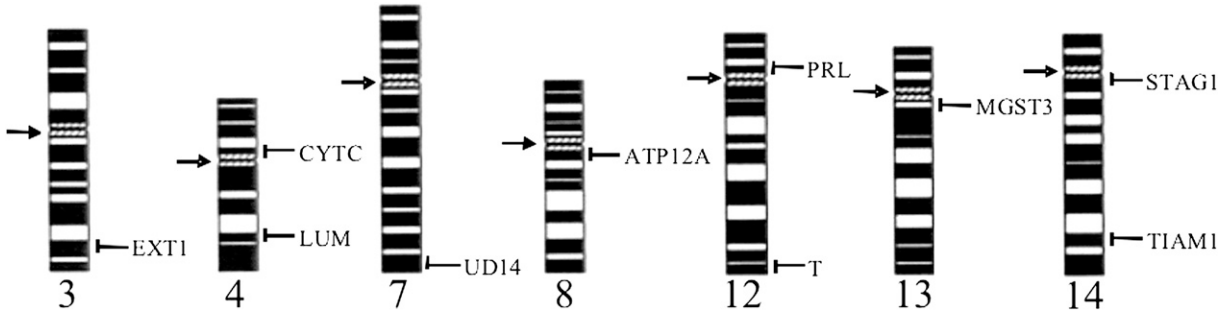

FIGURE 2.—Chromosome locations of the 10 autosomal regions used in this study. Arrows indicate the position of the centromere for each chromosome. Modified from CHANTRY-DARMON *et al.* (2005).

Genetic differentiation and gene flow: We assessed genetic differentiation in several ways. We estimated the fixation index (F_{ST}) from which we inferred the number of migrants per generation (Nm), and we estimated the net nucleotide divergence (D_a, NEI 1987) between *O. c. algirus* and *O. c. cuniculus*. In addition, we calculated the average pairwise differences (D_{xy}, NEI 1987) between all rabbit samples and *L. granatensis*. We also calculated fixation indexes within and between subspecies in an analysis of molecular variance (AMOVA) (EXCOFFIER *et al.* 1992), using the Arlequin genetics package (SCHNEIDER *et al.* 2000). Another way of looking at differentiation between groups is from a genealogical perspective. Phylogenetic relationships among alleles for each locus were reconstructed with the neighbor-joining method (SAITOU and NEI 1987), using the program MEGA4 (TAMURA *et al.* 2007). Bootstrap values were obtained after 1000 replicates (FELSENSTEIN 1985).

Patterns of shared variation (see RESULTS) may be explained by shared ancestral polymorphism or recent gene flow; to distinguish between these hypotheses we used the LD method developed by MACHADO *et al.* (2002). This test looks at the difference (x = DSS − DSX) between the amount of LD that is observed among pairs of shared polymorphisms (DSS) and the average estimate among pairs of sites where one is shared between species and the other is an exclusive polymorphism (DSX). If shared polymorphisms are due to postdivergence gene flow, x will tend to be positive because of the short time available for recombination to break down associations in the recipient species. This test was applied using D' (LEWONTIN 1964) as a measure of LD. *L. granatensis* was used to polarize LD. Statistical significance of the observed values of x was obtained by simulating data sets using the program WH (WANG *et al.* 1997).

Divergence population genetics: We investigated an isolation-with-migration (IM) model and estimated relevant demographic parameters using the programs IM and IMa (NIELSEN and WAKELEY 2001; HEY and NIELSEN 2004, 2007). IM and IMa are based on Markov chain Monte Carlo simulations of genealogies and they estimate the posterior probability density function for six model parameters (the population-split time, effective population size for the ancestral and current populations, and migration rates in both directions). Although these six parameters may give a good approximation to the demographic history of two diverging populations, there are also several important population genetic assumptions underlying the method. First, the method assumes that there are no other unsampled populations exchanging genes with the populations under study or with the ancestral population. Second, it assumes selective neutrality. Third, the model assumes a particular mutational model (in this case, the infinite-sites model; KIMURA 1969). Finally, it assumes free recombination between loci and no recombination or gene conversion within loci. Because of this last assumption, we used the program IMgc (WOERNER *et al.* 2007) to obtain the longest region without four gametic types for each locus. We simulated the data three times to check for convergence. To facilitate mixing of the Markov chains we used Metropolis coupling. In this Metropolis-coupled version of the algorithm multiple linked chains are run simultaneously and disparate starting points and heating are used to improve sampling efficiency. For each simulation we ran 10 million steps, using 15 chains and a geometric heating scheme. Output from IM is expressed in units of $4N_e\mu$ and $t_0\mu$, where μ is the mutation rate per generation and t is the divergence time in generations. To estimate the effective population size in numbers of individuals for each subspecies (N_e) and the time since population split in years, we calculated μ for each locus, assuming that the divergence time for the Oryctolagus–Lepus comparison is 11.8 MYA (MATTHEE *et al.* 2004). We then estimated the neutral mutation rate as $\mu = D/2t_1$, where D is the estimated D_a (NEI 1987) between Oryctolagus and Lepus for each locus. We assumed a generation time of 1 year and we calculated the geometric mean of the locus-specific mutation rates.

RESULTS

Multilocus patterns of polymorphism and recombination: We collected polymorphism data in *O. c. algirus* and *O. c. cuniculus* for 10 autosomal regions. This data set contained ~12 kb of noncoding DNA sequence for 52–72 chromosomes in *O. c. algirus* and 56–80 chromosomes in *O. c. cuniculus*. The length of each aligned locus varied between 925 and 1488 bp. Outgroup sequence of *L. granatensis* was obtained for 9 of the 10 loci. We could not amplify UD14 in Lepus despite using multiple combinations of primers.

Levels of nucleotide variation were high both in the total sample and within each subspecies (Table 2). In the total sample, values of π (NEI 1987) ranged from 0.383 to 1.373% (average $\pi = 0.836$). The values of θ_w (WATTERSON 1975) were even higher, ranging from 1.074 to 2.264% (average $\theta_w = 1.227$). Within subspecies, the genetic diversity was higher in *O. c. algirus* (average $\pi = 0.746$, average $\theta_w = 1.100$) than in *O. c. cuniculus* (average $\pi = 0.662$, average $\theta_w = 0.848$). This difference was significant for θ_w (Wilcoxon's signed rank test, $P = 0.005$) but not for π (Wilcoxon's signed rank test, $P = 0.1141$).

TABLE 2

Analyses of polymorphism, frequency spectrum tests of neutrality, and recombination in *O. c. algirus* (alg) and *O. c. cuniculus* (cun)

			Polymorphism				Frequency spectrum tests of neutrality		Recombination			Divergence:		
Locus	Pos[a]	n^b	L^c	S^d	π (%)	θ (%)	$D_T{}^e$	$D_{FL}{}^f$	γ(%)[g]	ρ(%)[h]	Rm[i]	$D_a{}^j$ (%)	γ/θ	θ/D_a
ATP12A	Cen													
	All	138	1459.5	83	0.832	1.034	−0.617	−0.131	0.87	0.41	5		0.84	0.25
	alg	72	1458.4	67	0.567	0.948	−0.432	0.313	1.19	0.53	5	4.134	1.26	0.23
	cun	66	1460.7	51	0.236	0.734	−0.757	−0.782	0.02	0.16	3		0.03	0.18
CYTC	Cen													
	All	124	1016.3	68	0.756	1.241	−1.237	−1.355	0.61	0.80	7		0.49	0.26
	alg	56	1015.6	49	0.849	1.050	−0.651	0.159	0.62	0.53	4	4.774	0.59	0.22
	cun	68	1017	42	0.644	0.862	−0.829	−1.378	0.32	0.79	5		0.37	0.18
MGST3	Cen													
	All	120	911.8	58	1.041	1.187	−0.387	−1.905*	1.50	1.42	9		1.26	0.25
	alg	52	909.2	45	0.735	1.095	−1.123	−1.516	0.89	1.73	5	4.710	0.81	0.23
	cun	68	914.1	36	0.833	0.822	0.240	0.161	1.34	1.25	8		1.62	0.17
PRL	Cen													
	All	138	1249.7	85	0.635	1.236	−1.542*	−1.029	0.66	0.24	3		0.53	0.29
	alg	62	1246.9	57	0.510	0.973	−1.610*	−1.291	0.43	0.65	3	4.314	0.44	0.23
	cun	76	1252	45	0.702	0.733	−0.137	1.620*	0.76	0.22	3		1.04	0.17
STAG1	Cen													
	All	146	1369.2	72	0.764	0.946	−0.602	−1.862	0.03	0.00	3		0.04	0.27
	alg	68	1334.5	49	0.334	0.767	−1.871*	−2.903**	0.03	0.00	2	3.447	0.04	0.22
	cun	78	1401.2	42	0.244	0.608	−1.934**	−4.583***	0.02	0.00	2		0.04	0.18
EXT1	Tel													
	All	146	1212.0	68	0.383	1.010	−1.936***	−3.659***	1.09	1.82	4		1.08	0.30
	alg	66	1203.5	47	0.457	0.821	−1.469*	−2.030*	1.55	2.09	4	3.299	1.89	0.25
	cun	80	1219.1	39	0.298	0.646	−1.724**	−3.049***	0.87	0.74	5		1.35	0.20
LUM	Tel													
	All	130	1177.7	74	0.978	1.155	−0.486	−0.228	2.57	0.98	8		2.23	0.30
	alg	58	1185.8	60	1.062	1.093	−0.097	−0.127	3.13	1.12	8	3.858	2.86	0.28
	cun	72	1171.3	56	0.849	0.986	−0.461	0.130	1.77	0.57	8		1.80	0.26
T	Tel													
	All	134	869	108	1.371	2.271	−1.273***	−1.563	8.41	10.62	17		3.70	0.32
	alg	62	845.7	84	1.351	2.115	−1.243***	−0.815*	9.56	28.57	16	7.083	4.52	0.30
	cun	72	889.7	68	1.329	1.577	−0.527	−1.343	5.62	3.32	9		3.56	0.22
TIAM1	Tel													
	All	110	1227.8	72	0.759	1.112	−1.021	−2.421	3.07	3.76	12		2.76	0.33
	alg	54	1219.7	68	0.792	1.223	−1.223	−2.155	3.23	7.70	10	3.318	2.64	0.37
	cun	56	1235.8	38	0.710	0.669	0.203	0.556	3.21	1.80	8		4.80	0.20

(*continued*)

TABLE 2

(Continued)

Locus	Pos[a]	n^b	Polymorphism				Frequency spectrum tests of neutrality		Recombination			Divergence:		
			L^c	S^d	π (%)	θ (%)	D_T^e	D_{FL}^f	γ(%)[g]	ρ(%)[h]	Rm[i]	D_a^j (%)	γ/θ	θ/D_a
UD14	Tel													
	All	144	1208.9	72	0.836	1.074	−0.695	−0.358	6.10	7.15	11		5.68	NA
	alg	64	1217	50	0.806	0.869	−0.241	0.333	2.85	2.44	12	NA	3.28	NA
	cun	80	1202.5	50	0.774	0.840	−0.254	0.460	7.09	5.92	11		8.45	NA

Statistical significance for the frequency spectrum tests of neutrality was obtained using coalescent simulations (*$P < 0.05$; **$P < 0.01$; ***$P < 0.005$).

[a] Physical position in the chromosome. Cen, centromere; Tel, telomere.

[b] Sample size is number of chromosomes.

[c] Average length (bp) of the sequence.

[d] Number of polymorphic sites.

[e] Tajima's (1989) *D*.

[f] Fu and Li's (1993) *D**.

[g] Maximum-likelihood estimate of the population recombination parameter between adjacent sites (Hey and Wakeley 1997).

[h] Composite-likelihood method to estimate the population recombination parameter between adjacent sites, employing a finite-sites mutation model (Hudson 2001).

[i] Minimum number of recombination events in the history of the sample (Hudson and Kaplan 1985).

[j] D_{xy} is the average nucleotide divergence (Nei 1987) between all rabbit haplotypes and *Lepus granatensis*.

A high proportion of rare polymorphisms was apparent at each locus, as reflected in the strongly negative Tajima's *D* (Tajima 1989) and Fu and Li's *D** (Fu and Li 1993) (Table 2). This was observed both for the complete data set (average Tajima's $D = -0.980$, average Fu and Li's $D^* = -1.451$) and for each subspecies separately (*O. c. algirus*, average Tajima's $D = -0.996$, average Fu and Li's $D^* = -1.000$; *O. c. cuniculus*, average Tajima's $D = -0.618$, average Fu and Li's $D^* = -0.821$). Significantly negative values of one or both statistics were observed at several individual loci (Table 2). We also observed significantly negative values of Tajima's *D* for each subspecies in multilocus tests ($P < 0.0001$ for *O. c. algirus* and $P = 0.024$ for *O. c. cuniculus*). A 10-locus HKA test (Hudson *et al.* 1987) indicated no significant heterogeneity in levels of polymorphism and divergence within subspecies as well as in the global data set ($P > 0.10$ for all three comparisons).

We found some evidence of recombination at each locus ($Rm > 1$ for each), although there was considerable variation among loci in estimates of the population recombination parameter (Table 2). *O. c. algirus* exhibited slightly but not significantly higher average values of the population recombination parameter than *O. c. cuniculus* for both γ and ρ ($\gamma = 2.35\%$ and $\rho = 4.54\%$ for *O. c. algirus*; $\gamma = 2.10\%$ and $\rho = 1.48\%$ for *O. c. cuniculus*; Wilcoxon's signed rank test, $P > 0.05$ for both). This trend is consistent with the higher values of nucleotide variation observed in *O. c. algirus* and may reflect a higher N_e in this subspecies. We also evaluated LD within each locus, using a chi-square test for independence between sites. In *O. c. algirus*, 1464 of 6771 pairwise comparisons (21.6%) between nonsingleton pairs of polymorphisms within loci showed statistically significant LD; with the Bonferroni correction for multiple comparisons, there were 529 (7.8%) significant associations. In *O. c. cuniculus* we found a higher percentage of within-locus nonrandom associations between pairs of polymorphic sites: 1609 of 5046 pairwise comparisons (31.9%) were significant, and 739 (14.6%) were significant with the Bonferroni correction.

The current genetic (Chantry-Darmon *et al.* 2006) and physical maps for the rabbit are not sufficiently detailed to estimate regional variation in the rate of crossing over. To test the hypothesis that recombination is suppressed near centromeres compared to telomeres in rabbits, as seen in other mammals (*e.g.*, Kong *et al.* 2002), we compared the population recombination parameter for loci near centromeres and loci near telomeres. The mean values of both γ and ρ calculated for both subspecies together were significantly lower in centromeric loci ($\gamma = 0.73\%$, $\rho = 0.57\%$) than in telomeric loci ($\gamma = 4.25\%$, $\rho = 4.87\%$; Mann–Whitney *U*-test, $P = 0.016$ in both tests). Within *O. c. algirus*, both estimators were significantly lower in centromeric loci ($\gamma = 0.63\%$, $\rho = 0.69\%$) compared to telomeric loci ($\gamma = 3.90\%$, $\rho = 8.18\%$; Mann–Whitney *U*-test, $P < 0.05$ for

TABLE 3

Genetic differentiation and migration between *O. c. algirus* (alg) and *O. c. cuniculus* (cun) at 10 autosomal and 4 X-linked loci

			Differentiation		Fixation indexes			Migration rates		
	Chr	Pos.[a]	F_{ST} (%)[b]	D_a (%)[c]	ϕ_{ct}[d]	ϕ_{st}[e]	ϕ_{sc}[f]	Nm (F_{ST})[g]	$2Nm_{alg}$ (IM)[h]	$2Nm_{cun}$ (IM)[i]
ATP12A	8	Cen	27.2	0.260	0.26	0.32	0.08	0.670	1.174	0.539
CYTC	4	Cen	5.0	0.039	0.04	0.11	0.07	4.778	9.315	0.137
MGST3	13	Cen	35.6	0.447	0.33	0.43	0.14	0.452	0.225	0.159
PRL	12	Cen	6.0	0.038	0.01	0.19	0.18	3.952	0.874	2.548
STAG1	14	Cen	76.9	0.959	0.78	0.79	0.02	0.075	0.175	0.011
EXT1	3	Tel	6.9	0.028	0.06	0.09	0.04	3.396	7.667	0.116
LUM	4	Tel	6.4	0.065	0.04	0.11	0.07	3.654	0.774	3.964
T	12	Tel	4.5	0.064	0.05	0.12	0.08	5.250	5.120	0.666
TIAM1	14	Tel	2.3	0.017	0.01	0.06	0.05	10.844	16.109	0.476
UD14	7	Tel	10.8	0.095	0.09	0.17	0.09	2.071	3.921	0.412
MSN[j]	X	Cen	82.9	0.786	0.84	0.86	0.11	0.070	0.008	0.012
SMCX[j]	X	Cen	68.0	0.531	0.64	0.80	0.43	0.160	0.008	0.404
HPRT1[j]	X	Tel	2.2	0.027	0.02	0.07	0.05	14.990	3.470	0.951
PHKA2[j]	X	Tel	2.7	0.008	0.02	0.07	0.05	12.220	10.215	0.281

[a] Physical position in the chromosome. Cen, centromere; Tel, telomere.
[b] Calculated using the method proposed by Hudson *et al.* (1992).
[c] The net nucleotide distance per base pair (Nei 1987) between *O. c. algirus* and *O. c. cuniculus*.
[d] The fixation index for the amount of variation segregating between *O. c. algirus* and *O. c. cuniculus*.
[e] The fixation index for the amount of variation segregating within each population.
[f] The fixation index for the amount of variation segregating among populations within each subspecies.
[g] The estimate of Nm was calculated using the expression $F_{ST} = 1/(1 + 4Nm)$ for autosomal and $F_{ST} = 1/(1 + 3Nm)$ for X-linked loci.
[h] ML estimates of population migration rate from *O. c. cuniculus* to *O. c. algirus* using the IM software.
[i] ML estimates of population migration rate from *O. c. algirus* to *O. c. cuniculus* using the IM software.
[j] Data from Geraldes *et al.* (2006).

each). Within *O. c. cuniculus*, γ was significantly lower in centromeric loci ($\gamma = 0.49\%$) compared to telomeric loci ($\gamma = 3.71\%$; Mann–Whitney *U*-test, $P = 0.008$) but the result was not significant for ρ (ρ centromeric = 0.48%, ρ telomeric = 2.47%; Mann–Whitney *U*-test, $P = 0.095$). Mean *Rm* values were also significantly lower at centromeric loci relative to telomeric loci in both *O. c. algirus* (*Rm* centromeric = 3.8, *Rm* telomeric = 10; Mann–Whitney *U*-test, $P = 0.032$) and *O. c. cuniculus* (*Rm* centromeric = 4.2, *Rm* telomeric = 8.2; Mann–Whitney *U*-test, $P = 0.032$), although these differences were not significant in the overall sample (*Rm* centromeric = 5.4, *Rm* telomeric = 10.4; Mann–Whitney *U*-test, $P = 0.075$).

Under neutrality, the ratio ρ/θ gives a measure of the relative contribution of recombination and mutation in a given locus ($\rho/\theta = 4N_e c/4N_e\mu = c/\mu$). The mean values for centromeric loci ($\rho_{Sites}/\theta_w = 0.63$) were significantly lower than those for telomeric loci ($\rho_{Sites}/\theta_w = 3.09$; Mann–Whitney *U*-test, $P = 0.016$). The lower ratio at centromeric loci could, in principle, have two explanations: it could be due to either stronger LD or higher mutation rates close to the centromeres, or some combination of these factors. We estimated mutation rates (μ) per generation per site for centromeric loci ($\mu = 1.622 \times 10^{-9}$) and for telomeric loci ($\mu = 1.662 \times 10^{-9}$). The similarities of these estimates (Mann–Whitney *U*-test, $P = 0.556$) suggest that lower recombination, rather than higher mutation, is responsible for the patterns seen at centromeric loci.

We were also interested in evaluating the possibility that regions of low recombination might exhibit lower levels of polymorphism, as seen in other species (*e.g.*, Begun and Aquadro 1992). To test this idea, we divided θ_w by divergence to Lepus (D_{xy}) to correct for possible mutation rate differences among regions and compared centromeric and telomeric loci. We did this using θ_w from the entire species as well as from the average within each subspecies. When θ_w was calculated from both subspecies together, the ratio θ_w/D_{xy} was slightly, but significantly lower at centromeric loci (0.26) compared to telomeric loci (0.31) (Mann–Whitney *U*-test, $P = 0.016$). When θ_w was calculated from the average within each subspecies, the ratio θ_w/D_{xy} was also significantly lower in centromeric loci compared to telomeric loci both in *O. c. algirus* (centromeric $\theta_w/D_{xy} = 0.23$; telomeric $\theta_w/D_{xy} = 0.30$; Mann–Whitney *U*-test, $P = 0.016$) and *O. c. cuniculus* (centromeric $\theta_w/D_{xy} = 0.18$; telomeric $\theta_w/D_{xy} = 0.22$; Mann–Whitney *U*-test, $P = 0.016$). Additionally, we tested the neutral prediction of equal ratios of polymorphism to divergence between centromeric loci and telomeric loci, using the HKA test. We could not reject a neutral model using either polymorphism within the entire species or that within either subspecies ($P > 0.10$ for each). These analyses suggest at most a modest reduction in levels of polymorphism at centromeric compared

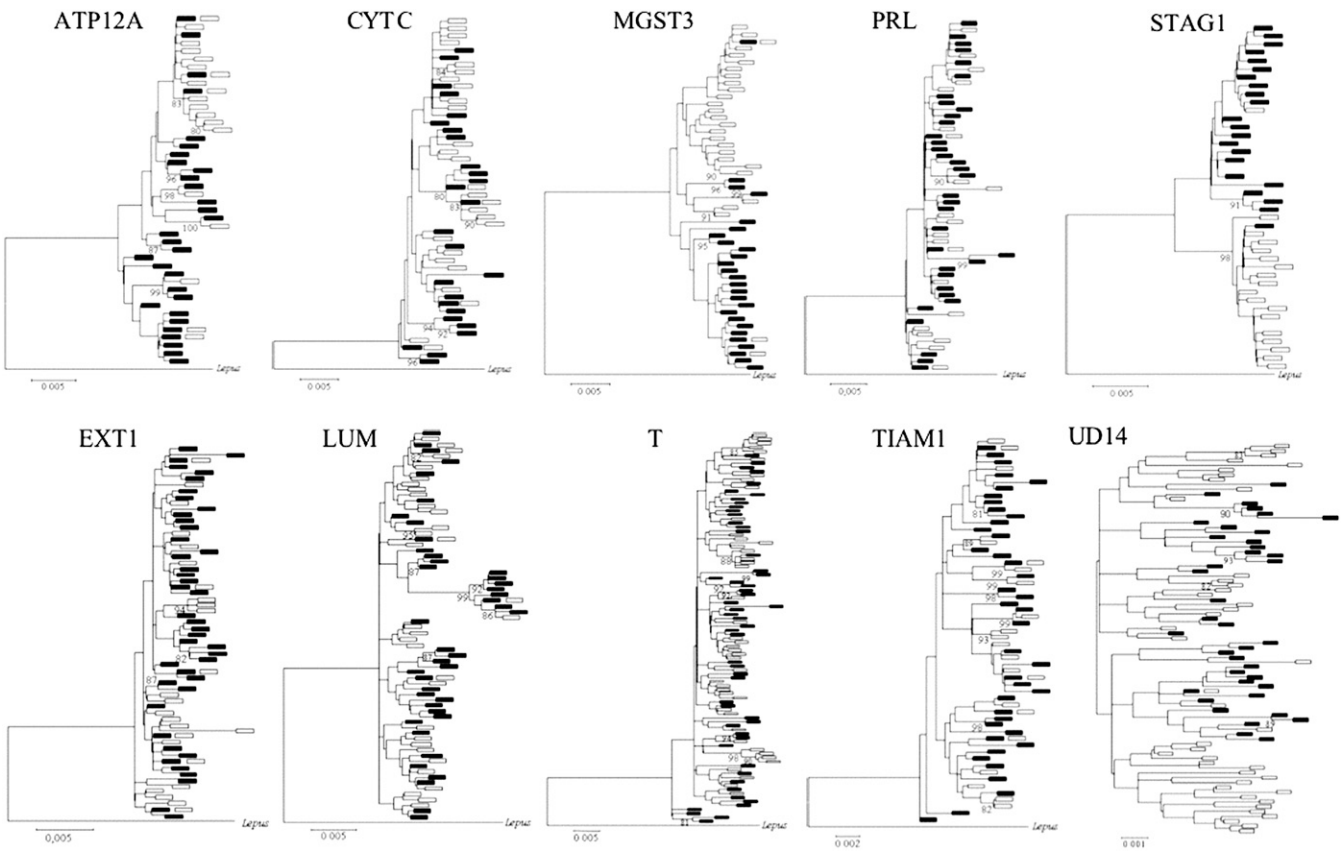

Figure 3.—Haplotype variation illustrated as neighbor-joining trees for each of the 10 autosomal loci. Because recombination events were detected for all loci in our data set, the trees represent overall similarity between the haplotypes. Portions of each gene may have different genealogies. Haplotypes found in *O. c. algirus* and *O. c. cuniculus* are followed by a solid and an open box, respectively, and shared haplotypes by both solid and open boxes. *Lepus granatensis* was used as an outgroup except in UD14. Bootstrap values ≥80 are shown next to branches.

to telomeric loci. Finally, we compared values of Tajima's *D* and Fu and Li's *D** among centromeric and telomeric loci. Both statistics were less negative at centromeric loci (Tajima's $D = -0.877$, Fu and Li's $D^* = -1.256$) compared to telomeric loci (Tajima's $D = -1.082$, Fu and Li's $D^* = -1.646$), although not significantly different (Mann–Whitney *U*-test, $P > 0.5$ for each).

Divergence and LD test of gene flow: Patterns of differentiation were highly variable among loci (Table 3). None of the telomeric loci showed strong differentiation between subspecies. The centromeric loci, in contrast, were heterogeneous. Two centromeric loci (CYTC and PRL) showed little differentiation, while three centromeric loci (ATP12A, MGST3, and STAG1) showed high levels of differentiation. We note that these patterns are not associated with the allele-frequency distribution as measured by Tajima's *D* (Table 2). For example, we observed loci showing high levels of differentiation and loci showing low levels of differentiation among those showing no significant departures from a neutral equilibrium distribution (Tables 2 and 3). The mean F_{ST} for centromeric loci was 30.1% and the mean value for telomeric loci was 6.18%. In particular, F_{ST}-based inferences of *Nm* at three loci were $<<1$. All three of these loci are near centromeres (ATP12A, MGST3, and STAG1). In contrast, the inferred number of migrants per generation for all telomeric loci was $>>1$. Similarly, AMOVA analyses (Excoffier *et al.* 1992) showed that in centromeric genes a higher average proportion of the genetic variation is partitioned between subspecies (28.4%) while for all telomeric loci only a small amount (<9%) of the variation segregated between subspecies. The mean net divergence (Nei 1987) for centromeric loci ($D_a = 0.349\%$) was considerably but not significantly higher than that for telomeric loci ($D_a = 0.054\%$; Mann–Whitney *U*-test, $P > 0.05$).

Another way of examining differentiation is by looking at the phylogeny of alleles for each gene (Figure 3). Striking contrasts were observed among loci. While for two of the centromeric loci (STAG1 and MGST3) there was clear phylogenetic differentiation among subspecies, for the other eight genes there was little or no geographic pattern in the clustering of alleles.

The lack of concordance between phylogeny and geography in recently diverged taxa may be the consequence of either shared ancestral polymorphism or gene flow following species divergence. We evaluated

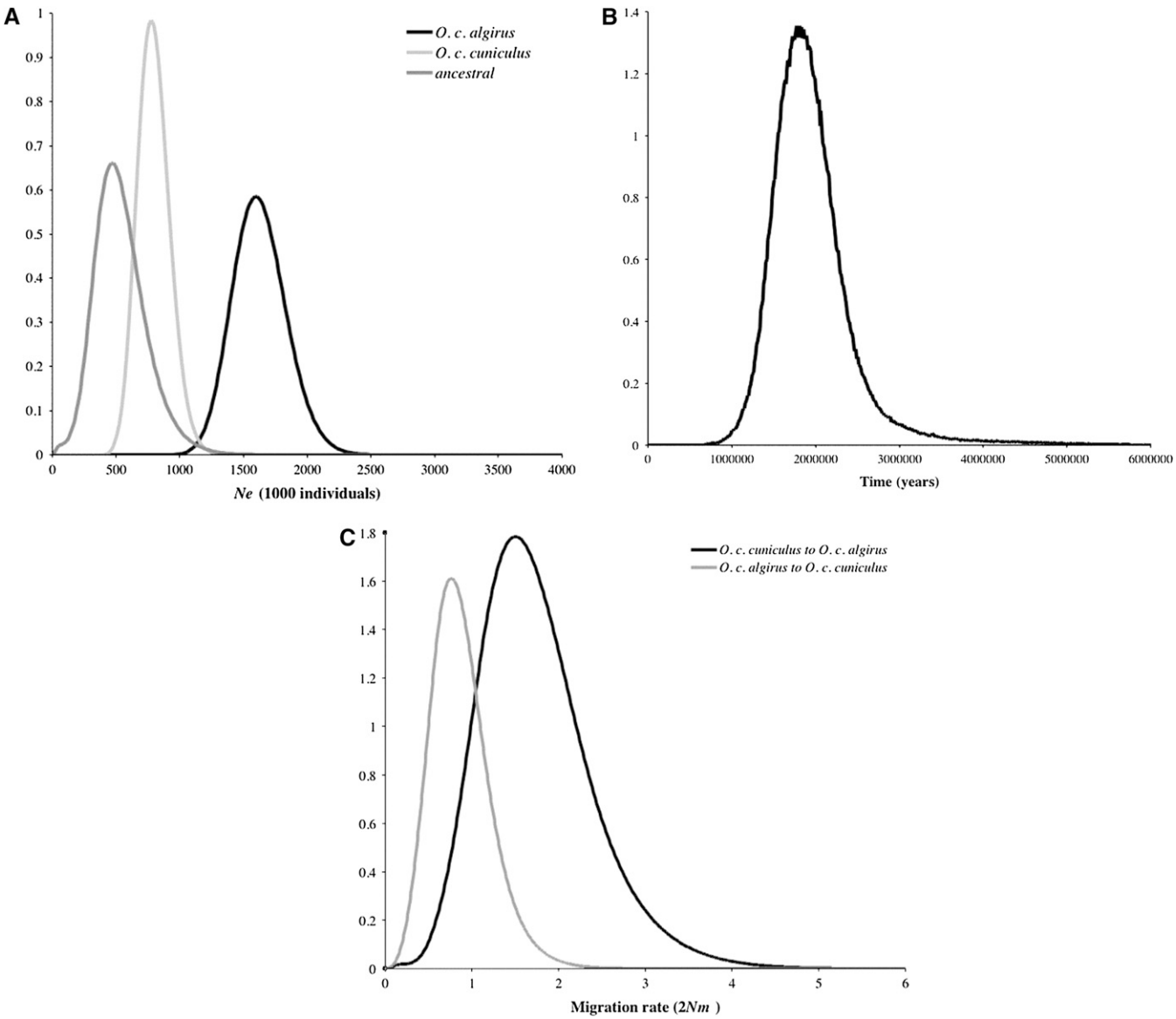

Figure 4.—Marginal posterior probability distributions (*y*-axis) for model parameters (*x*-axis). (A) Effective population size of *O. c. algirus*, *O. c. cuniculus*, and the ancestral population in numbers of thousands of individuals. (B) Divergence time between the two subspecies in years. (C) Migration rates in both directions in number of genes per year.

a strict allopatric mode of speciation with no gene flow, using the test of Machado *et al.* (2002) based on patterns of intragenic LD as implemented in the program WH (Wang *et al.* 1997). The maximum input for WH is a total sample size of 32 chromosomes per locus; therefore, we generated three random data sets with the appropriate sample size, and three independent simulations were performed. Supplemental Table 2 shows the observed and simulated values of x for each locus, calculated using D' (Lewontin 1964). Across the different runs the mean observed values of x were significantly elevated relative to simulated values for *O. c. algirus*. In *O. c. cuniculus*, the mean observed values of x were significantly elevated relative to simulated values for two of the three runs. For each locus separately, the observed values of x were consistently positive in the two subspecies and in a number of genes were significant in at least one of the independent runs (supplemental Table 2). This pattern can be visualized by looking at the genealogies in Figure 3 or the tables of polymorphism in supplemental Figure 1, where identical haplotypes are shared between subspecies. Taken together, these results suggest that gene flow has occurred in both directions with higher gene flow from *O. c. cuniculus* to *O. c. algirus*.

Estimates of ancestral demographic parameters using an isolation-with-migration model: Using a divergence-with-gene-flow framework (Hey and Nielsen 2004, 2007), we obtained maximum-likelihood estimates (MLE) of effective population sizes for both extant and ancestral populations, migration rates in both directions, and time of divergence. We were able to achieve adequate mixing of the Markov chains, and independent runs

TABLE 4

ML estimates (MLE) and 90% posterior density intervals (HPD) of demographic parameters between *O. c. algirus* (alg) and *O. c. cuniculus* (cun)

	$N_{e_{alg}}$ [a]	$N_{e_{cun}}$ [a]	$N_{e_{ancestral}}$ [a]	$2Nm_{alg}$ [b]	$2Nm_{cun}$ [c]	t [d]
MLE	1,598,056	780,367	469,122	1.498	0.762	1,837,126
Lower 90% HPD	1,285,287	587,393	225,025	0.726	0.322	1,282,637
Higher 90% HPD	1,979,435	1,001,675	858,148	2.741	1.400	2,525,667

[a] Effective population size (N_e) in numbers of individuals.
[b] The population migration rate ($2Nm$) into *O. c. algirus* from *O. c. cuniculus*.
[c] The population migration rate ($2Nm$) into *O. c. cuniculus* from *O. c. algirus*.
[d] The time since population split (t) in number of years.

converged on approximate marginal posterior probability distributions for all parameters incorporated in the model (Figure 4). N_e for *O. c. algirus* (1,598,056) and N_e for *O. c. cuniculus* (780,367) were estimated to be larger than N_e for the ancestral population (469,122) (Table 4). We estimated the time since divergence between the two subspecies to be ~1.8 MYA. Both migration rate parameters inferred from the model indicate moderate to high levels of gene flow. A model of strict allopatry with no gene flow was rejected when compared in a likelihood framework to a model allowing gene flow in both directions ($P < 0.001$). Inferred gene flow was higher from *O. c. cuniculus* to *O. c. algirus* ($2Nm = 1.498$) than in the opposite direction ($2Nm = 0.762$). One of the assumptions of this method is selective neutrality. To verify whether the results were affected by skews in the allele-frequency spectrum, we redid the IM analysis excluding those loci that showed significant deviations from neutral expectations. The results obtained were nearly identical (*O. c. algirus* $N_e = 1{,}278{,}610$; *O. c. cuniculus* $N_e = 825{,}743$; ancestral population $N_e = 394{,}255$; $t = 1{,}943{,}816$; *O. c. algirus* $2Nm = 2.040$; *O. c. cuniculus* $2Nm = 0.660$). We also estimated population migration rates for each locus in each direction (Table 3). Not all of these estimates were recovered with confidence either because the tails of the distributions did not reach zero or were relatively flat. Nonetheless, the estimates were remarkably consistent with inferences of gene flow from estimates of F_{ST}. Estimates of $2Nm$ were generally lower for loci near centromeres compared to loci near telomeres (Table 3). Remarkably, estimates of $2Nm$ were >1 for each telomeric locus in at least one direction.

DISCUSSION

We studied sequence variation at 10 autosomal loci in wild populations of the European rabbit (*O. cuniculus*) and found that (1) levels of nucleotide variation are high and reflect large effective population sizes, (2) these subspecies diverged ~1.8 MYA with subsequent high levels of gene flow in both directions, and (3) centromeric loci are generally more differentiated and show higher LD than telomeric loci. This pattern is in agreement with recent speciation models that propose that regions of reduced recombination facilitate divergence in the presence of gene flow.

Levels and patterns of genetic variation and LD at autosomal loci in the European rabbit: This is the first large-scale survey of sequence polymorphism at multiple autosomal loci in the European rabbit. We observed surprisingly high levels of nucleotide diversity: π was 0.746% in *O. c. algirus* and 0.662% in *O. c. cuniculus*. Nucleotide diversity in humans is on the order of 0.1% (Aquadro *et al.* 2001). In mice nucleotide diversity is approximately twice that seen in humans (Baines and Harr 2007). Surprisingly, the values reported here for two rabbit subspecies are closer to observed levels of nucleotide diversity in *Drosophila melanogaster* (Moriyama and Powell 1996). In rabbits, estimates of N_e were 1,600,000 for *O. c. algirus* and 780,000 for *O. c. cuniculus*. These estimates are more than an order of magnitude higher than in humans (Takahata 1993) and several times higher than in mice (Salcedo *et al.* 2007). It is important to bear in mind that estimates of N_e depend on estimates of μ that should be considered approximate owing to the uncertainty of generation length and time of divergence between Lepus and Oryctolagus.

The higher N_e estimated for *O. c. algirus* compared to *O. c. cuniculus* is reflected in higher levels of heterozygosity in the former. Our results are consistent in this regard with previous studies using mtDNA (Branco *et al.* 2002) and X-linked sequences (Geraldes *et al.* 2006), as well as with studies using allozyme (Ferrand and Branco 2007) and microsatellite variation (Queney *et al.* 2001). Additionally, levels of intralocus LD in our data set are lower in *O. c. algirus* than in *O. c. cuniculus*. Since LD at neutral sites is governed by $1/(4N_e c)$, this also suggests that N_e in *O. c. algirus* is larger.

We also estimated N_e of the extant subspecies to be much higher than N_e of the ancestral population. Consistent with this change in N_e, we observed a significant skew in the frequency distribution of polymorphisms toward rare variants compared to the standard neutral expectations. The pattern is observed across loci in both subspecies and is suggestive of a common history of population growth, congruent with the phylogeograph-

ical scenario of a postglacial range expansion of the European rabbit proposed by Branco *et al.* (2002).

Our experimental design allowed us to compare levels of polymorphism and LD between different chromosomal regions. Levels of LD were strongly associated with physical location, with centromeric loci showing higher levels of LD than telomeric loci. This suggests that recombination is most probably suppressed at centromeric loci relative to telomeric loci as observed for other species (*e.g.*, Kong *et al.* 2002). Further support for suppressed recombination near centromeres derives from the ratio ρ/θ. Since $\rho = 4N_e c$ and $\theta = 4N_e \mu$, this ratio controls for differences in effective population size between regions. The observed ratio ρ/θ for telomeric loci is 3.09 and is similar to estimates in other outcrossing species such as Drosophila (Haddrill *et al.* 2005) and wild *Zea mays* (Wright *et al.* 2005). Strikingly, the ratio ρ/θ is sharply reduced at centromeric loci ($\rho/\theta = 0.63$) despite similar estimates of mutation rate for both centromeric and telomeric loci.

Selection at linked sites may lower nucleotide diversity in genomic regions that experience less recombination (Maynard Smith and Haigh 1974; Charlesworth *et al.* 1993). Consistent with this prediction we observed slightly lower polymorphism scaled to divergence in centromeric loci compared to telomeric loci. This adds to a growing body of literature in several other species that shows lower levels of genetic variation in regions of reduced recombination (*e.g.*, Begun and Aquadro 1992; Nachman *et al.* 1998).

The history of divergence and gene flow between *O. c. algirus* and *O. c. cuniculus*: By using multiple loci sampled across the genome it is possible to estimate parameters that provide insight into the history of diverging populations. Using an isolation-with-migration framework we date the initial split between the two rabbit subspecies to be ~1.8 MYA. This estimate is consistent with published mtDNA (Branco *et al.* 2000) and X chromosome data (Geraldes *et al.* 2006) that place the divergence between the two subspecies during the Quaternary glaciations (Branco *et al.* 2000; Geraldes *et al.* 2006). This age of separation mirrors several estimates involving other Iberian taxa (reviewed in Gómez and Lunt 2007) and probably results from the influence of common historical biogeographic events (*e.g.*, glaciations) on Iberian biota (Hewitt 2001).

Although the estimated divergence between *O. c. algirus* and *O. c. cuniculus* is old, several lines of evidence point to the occurrence of gene flow between these groups. First, by fitting an isolation–migration model we estimated high levels of gene flow in both directions, and these estimates were significantly different from zero. Second, we found significantly elevated values of the LD statistic x (Machado *et al.* 2002), reflecting stronger LD than expected in both subspecies compared to an isolation model with no gene flow. Third, there is full haplotype sharing at a number of the surveyed regions (Figure 3 and supplemental Figure 1). Given the high levels of recombination observed at some of the loci, identical haplotypes are probably due to recent instances of gene flow. The IM analysis also suggests that gene flow is asymmetric. Interestingly, higher gene flow seems to occur from the smaller population, *O. c. cuniculus*, to the larger population, *O. c. algirus*. This could be explained by higher vagility in *O. c. cuniculus*, by an asymmetric mate preference between the two subspecies, or by intrinsic incompatibilities in which backcrosses may have higher fitness in one direction than the other.

Mosaic nature of the rabbit genome with respect to gene flow: The most notable aspect of our data is the heterogeneity observed among loci in patterns of differentiation between the two subspecies. This heterogeneity demonstrates the semipermeable nature of the rabbit hybrid zone and supports the idea of a mosaic genome with respect to gene flow (Geraldes *et al.* 2008). In particular, we observe strong differentiation at three of five centromeric loci, which experience less recombination, and little differentiation at all telomeric loci, which experience greater recombination.

This observation fits the predictions of several recent speciation models that suggest that regions of low crossing over might harbor a disproportionate share of loci involved in isolation (Noor *et al.* 2001; Rieseberg 2001; Navarro and Barton 2003). For example, in the model of Noor *et al.* (2001), alleles contributing to isolation initially arise throughout the genome, but, following secondary contact, are purged more easily if they reside in recombining regions but less easily if they reside in regions where recombination is suppressed. This model was based on experiments in which genes contributing to isolation between *D. persimilis* and *D. pseudoobscura* were mapped to small chromosomal rearrangements that differed between these species. Consistent with this model, Machado *et al.* (2002, 2007) showed a reduction in introgression in rearranged regions (with little recombination) compared to colinear regions (which can recombine freely). In sunflowers, some studies have revealed that rates of introgression were significantly reduced in rearranged regions relative to colinear regions (Rieseberg *et al.* 1999), while other studies failed to find such a pattern (Strasburg and Rieseberg 2008). Additionally, studies of the M and S forms of *Anopheles gambiae*, where differentiation appears to be limited to a few regions of the genome (Turner *et al.* 2005), have shown that two of three areas of reduced introgression map close to centromeres and recombination has been shown to be suppressed in those regions (Stump *et al.* 2005; Slotman *et al.* 2006). We also see less introgression in regions with little recombination, and the data presented here are consistent with similar data from the rabbit X chromosome (Geraldes *et al.* 2006). Overall, our observations in rabbit are consistent with models suggesting

that regions of reduced crossing over favor divergence in the face of gene flow (Noor *et al.* 2001; Rieseberg 2001; Navarro and Barton 2003). Importantly, however, not all centromeric regions show reduced introgression. Our results thus help identify particular genomic regions that may contribute to reproductive isolation (*i.e.*, near the centromeres of chromosomes 8, 13, and 14).

A second potential explanation for the higher differentiation in low-recombination regions is the increased likelihood of these regions being in LD with alleles contributing to reproductive isolation (rather than an enrichment for "isolation genes" in such regions). In other words, if genes contributing to isolation are evenly distributed across the genome, but recombination rates vary, the likelihood of a nonrandom association with an isolation gene is greater in regions of lower crossing over. According to this model, isolation genes might occur throughout the genome, but their effects will extend over a greater genomic distance when they occur in regions of low recombination.

It should be possible to distinguish between these two models using introgression lines (*e.g.*, consomic strains). If genes contributing to isolation are found only in low-recombination regions, then only those regions should cause isolation phenotypes when introgressed. In contrast, if genes contributing to isolation are found throughout the genome, some introgression lines containing high-recombination regions should show isolation phenotypes.

A third explanation for the higher differentiation in low-recombination regions is that rates of lineage sorting differ between centromeric and telomeric loci. We observed a slightly lower θ/D_{xy} ratio at centromeric loci compared to telomeric loci, and this may reflect a smaller effective population size for centromeric loci. This effect is modest and probably cannot account for all of the differences in patterns of differentiation among loci. For example, the IMa analysis shows that gene flow has occurred, and the IM analysis provides evidence of lower gene flow at centromeric loci compared to telomeric loci (Table 3). Although ancestral polymorphism may still segregate it cannot explain why some regions remain differentiated despite high levels of admixture.

Finally, although we do not know whether the studied regions are in LD with the centromeres, we note that the observation of increased differentiation near centromeres might also be consistent with recent models of centromere evolution (Henikoff *et al.* 2001; Malik and Henikoff 2002). These authors suggest that rapid evolution of centromeric sequences could lead to higher rates of chromosomal nondisjunction in hybrids and reduced fertility. This model makes a simple, testable prediction: rates of nondisjunction are expected to be higher in F_1's from crosses between subspecies compared to crosses within subspecies. Experiments to test this prediction are currently underway.

We thank R. Villafuerte and M. Delibes-Mateos for providing rabbit samples, P. Tarroso for help with the map in Figure 1, members of the Nachman lab for valuable discussions, and David Rand and two anonymous reviewers for valuable comments on the manuscript. This work was partially supported by a Fundação para a Ciência e a Tecnologia Ph.D. grant to M.C. (SFRH/BD/23786/2005), by a Research Project to N.F. (PTDC/BIA-BDE/72304/2006), and by National Science Foundation and National Institutes of Health grants to M.W.N.

LITERATURE CITED

Aquadro, C. F., V. Bauer DuMont and F. A. Reed, 2001 Genome-wide variation in the human and fruitfly: a comparison. Curr. Opin. Genet. Dev. **11:** 627–634.

Baines, J. F., and B. Harr, 2007 Reduced X-linked diversity in derived populations of house mice. Genetics **175:** 1911–1921.

Barton, N. H., 1979 Gene flow past a cline. Heredity **43:** 333–339.

Barton, N. H., and G. M. Hewitt, 1989 Adaptation, speciation and hybrid zones. Nature **341:** 497–503.

Begun, D. J., and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. Nature **356:** 519–520.

Bengtsson, B. O., 1979 Theoretical models of speciation. Zool. Scr. **8:** 303–304.

Boursot, P., J.-C. Auffray, J. Britton-Davidian and F. Bonhomme, 1993 The evolution of house mice. Annu. Rev. Ecol. Syst. **24:** 119–152.

Branco, M., N. Ferrand and M. Monnerot, 2000 Phylogeography of the European rabbit (Oryctolagus cuniculus) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome b gene. Heredity **85**(4): 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 genetic variation. Evol. Int. J. Org. Evol. **56:** 792–803.

Britton-Davidian, J., F. Fel-Clair, J. Lopez, P. Alibert and P. Boursot, 2005 Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. Biol. J. Linn. Soc. Lond. **84:** 379–393.

Butlin, R. K., 2005 Recombination and speciation. Mol. Ecol. **14:** 2621–2635.

Caisse, M., and J. Antonovics, 1978 Evolution in closely adjacent plant populations. IX. Evolution of reproductive isolation in clinal populations. Heredity **40:** 371–384.

Chantry-Darmon, C., M. Bertaud, C. Urien, S. Chadi-Taourit, M. Perrocheau *et al.*, 2005 Expanded comparative mapping between man and rabbit and detection of a new conserved segment between HSA22 and OCU4. Cytogenet. Genome Res. **111:** 134–139.

Chantry-Darmon, C., C. Urien, H. de Rochambeau, D. Allain, B. Pena *et al.*, 2006 A first-generation microsatellite-based integrated genetic and cytogenetic map for the European rabbit (Oryctolagus cuniculus) and localization of angora and albino. Anim. Genet. **37:** 335–341.

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

Choo, K. H., 1998 Why is the centromere so cold? Genome Res. **8:** 81–82.

Endler, J. A., 1977 *Geographic Variation, Speciation, and Clines.* Princeton University Press, Princeton, NJ.

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

Ewing, B., L. Hillier, M. C. Wendl and P. Green, 1998 Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:** 175–185.

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.

Feder, J. L., J. B. Roethele, K. Filchak, J. Niedbalski and J. Romero-Severson, 2003 Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. Genetics **163:** 939–953.

Felsenstein, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evol. Int. J. Org. Evol. **39:** 783–791.

Ferrand, N., and M. Branco, 2007 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 Southern European Refugia*, edited by S. Weiss and N. Ferrand. Springer, Amsterdam.

Fu, Y. X., and W. H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

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

Geraldes, A., M. Carneiro, M. Delibes-Mateos, R. Villafuerte, M. W. Nachman *et al.*, 2008 Reduced introgression of the Y chromosome between subspecies of the European rabbit (Oryctolagus cuniculus) in the Iberian Peninsula. Mol. Ecol. **17:** 4489–4499.

Gómez, A., and D. H. Lunt, 2007 Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula, pp. 155–188 in *Phylogeography of Southern European Refugia*, edited by S. Weiss and N. Ferrand. Springer, Amsterdam.

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

Haddrill, P. R., K. R. Thornton, B. Charlesworth and P. Andolfatto, 2005 Multilocus patterns of nucleotide variability and the demographic and selection history of Drosophila melanogaster populations. Genome Res. **15:** 790–799.

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.

Harrison, R. G., 1990 Hybrid zones: windows on evolutionary process. Oxf. Surv. Evol. Biol. **7:** 69–128.

Henikoff, S., K. Ahmad and H. S. Malik, 2001 The centromere paradox: stable inheritance with rapidly evolving DNA. Science **293:** 1098–1102.

Hewitt, G. M., 2001 Speciation, hybrid zones and phylogeography—or seeing genes in space and time. Mol. Ecol. **10:** 537–549.

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 R. Nielsen, 2007 Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. USA **104:** 2785–2790.

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

Hudson, R. R., 2001 Two-locus sampling distributions and their application. Genetics **159:** 1805–1817.

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. Aguade, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

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

Kimura, M., 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics **61:** 893–903.

Kirkpatrick, M., and N. Barton, 2006 Chromosome inversions, local adaptation and speciation. Genetics **173:** 419–434.

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 The 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.

Machado, C. A., T. S. Haselkorn and M. A. Noor, 2007 Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. Genetics **175:** 1289–1306.

Malik, H. S., and S. Henikoff, 2002 Conflict begets complexity: the evolution of centromeres. Curr. Opin. Genet. Dev. **12:** 711–718.

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 favourable gene. Genet. Res. **23:** 23–35.

Mayr, E., 1963 *Animal Species and Evolution*. Harvard University Press, Cambridge, MA.

McVean, G., P. Awadalla and P. Fearnhead, 2002 A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics **160:** 1231–1241.

Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in Drosophila. Mol. Biol. Evol. **13:** 261–277.

Nachman, M. W., V. L. Bauer, S. L. Crowell and C. F. Aquadro, 1998 DNA variability and recombination rates at X-linked loci in humans. Genetics **150:** 1133–1141.

Navarro, A., and N. H. Barton, 2003 Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evol. Int. J. Org. Evol. **57:** 447–459.

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

Nickerson, D. A., V. O. Tobe and S. L. Taylor, 1997 PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. **25:** 2745–2751.

Nielsen, R., and J. Wakeley, 2001 Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics **158:** 885–896.

Noor, M. A., 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.

Noor, M. A., D. A. Garfield, S. W. Schaeffer and C. A. Machado, 2007 Divergence between the *Drosophila pseudoobscura* and *D. persimilis* genome sequences in relation to chromosomal inversions. Genetics **177:** 1417–1428.

Patterson, N., D. J. Richter, S. Gnerre, E. S. Lander and D. Reich, 2006 Genetic evidence for complex speciation of humans and chimpanzees. Nature **441:** 1103–1108.

Putnam, A. S., J. M. Scriber and P. Andolfatto, 2007 Discordant divergence times among Z-chromosome regions between two ecologically distinct swallowtail butterfly species. Evol. Int. J. Org. Evol. **61:** 912–927.

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

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.

Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon *et al.*, 2003 Sex chromosome evolution and speciation in Ficedula flycatchers. Proc. Biol. Sci. **270:** 53–59.

Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Salcedo, T., A. Geraldes and M. W. Nachman, 2007 Nucleotide variation in wild and inbred mice. Genetics **177:** 2277–2291.

Sambrook, J., and D. W. Russell, 2001 *Molecular Cloning: A Laboratory Manual*, Ed. 3. 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.

Sharples, C. M., J. E. Fa and D. J. Bell, 1996 Geographic variation in size in the European rabbit *Oryctolagus cuniculus* (Lagomorpha: Leporidae) in western Europe and North Africa. Zool. J. Linn. Soc. **117:** 141–158.

Slotman, M. A., L. J. Reimer, T. Thiemann, G. Dolo, E. Fondjo *et al.*, 2006 Reduced recombination rate and genetic differentiation between the M and S forms of *Anopheles gambiae* s.s. Genetics **174:** 2081–2093.

Stephens, M., and P. Donnelly, 2003 A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. **73:** 1162–1169.

Stephens, M., N. J. Smith and P. Donnelly, 2001 A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. **68:** 978–989.

Strasburg, J. L., and L. H. Rieseberg, 2008 Molecular demographic history of the annual sunflowers Helianthus annuus and H. petiolaris—large effective population sizes and rates of long-term gene flow. Evolution **62:** 1936–1950.

Stump, A. D., M. C. Fitzpatrick, N. F. Lobo, S. Traore, N. Sagnon *et al.*, 2005 Centromere-proximal differentiation and speciation in Anopheles gambiae. Proc. Natl. Acad. Sci. USA **102:** 15930–15935.

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

Takahata, N., 1993 Allelic genealogy and human evolution. Mol. Biol. Evol. **10:** 2–22.

Tamura, K., J. Dudley, M. Nei and S. Kumar, 2007 MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:** 1596–1599.

Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau *et al.*, 2008 Genome-wide patterns of gene flow across a house mouse hybrid zone. Genome Res. **18:** 67–76.

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

Villafuerte, R., 2002 *Oryctolagus cuniculus*, pp. 464–466 in *Atlas de los Mamíferos Terrestres de España*, edited by L. J. Palomo and J. Gisbert. Dirección General de Conservación de la Naturaleza-SECEM-SECEMU, Madrid.

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.

Woerner, A. E., M. P. Cox and M. F. Hammer, 2007 Recombination-filtered genomic datasets by information maximization. Bioinformatics **23:** 1851–1853.

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

Wright, S. I., I. V. Bi, S. G. Schroeder, M. Yamasaki, J. F. Doebley *et al.*, 2005 The effects of artificial selection on the maize genome. Science **308:** 1310–1314.

Wu, C.-I., 2001 The genic view of the process of speciation. J. Evol. Biol. **14:** 851–865.

Yatabe, Y., N. C. Kane, C. Scotti-Saintagne and L. H. Rieseberg, 2007 Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. Genetics **175:** 1883–1893.

Communicating editor: D. M. Rand