

The Genetic Structure of Domestic Rabbits

Miguel Carneiro,^{*,1,2,3} Sandra Afonso,¹ Armando Geraldes,⁴ Hervé Garreau,⁵ Gerard Bolet,⁵ Samuel Boucher,⁶ Aurélie Tircazes,⁵ Guillaume Queney,⁷ Michael W. Nachman,³ and Nuno Ferrand^{1,2}

¹CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Vairão

²Departamento de Biologia da Faculdade de Ciências, Universidade do Porto, Porto, Portugal

³Department of Ecology and Evolutionary Biology, University of Arizona

⁴Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada

⁵Institut national de la recherche agronomique, UR 631 Station d'amélioration génétique des animaux, Castanet-Tolosan Cedex, France

⁶Labovet Conseil, réseau Cristal, Les Herbiers, France

⁷Antagene, Wildlife Genetics Laboratory, Limonest, Lyon, France

*Corresponding author: E-mail: miguel.carneiro@mail.icav.up.pt.

Associate editor: Naoko Takezaki

Abstract

Understanding the genetic structure of domestic species provides a window into the process of domestication and motivates the design of studies aimed at making links between genotype and phenotype. Rabbits exhibit exceptional phenotypic diversity, are of great commercial value, and serve as important animal models in biomedical research. Here, we provide the first comprehensive survey of nucleotide polymorphism and linkage disequilibrium (LD) within and among rabbit breeds. We resequenced 16 genomic regions in population samples of both wild and domestic rabbits and additional 35 fragments in 150 rabbits representing six commonly used breeds. Patterns of genetic variation suggest a single origin of domestication in wild populations from France, supporting historical records that place rabbit domestication in French monasteries. Levels of nucleotide diversity both within and among breeds were $\sim 0.2\%$, but only 60% of the diversity present in wild populations from France was captured by domestic rabbits. Despite the recent origin of most breeds, levels of population differentiation were high ($F_{ST} = 17.9\%$), but the majority of polymorphisms were shared and thus transferable among breeds. Coalescent simulations suggest that domestication began with a small founding population of less than 1,200 individuals. Taking into account the complex demographic history of domestication with two successive bottlenecks, two loci showed deviations that were consistent with artificial selection, including GPC4, which is known to be associated with growth rates in humans. Levels of diversity were not significantly different between autosomal and X-linked loci, providing no evidence for differential contributions of males and females to the domesticated gene pool. The structure of LD differed substantially within and among breeds. Within breeds, LD extends over large genomic distances. Markers separated by 400 kb typically showed r^2 higher than 0.2, and some LD extended up to 3,200 kb. Much less LD was found among breeds. This advantageous LD structure holds great promise for reducing the interval of association in future mapping studies.

Key words: domestication, rabbit, nucleotide diversity, bottleneck, artificial selection, linkage disequilibrium.

Introduction

The domestication of both plants and animals is one of the most notable “experiments” ever conducted in biology. The modification of genomes under artificial selection revolutionized human societies and has attracted the attention of biologists for two main reasons. First, the close association of domestic species and humans has motivated genetic and archaeological studies aimed at better understanding the historical and cultural conditions, as well as the biological requirements, underlying the transformation of a wild species into its domesticated relative. Second, the immense phenotypic diversity that commonly segregates in domestic species provides exceptional opportunities to establish specific genotype/phenotype associations and to study the general mechanisms by which genetic variation governs phenotypic diversity.

The domestic rabbit is one of the most recently domesticated species (most likely within the last 1,500 years; see below) and is characterized by an exceptionally high phenotypic diversity with more than 200 breeds recognized worldwide (Whitman 2004). Breeds vary extensively in weight, body conformation, fur type, coat color, and ear length, and this visible morphological variation dramatically exceeds the phenotypic diversity of their wild counterparts (fig. 1). Size variation in domestic rabbits is greater than that in the entire Leporidae family. Rabbit breeds also differ extensively in litter size, growth rate, and behavior. This immense phenotypic diversity is reflected in a wide variety of commercial and laboratory uses (Weisbroth 1974; Lindsey and Fox 1994; Lebas et al. 1997; Bosze et al. 2003; Fan and Watanabe 2003; Houdebine et al. 2009; Rogel-Gaillard et al. 2009). Commercial uses include

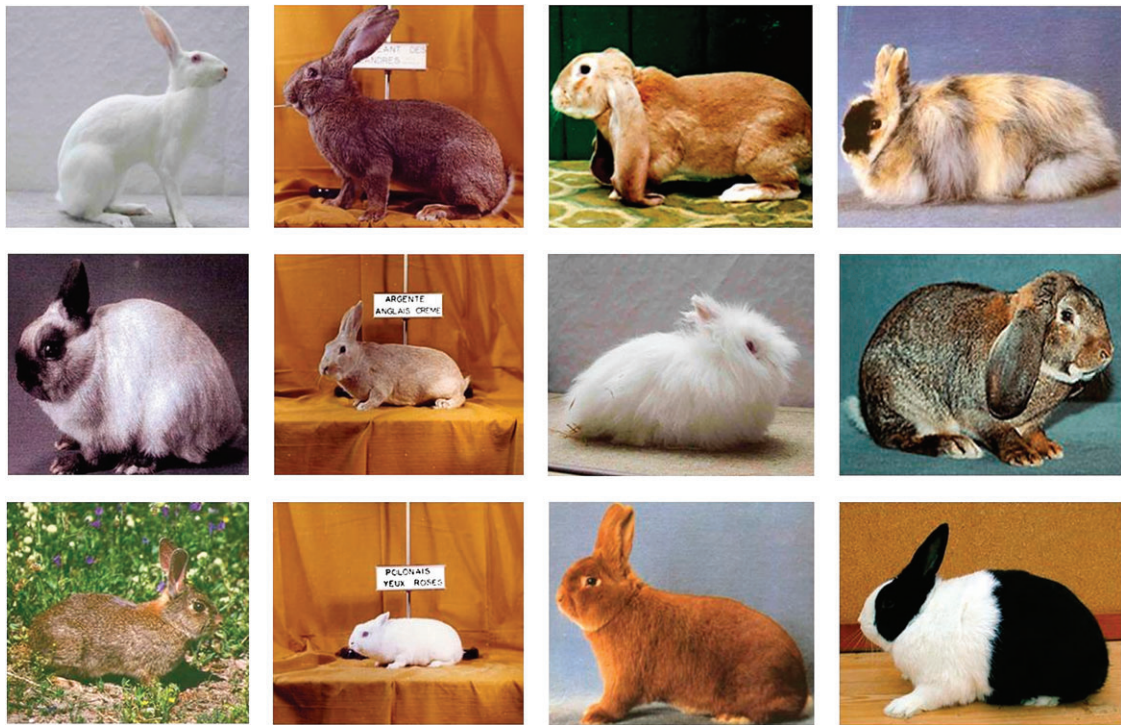


FIG. 1. Phenotypic variation in domestic and wild rabbits. The bottom left picture shows the typical phenotype of a wild rabbit from the subspecies *O. c. cuniculus* (kindly provided by P.C. Alves). All the other pictures illustrate the phenotypic diversity observed for numerous traits in domestic rabbits (kindly provided by François Lebas, <http://www.cuniculture.info>, and references therein).

the production of meat, fur, wool, and therapeutic proteins, and numerous breeds are raised specifically as pets. Moreover, rabbits have many hereditary diseases common to humans (e.g., aortic arteriosclerosis, cataracts, hypertension, hypertrophic cardiomyopathy, epilepsy, spina bifida, osteoporosis, and many more), making them a valuable model in both biomedical and fundamental research. The rabbit is also commonly used in studies of *in vitro* fertilization, embryology, organogenesis, and toxicology.

The European rabbit (*Oryctolagus cuniculus*) is the single recognized progenitor of domestic rabbits. This species is native to the Iberian Peninsula, where two subspecies that diverged ~ 1.8 Ma (Carneiro et al. 2009) are found: *O. c. algirus* is present in the southwestern Iberian Peninsula, whereas *O. c. cuniculus* is present in the northeastern Iberian Peninsula and in France (supplementary fig. 1, Supplementary Material online). These subspecies are well differentiated genetically, although there is good evidence for gene flow between them following secondary contact since the Pleistocene, such that at some loci alleles from *O. c. algirus* are found in *O. c. cuniculus* and vice versa (e.g., Carneiro et al. 2009). This secondary contact likely predated the domestication of rabbits by thousands of years. France was colonized through natural dispersal from Northern Spain likely after the last glacial maximum (Queney et al. 2001). Rabbits have also been introduced by people to many places throughout the world (Flux and Fullagar 1983). Rabbits were first transported around the Mediterranean by Phoenician traders, later introduced in the British Isles and other islands in the northeast Atlantic in the middle ages, and worldwide after the 18th century (e.g.,

Australia, Chile, New Zealand, North America, and South Africa). In spite of its recent domestication, conflicting historical records suggest different geographic origins for the domestic rabbit. Some historical records indicate that rabbits were first kept in captivity in the Iberian Peninsula during the Roman occupation in the first century BC, where they were maintained in large enclosures for meat production (Clutton-Brock 1999). Because no selective breeding was apparently implemented, this might not be considered a true domestication. Other historical records, however, suggest that true domestication, including taming and selective breeding for traits of interest, likely began about AD 600 in French monasteries, following a decree by Pope Gregory the Great that new born rabbits were not considered meat and so could be eaten during Lent (Clutton-Brock 1999; Callou 2003; Whitman 2004). Unlike many other domesticated species (Bruford et al. 2003; Dobney and Larson 2006), the availability of wild populations allows a direct comparison with the domesticated population, and limited genetic evidence using mitochondrial DNA, Y chromosome, and microsatellite data, tentatively suggests a unique source for rabbit domestication in France (Hardy et al. 1995; Queney et al. 2002; Geraldès et al. 2005; reviewed in Ferrand and Branco 2007). It is also clear from this small set of markers that domestic rabbits harbor less genetic diversity than their wild counterparts. After the recorded initiation of the rabbit domestication in French monasteries, it was not until much later that most breeds originated. By the 16th century, several varieties of different sizes and coat colors were recorded in France, Italy, Flanders, and England, but the development of the majority

of breeds dates back no further than 200 years (Lebas et al. 1997; Whitman 2004).

Despite its scientific and economic value, large-scale efforts to understand the impact of the domestication process upon the rabbit genome are lacking. To fill this gap, we provide the first comprehensive report on levels and patterns of nucleotide polymorphism in domestic rabbits. We investigated the demographic history of domestication by comparing genetic variation between wild populations and a representative set of 14 breeds at multiple autosomal and X-linked loci. To guide future mapping studies, we also provide the first description of the scale of linkage disequilibrium (LD) in rabbits by resequencing 35 fragments in six different breeds. We focus on five main questions. 1) Where were rabbits domesticated? 2) What was the impact of the domestication process on levels and patterns of genetic variation within and among domestic rabbit breeds? 3) What was the magnitude of the domestication bottleneck and what proportion of the genetic variability present in wild rabbits was captured by the rabbit domestication process? 4) Do patterns of variation at autosomal and X-linked loci suggest unequal contributions of males and females to the domesticated gene pool? 5) What is the extent of LD among and within rabbit domestic breeds?

Materials and Methods

Sampling Strategy

Our sampling design was divided into two main parts. First, because this work was largely motivated by a comparison between domestic and wild rabbits, a set of individuals was selected to represent a broad range of the genetic diversity found both in domestic and in wild populations. To maximize genetic diversity among domestic rabbits and because there is not enough genetic data to infer relationships between breeds, we selected breeds that represent a wide range of phenotypes and for which historical records indicate an old and mostly unrelated origin. The rationale behind this choice is that 1) high phenotypic divergence may reflect higher genetic divergence and 2) older breeds may represent reservoirs of genetic diversity. In fact, recently created breeds were mostly developed through crosses between older breeds with subsequent selection (Whitman 2004). One or two individuals were sampled for each of the following 14 breeds (supplementary table 1, Supplementary Material online): Belgian Hare, Champagne Silver, Chinchilla (standard), English Lop, English Spot, Fauve de Bourgogne, French Angora, Flemish Giant, Himalayan, Hungarian Giant, New Zealand White, Polish, Thuringer, and Vienna White. By using a small number of individuals for a set of old and phenotypically divergent breeds, we expected to 1) approximate the ancestral genetic diversity present in domestic rabbits before modern breeding practices were implemented and 2) sample lineages that represent most of the genetic variation captured in the early years of domestication. In the wild, we surveyed a total of 15 individuals from five localities in France

(supplementary table 1, Supplementary Material online), and we also combined published sequence data for representative samples of diverse geographic localities corresponding to the range of both rabbit subspecies in the Iberian Peninsula and for a *Lepus granatensis* sample, which was used as outgroup in this study (supplementary fig. 1, Supplementary Material online; Carneiro et al. 2009; Carneiro et al. 2010). Thus, we compared four groups: *Oryctolagus cuniculus algirus* ($n = 10$, OCA), *O. c. cuniculus* from Iberian Peninsula ($n = 12$, OCCIP), *O. c. cuniculus* from France ($n = 15$, OCCFR), and domesticated rabbits ($n = 25$, DOM).

Second, to investigate patterns of genetic variation and LD within and among rabbit breeds, we sampled 25 individuals for each of six different breeds (supplementary table 1, Supplementary Material online): Champagne Silver, English Spot, French Angora, French Lop, New Zealand White (INRA 1077 strain), and Rex. To help guide future research in rabbits, we included breeds that reflect the various uses of domestic rabbits and are among the most used breeds in agronomic and scientific research. Angora and Rex are primarily used for wool and fur. Champagne Silver, French Lop, and English Spot are used as pets or exhibition animals; however, the first two are dual-purpose breeds and are often used for meat production. New Zealand White is the most widely used breed in laboratory research and intensive meat productions.

Selection of Loci

To determine levels and patterns of nucleotide variation in domestic and French wild rabbits, we amplified and resequenced nine autosomal and seven X-linked introns (supplementary table 2, Supplementary Material online). These were chosen from a set of loci used in previous studies on wild rabbits from the Iberian Peninsula (Carneiro et al. 2009; Carneiro et al. 2010), allowing direct comparison of all four groups (OCA, OCCIP, OCCFR, and DOM). Loci were chosen without any a priori knowledge of functional relevance to domestication traits. Autosomal loci were chosen from different chromosomes or far apart on the same chromosome (chromosomes 4 and 14). On the X, loci were chosen at regular intervals along the chromosome.

To study LD and to compare patterns of genetic variation within and among breeds, we generated sequences for five fragments in each of seven genomic regions for a total of 35 fragments (supplementary table 2, Supplementary Material online). Regions were selected on different chromosomes far from centromeric regions, where recombination has been shown to be highly reduced for several species (e.g., Kong et al. 2002). For each of the seven regions, five different segments of 250 bp to 700 bp in length were sequenced at distances of 50, 400, 1,200 and 3,200 kb in one direction from the initial selected fragment. Primers were designed based on the sequence of the rabbit genome and are provided in supplementary table 3, Supplementary Material online, along with polymerase chain reaction conditions. Sequencing was carried out from both directions for

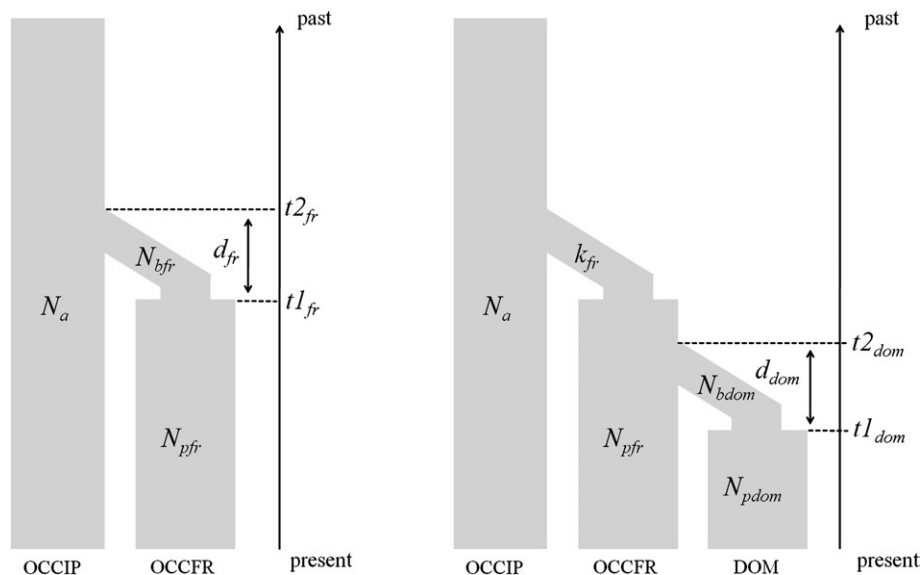


Fig. 2. Schematic representation of the coalescent models used to represent the main events of the demographic history of French wild rabbits (left) and domestic rabbits (right). See Materials and Methods for a detailed description of all parameters and assumptions.

the nine autosomal and seven X-linked loci and from a single direction for the LD study using an ABI 3700 automated sequencer.

Data Analysis

Resequencing Data Set

Sequence traces were base called, trimmed, assembled into contigs, and scanned for heterozygotes 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 provided by August Woerner. All traces and heterozygote calls were visually inspected and edited when necessary. Sequence data have been submitted to GenBank (accession numbers: HM155112–HM155727 and HM454302–HM459434).

Haplotype Inference, Levels of Variation, Tests of Neutrality, and Gene Genealogies

Intralocus haplotypes were inferred using the computer program PHASE 2.1 (Stephens et al. 2001; Stephens and Donnelly 2003). This Bayesian approach assumes a neutral coalescent within a single unstructured population of constant size and accommodates recombination. We applied the algorithm independently for OCA, OCCIP, OCCFR, and DOM. For X-linked loci, we specified that the phase of male haplotypes was known.

Most population genetic parameters were obtained using the computer program SITES (Hey and Wakeley 1997). We estimated the number of segregating sites (S), number of haplotypes (n haps), Hudson's minimum number of recombination events (R_m , Hudson and Kaplan 1985), and the neutral mutation parameter using Watterson's θ_w (Watterson 1975) and π (Nei 1987). We summarized the level of genetic differentiation between populations by estimating the fixation index (F_{ST} , Hudson et al. 1992), the net nucleotide divergence (D_a , Nei 1987), and the average pairwise differences between populations (D_{xy} , Nei 1987).

Hardy–Weinberg equilibrium (HWE) was evaluated using the function “HWE.exact” implemented in the “genetics” package (Warnes and Leisch 2005) in R statistical software (R Development Core Team 2009). The “diseq” function was used to estimate the within-population inbreeding coefficient, which is defined as the correlation between alleles within an individual (Weir et al. 2004).

The frequency distribution of polymorphic sites was compared with neutral expectations using Tajima's D (Tajima 1989). To generate the null distribution of this statistic, we used coalescent simulations using DnaSP 4.50.3 (Rozas et al. 2003). We ran 10^4 coalescent simulations of the standard neutral model conditioned on the sample size and the observed estimates of θ_w . To detect deviations in the number of observed haplotypes compared with the standard neutral model, we used Strobeck's S statistic (Strobeck 1987). The neutral prediction of a constant ratio of polymorphism to divergence was assessed with the Hudson–Kreitman–Aguadé (HKA) test (Hudson et al. 1987). We performed multilocus tests using the HKA program by Jody Hey.

Evolutionary relationships among alleles were estimated using median-joining networks (Bandelt et al. 1999), as implemented in the software NETWORK 4.2.0.1.

Demographic Inference Using Coalescent Simulations

We used computer simulations of the coalescent process to investigate the demographic history of domestic rabbits. The simulations were performed using the algorithm described by Hudson (1990) and implemented in the program ms (Hudson 2002). Statistics were computed using msstats (Thornton 2003).

Diagrams and parameters of the demographic models used in simulations are depicted in figure 2 and are similar to those applied to other domesticated species (Eyre-Walker et al. 1998; Tenaillon et al. 2004; Wright et al. 2005; Haudry et al. 2007; Zhu et al. 2007). According to historical records and genetic evidence (see Introduction

and Results and Discussion), domestic rabbits apparently originated from OCCFR; therefore, we assumed this population as the source of DOM. The purpose of our simulations was to infer the severity of the domestication bottleneck ($k_{\text{dom}} = N_{\text{bdom}}/d_{\text{dom}}$), where N_{bdom} is the size of the bottlenecked population and d_{dom} the duration of the bottleneck in generations. Due to the recent colonization of France—likely after the last glacial maximum—and consequent bottleneck (see Introduction and Results and Discussion), we incorporated this demographic event in our simulations (k_{fr}). The model mimicking the colonization of France consisted of two populations and approximates a bottleneck event from a single source population in which an ancestral population of constant size (N_a) gives rise at time $t_{2\text{fr}}$ to a small founder population (N_{bfr}). More recently, the bottlenecked population at time $t_{1\text{fr}}$ ($d_{\text{fr}} = t_{2\text{fr}} - t_{1\text{fr}}$) instantaneously expands into its current size (N_{pfr}). The model used to estimate the severity of the bottleneck effect in the domestication event of DOM is similar and described by similar parameters (N_{bdom} , N_{pdom} , d_{dom} , and k_{dom}) but included a third population (DOM) and incorporated the estimated parameter k_{fr} for the formation of the OCCFR population from the first model. This model can be briefly summarized as two consecutive bottlenecks followed by growth during the colonization of France by wild rabbits and at domestication. In both models, populations remain isolated after the initial split.

We used a number of assumptions. First, neutral mutation rates were determined for each locus from the relationship $\mu = D/2T$, where D is the estimated D_{xy} for the *Oryctolagus/Lepus* comparison (values available in Carneiro et al. 2010) and T is the divergence time for the separation between the two genera, which is thought to have occurred 11.8 Ma (Matthee et al. 2004). For UD14, we had no outgroup sequence and μ was calculated as the average among the other autosomal loci. Second, given the estimated neutral mutation rate, N_a prior to the split between OCCIP and OCCFR was estimated for each locus from the population mutation parameter (θ) estimated for OCCIP. Note that $\theta = 3N_e\mu$ for X-linked loci and $\theta = 4N_e\mu$ for autosomal loci. Third, we assumed a split time between OCCIP/OCCFR and OCCFR/DOM of 10,000 (after the last glacial maximum) and 1,400 years (the initial domestication event), respectively (see Introduction). We assumed a generation time of 1 year (Soriguer 1983). Although split times should be considered approximate, we performed additional simulations for a subset of the loci varying the split time 3-fold in both directions and found that estimates of k_{fr} and k_{dom} for each locus remained similar (data not shown). This result also indicates that slight uncertainties in mutation rate estimates and generation time are not likely to change our conclusions, which is not surprising because this is a short time scale for mutation to have a substantial impact. Fourth, similarly to N_a , values of N_{pfr} and N_{pdom} were derived from the population mutation parameter. We further varied N_{pfr} and N_{pdom} by one order of magnitude in both directions and the results remained largely unaltered (data not shown). Fifth, we included intragenic recombination in

all simulations. The population recombination parameter ($\rho = 4N_e r$) was estimated for each locus from the OCCIP data using γ (Hey and Wakeley 1997). γ is a maximum likelihood (ML) estimator developed using a coalescent model for a sample of four DNA sequences with recombination. Finally, it has been previously shown for similar protocols that the parameter k is a compound parameter and N_b and d are positively correlated (Eyre-Walker et al. 1998; Tenaillon et al. 2004). This tendency was also observed in our simulations. For this reason, we present the results of d_{fr} and d_{dom} fixed at 500 generations for which we used several values of N_{bfr} and N_{bdom} so that the parameters k_{fr} and k_{dom} ranged between 0.1 and 16 with an interval of 0.1.

To condition the simulations on the observed data, we used rejection sampling based on three summary statistics: the average pairwise nucleotide divergence per sequence (π), the number of segregating sites (S), and the number of haplotypes ($nhaps$). By conditioning the simulated values on the observed data, we avoid biasing our estimates by features of the demographic history that are not considered in our model. For the demographic scenario reproducing the colonization of France, we fitted the simulated data to the observed sequence data of OCCIP, whereas for the domestication model, we simulated the OCCIP as the ancestral population and forced the simulations to the observed data of OCCFR. Simulations were recorded if the values obtained were within 20% of the observed values and run until 10,000 genealogies were obtained.

To find the best-fitting parameter value of k_{fr} and k_{dom} , the likelihood value for each locus was estimated as the proportion of 10,000 simulations whose summary statistics (S , π , and $nhaps$) were contained within 20% of the observed values. The ML value for the global data set was estimated only for k_{dom} as the product of the likelihood for each locus. We calculated 95% confidence intervals (CIs) around the ML estimate of k_{dom} by determining the value for which the log-likelihood values were two log-likelihood units lower than the best estimate. Finally, we used the multilocus ML value of k_{dom} to perform additional coalescent simulations at each locus to investigate if some loci deviated significantly from the neutral demographic model inferred. We generated an empirical distribution with 10,000 replicates and asked whether levels of diversity for a specific locus, measured as S , π , and $nhaps$, were lower than what would be expected ($P < 0.05$) given the inferred multilocus estimate of k_{dom} .

LD Analysis

Polymorphic single nucleotide polymorphisms (SNPs) or insertion and deletion (indel) markers having more than 10% of missing data and/or minor allelic frequencies within a breed lower than 10% were excluded from the analysis. LD was measured using the genotypic r^2 (Weir et al. 2004), which does not rely on any assumption about the allelic frequencies in the sample. This statistic was computed for all pairwise comparisons between markers using the

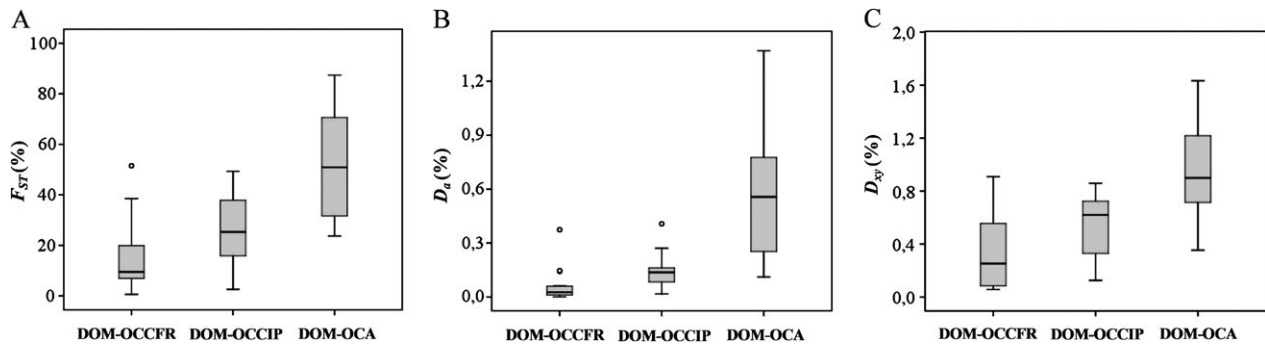


FIG. 3. Box plots depicting genetic differentiation between domestic rabbits and each of the wild rabbit populations considered in this study. Levels of genetic differentiation between domestic rabbits and wild rabbits were estimated using the fixation index (F_{ST} , A), the net nucleotide divergence (D_a , B), and the average pairwise differences between populations (D_{xy} , C). Levels of genetic differentiation suggest a closer proximity between domestic rabbits (DOM) and French wild rabbits (OCCFR).

“LD” function as implemented in R “genetics” package (Warnes and Leisch 2005).

Historical effective population size estimates were obtained from the expectation of LD decay with genetic distance using the model $E(r^2) = (1/4N_e c + 1) + (1/n)$ (Sved 1971), where $E(r^2)$ is the average of the r^2 values between markers across the seven genomic regions for a given distance, N_e is the effective population size at t generations in the past, c is the genetic distance between markers in Morgans, and n is the number of chromosomes sampled. We assumed $t = 1/(2c)$ (Hayes et al. 2003) and an average recombination rate of $1 \text{ cM} = 1 \text{ Mb}$ based on the total genetic map length in rabbits (2766.6 cM excluding chromosomes 20, 21, and X) and a genome size of approximately 3 Gb (Chantry-Darmon et al. 2006).

Results and Discussion

A Single Geographical Origin of Domestication

Several previous studies in both plants and animals have documented multiple centers of domestication for the same species (Diamond 2002; Bruford et al. 2003). To address the origins of rabbit domestication, we resequenced nine autosomal and seven X-linked introns varying in length between 705 bp and 1,314 bp (supplementary table 2, Supplementary Material online). In total, we obtained ~ 16.1 kb of sequence per individual (~ 10.8 kb for autosomal loci and ~ 5.3 kb for X-linked loci). Levels of differentiation and gene genealogies were contrasted between native range wild rabbits and a representative set of 14 domestic rabbit breeds (supplementary table 1, Supplementary Material online). Throughout this study, we considered four populations: OCA, OCCIP, OCCFR, and DOM.

Various aspects of the data were consistent in revealing a closer relationship between OCCFR and DOM than between DOM and wild rabbits from the Iberian Peninsula (OCA and OCCIP). First, genetic differentiation between DOM and OCCFR ($F_{ST} = 15.0\%$, $D_a = 0.060\%$, $D_{xy} = 0.341\%$, fig. 3; supplementary table 4, Supplementary Material online) was significantly lower than that between DOM and OCA ($F_{ST} = 52.1\%$, $D_a = 0.556\%$, $D_{xy} = 0.979\%$) or DOM and OCCIP ($F_{ST} = 25.6\%$, $D_a =$

0.142% , $D_{xy} = 0.547\%$) ($P < 0.025$ in all cases by Wilcoxon signed-rank test). This accords with the findings for microsatellites (Queney et al. 2002). Second, a simple visual inspection of the gene genealogies in figure 4 also suggests that haplotypes in DOM are a subset of those in OCCFR. For example, 78.2% of the haplotypes detected in DOM were shared with OCCFR, whereas only 49.1% and 18.2% were shared with OCCIP and OCA, respectively. In fact, of the 45 haplotypes detected in DOM that were shared with wild populations, a single haplotype in UD14 was not shared with OCCFR, but with OCCIP. Moreover, several haplotypes that were present in DOM were exclusively found in OCCFR, and this was observed for multiple genes (ATP12A, CYTC, MGST3, STAG1, UD14, DIAPH2, and PGK1). These results remained unaltered when we restricted the haplotype analysis to PHASE confidence probability thresholds higher than 0.8.

Our results further indicate that the subspecies *O. c. cuniculus* appears to be the only direct source of all domestic rabbits (i.e., there is no evidence that domestic rabbits were derived directly from OCA). In agreement with previous studies of mitochondrial and Y chromosome sequence variation (Hardy et al. 1995; Geraldès et al. 2005), no alleles were shared by the subspecies at one autosomal (STAG1) and two X-linked loci (F9 and POLA1) in our data set, and in every case, all DOM haplotypes belong to the lineage typically found in OCCIP and OCCFR (fig. 4). We note that other taxa are also unlikely contributors to rabbit domestication because the European rabbit is the sole species in the genus *Oryctolagus* and is almost certainly reproductively isolated from other leporids. The most closely related genera (*Caprolagus*, *Bunolagus*, and *Pentalagus*) are thought to have diverged from *Oryctolagus* at least 8.5 Ma (Matthee et al. 2004).

Nonetheless, OCA was likely an indirect contributor to the domestic rabbit gene pool through introgressive hybridization across the rabbit hybrid zone. For example, some gene genealogies illustrated in figure 4 (DIAPH2, KLHL4, LUM, and MAOA) consisted of two distinct haplogroups separated by a long internal branch. These genealogies had a similar shape compared with genealogies showing no shared variation between subspecies but the

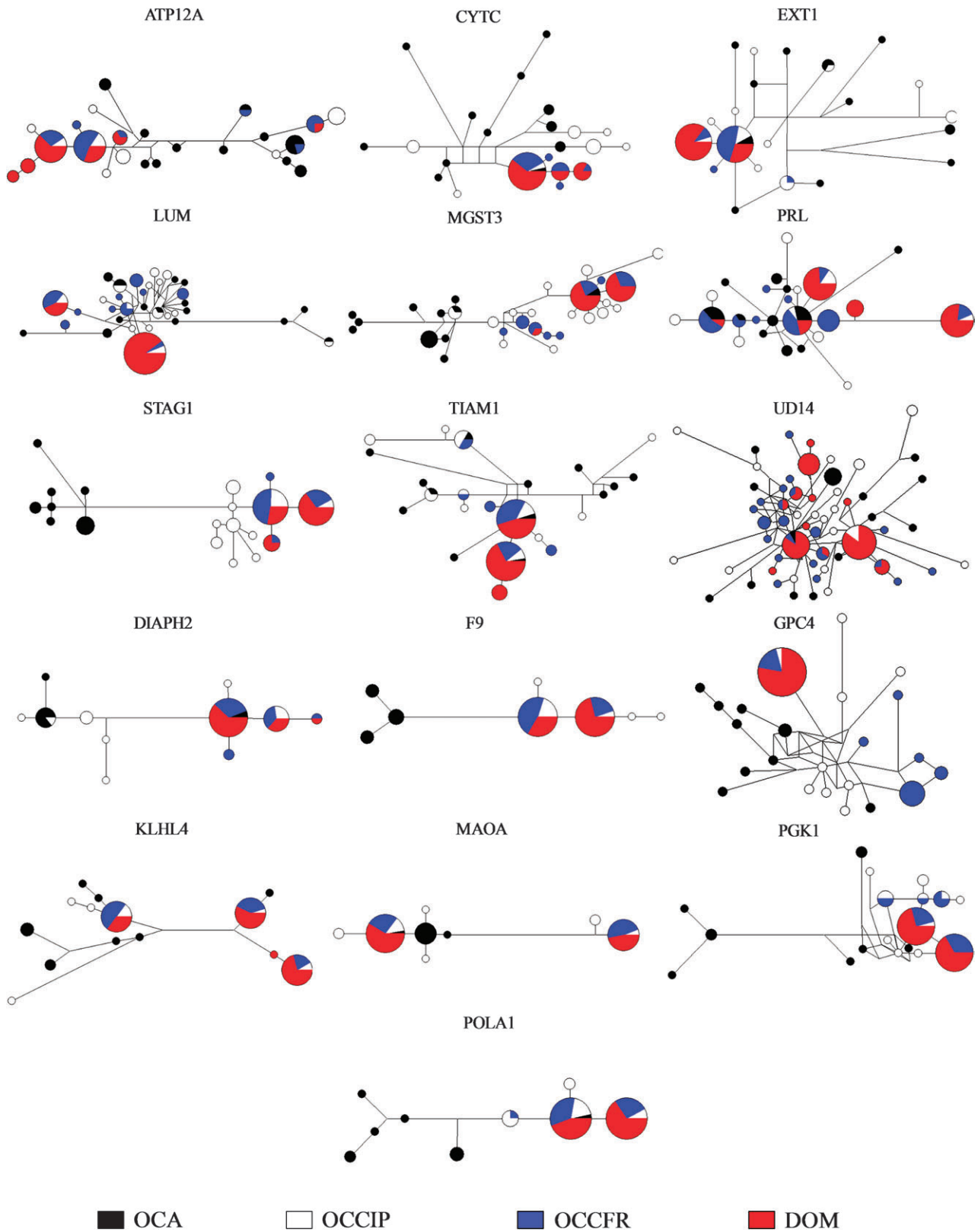


FIG. 4. Median-joining haplotype networks representing the evolutionary relationships among alleles found in both domestic and wild rabbits. Individuals with missing data were excluded. Size of the circles is proportional to the frequency of each haplotype. Population groups are denoted by black for OCA, white for OCCIP, blue for OCCFR, and red for DOM. Note the substantial haplotype sharing between DOM and OCCFR and the absence of OCA haplotypes in DOM for all genes showing fixed variation between subspecies.

Table 1. Summary of the analyses of polymorphism, frequency spectrum tests of neutrality and recombination in OCA, OCCIP, OCCFR, and DOM for nine autosomal and seven X-linked loci.

Locus	Population	Mean π (%) ^a	Mean θ_w (%) ^b	Mean D_T ^c	Mean R_m ^d
Autosomal	OCA	0.721	0.850	-0.645	3.1
	OCCIP	0.658	0.727	-0.369	4.0
	OCCFR	0.348	0.375	-0.368	2.1
	DOM	0.188	0.191	0.243	0.7
X-linked	OCA	0.554	0.557	-0.065	1.0
	OCCIP	0.581	0.596	-0.236	1.1
	OCCFR	0.394	0.283	1.012	0.4
	DOM	0.205	0.131	1.164	0.0
Overall	OCA	0.648	0.722	-0.391	2.2
	OCCIP	0.625	0.670	-0.311	2.8
	OCCFR	0.368	0.335	0.236	1.4
	DOM	0.195	0.165	0.612	0.4

NOTE.—Unweighted average values across 9 autosomal loci, 7 X-linked loci, or 16 combined loci are given in each row.

^a The average number of pairwise differences in a sample (Nei 1987).

^b The proportion of segregating sites in a sample (Watterson 1975).

^c Tajima's D (1989).

^d Minimum number of recombination events in the history of the sample (Hudson and Kaplan 1985).

fact that both lineages showed no correspondence with geography (i.e., shared between both subspecies) is suggestive of extensive admixture after secondary contact for some regions of the genome. In fact, high levels of gene flow between OCA and OCCIP for each of these loci has been inferred using an isolation-with-migration model (Carneiro et al. 2009; Carneiro et al. 2010). Interestingly, DOM and OCCFR harbor both divergent lineages in KLHL4 and MAOA offering strong evidence that at some loci OCA haplotypes introgressed as far as France through natural hybridization and were subsequently incorporated in the domesticated gene pool. In other words, extensive gene flow between subspecies—probably after the last glacial maximum but before domestication—resulted in high levels of introgression for some portions of the genome. These introgressed regions were thus already present in wild rabbit populations before the advent of domestication.

Several observations suggest that rabbit domestication is unusual compared with most domesticated mammals. First, a single French origin for domestic rabbits seems likely and is supported by a combination of historical records (Clutton-Brock 1999; Callou 2003; Whitman 2004) and genetic data from this and previous studies (reviewed in Ferrand and Branco 2007). This simple domestication scenario stands in sharp contrast with most animal species (e.g., cattle, donkeys, goats, horses, pigs, and sheep) that are characterized by independent domestication events either from multiple geographical regions or even from multiple subspecies or species (Bradley et al. 1996; Vila et al. 1997; Luikart et al. 2001; Hiendleder et al. 2002; Jansen et al. 2002; Beja-Pereira et al. 2004; Larson et al. 2005). Second, in contrast to many other domesticated species whose wild progenitors are thought to be extinct (e.g., horse and cattle) or for which the identity of all wild species that contributed genetically to the domesticated population is not clear (e.g., sheep, goat, and cat; Dobney and Larson 2006), rabbits have a known and extant progenitor species. The third important difference relates to the geographical area of domestication. The rabbit is the only mammalian species

that has been exclusively domesticated in Western Europe in contrast to most animal domestication events that occurred in three core areas: southwest Asia, eastern Asia, and the Americas (Clutton-Brock 1999; Diamond 2002; Bruford et al. 2003). Finally, rabbit domestication was initiated within the last 1,500 years (Clutton-Brock 1999; Callou 2003; Whitman 2004). This is recent compared with most other mammals that are thought to have been domesticated more than 5,000 years ago, and domestication of dogs may have started as early as 14,000 years ago (Clutton-Brock 1999; Diamond 2002; Bruford et al. 2003). The much later date of rabbit domestication may have been associated with a more focused and deliberate effort toward domestication (Zeder 2006) and may in part explain why the rabbit domestication process is in several aspects distinct from other domesticated mammals (see below).

Levels of Genetic Variation, Strong Breed Structure, and Departures from Mutation-Drift Equilibrium

Levels of nucleotide diversity (π) in DOM averaged 0.188% and 0.205% for autosomal and X-linked loci, respectively (table 1) but there was substantial heterogeneity among loci with π ranging from 0% to 0.760% (supplementary table 5, Supplementary Material online). Although the average values of π for DOM were similar to that in other domesticated species such as cattle ($\pi = 0.140$ – 0.247% , Gibbs et al. 2009) and dogs ($\pi = 0.1\%$, Brouillette et al. 2000), they were significantly lower than the corresponding values for each of the wild rabbit populations considered in this study, for both the X chromosome and the autosomes ($P < 0.025$ in all cases by Wilcoxon signed-rank test; supplementary table 6, Supplementary Material online). The same was true for θ_w and for the number of haplotypes, as seen in the haplotype networks depicted in figure 4. The mean number of recombination events (R_m) was also substantially lower in DOM than in OCA, OCCIP, and OCCFR (table 1). Consistent with the recent colonization of France, genetic variation in OCCFR was largely a subsample from OCCIP. Overall, diversity parameters in domestic rabbits

Table 2. Average pairwise F_{ST} values (%) among six breeds of domestic rabbits.

Breed	CS	ES	FA	FL	NZ
CS	—				
ES	19.4	—			
FA	10.8	22.0	—		
FL	10.4	10.8	14.1	—	
NZ	25.6	32.9	24.2	22.4	—
RX	12.5	13.4	17.4	7.5	26.1

NOTE.—Unweighted average values across 35 fragments in seven genomic regions. Champagne Silver (CS), English Spot (ES), French Angora (FA), French Lop (FL), New Zealand White (NZ), and Rex (RX).

are consistent with a domestication bottleneck as observed for numerous domesticated species (Wright et al. 2005; Hamblin et al. 2006; Liu and Burke 2006; Haudry et al. 2007; Zhu et al. 2007; Muir et al. 2008; Gray et al. 2009). Importantly, levels of genetic diversity indicate that the ongoing SNP discovery projects for rabbits will result in a large collection of polymorphic SNPs likely to be associated with most genes in the genome.

Next, we examined how genetic variation was organized within and among breeds. We resequenced five fragments from each of seven genomic regions, for a total of 35 fragments of several hundred base pair each (see Methods). This design was implemented to ascertain the structure of LD, as described below. Twenty-five individuals were sequenced in each of six breeds. Consistent with global levels of diversity among breeds, mean levels of genetic diversity within breeds were high, with values of π varying between 0.215% for English Spot and 0.313% for French Angora (summarized in [supplementary table 7](#) and detailed in [supplementary table 8](#), [Supplementary Material](#) online). Pairwise F_{ST} values among breeds (averaged across loci) varied between 7.5% and 32.9% ([table 2](#)), indicating that most of the total genetic variation was within breeds as opposed to between breeds. However, despite the recent origin of most breeds, mean levels of population differentiation in pairwise comparisons ($F_{ST} = 17.9\%$) were high, suggesting that breeds have been generally maintained as closed gene pools. Of additional relevance for mapping studies, polymorphisms exclusive to a single breed were rare (7.9% of the total number of markers excluding singletons), and thus most genetic markers should be informative and transferable across breeds. As expected given this population structure, when all six breeds were considered together 89.6% of the markers showed significant deviations from HWE ($P < 0.05$; [supplementary table 7](#), [Supplementary Material](#) online). In contrast, when we looked at each breed separately, a much smaller percentage of loci showed deviations from Hardy–Weinberg (HW) expectations. The inbreeding coefficient estimated from HW deviations and averaged across markers was highly variable among breeds: 0.25 for Champagne Silver, 0.18 for English Spot, -0.05 for French Angora, 0.16 for French Lop, 0.09 for New Zealand White, and 0.14 for Rex ([supplementary table 7](#), [Supplementary Material](#) online).

Finally, we tested for departures from a neutral equilibrium model in order to assess the impact of the domes-

tication process on patterns of genetic variation. Our intent was not to detect selection at particular loci but to provide a preliminary assessment of population-level changes that may have accompanied the domestication process and to guide the demographic modeling presented below. We observed a large shift in the frequency distribution of polymorphisms between wild and domestic rabbits ([table 1](#)). In wild populations from the Iberia Peninsula, the observed mean values of Tajima's D ($D = -0.391$ for OCA and $D = -0.311$ for OCCIP) were negative indicating a skew toward low-frequency variants, although not significantly different from zero when compared with a multilocus simulated distribution ($P > 0.1$ in both cases). In contrast, both OCCFR and DOM were characterized by the opposite pattern. In OCCFR, Tajima's D was slightly positive but also not significantly different from zero ($D = 0.236$, $P > 0.1$). Strikingly, the mean Tajima's D for DOM was positive and significantly different from zero ($D = 0.612$, $P < 0.05$), indicative of a paucity of low-frequency polymorphisms. This pattern is consistent with a population contraction because low-frequency variants are quickly lost when a population undergoes a substantial reduction in size. Tajima's D values for each locus ranged from -2.073 to 2.467 ([supplementary table 5](#), [Supplementary Material](#) online), with some loci showing significant skews in the allelic frequency spectrum for both positive (ATP12A and MGST3) and negative values (UD14, KLHL4, MAOA, and PGK1). Strobeck's S indicates that fewer haplotypes were observed than expected under an equilibrium model for four loci in DOM (LUM, PRL, KLHL4, and MAOA), but for only two loci in OCCFR (KLHL4 and MAOA), and none in both OCA and OCCIP ([supplementary table 5](#), [Supplementary Material](#) online). Likewise, multilocus HKA tests indicated departures from a neutral equilibrium model in DOM ($P = 0.000$) and OCCFR ($P = 0.002$) but not in OCA ($P = 0.761$) or OCCIP ($P = 0.611$; [supplementary fig. 2](#), [Supplementary Material](#) online). In DOM, the sequential elimination of five loci (MAOA, ATP12A, PRL, MGST3, and KLHL4) was necessary to produce a data set that was not significantly different from neutral expectations ($P = 0.05$). Similar deviations from neutrality were inferred from polymorphism within breeds ([supplementary table 8](#), [Supplementary Material](#) online) demonstrating that breed structure alone cannot account for these patterns.

Although the effects of selection cannot be discarded (see below), it is unlikely that all these regions have been under selection during rabbit domestication. Instead, many of the observed deviations from a neutral equilibrium model are likely the result of the effects of consecutive bottlenecks (the colonization of France, the initial stage of domestication, and subsequent establishment of breeds) coupled with ancient population structure in the wild populations from which domestic rabbits were derived. These findings highlight the challenge of detecting signatures of artificial selection in rabbits using classical tests of neutrality based on equilibrium models. To circumvent this problem, we modeled the domestication process, as described below.

Severe Bottleneck in the Early Stages of Domestication, Small Effective Population Sizes Within Breeds and Detection of Artificial Selection Using a Nonequilibrium Model

We quantified the amount of genetic diversity lost at the onset of rabbit domestication relative to the source population of wild rabbits in France using the following quantities: $1 - \pi_{\text{DOM}}/\pi_{\text{OCCFR}}$ and $1 - \theta_{\text{wDOM}}/\theta_{\text{wOCCFR}}$. The extent of loss of diversity varied greatly among loci (supplementary fig. 3, Supplementary Material online). Removing GPC4 and LUM, which deviate significantly from expected values given the magnitude of the domestication bottleneck (see below), domestic rabbits have lost 37.1% of genetic diversity based on π and 43.8% based on θ_w , on average. The slight difference between the two values is likely explained by the different sensitivity of the two estimators to low-frequency variants.

Rabbit domestication appears to have resulted in a greater loss of genetic diversity than seen in other domesticated mammals. For example, Gray et al. (2009) reported a modest reduction of genetic diversity in dogs (5%), indicating that they probably underwent a smaller population contraction as a result of domestication. In a similar fashion, most studies in pigs have also suggested that domestication has not produced a detectable decrease in variability (Ojeda et al. 2006, 2008). Although evaluation of the amount of variation captured in most domesticated animals is hampered by the fact that wild progenitors are extinct or not readily available, the two available examples described above suggest that rabbit domestication led to a great loss of genetic diversity when compared with other mammals. A single origin of domestication in rabbits may contribute to some of the observed differences in the amount of diversity captured from wild populations. The rabbit domestication process is in this regard more similar to several examples in plants for which comparable losses of variability associated with domestication have been reported (e.g., Wright et al. 2005; Burke et al. 2007).

Given that multiple features of the data support a domestication bottleneck, we conducted simulations to estimate its severity, $k_{\text{dom}} = N_{\text{bdom}}/d_{\text{dom}}$, where N_{bdom} is the size of the bottlenecked population and d_{dom} is the duration in generations of the initial bottleneck. Smaller values of k_{dom} indicate more severe reductions in genetic diversity. In our model (fig. 2), the ancestral population undergoes two consecutive bottlenecks, corresponding to those in the postglacial colonization of France (k_{fr}) and the initiation of domestication (k_{dom}). Following each bottleneck, the model incorporates an instantaneous increase in population size. We conducted coalescent simulations varying k_{fr} and k_{dom} and used a goodness of fit analysis comparing summary statistics computed on the observed data to those computed on simulated data.

The number of segregating sites (S) and haplotypes ($nhaps$) produced nonflat unimodal likelihood distributions for most loci, whereas for π we obtained somewhat flat curves on larger values of k_{dom} and k_{fr} (supplementary

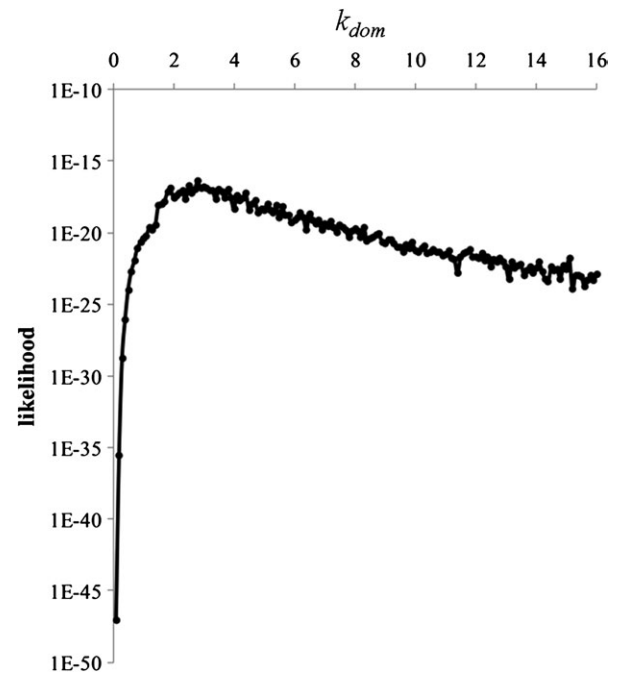


FIG. 5. Diagram representing the multilocus ML estimate for the strength of the population bottleneck (k). Multilocus ML (y axis, logarithmic scale) of the parameter k (x axis) was estimated as the product of the likelihood for each locus. The best-fitting parameter value of k for each locus was estimated as the proportion of 10,000 simulations whose summary statistics (S , $nhaps$, and π) were within 20% of the observed values. GPC4 and LUM were not included.

figs. 4 and 5, Supplementary Material online). Because single summary statistics may fail to capture important aspects of the data, we estimated the multilocus k_{dom} using a combination of all three statistics. To avoid biasing our estimates by loci that deviated significantly in levels of diversity from the expected values given the magnitude of the bottleneck, presumably due to selection, we performed additional simulations and asked whether the observed loss of diversity for each locus was compatible with the multilocus estimate of k_{dom} . Loci that showed significant deviations at $P < 0.05$ were eliminated sequentially and k_{dom} was reestimated iteratively until no significant result was obtained. This procedure led to the removal of two loci (GPC4 and LUM), suggesting that these loci may be subject to selection as discussed in more detail below. The ML estimate of k_{dom} for the remaining 15 loci was 2.8 (two log-likelihood support limits = 1.6–6.6; fig. 5) and, as previously shown (Eyre-Walker et al. 1998; Tenaillon et al. 2004), d_{dom} and N_{bdom} were positively correlated with different combinations of values producing identical estimates (data not shown). Using historical records, we can put an upper limit on the duration of the domestication bottleneck. By the 16th century, several varieties were recognized suggesting that rabbit domestication was already finalized at this time. Given the relationship $k_{\text{dom}} = N_{\text{bdom}}/d_{\text{dom}}$ and assuming a maximum of 900 generations (one generation a year) for the duration of the rabbit domestication bottleneck (d_{dom}), we estimate that $\sim 1,200$ individuals (2,400 chromosomes) can explain the amount of genetic diversity captured in the

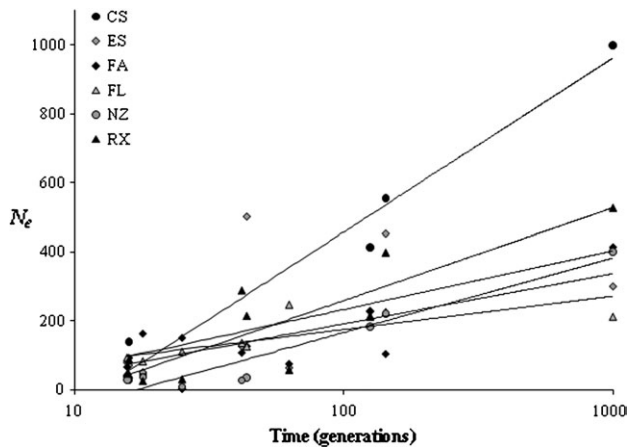


FIG. 6. Estimates of effective population size (N_e) over time from the present (generations). The relationship between effective population size (y axis) and generation time from the present (x axis) was estimated from LD data using the model of Sved (1971). Champagne Silver (CS), English Spot (ES), French Angora (FA), French Lop (FL), New Zealand White (NZ), and Rex (RX).

initial stages of rabbit domestication. Conversely, sequence diversity in domestic rabbits can be explained by a bottleneck of a shorter duration and lower numbers of individuals. For example, assuming a bottleneck of 10 generations, as few as 14 individuals could, in principle, account for the genetic diversity captured at rabbit domestication. These results of simulation indicate that the domestication of rabbits was accompanied by a fairly severe bottleneck of 1,200 individuals or less, which is much smaller than the effective populations size of wild rabbits in France (OCCFR, $N_e = 538,239$) or in Northern Spain (OCCIP, $N_e = 1,043,467$) inferred using a simple mutation-drift expectation ($\theta = 4N_e\mu$, where θ is the average θ_w value for all autosomal loci, μ is the average mutation rate, and $\mu = 1.74 \times 10^{-9}$ was obtained from the comparison between rabbits and *Lepus*).

We used the extent and magnitude of LD in seven genomic regions as an additional tool for demographic inference. Using a model that describes the expectation for the amount of LD as a function of genetic distance and effective population size (Sved 1971; Hayes et al. 2003), we explored the trend in effective population size change in time from r^2 values at different genomic distances for six different breeds. Distance bins used in our sampling design, which ranged from 50 kb to 3,200 kb, corresponded to a time interval of ~ 15 to 1,000 generations in the past (see Methods). We inferred a decline in effective population size for all breeds in the time period considered (fig. 6). This decline seems to have been stronger in recent generations and the most recent effective population size estimates ranged from 28 individuals in the New Zealand White strain to 93 in the French Lop breed. The observation of strong breed structure achieved in a short period of time in rabbits concurs with this idea. Investigations of effective population size variation through time using LD information in other domestic species, such as cattle and chickens,

Table 3. Proportion of simulations with lower values than the observed data assuming the multilocus estimate of the magnitude of the bottleneck ($k_{\text{dom}} = 2.8$).

Locus	S^a	$nhaps^b$	π^c
ATP12A	0.663	0.938	0.096
CYTC	0.746	0.842	0.588
EXT1	0.436	0.555	0.532
LUM	0.022	0.001	0.005
MGST3	0.611	0.153	0.010
PRL	0.755	0.611	0.706
STAG1	0.714	0.850	0.437
TIAM1	0.113	0.772	0.130
UD14	0.211	0.986	0.606
DIAPH2	0.875	0.295	0.751
F9	0.865	0.868	0.441
GPC4	0.003	0.003	0.003
KLHL4	0.678	0.740	0.644
MAOA	0.554	0.872	0.447
PGK1	0.061	0.214	0.053
POLA1	0.115	0.387	0.161

^a Number of segregating sites.

^b Number of haplotypes.

^c Average number of pairwise differences in a sample (Nei 1987).

have also showed recent declines in effective sizes within breeds, suggesting this to be a common trend among domesticated animals (Gibbs et al. 2009; Megens et al. 2009).

We used the multilocus estimate of k_{dom} to ask whether there has been a greater than expected loss of diversity in DOM relative to OCCFR for specific loci, a pattern consistent with positive selection. This method for identifying genomic regions under selection does not rely on equilibrium conditions and looks for outliers against a null model that specifically incorporates the two bottlenecks depicted in figure 2. Levels of variation at most loci did not differ significantly from neutral expectations under this nonequilibrium demographic model (table 3). However, GPC4 and LUM showed a greater reduction in diversity for all summary statistics beyond what could be ascribed to the multilocus estimate of k_{dom} alone, suggesting that these regions may have been subject to selection during the domestication process. GPC4 is of special interest because it is the second most variable locus in OCCFR but invariant in DOM. Moreover, GPC4 had the highest F_{ST} value between DOM and OCCFR (supplementary table 4, Supplementary Material online). Mutations in GPC3 and GPC4, which are located next to each other, have been implicated in pre- and postnatal overgrowth in humans (Pilia et al. 1996), and regulatory variation in GPC3 is responsible for a great deal of the response to selection on growth in mice (Oliver et al. 2005). Interestingly, increased size was likely one of the first traits to be selected in domestic rabbits (Clutton-Brock 1999), raising the intriguing possibility that GPC4 was itself the target of selection. Alternatively, it is also possible that GPC4 was not the direct target of selection and that it is reflecting patterns of selection on linked genes.

This preliminary inquiry into the signature of artificial selection demonstrates the utility and power of nonequilibrium demographic models for identifying candidate “domestication genes” (see Tenailon and Tiffin 2008).

Moreover, this preliminary screen suggests that artificial selection may have had a strong impact on levels of diversity throughout a large fraction of the domestic rabbit genome ($2/16 = 12.5\%$), in addition to the genome-wide effect of the domestication bottleneck. Future studies targeting many genes throughout the genome will provide a powerful means for identifying genes underlying the early phenotypic transition from wild to domesticated forms. A number of excellent studies in domesticated plants (Wright et al. 2005; Yamasaki et al. 2005, 2007; Chapman et al. 2008), and most recently in chickens (Rubin et al. 2010), have used similar approaches with great success.

Similar Contribution of Males and Females to the Domestic Rabbit Gene Pool

To ask whether males and females contributed differently to the domestic rabbit gene pool, we contrasted the amount of diversity captured from French wild rabbits between autosomal and X-linked loci. Strong sex bias, with fewer males than females contributing genetically to the domestication process and breed formation, has been suggested for domesticated species such as horses (Lindgren et al. 2004; Lau et al. 2009) and dogs (Sundqvist et al. 2006). If there was a higher contribution of females when domestication started, then we should expect a relatively higher amount of genetic diversity captured for the X chromosome when compared with the autosomes. The opposite pattern should be expected if males had a higher contribution.

Neither pattern was observed. Levels of variation captured during domestication (after excluding both GPC4 and LUM; [supplementary table 5, Supplementary Material](#) online) were similar and not significantly different for loci residing on the autosomes ($\pi_{\text{DOM}}/\pi_{\text{OCCFR}} = 65.1\%$ and $\theta_{\text{DOM}}/\theta_{\text{OCCFR}} = 53.2\%$) compared with loci residing on the X chromosome ($\pi_{\text{DOM}}/\pi_{\text{OCCFR}} = 60.0\%$ and $\theta_{\text{DOM}}/\theta_{\text{OCCFR}} = 60.2\%$; $P > 0.30$ in both cases by Mann–Whitney U test). We further compared the X chromosome and autosomes in two ways. First, we performed an HKA test (Hudson et al. 1987) using divergence to *Lepus*. Comparison of polymorphism and divergence levels between X-linked and autosomal loci uncovered no significant difference in levels of diversity ($P = 0.81$). Second, we compared multilocus estimates of k_{dom} between autosomal and X-linked loci using the demographic model described above. The estimate for autosomal loci ($k_{\text{dom}} = 2.8$; two log-likelihood support limits = 1.6–10.0) was slightly higher than that for X-linked loci ($k_{\text{dom}} = 1.9$; two log-likelihood support limits = 0.6–12.0), however, due to the low number of genes in each category, the two estimates have wide 95% CI and overlap considerably. Thus, there is no evidence in these data for a highly unequal contribution of males and females to the domestic rabbit gene pool. However, we cannot rule out the possibility of small differences in the initial sex ratio with these data nor can we evaluate any potential sex ratio biases in the formation of individual breeds.

Advantageous LD Structure for Genetic Mapping

To characterize LD as a function of physical distance, we resequenced seven autosomal regions for 25 individuals for each of six breeds. We resequenced five fragments at distances of 50 kb, 400 kb, 1.2 Mb, and 3.2 Mb from the fragment initially selected for each region. The use of complete resequencing data has the advantage of avoiding ascertainment biases in the composition of SNPs toward common alleles. By sampling the total spectrum of mutations for each breed, we increased the probability of sampling markers whose patterns of LD will be representative of the genome at large. We detected a total of 267 SNPs and indel markers and 74.5% had a minor allelic frequency of 10% in at least one breed.

All breeds showed a decrease in r^2 with distance ([fig. 7](#)). Within breeds, at distances of 50 kb r^2 values ranged from 0.48 to 0.70 (average across breeds = 0.59), and at distances of 400 kb ranged from 0.21 to 0.34 (average across breeds = 0.25). Within breeds, the average distance at which r^2 equals 0.5 (i.e., “half-length of LD”; Reich et al. 2001) was ~ 98 kb. In some breeds, moderately strong associations ($r^2 > 0.2$) extended well beyond 1 Mb. For example, in New Zealand White, most SNPs 1,200 kb apart showed r^2 higher than 0.3, consistent with a recent history of strong selection in this strain. Notably, LD did not decline to background levels in four of the six breeds (Champagne Silver, English Spot, French Angora, and New Zealand White) within the 3,200 kb considered here—the results of Mann–Whitney U test between marker pairs distancing 3,200 kb and marker pairs located on different chromosomes were significant at $P < 0.001$ —indicating that residual associations can persist for very long distances in some breeds. Given that the breeds were chosen to reflect the main uses of domestic rabbits, the overall findings are likely to be extendable to most breeds. By extrapolating the extent of LD reported here to the whole genome and assuming a threshold of 0.3 for r^2 , we predict that less than 30,000 evenly spaced markers should suffice for efficient LD screening within rabbit breeds.

In contrast, when all breeds are considered together, LD decayed at a faster rate. Among breeds, LD had a “half-length” of ~ 20 kb, and at a distance of 400 kb r^2 averaged only 0.13 ([fig. 7](#)). The fact that we observed higher LD within rabbit breeds coupled with lower LD when different breeds were combined is indicative of low haplotype sharing across breeds and reflects the separation of the breeds for dozens of generations. This important result suggests that using multiple breeds and a dense set of markers in a region of association holds great promise for interval reduction in fine-scale mapping studies in rabbits. Such a strategy has been applied with outstanding success in dogs (Karlsson et al. 2007; Sutter et al. 2007; Parker et al. 2009) and may help overcome the reduced resolution that results from extensive LD within rabbit breeds. Moreover, rabbits are particularly amenable for laboratory rearing and a combination of linkage and LD mapping approaches may help increase the power for finding genes underlying complex traits (Payseur and Place 2007).

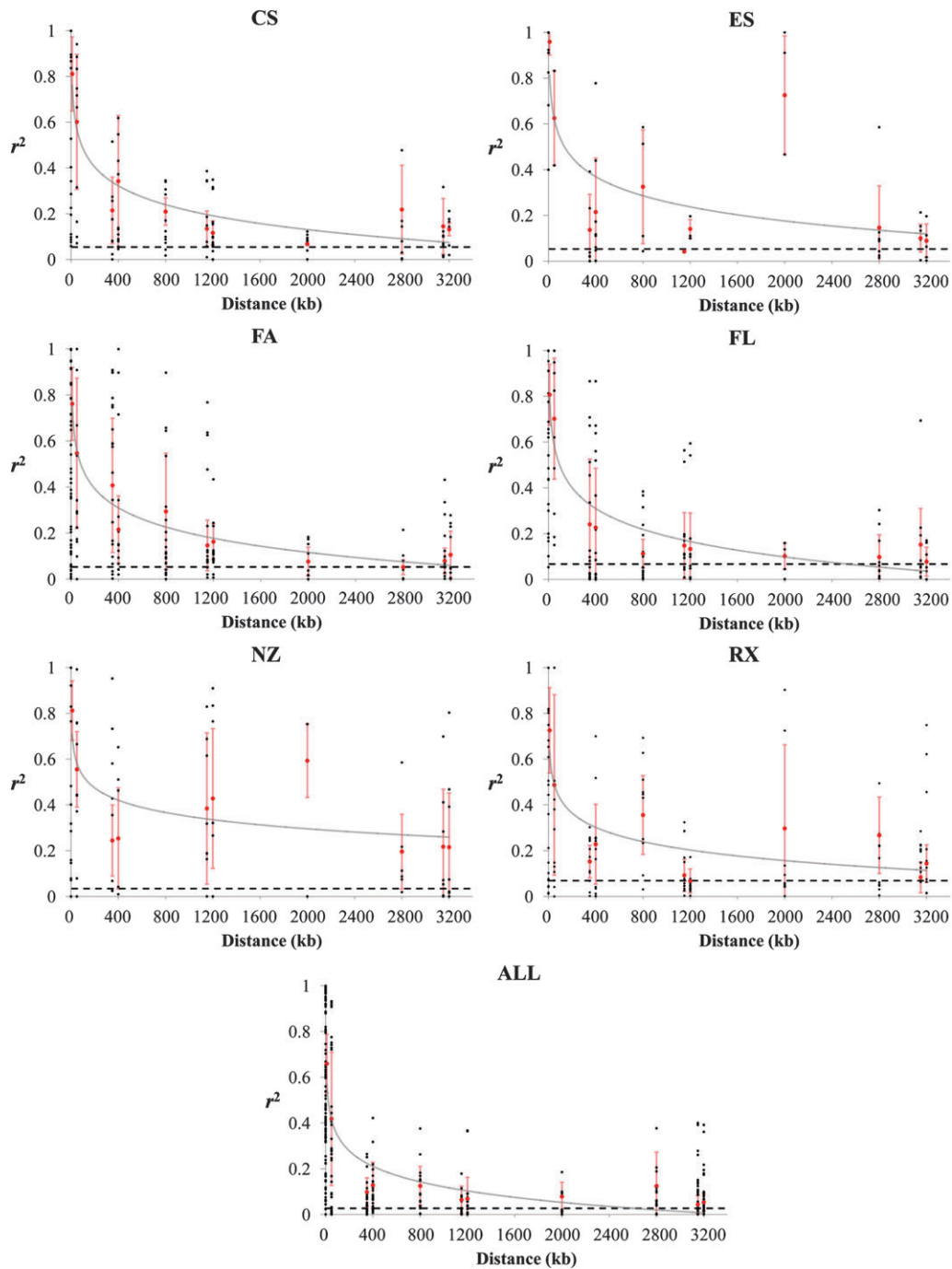


Fig. 7. Plots of LD decay within and among domestic breeds for seven genomic regions. LD was measured using the genotypic r^2 (y axis) and plotted against physical distance in kilobase pairs (x axis). The solid lines represent a logarithmic trend line fitting the mean of the means r^2 values for each region and for each distance bin. The corresponding standard deviation is indicated by vertical bars. The dashed lines in each panel represent the mean r^2 values between unlinked markers. The first six panels show LD within breeds: Champagne Silver (CS), English Spot (ES), French Angora (FA), French Lop (FL), New Zealand White (NZ), and Rex (RX). The bottom panel shows LD considering all six breeds together.

The results presented here demonstrate that domestic rabbits harbor considerable genetic variation both within and among breeds. Moreover, the structure of this variation is such that LD extends for considerable distances within breeds yet decays more rapidly among breeds, a structure that may be advantageous for association mapping studies. The remarkable phenotypic diversity segregating in domestic rabbits makes this species well suited for

understanding the mechanisms underlying phenotypic diversification. In addition, the expanding list of traits of medical interest also holds great promise for the continued development of the rabbit as a biomedical model. Our data suggest that this species, which was first domesticated in France within the last 1,500 years, will be a useful resource for identifying genetic variants that directly affect function in traits of both medical and agronomic interest.

Supplementary Materials

Supplementary tables 1–8 and figures 1–5 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

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 research projects to N.F. (POCI/CVT/61590/2004; PTDC/BIA-BDE/72304/2006; PTDC/BIA-BDE/72277/2006), and by National Science Foundation and National Institutes of Health grants to M.W.N. We are grateful to the Fédération Française de Cuniculture, Union des Sélectionneurs Avicoles du Sud Ouest, Société central d'Aviculture de France, Union des Aviculteurs du Haut-Rhin, Groupement Avicole Pyrénées Atlantiques et Landes, Lapin Club du Sud-Ouest, J. Ruesche, D. Allain, and G. Saleil, for their kind participation in the breeds sampling. We thank CIBIO, especially F. Sequeira and Nachman lab members for valuable discussions, and Naoko Takezaki, and three anonymous reviewers for valuable comments to the manuscript.

References

- Bandelt HJ, Forster P, Rohlf A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 16:37–48.
- Beja-Pereira A, England PR, Ferrand N, Jordan S, Bakhtiet AO, Abdalla MA, Mashkour M, Jordana J, Taberlet P, Luikart G. 2004. African origins of the domestic donkey. *Science* 304:1781–1781.
- Bosze Z, Hiripi L, Carnwath JW, Niemann H. 2003. The transgenic rabbit as model for human diseases and as a source of biologically active recombinant proteins. *Transgenic Res.* 12:541–553.
- Bradley DG, MacHugh DE, Cunningham P, Loftus RT. 1996. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci U S A.* 93:5131–5135.
- Brouillette JA, Andrew JR, Venta PJ. 2000. Estimate of nucleotide diversity in dogs with a pool-and-sequence method. *Mamm Genome.* 11:1079–1086.
- Bruford MW, Bradley DG, Luikart G. 2003. DNA markers reveal the complexity of livestock domestication. *Nat Rev Genet.* 4:900–910.
- Burke JM, Burger JC, Chapman MA. 2007. Crop evolution: from genetics to genomics. *Curr Opin Genet Dev.* 17:525–532.
- Callou C. 2003. De la garenne au clapier: étude archéozoologique du lapin en Europe Occidentale. *Memoir Mus Natl Hist.* 2003:1–352.
- Carneiro M, Blanco-Aguilar JA, Villafuerte R, Ferrand N, Nachman MW. 2010. Speciation in the European rabbit (*Oryctolagus cuniculus*): islands of differentiation on the X chromosome and autosomes. *Evolution* 64:3443–3460.
- Carneiro M, Ferrand N, Nachman MW. 2009. Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* 181:593–606.
- Chantry-Darmon C, Urien C, de Rochambeau H, et al. (14 co-authors). 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.
- Chapman MA, Pashley CH, Wenzler J, Hvala J, Tang SX, Knapp SJ, Burke JM. 2008. A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *Plant Cell.* 20:2931–2945.
- Clutton-Brock J. 1999. A natural history of domesticated mammals. Cambridge: Cambridge University Press.
- Diamond J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418:700–707.
- Dobney K, Larson G. 2006. Genetics and animal domestication: new windows on an elusive process. *J Zool.* 269:261–271.
- Ewing B, Green P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186–194.
- Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8:175–185.
- Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS. 1998. Investigation of the bottleneck leading to the domestication of maize. *Proc Natl Acad Sci U S A.* 95:4441–4446.
- Fan JL, Watanabe T. 2003. Transgenic rabbits as therapeutic protein bioreactors and human disease models. *Pharmacol Therapeut.* 99:261–282.
- Ferrand N, Branco M. 2007. The evolutionary history of the European rabbit (*Oryctolagus cuniculus*): major patterns of population differentiation and geographic expansion inferred from protein polymorphism. In: Weiss S, Ferrand N, editors. *Phylogeography of Southern European Refugia*. Amsterdam (The Netherlands): Springer. p. 207–235.
- Flux J, Fullagar P. 1983. World distribution of the rabbit (*Oryctolagus cuniculus*). *Acta Zool Fennica.* 174:75–77.
- Geraldes A, Rogel-Gaillard C, Ferrand N. 2005. High levels of nucleotide diversity in the European rabbit (*Oryctolagus cuniculus*) SRY gene. *Anim Genet.* 36:349–351.
- Gibbs RA, Taylor JF, Van Tassel CP, et al. (91 co-authors). 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324:528–532.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195–202.
- Gray MM, Granka JM, Bustamante CD, Sutter NB, Boyko AR, Zhu L, Ostrander EA, Wayne RK. 2009. Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics.* 181:1493–1505.
- Hamblin MT, Casa AM, Sun H, Murray SC, Paterson AH, Aquadro CF, Kresovich S. 2006. Challenges of detecting directional selection after a bottleneck: lessons from *Sorghum bicolor*. *Genetics* 173:953–964.
- Hardy C, Callou C, Vigne JD, Casane D, Dennebouy N, Mounolou JC, Monnerot M. 1995. Rabbit mitochondrial DNA diversity from prehistoric to modern times. *J Mol Evol.* 40:227–237.
- Haudry A, Cenci A, Ravel C, et al. (11 co-authors). 2007. Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Mol Biol Evol.* 24:1506–1517.
- Hayes BJ, Visscher PM, McPartlan HC, Goddard ME. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res.* 13:635–643.
- Hey J, Wakeley J. 1997. A coalescent estimator of the population recombination rate. *Genetics* 145:833–846.
- Hiendleder S, Kaupe B, Wassmuth R, Janke A. 2002. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc Biol Sci.* 269:893–904.
- Houdebine LM, Jolivet G, Ripoli PJ. 2009. Transgenic rabbits to prepare pharmaceutical proteins. In: Houdebine LM, Fan J, editors. *Rabbit Biotechnology: rabbit genomics, transgenesis, cloning and models*. New York: Springer. p. 65–75.
- Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18:337–338.

- Hudson RR. 1990. Gene genealogies and the coalescent process. In: Antonovics J, Futuyma D, editors. *Oxford surveys in evolutionary biology*. Oxford: Oxford University Press. p. 1–44.
- Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111:147–164.
- Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153–159.
- Hudson RR, Slatkin M, Maddison WP. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132:583–589.
- Jansen T, Forster P, Levine MA, Oelke H, Hurler M, Renfrew C, Weber J, Olek K. 2002. Mitochondrial DNA and the origins of the domestic horse. *Proc Natl Acad Sci U S A*. 99:10905–10910.
- Karlsson EK, Baranowska I, Wade CM, et al. (20 co-authors). 2007. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet*. 39:1321–1328.
- Kong A, Gudbjartsson DF, Sainz J, et al. (16 co-authors). 2002. A high-resolution recombination map of the human genome. *Nat Genet*. 31:241–247.
- Larson G, Dobney K, Albarella U, et al. (13 co-authors). 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*. 307:1618–1621.
- Lau AN, Peng L, Goto H, Chemnick L, Ryder OA, Makova KD. 2009. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. *Mol Biol Evol*. 26:199–208.
- Lebas F, Coudert P, Rouvier R, Rochambeau H. 1997. The rabbit: husbandry, health and production. Rome: FAO Animal Production and Health Series No 21, FAO Publications.
- Lindgren G, Backstrom N, Swinburne J, Hellborg L, Einarsson A, Sandberg K, Cothran G, Vila C, Binns M, Ellegren H. 2004. Limited number of patrilineal lines in horse domestication. *Nat Genet*. 36:335–336.
- Lindsey J, Fox R. 1994. The biology of the laboratory rabbit. In: Manning P, Ringler D, Newcomer C, editors. *Inherited disease and variations*. San Diego (CA): Academic Press. p. 293–319.
- Liu A, Burke JM. 2006. Patterns of nucleotide diversity in wild and cultivated sunflower. *Genetics* 173:321–330.
- Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, Taberlet P. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci U S A*. 98:5927–5932.
- Matthee CA, van Vuuren BJ, Bell D, Robinson TJ. 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.
- Megens HJ, Crooijmans RP, Bastiaansen JW, et al. (12 co-authors). 2009. Comparison of linkage disequilibrium and haplotype diversity on macro- and microchromosomes in chicken. *BMC Genet*. 10:86.
- Muir WM, Wong GK, Zhang Y, et al. (16 co-authors). 2008. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci U S A*. 105:17312–17317.
- Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
- Nickerson DA, Tobe VO, Taylor SL. 1997. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res*. 25:2745–2751.
- Ojeda A, Huang LS, Ren J, Angiolillo A, Cho IC, Soto H, Lemus-Flores C, Makuza SM, Folch JM, Perez-Enciso M. 2008. Selection in the making: a worldwide survey of haplotypic diversity around a causative mutation in porcine IGF2. *Genetics* 178:1639–1652.
- Ojeda A, Rozas J, Folch JM, Perez-Enciso M. 2006. Unexpected high polymorphism at the FABP4 gene unveils a complex history for pig populations. *Genetics* 174:2119–2127.
- Oliver F, Christians JK, Liu X, Rhind S, Verma V, Davison C, Brown SD, Denny P, Keightley PD. 2005. Regulatory variation at glypican-3 underlies a major growth QTL in mice. *PLoS Biol*. 3:e135.
- Parker HG, VonHoldt BM, Quignon P, et al. (17 co-authors). 2009. An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science*. 325:995–998.
- Payseur BA, Place M. 2007. Prospects for association mapping in classical inbred mouse strains. *Genetics* 175:1999–2008.
- Pilia G, Hughes-Benzie RM, MacKenzie A, Baybayan P, Chen EY, Huber R, Neri G, Cao A, Forabosco A, Schlessinger D. 1996. Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat Genet*. 12:241–247.
- Queney G, Ferrand N, Weiss S, Mougél F, Monnerot M. 2001. Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Mol Biol Evol*. 18:2169–2178.
- Queney G, Vachot AM, Brun JM, Dennebouy N, Mullsant P, Monnerot M. 2002. Different levels of human intervention in domestic rabbits: effects on genetic diversity. *J Heredity*. 93:205–209.
- R Development Core Team. 2009. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. [cited 2009 April 10]. Available from: <http://www.R-project.org>
- Reich DE, Cargill M, Bolck S, et al. (11 co-authors). 2001. Linkage disequilibrium in the human genome. *Nature* 411:199–204.
- Rogel-Gaillard C, Ferrand N, Hayes H. 2009. Rabbit. In: Cockett NE, Kole C, editors. *Genome mapping and genomics in domestic animals*. Heidelberg: Springer. p. 165–230.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.
- Rubin CJ, Zody MC, Eriksson J, et al. (19 co-authors). 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*. 464:587–591.
- Soriguer RC. 1983. El conejo: papel ecológico y estrategia de vida en los ecosistemas mediterráneos. *XV Congr Int Fauna Cinegética y Silvestre Trujillo Spain*. 517–542.
- Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*. 73:1162–1169.
- Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 68:978–989.
- Strobeck C. 1987. Average number of nucleotide differences in a sample from a single subpopulation—a test for population subdivision. *Genetics* 117:149–153.
- Sundqvist AK, Björnerfeldt S, Leonard JA, Hailer F, Hedhammar A, Ellegren H, Vila C. 2006. Unequal contribution of sexes in the origin of dog breeds. *Genetics* 172:1121–1128.
- Sutter NB, Bustamante CD, Chase K, et al. (21 co-authors). 2007. A single IGF1 allele is a major determinant of small size in dogs. *Science* 316:112–115.
- Sved JA. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor Popul Biol*. 2:125–141.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tenaillon MI, Tiffin PL. 2008. The quest for adaptive evolution: a theoretical challenge in a maze of data. *Curr Opin Plant Biol*. 11:110–115.
- Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS. 2004. Selection versus demography: a multilocus investigation of the domestication process in maize. *Mol Biol Evol*. 21:1214–1225.

- Thornton K. 2003. Libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* 19:2325–2327.
- Vila C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK. 1997. Multiple and ancient origins of the domestic dog. *Science*. 276:1687–1689.
- Warnes G, Leisch F. 2005. Genetics: population genetics. R package version 1.2.0. [cited 2009 October 10]. Available from: <http://cran.r-project.org>
- Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol.* 7:256–276.
- Weir BS, Hill WG, Cardon LR. 2004. Allelic association patterns for a dense SNP map. *Genet Epidemiol.* 27:442–450.
- Weisbroth SH. 1974. Neoplastic diseases. In: Weisbroth SH, Flatt RE, Krauss AL, editors. *Biology of the laboratory rabbit*. New York: Academic Press. p. 332–369.
- Whitman BD. 2004. *Domestic rabbits and their histories*. Leawood (Kansas): Leathers Publishing.
- Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS. 2005. The effects of artificial selection on the maize genome. *Science* 308:1310–1314.
- Yamasaki M, Tenaillon MI, Bi IV, Schroeder SG, Sanchez-Villeda H, Doebley JF, Gaut BS, McMullen MD. 2005. A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell*. 17:2859–2872.
- Yamasaki M, Wright SI, McMullen MD. 2007. Genomic screening for artificial selection during domestication and improvement in maize. *Ann Bot.* 100:967–973.
- Zeder MA. 2006. Central questions in the domestication of plants and animals. *Evol Anthropol.* 15:105–117.
- Zhu QH, Zheng XM, Luo JC, Gaut BS, Ge S. 2007. Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice. *Mol Biol Evol.* 24:875–888.