A relict bank vole lineage highlights the biogeographic history of the Pyrenean region in Europe

VALÉRIE DEFFONTAINE,* † RONAN LEDEVIN,§ MICHAËL C. FONTAINE,¶ ** JEAN-PIERRE QUÉRÉ, † SABRINA RENAUD, § ROLAND LIBOIS* and JOHAN R. MICHAUX† ††

*Unité de recherches zoogéographiques, University of Liège, Bât. B22, Boulevard du Rectorat, Sart Tilman, 4000 Liège, Belgium, †INRA, UMR 1064 CBGP, Campus international de Baillarguet, CS30016, F-34988 Montferrier sur Lez cedex, France, ‡Department of Ecology and Evolutionary Biology, University of Glasgow, Graham Kerr Building, Room 314, University Avenue, G128QQ Glasgow, UK, §UMR 5125 PEPS, CNRS, Université Lyon 1, Campus de la Doua, Bâtiment Géode, 69622 Villeurbanne Cedex, France, ¶MARE Center, Laboratory for Oceanology, University of Liège, Bât. B22, Boulevard du Rectorat, Sart Tilman, 4000 Liège, Belgium, **Laboratoire d’Ecologie, Systématique et Evolution, Université Paris-Sud, Bât. 360, Orsay, France, ††Laboratoire de génétique des microorganismes, University of Liège, Bât. B22, Boulevard du Rectorat, Sart Tilman, 4000 Liège, Belgium

Abstract

The Pyrenean region exhibits high levels of endemism suggesting a major contribution to the phylogeography of European species. But, to date, the role of the Pyrenees and surrounding areas as a glacial refugium for temperate species remains poorly explored. In the current study, we investigated the biogeographic role of the Pyrenean region through the analyses of genetic polymorphism and morphology of a typical forest-dwelling small mammal, the bank vole (Myodes glareolus). Analyses of the mitochondrial cytochrome b gene and the third upper molar (M3) show a complex phylogeographic structure in the Pyrenean region with at least three distinct lineages: the Western European, Spanish and Basque lineages. The Basque lineage in the northwestern (NW) Pyrenees was identified as a new clearly differentiated and geographically localized bank vole lineage in Europe. The average M3 shape of Basque bank voles suggests morphological differentiation but also restricted genetic exchanges with other populations. Our genetic and morphological results as well as palaeo-environmental and fossils records support the hypothesis of a new glacial refugium in Europe situated in the NW Pyrenees. The permissive microclimatic conditions that prevailed for a long time in this region may have allowed the survival of temperate species, including humans. Moreover, local differentiation around the Pyrenees is favoured by the opportunity for populations to track the shift of the vegetation belt in altitude rather than in latitude. The finding of the Basque lineage is in agreement with the high level of endemic taxa reported in the NW Pyrenees.

Keywords: bank vole, Basque country, glacial refugia, mitochondrial DNA, molar morphology, phylogeography, Pyrenees

Received 2 September 2008; revised 8 January 2009; accepted 22 January 2009

Introduction

Survival of European temperate species during the glacial periods of the Pleistocene was conditioned by the ability of populations to track favourable habitats. Southern Europe and its three Mediterranean peninsulas have thus been traditionally considered as core areas for the survival of temperate plants and animals, by offering a southern way to escape to from the prevailing steppic and cold conditions (Taberlet et al. 1998). Southern differentiation of temperate but also cold-adapted species were driven by the presence of major geographic barriers (Mediterranean sea, Pyrenees, Alps), environmental heterogeneities and fragmentation of temperate forests and shrubby ecosystems during glacial and interglacial events (Blondel 1995; Garcia-Barros et al. 2002). It is now well established that each Mediterranean...
peninsula formed multiple glacial refugia instead of a single broad one (the concept of ‘refugia within refugia’, Gómez & Lunt 2007). This idea is consistent with the identification of multiple mitochondrial DNA (mtDNA) lineages evidenced for many Iberian and Italian species (Paulo et al. 2001; Allegrucci et al. 2005; Vila et al. 2005; Bella et al. 2007; Terrab et al. 2008). Many lineages have been described in the Iberian peninsula, and the location of the corresponding glacial refugia was recently reviewed by Gómez & Lunt (2007). These authors compared the phylogeographic patterns of Iberian species and reported seven putative Iberian glacial refugia for terrestrial taxa: in southern Spain, along the Spanish Mediterranean coast, in central Spain, in central Portugal and also in mountain areas such as in the Cantabrians and in the Pyrenees.

Phylogeographic studies, palaeontological and palynological records have provided compelling evidences that mountain areas such as the Carpathians and Pindus mountains in the Balkan region served as glacial refugia for temperate species and cold-tolerant species (Willis et al. 2000; Tzedakis et al. 2002; Babik et al. 2004; Deffontaine et al. 2005; Kotlík et al. 2006; Sommer & Nadachowski 2006; Schmitt & Haubrich 2008). However, the role of the Pyrenees and surrounding areas as a glacial refugium for temperate species remains poorly explored to date, despite its classification as a major regional unit of endemism (García-Barros et al. 2002). This mountain range has an east–west orientation and peaks at around 2000 m. It is mostly regarded as a barrier to postglacial re-colonization or as a postglacial contact zone of lineages rather than as a glacial refugium (see reviews by Hewitt 1999 and Guillaume et al. 2000, as well as Hewitt 2000). The Pyrenees were logically designed as a putative refugium region for cold-adapted species such as the bird Tetrao urogallus (Duriez et al. 2006), the plants Pinus sylvestris and Plantago media (Sinclair et al. 1999 and Soranzo et al. 2000, both as cited by Gómez & Lunt 2007) and the mountain-forest butterfly Erebia euryale (Schmitt & Haubrich 2008). However, only two mammalian phylogeographic studies have addressed this issue. In humans, González et al. (2006) demonstrated an ancestral mitochondrial subhaplogroup inside the Basque population located in the northwestern (NW) Pyrenees. Another study conducted on Iberian populations of the field vole (Microtus agrestis) using the mtDNA cytochrome b (cyt b) marker evidenced two southern, highly differentiated phylogroups related to central Portugal and northern Spain glacial refugia (Jaarola & Searle 2004). More surprising was the Pyrenean phylogeography of the bushcricket (Ephippiger ephippiger). This species is locally distributed in Western Europe, and the examination of mitochondrial gene polymorphisms suggested that some populations could have survived the glacial stages in the lower regions of the Pyrenean valleys (Spooner & Ritchie 2007). In this framework, what was the biogeographic role of the Pyrenean area for other temperate species during the climatic fluctuations of the Pleistocene? Was the mountain range predominantly a geographic barrier for re-colonizing lineages isolated in other refugia, or was it a zone of survival in a mosaic landscape, leading to a high local endemism today?

We addressed this question by more thoroughly investigating bank vole populations in this area. This species is particularly well appropriate to disentangling the role of Pyrenean mountains in this context by being a typical temperate forest species. The rodent has a wide distribution range in Europe and Russia and is currently found in the northern part of the Iberian, Italian and Balkan peninsulas (Shenbrot & Krasnov 2005). Its distribution range in the Iberian peninsula encompasses the northern and southern parts of the Pyrenees as well as the Cantabrians, but it does not spread farther to the south (Blanco 1998; Fig. 1a). 

![Fig. 1 (A) Distribution of the bank vole (Myodes glareolus) sampled around the Pyrenees and the groups revealed by the genetic and morphological analyses. The numbers refer to the localities analysed: 1, Spain, Asturias; 2, Spain, Navarre; 3, France, Armendarits; 4, France, Lantabat; 5, France, Ponson de Bat; 6, France, Néouvielle; 7, Spain, Granollers; 8, France, Py Mantet; 9, France, Hérépian; 10, France, Montpellier; 11, France, Saint Aignan; 12, France, Cadouin; 13, France, Paussac; 14, France, Bourdeilles; 15, Italy, Pietraporzio. (B) Palaeo-environmental reconstruction of the Pyrenean region during the last glacial maximum (18 000 years ago) (adapted from Antoine et al. 1999). See text for details and references.](image)
past, the species was classified under the genus *Clethrionomys*, but more recently it has been argued that *Myodes* should take precedence (Wilson & Reeder 2005). Previous phylo-

geographic studies have shown a high level of polymorphism in mtDNA genome in Europe for this species, with the presence of four continental lineages (Western European, Eastern European, Ural and Carpathian) and three Medi-

terranean lineages (Iberian, Italian and Balkan) (Defontaine *et al.* 2005; Kotlík *et al.* 2006). These bank vole lineages started to diverge during the Upper Pleistocene (~300 000 years ago) inside several glacial refugia. The present study focuses on the bank vole populations associated with the Iberian region and investigates the biogeographic role of the Pyrenean region in the complex phylogeographic history of this species.

**Materials and methods**

**Samples**

The genetic polymorphism of the mtDNA *cyt b* gene was analysed for 292 bank voles sampled throughout Europe (Supporting information, Table S1). Two hundred and thirty-eight sequences were obtained from previous published studies (Conroy & Cook 1999; Dekonenko *et al.* 2003; Cook *et al.* 2004; Defontaine *et al.* 2005; Essbauer *et al.* 2006; Kotlík *et al.* 2006). We did not include in the analyses partial published sequences of cyt b. The 54 new sequences of bank voles *cyt b* used here originated from 17 localities including five sampling points in the Pyrenean massif or its foothills (Armendariz, Ponson de Bat, Py Mantet, Navarre, Néouvielle; Fig. 1a), six in surrounding regions of the Pyrenees (Asturias, Granollers, Montpellier, Herepian, Cadouin, Saint Aignan; Fig. 1a) and six in several regions of the eastern part of the bank vole distribution range (Romania, Russia). The morphological analyses were conducted on 113 bank vole skulls from 15 Eurasian localities including Pyrenean and Iberian localities (Supporting information, Table S2).

**Laboratory procedures**

Total genomic DNA was extracted from ethanol-preserved tissues using the DNeasy Blood & Tissue Kit (Qiagen). The *cyt b* gene was analysed using the primers and protocols published in Defontaine *et al.* (2005) that produced unambiguous results. A single strand was sequenced on ABI 3700 automated DNA sequencer (PE, Applied Biosystems, CA) using the manufacturer’s protocol. The sequences were aligned using the Clustal W algorithm (Thompson *et al.* 1994) in *bioedit* 7.0.5.2 (Hall 1999). We have checked the dataset for pseudogenes by looking for stop codons in the translated *cyt b* sequences using *dnasp* 4.0 (Rozas *et al.* 2003).

**Phylogenetic and phylogeographic analyses**

First, a haplotype data file was generated using *macclade* 3.07 (Maddison & Maddison 1992). Phylogenetic reconstruc-

tions were performed by a distance method using the neighbour-joining algorithm (NJ, Saitou & Nei 1987) in *paup* 4.0b10 (Swofford 2000) and by the maximum-likelihood criterion (ML & Felsenstein 1981) implemented in the *phyml* program (Guindon & Gascuel 2003). Three northern red-

backed vole sequences (*Myodes rutilus*, GenBank Accession nos AB072224, AF272638, AY309428, Cook *et al.* 2001, 2004; Iwasa *et al.* 2002) and two grey red-backed vole sequences (*M. rufocanus*, GenBank Accession nos AF429815, AF429816; Dekonenko *et al.* 2003) were also included in the dataset. *M. rufocanus* was used as outgroup in the phylogenetic reconstructions. Using alternative outgroups (Kotlík *et al.* 2006; Lebedev *et al.* 2007) did not improve the results shown here. The program *modeltree* 3.0 (Posada & Crandall 1998) was used to determine the most suitable substitution model for the *cyt b* dataset studied. The robustness of the trees was assessed by bootstrap re-sampling (BS) (10 000 random replications for NJ and ML analyses; Felsenstein 1985). A Bayesian phylogeny reconstruction approach (Yang & Rannala 1997; Huelsenbeck *et al.* 2001) was also used, implemented in *mrbayes* 2.01 (Huelsenbeck *et al.* 2001). Metropolis-coupled Markov chain Monte Carlo (MCMC) sampling was performed with five chains run for 1 000 000 iterations, using default model parameters as starting values. Bayesian posterior probabilities (BPP) were obtained from the 50% majority rule consensus of trees sampled every 20 generations, discarding the trees obtained before the chains reached stationary distribution (‘burn-in’, empirically determined by checking of likelihood values). We used the approximate likelihood-ratio test (aLRT) implemented in *aLRT PHYML* to test the branches support (Guindon & Gascuel 2003; Anisimova & Gascuel 2006). The aLRT statistics assesses whether a branch provides a significant gain in likelihood, in comparison with the hypothesis corresponding to the second-best nearest-neighbour interchange (NNI) configuration around the branch of interest. We examined two support options [Shimodaira Hasegawa (SH)-like support and Chi-squared-based support] and analysed 100 data sets using the TN93 model of evolution. We chose the tree with the higher log-likelihood value.

Second, and for the sake of clarity, the genetic structure of Pyrenean populations was examined using a smaller dataset (49 individuals) that encompassed samples from the localities underlined in bold in the Supporting information, Table S1. A spatial analysis of molecular variance was conducted using the program *samova* 1.0 (Dupanloup *et al.* 2002) for two to five groups. The two localities Pietraporzio and Cadouin were left out of the *samova* analyses because these regions were represented only by a single specimen. We selected the partition that maximized the among-group
A network of haplotypes was constructed using the median-joining method (MJ, NETWORK 4.5 program; Bandelt et al. 1999). A post-processing calculation (the ‘MP’ calculation option) was applied in NETWORK 4.5 to eliminate superfluos (non-parsimonious) links of the network (Polzin & Daneshmand 2003). The demographic history of the newly defined bank vole lineage was inferred using a pairwise mismatch distribution analysis between individuals in ARLEQUIN 3.1. The mismatch distribution was calculated under a model of population growth–decline: multimodal distributions would be consistent with demographic stability while sudden expansion would generate a unimodal pattern (Slatkin & Hudson 1991). We tested the smoothness of the observed pairwise differences distribution with the null hypothesis of population expansion using the sum of squared deviation (SSD) (Excoffier et al. 2005) and the raggedness statistic (r) (Harpending 1994). We also calculated Fu’s Fs statistic (Fu 1997), which is considered as a powerful statistical test to detect the population demographic expansion. These statistical tests and their significances were evaluated by 1000 random permutations in ARLEQUIN 3.1. Haplotype (h) and nucleotide (π) diversities (Nei 1987) and their standard deviations (± SD; Tajima 1993) were estimated using DNASP 4.0. We specifically carried out these analyses on the lineages defined around the Pyrenees.

We used the McDonald–Kreitman test (McDonald & Kreitman 1991) in DNASP 4.0 to test whether our cyt b dataset departed significantly from neutral expectations. A Fisher’s exact test was conducted to determine whether the ratio of synonymous to non-synonymous substitutions differed between two categories: polymorphisms that are variable within bank vole species and a Tien Shan red-backed vole (Myodes centralis) species and polymorphisms that distinguish these two species (i.e. fixed differences). We chose M. centralis instead of M. rutilus or M. rufocanus because this species displayed a sufficient amount of non-synonymous substitutions compared to the bank vole to perform the exact test (Kotlík et al. 2006).

Molecular clock analyses

We conducted a likelihood-ratio test (LRT) in PAUP 4.0b10 to test if the data violate the assumption of a constant molecular clock. We also used the program RRTRREE to test if different substitution rates are present among phylogenetically defined cyt b lineages (Robinson et al. 1998). Under the assumption of a molecular clock, the times of bank vole lineage differentiation were estimated using the genetic divergences among lineages. An individual matrix of genetic distances (in percentage K, P) was calculated in PAUP 4.0b10. The within- and between-lineage divergences and their SDs were calculated in Microsoft Office Excel 2003. We used the formula of Edwards (1997) to correct the K, P genetic divergences for ancestral polymorphism. As calibration point, we used the separation time between M. glareolus and M. rutilus estimated at 2.5 million years ago (Mya) based on palaeontological records (Lebedev et al. 2007).

Morphological analyses

A morphological marker was also used as an alternative means to evidence a possible phenotypic variation in the bank vole populations. Bank vole skulls were analysed for the third upper molar (M3) that varies within M. glareolus (Bauchau & Chaline 1987).

For each animal, the two-dimensional outline of the left M3 (or the mirror image of the right M3) was digitized, sampling the coordinates of 64 points at equal curvilinear distance using the image analysing software OPTIMAS 6.5. Only mature specimens with complete eruption of the M3 were considered.

The molars of the arvicoline rodents are composed of successive triangles leading to a complex shape. The Elliptic Fourier transform is an adequate method to describe such outlines (Navarro et al. 2004). This method is based on the separate Fourier decompositions of the incremental changes of the x- and y-coordinates as a function of the cumulative length along the outline (Kuhl & Giardina 1982). Any harmonic corresponds to four coefficients: A, and B, for x, and C, and D, for y, defining an ellipse in the xy-plane. The coefficients of the first harmonic, describing the best-fitting ellipse of any outline, are used to standardize the size and orientation of the object. These coefficients therefore correspond to residuals after standardization and should not be included in following statistical analyses (Crampton 1995; Renaud et al. 1996). However, the information about the minor axis of the ellipse, corresponding to its elongation and hence to the general elongation of the object, should still be contained in the residuals coefficients, namely the coefficient D1.

A visual inspection showed that the overall shape of the M3 was adequately described retaining the coefficients of the first 10 harmonics, in agreement with previous studies on arvicoline teeth (Navarro et al. 2004). The shape of each M3 was thus described by 37 variables: nine harmonics (10 retained minus the first) per four coefficients, plus the coefficients D1.

These variables were standardized for size variation and contained shape information only. Due to their large number, a multivariate analysis was necessary to reveal the main patterns of morphological differentiation to emerge. A principal component analysis (PCA), including all the specimens, was performed on the variance–covariance matrix of the 37 Fourier coefficients in order to evidence the patterns of morphological differentiation without using the information about mtDNA lineages or populations.
The significance of the difference among the Western European, Eastern European, Spanish and Basque lineages was further tested using a multivariate analysis of variance (Manova). Using a discriminant analysis on the same data set, the specimens from the Basque locality Lantabat and those from the bank vole from Armendarits displaying a Spanish mtDNA have been attributed to one of the four documented lineages. This attribution is based on the Mahalanobis distance of each specimen to the group mean of each lineage that allows the calculation of probabilities for group membership.

Results

Phylogenetic and phylogeographic analyses

A total of 239 cyt b haplotypes (including 25 new haplotypes) was identified among the 294 Myodes sp. specimens. Some sequences published in GenBank were found to be identical to others. Out of the 994 base pairs (bp), 243 were variable and 175 parsimony-informative. The average transitions/transversions ratio was equal to 8.16 and the base composition was of 25% of T, 30% of C, 31% of A and 14% of G. No stop codons were found in the translation of the cyt b sequences used in this study supporting the absence of pseudogenes in our dataset.

Phylogenetic trees were rooted with cyt b sequences of Myodes rufocanus that provided the best supported phylogenies for the present bank vole cyt b dataset. By choosing a more distant monophyletic outgroup, we ensured that basal clades of Myodes glareolus phylogeography were resolved. The hierarchical likelihood ratio tests (HLRTS) and the Akaike information criterion (AIC) conducted in modeltest 3.0 respectively pointed the K81P (K81) model of substitution (Kimura 1981) and the transitional model (TIM) of substitution as the best suited model for the present dataset. The TIM model is not an available option for model substitution in the available programs of phylogenetic reconstructions. The K81P model (already used in previous publications) was preferred for this reason. The proportion of invariant sites was equal to 0.49 and the gamma distribution shape parameter to 0.74. The Bayesian tree presents the relationship between the bank vole haplotypes (Fig. 2). We did not observed any differences in the clades' support values using another model of evolution (Hasegawa–Kishino–Yano model from Hasegawa et al. 1985, suggested in Kotlik et al. 2006). The bank vole lineages previously described by Deffontaine et al. 2005 (Spanish, Italian, Balkan, Western European, Eastern European) and by Kotlik et al. 2006 (Carpathian) were grouped together in a major monophyletic clade (BS: 71, 62 and BPP: 0.78). These lineages have delimited distribution range in Spain and southern France (Spanish lineage), in Italy (Italian lineage), in the Balkan and Turkey (Balkan lineage), in Western Europe, Eastern Europe and in northern Italy (Western lineage), in Eastern Europe and Russia (Eastern lineage), and in the Carpathians (Carpathian lineage). The bank voles from the Ural lineage clustered together with the haplotypes of M. rutilus. This particularity was described by Tegelström (1987) and Deffontaine et al. (2005) as a case of unidirectional introgression of the northern red-backed vole mtDNA into bank voles from northern Europe and the Ural mountains. A new lineage clustered apart from the others and included bank voles from the French Basque country and surrounding regions (Fig. 2). This lineage, called ‘Basque’ lineage, had a basal position in the phylogenetic trees compared to other bank vole lineages and was highly supported (BS: 92, 76 and BPP: 0.88). The aLRT analysis supported the monophyly of the Basque lineage and its basal position in the phylogenetic tree (support values for the SH and $\chi^2 \geq 0.95$). The position of the Basque lineage in the different tree reconstructions was stable (basal to other bank vole lineages) unlike the other bank vole lineages (Spanish, Italian, Balkan, Western European, Eastern European). The support values for these lineages were slightly lower than those previously reported in Deffontaine et al. (2005). This may have resulted from the higher number of sequences used in this study, which introduced a lot of polymorphisms and hid the synapomorphies, especially in a shallow phylogeny. The absence of strong monophyly of the Carpathian lineage and its inclusion in the Eastern European lineage was an unexpected result in our analysis. Kotlik et al. (2006) defined this lineage as monophyletic and close to the Western European lineage. But these authors used a sample that mainly focused on the Carpathians. It might be possible that some genetic relationships between the bank voles from the Carpathians and those from the other parts of the distribution range in Europe and Russia were missed in their study.

Population structure and diversity in the Pyrenean region

Only samples from the Pyrenean region were considered for population structure analysis and haplotype network reconstructions. The samova analysis evidenced three populations around the Pyrenean region ($\phi_{CT} = 84,26\%$) corresponding to the clustering of the following localites: [Armendarits, Ponson de Bat], [Néouvielle, Py Mantet, Montpellier, Navarre, Asturias, Granollers] and [Bourdeilles]. The locality of Armendarits (see Fig. 1a) carries animals with haplotypes from the Basque and Spanish lineages and might correspond to the contact zone between the two lineages. The MJ analysis evidenced three differentiated groups (separated by at least eight mutations) around the Pyrenees. They correspond to the Basque, Spanish and Western European lineages (Fig. 3).

The genetic features of the new bank vole Basque lineage were compared to those previously identified in Deffontaine...
Fig. 2 Bayesian phylogenetic tree of the 239 *Myodes* mtDNA haplotypes. Bootstrap values (BS, in %) obtained using the neighbour-joining (NJ) and maximum-likelihood (ML) analyses, as well as Bayesian posterior probabilities (BPP) are shown for the major phylogroups. See Supporting information, Table S1, for the haplotype designations.
et al. (2005). The mismatch distribution analysis, with unimodal distributions, supported the hypothesis of population expansion for the Western European, Eastern European and Italian lineages. The Balkan, Spanish and Basque lineages, on the other hand, displayed multimodal distributions suggesting long-term population stability (Fig. 4, in Deffontaine et al. 2005). In parallel, the raggedness index $r$ of the Basque lineage was greater than 0.05 which is consistent with constant population size ($r = 0.098$, probability = 0.69; Harpending et al. 1993). The SSD calculation was not very informative and did not support any departure from the null hypothesis of a population in expansion ($SSD = 0.03$, probability = 0.45). The Fu’s Fs statistic, which is sensitive to demographic expansion, showed no negative values for the Basque lineage, similarly suggesting demographic stability ($F = 0.72$, probability = 0.68). The Basque lineage displayed a nucleotide genetic diversity of $\pi = 0.24 \pm 0.16\%$ and a haplotype genetic diversity of $h = 0.705 \pm 0.112$. These values are similar to those reported in Deffontaine et al. (2005) for the Spanish ($\pi = 0.21 \pm 0.05\%$, $h = 0.758 \pm 0.08$) and the Italian lineages ($\pi = 0.18 \pm 0.04\%$, $h = 0.786 \pm 0.110$). The mean genetic distance of the Basque lineage to the other lineages (0.92% ± 0.30) highlighted a rather large genetic divergence inside the shallow phylogeny of the species, which was of the same order of magnitude as those of the other bank vole lineages (Deffontaine et al. 2005). This would correspond to a time of genetic differentiation not later than during the late Pleistocene. Finally, the McDonald–Kreitman test showed no significant departure of the present bank vole cyt b dataset to the neutral expectation.

**Time of lineages differentiation**

The LRT showed no significant differences between the likelihoods of the models with and without molecular clock constraints ($\chi^2 = 0.0212$, d.f. = 1, probability = 0.884). The rrRtree test was also not significant for all pairs of
lineages compared ($P > 0.2$), suggesting the absence of rate heterogeneity among them and the clock-like evolution of the present dataset. The molecular clock was estimated at $4.3 \pm 0.24\% [4.54–4.06]$ K$_{P}$ distance per million years (My). This estimate made use of the most recent calibration date in the *Myodes* phylogeny and is higher than those reported in Deffontaine *et al.* (2005), which was $3.6\%$ K$_{P}$ distance per My. Times of lineage differentiation were calculated on the basis of this new evolution rate. Therefore, the separation time between all lineages and between the Basque lineage and the other bank vole lineages were estimated to have been during the Upper Pleistocene (between 0.19 and 0.56 Mya, excluding the Ural lineage, and between 0.46 and 0.56 Mya, respectively).

**Morphological analyses**

A first morphometric analysis relied on descriptive multivariate statistics, using a PCA to display the principal axes of morphological variation of the M$_{3}$. Fig. 5 shows the morphospace corresponding to the first three principal components of a PCA on the Fourier coefficients of the outline of the third upper molars (M$_{3}$). Average values per locality are represented, with symbols indicating the genetic lineage dominating this locality. No mixing has been evidenced in the selected western, eastern and Spanish localities, but the Basque locality Armendarits [PB.LA], dominated by the Basque lineage, includes a ‘Spanish’ specimen. No genetic identification is available for the specimens of the other Basque locality Landabat [PB.LA]. Reconstructed outlines on the top visualize the average shape of the main groups. See Supporting information, Table S2, for samples designations.

![Fig. 5 Morphological variation of the third upper molar (M$_{3}$) of the bank vole (*Myodes glareolus*) in Europe, estimated by the first three principal components of a principal component analysis on the Elliptic Fourier coefficients of the outline of the third upper molars (M$_{3}$). Average values per locality are represented, with symbols indicating the genetic lineage dominating this locality. No mixing has been evidenced in the selected western, eastern and Spanish localities, but the Basque locality Armendarits [PB.LA], dominated by the Basque lineage, includes a ‘Spanish’ specimen. No genetic identification is available for the specimens of the other Basque locality Landabat [PB.LA]. Reconstructed outlines on the top visualize the average shape of the main groups. See Supporting information, Table S2, for samples designations.](image-url)
Spanish and one to the Western European group. In most of the cases (6/7), this attribution relied on a high probability (> 0.85). Furthermore, the Armendarits individual with a Spanish mtDNA was attributed to the Basque group (probability = 0.96).

Discussion

The bank vole phylogeography in the Pyrenees and surrounding regions

The present phylogenetic and morphological results demonstrate the existence of a new, well-differentiated bank vole population, geographically localized in the French Basque country (Armendarits and Lantabat localities) and in the French department of the Atlantic Pyrenees (Ponson de Bat locality): the 'Basque' lineage. According to this new finding, the Pyrenean region and central France house at least three bank vole lineages: (i) the Western European lineage that currently extends in northern and central parts of France but does not reach the Pyrenees; (ii) the Spanish lineage that currently encompasses the Cantabrian mountains, the southern part of the Pyrenees, the eastern Pyrenees and southern France and NW Italy along the Mediterranean coast; and (iii) the new Basque lineage that is currently surrounded by the Atlantic Ocean and the Western European and Spanish lineages (Fig. 1a). Although these lineages cover different geographic areas, they cannot be considered completely allopatric. First, one mtDNA contact zone was found in the Basque locality of Armendarits. Second, the M² shape of Basque bank voles appears as intermediate between the Spanish and the Western European ones, suggesting morphological differentiation but also genetic exchanges between populations.

These lineages seem to have different demographic histories. The Western European lineage differentiated 0.25–0.30 Mya and re-colonized Western Europe at the end of the last glaciation from a refugium situated in central Europe. This scenario was suggested on the basis of the decreasing nucleotide diversity gradient from central European populations to more northern populations (Deffontaine et al. 2005). On the other hand, the Spanish and Basque lineages display low values of haplotype and nucleotide diversity. These results contrast with the high genetic diversity of the Balkan lineage reported in Deffontaine et al. (2005). The low levels of genetic diversity of the former lineages can be explained by the small effective population sizes of these geographically restricted lineages and/or by severe population bottlenecks during the glaciations (Avise 2000). Indeed, during the glaciations, the Spanish peninsula underwent more arid and cold climatic events than the Balkan peninsula (Petit et al. 2003; Pinho et al. 2007), leading to the disappearance of many forest habitats suitable for the bank vole. Other temperate species, like the lizard Podarcis bocagei, also showed a loss of genetic diversity in northern Spain (Pinho et al. 2007). Despite its low levels of genetic diversity, the Basque lineage represents the most differentiated bank vole mtDNA lineage described to date. These results are reminiscent of a previous study on human populations that identified a Basque mtDNA lineage as the oldest lineage in Europe (González et al. 2006).

What can the history of this bank vole Basque lineage tell us about the survival of temperate species during past climatic changes?

Insight into the Basque lineage’s phylogeographic history

A previous phylogeographic study on bank voles has shown that globally shallow gene trees are likely the result of strong population bottlenecks during the Quaternary glaciations (Deffontaine et al. 2005). The bank vole phylogeography is associated with ‘Category III’ as defined by Avise (2000), which corresponds to closely related but geographically localized haplotypes. The multiple modes of the mismatch distribution and other statistics regarding demographic history suggest that the Basque population size would have remained stable for a long time compared to the Spanish and ‘continental’ lineages during these strong climatic fluctuations. Moreover, the complete lineage sorting of a maternal germ line and the emergence of reciprocal monophyly suggest a long-term isolation of populations (Avise 2004). However this hypothesis must be confirmed by other genetic markers. The basal position of the Basque lineage may suggest that it could be at the source of the radiation of all other bank vole lineages in Europe. However, this is not very likely given that the Basque haplotypes are geographically restricted to one area and are not paraphyletic in the different phylogeographic reconstructions (Parker & Kornfield 1997). It is also unlikely that the Basque lineage was a remainder of an old extinct lineage in Western Europe with a formerly larger distribution area, because Basque-like haplotypes were only found in the western part of the Pyrenees.

Hypothesis of a glacial refugium in the NW Pyrenees

The discovery of a bank vole lineage in the Basque region supports the hypothesis of glacial refugia for temperate species around the Pyrenees. We cannot totally rule out that the Basque individuals could be migrants from another glacial refugium where populations subsequently went extinct or that were not sampled in the present study. However, several arguments support the hypothesis of a glacial refugium in the NW Pyrenees where this lineage is currently distributed.

First, palaeo-environmental studies document that forest habitats were available for bank voles during the glaciations in the vicinity of the Pyrenees (Fig. 1b). Flint (1971)
and Brown & Gibson (1983) reported the presence of steppe-tundra in the northern part of the Pyrenees during the Quaternary glaciations, whereas Williams et al. (1998) indicated that the southern boundary of permafrost was in southern France. In the Aquitaine basin, in southwestern France, tundra formation was reported during the last glacial maximum (LGM, Drucker et al. 2003). In another study, Antoine et al. (1999) showed that the region situated between the Aquitaine basin and the French Basque country (to the southwestern of the ‘Landes’) was covered by aeolian sand and thus was inappropriate for forest-dwelling species (Fig. 1b). In the Pyrenees, the geological records of the last glaciation showed that piedmont glaciers reached the Aquitaine basin in the Pau and Ossau valleys, thus at c. 60 km west to Basque locality of Armendarits (Fig. 1b) (Pallás et al. 2006). The retreat of these Pyrenean ice sheets began before the LGMs (24 400 years before present for the Ossau glacier). The re-colonization of pine and temperate forests, however, was reduced by a phase of climate aridity during the following 14 000 years (Andrieu 1987). On the basis of this information and despite the lack of palaeo-environmental data for the French Basque country, we can hypothesize that this region was likely sand- and ice-free during the LGMs and that the proximity with the Atlantic Ocean would have brought a certain amount of humidity during the aridity phase mentioned above. Indeed, the French Basque country has an ideal geographic position in Europe to allow the development of a mosaic of habitats comprising forests.

Second, the NW Pyrenees are considered as a refuge region for many plants and animals species (García-Barros et al. 2002). For example, this region contains many archaeological records which highlight the presence of humans for at least the last 28 000 years (González et al. 2006). The current oldest European deciduous forests (like the Iraty forest and the Sare forest) are situated in the NW Pyrenees (Bourquin-Mignot & Girardclos 2001). These forests are known to shelter relict populations of species commonly found in central European deciduous forests, such as the woodpecker Dendrocopos leucotos lilfordi and the beetle Elelonoprius armatus (Kompanetsva & Tschigel 2000; Grangé 2001; Gorman 2004). However, it is clear that the old deciduous and coniferous forests observed nowadays in the NW Pyrenees do not correspond to the location of the forested glacial refugia. Indeed, during the LGM, several valleys at the east of the French Basque country were glaciated (see above). Vegetation records indicated that the major part of the Pyrenean massif was covered by steppic vegetation at an altitude of approximately 400 m (Antoine et al. 1999). Most of our study sites in the mountain were situated at an altitude between 900 m and 1500 m and was likely glaciated or steppic during the LGM. The major part of the Pyrenean mountains and bank voles, the glacial refugia would not have been situated at fairly high altitude but in the Basque foothills, close to the Atlantic ocean, at an altitude below 400 meters. This hypothesis is supported by the discovery of bank vole fossils dated back to the LGM (radiocarbon datings from 12 310 to 16 200 years ago (Peman 1985) in the Basque country near the Atlantic shoreline and at an altitude of only 230 meters (Fig. 1b). A recent study on the Pyrenean lineages of the bushcricket and of the mountain forest butterfly arrived at the same conclusion, that populations would have survived in the lower regions of the valleys during recent glaciations (Spoon & Ritchie 2007; Schmitt & Haubrich 2008). The valley systems offer a diversity of persisting, local, favourable microclimatic conditions and habitats and provide the populations with the opportunity to track the shift of the vegetation in altitude.

Post-glacial expansion through the Pyrenees

Previous demographic analyses showed that several bank vole lineages (Western European, Eastern European, Spanish and Italian) experienced post-glacial expansion after the LGM (Defontaine et al. 2005). The lineages meet in suture zones where, generally, gene flow occurs. The average M³ shape of Basque bank voles appears as intermediate between the Spanish lineage and that of the Western European group, suggesting restricted genetic exchanges between populations. In the reviews by Hewitt (1999, 2000), the Pyrenees are considered as one of the main suture zones in Europe. In the western part of the Pyrenees, two bank vole lineages were identified: the Basque lineage on the northern slope and the Spanish lineage on the southern slope. We hypothesize that the growth of the ice cap and steppe during glaciations has decreased the gene flow between the two lineages, promoting their differentiation.

The geographic distribution of the Basque lineage suggests that it did not contribute to the colonization of the entire Pyrenees and of northern Europe. One possible explanation is that the north–south orientation of the Pyrenean ridges may have blocked the expansion of the Basque vole population to the east. Alternatively, it is possible that the development of marshes after the deposits of aeolian sand in the southwestern Landes region could have blocked the northward expansion of the Basque vole population during the warming in the Holocene. These marshes were drained by humans in the 18th century by extensive planting of maritime pines (Pinus pinaster) (Sargos 2004). A third possible explanation is that the local adaptation of Basque
vole population to the Basque foothills environment may have prevented their expansion in less suitable habitats. Further studies using a larger sampling from the French Basque country and surrounding populations will provide a more detailed picture of the expansion processes of the Basque lineage and gene flow with other populations.

Conclusions

Our study strongly suggests a differentiated bank vole population in Europe, geographically localized in the NW part of the Pyrenean region. Genetic methods (mtDNA polymorphism) and morphological analyses (M² shape differentiation) highlight the singular phylogeographic history of the bank vole Basque population. Many phylogeographic studies assume that glacial refugia for temperate forest species were mainly located in southern Mediterranean peninsulas. Our study suggests a potential glacial refugium for a temperate forest species in mountain foothills and points to the biogeographic importance of the NW Pyrenees. Our results also provide another piece in the puzzle regarding the palaeo-environments of Western Europe during cold climatic phases by suggesting that the French Basque country was situated in the transition area between the steppe-tundra and mixed coniferous-deciduous biotopes, where forest-dwelling species were able to survive.

Finally, we hope that the discovery of this Basque lineage will draw close attention to the potential phylogeographic significance of this area. Other lineages may be discovered in the future inside the phylogeography of relict forest species of the Basque area (woodpeckers, beetles). This source of genetic diversity may be important for the conservation of species biodiversity in Europe.

Acknowledgements

We thank Dr R. Biek for his scientific comments and his help with improving the English grammar of the manuscript. The authors are grateful to Pr. G. Jalut, Dr M. Anisimova, Dr S.J.E. Baird, Dr O. Gasquet, M. Dao, B. Dumont, F. Santos, J.-L. Grangé and L. Soldati for their help or comments. We thank all persons who provided new tissue samples of bank voles: F. Catzeflis, J. Cassaing, C. Feliu, J. Goüy de Bellocq, E. Magnanou, C. Nieberding, M.G. Ramalinho and R. Rosoux. We thank four anonymous referees for their helpful suggestions. This publication has been funded under the EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, GOCE-CT-2003-010284 EDEN) and is officially catalogued by the EDEN Steering Committee as EDEN0140 and as a MARE publication n°164. The content of this paper does not represent the official position of the European Commission and is entirely the responsibility of the authors. This work was also funded by the Belgian FRS-FNRS (Fonds de la Recherche Scientifique, ‘crédit de fonctionnement’ discerned to V. Deffontaine). V. Deffontaine and M.C. Fontaine were supported by a Belgian research fellowship from the FRS-FNRS (‘mandat Aspirant’). The Postdoctoral fellowship of Valérie Deffontaine at the University of Glasgow was funded by the Human Frontier Science Program (grant No. ST-0086/2007-C) and the Wellcome Trust Value in People Award.

References


© 2009 Blackwell Publishing Ltd

Dekonenko A, Yakimenko V, Ivanov A et al. (2003) Genetic similarity of Puumala viruses found in Finland and western Siberia and of the mitochondrial DNA of their rodent hosts suggests a common evolutionary origin. Infection, Genetics and Evolution, 3, 245–247.


© 2009 Blackwell Publishing Ltd
Phylogeography of Bank Vole in Pyrenees


Pensof Publisher, Sofia.


Pensof Publisher, Sofia.


Pensof Publisher, Sofia.


Pensof Publisher, Sofia.


new haplotypes published in the present study. In bold: the samples selected to build the dataset on bank voles from the Pyrenean region.

Table S2  Geographic locations, latitude (in decimal degree North), longitude (in decimal degree East), codes, genetic lineage and total number of animals per locality (N) of *Myodes glareolus* third upper molars used in the morphological analyses.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.