The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2011, **102**, 11–21. With 3 figures

Italian Peninsula preserves an evolutionary lineage of the fat dormouse *Glis glis* L. (Rodentia: Gliridae)

SABRINA LO BRUTTO*, MAURIZIO SARÀ and MARCO ARCULEO

Dipartimento di Biologia Animale 'G. Reverberi', University of Palermo, via Archirafi 18, 90123 Palermo, Italy

Received 23 May 2010; accepted for publication 3 August 2010

The present study examines the population genetic structure of fifty-nine specimens of *Glis glis* (Linneaus, 1766) from thirteen localities in central Europe, sequencing a 400-bp segment of the mitochondrial cytochrome *b* (cyt *b*) gene and a 673-bp segment of the cytochrome *c* oxidase subunit I (COI) gene. The consensus tree obtained from Bayesian analysis revealed a robust dichotomy, showing two sister groups: one clade includes samples from a wide geographical area, extending from north-central Europe to northern Italy (major branch *sensu* Bilton), and the other comprises samples collected in central and southern Italy and in Sicily (Italian branch). According to the Tajima–Nei model, the two phylogroups were separated by a sequence divergence of 0.8% (cyt *b*) – 2.6% (COI), showing the COI gene to be more informative than cyt *b*. On a smaller geographical scale, the Italian clade was further substructured, displaying geographical differentiation along the Peninsula. The gene pool in this area was patchy; whereas populations from Sicily Island demonstrated fixed cyt *b* and COI haplotypes, assuming processes of isolation and selection. *Biological Journal of the Linnean Society*, 2011, **102**, 11–21.

ADDITIONAL KEYWORDS: Italy – mtDNA – phylogeography – rodents.

INTRODUCTION

Ice age events have produced effects on the differentiation among populations in relation to the latitude and the topography of lands covered by the species' range (Hewitt, 2004; Lister, 2004; Stewart *et al*., 2010); however, at the same time, individual species have responded to climatic fluctuations in different ways (i.e. species from the same communities dispersed diachronically in different directions and at various rates) (Graham *et al*., 1996; Stewart *et al*., 2010).

Many of the investigated species were present in a surface area from where they could spread in any direction (Graham *et al*., 1996; Stewart *et al*., 2010), whereas few comprised terrestrial mammals that could have dispersed only in a constrained direction, delimited by the peculiar outline of the territory; such as the case of species with a range covering geographical areas with peninsulas or islands.

In the light of this, phylogeographical studies on terrestrial mammal species necessitate a deeper knowledge of the phylogeography of species inhabiting the temperate European region, given its complex geomorphology as a result of the presence of peninsulas and large islands.

The Mediterranean region is one of the most irregular areas worldwide as a result of the presence of three Peninsulas that extend from the westernmost Iberian, to the central Italian and the easternmost Balkan Peninsula, and many large islands (e.g. Sardinia and Sicily), together with numerous archipelagos. In particular, the peninsulas projecting from north to south cover a range of latitude, along which features of the landscape and environment vary. Particular features of these Mediterranean lands assume major importance if we consider the Mediterranean Italy and Iberia peninsulas as a potential landbridge between the two continents of Europe and Africa,

*Corresponding author. E-mail: sabrilob@unipa.it

especially with respect to the phylogeography of the Palaearctic species (Lo Brutto, Arculeo & Sarà, 2004; Stöck *et al*., 2008; Médail & Diadema, 2009; Trucchi & Sbordoni, 2009).

Italy contains the richest assemblage of mammals of any European country: specific taxa are represented by clearly distinctive lineages often restricted to the southern parts of the Peninsula (Gippoliti & Amori, 2002). The need to collect phylogenetic information and biogeographical data relating to mammals in Italy has recently emerged in the literature (Bilton *et al*., 1998; Gippoliti & Amori, 2002; Hewitt, 2004; Michaux, Libois & Filippucci, 2005; Grill *et al*., 2009; Médail & Diadema, 2009; Trucchi & Sbordoni, 2009) because genetic diversity distributions and taxonomy at the species-rank level of several mammal taxa have still not been satisfactorily analyzed. The decision to study species covering a wide range (i.e. extensive territories from the continental mainland to islands) has led us to select the fat dormouse, *Glis glis* (Linneaus, 1766), a Palaearctic rodent, whose population genetics data from the literature are still only beginning to emerge (Hürner *et al*., 2010).

The fat dormouse (*Glis glis* L., Gliridae) is an arboreal, nocturnal rodent that inhabits deciduous and mixed forests of temperate and Mediterranean zones from northern Spain to the Caucasus mountains (Niethammer, 1990). Its diet consists of the fruit and seeds of trees and shrubs (Niethammer, 1990). The availability of food appears to be the major factor limiting reproduction because fat dormice do not reproduce in years when the hard mast crop is scarce or absent (Schlund, Scharfe & Ganzhorn, 2002; Pilastro, Tavecchia & Marin, 2003). Within its range, the fat dormouse hibernates for 7–9 months and reproduces from June to August, depending on local climate. Mating occurs at the beginning of July, and females produce only one litter a year, generally in August (Sciński & Borowski, 2008). Features such as its long adult life expectancy (6–12 years; Pilastro *et al*., 2003), intermittent reproduction, and a relatively short period of biological activity throughout the year distinguish the life history of this rodent from other small rodents. To assess the genetic structure of these populations (also with the intention to advance the intraspecific taxonomy), mitochondrial cytochrome *b* (cyt *b*) and cytochrome *c* oxidase subunit I (COI) sequencing analysis was performed on specimens of fat dormouse, mainly sampled in the Italian Peninsula and Sicily.

To date, different subspecies of *G. glis* have been described in Italy (Capizzi & Angelici, 2008): the nominal *Glis glis glis*, covering various Alpine provinces and probably the Po plain, whose range largely overlaps in the southern Alps with *Glis glis italicus* (Barrett-Hamilton, 1848). The latter is recorded throughout most of peninsular Italy and Sicily, whereas *Glis glis melonii* Thomas, 1907 has been described with reference to Sardinia. The fat dormouse found in Sicily has the typical *italicus* fur pattern but the smallest body size of the Italian populations; however, some individuals are comparable or even slightly larger than *G. g. melonii* (Milazzo, Falletta & Sarà, 2003).

MATERIAL AND METHODS

Sample collection

Tissue samples of fifty-nine *G. glis* L. specimens were collected from 13 populations located in four European countries (Fig. 1A) but mostly in the Italian Peninsula (Fig. 1B), in the period between 1998 and 2004. The tissues were kept in ethanol at –20 °C or frozen at –80 °C.

Polymerase chain reaction (PCR) and sequencing

Total DNA was extracted from a very small (approximately 20 mg) piece of tissue using the DNeasy Tissue Kit (Qiagen). Two fragments of the mitochondrial genome were amplified by PCR, using primers corresponding to the cyt *b* and COI genes: H15149 and L14724 (Irwin, Kocher & Wilson, 1991), LCO1490 and HCO2198 (Folmer *et al*., 1994). DNA was amplified in 50-μL reaction volumes which contained 2 units of *Taq* polymerase, 5 μL of 10 × reaction buffer [200 mM $(NH_4)SO_4$, 750 mM Tris HCl pH 8.8, 0.1% (v/v Tween), 4 μl of 25 mM $MgCl_2$, 10 μl of 200 mM dNTPs, 10 pmol of each primer and approximately 200 ng of DNA. The amplification conditions were: strand denaturation at 94 °C for 1 min; annealing at 50 °C for 1 min for COI; two-steps annealing at 48 °C for 45 s, five cycles, at 52 °C for 45 s, 30 cycles for cyt *b*; primer extension at 72 °C for 2 min, repeated for 35 cycles; always preceded by a single cycle of 5 min denaturation at 94 °C and followed by a single cycle of 10 min extension at 72 °C. The single bands were cut from the standard agarose gel, extracted and purified using a QIAquick Gel Extraction Kit (Qiagen) and the DNA sequenced in an automated sequencer (ABI PRISM 3700; Applied Biosystems).

Phylogenetic analysis

Sequences were before manually edited with CHROMAS software version 1.45 (Technelysium Pty Ltd) and then aligned by CLUSTAL-W (Thompson, Higgins & Gibson, 1994). Phylogenetic and molecular evolutionary analyses were conducted using ARLEQUIN, version 3.0 (Excoffier, Laval &

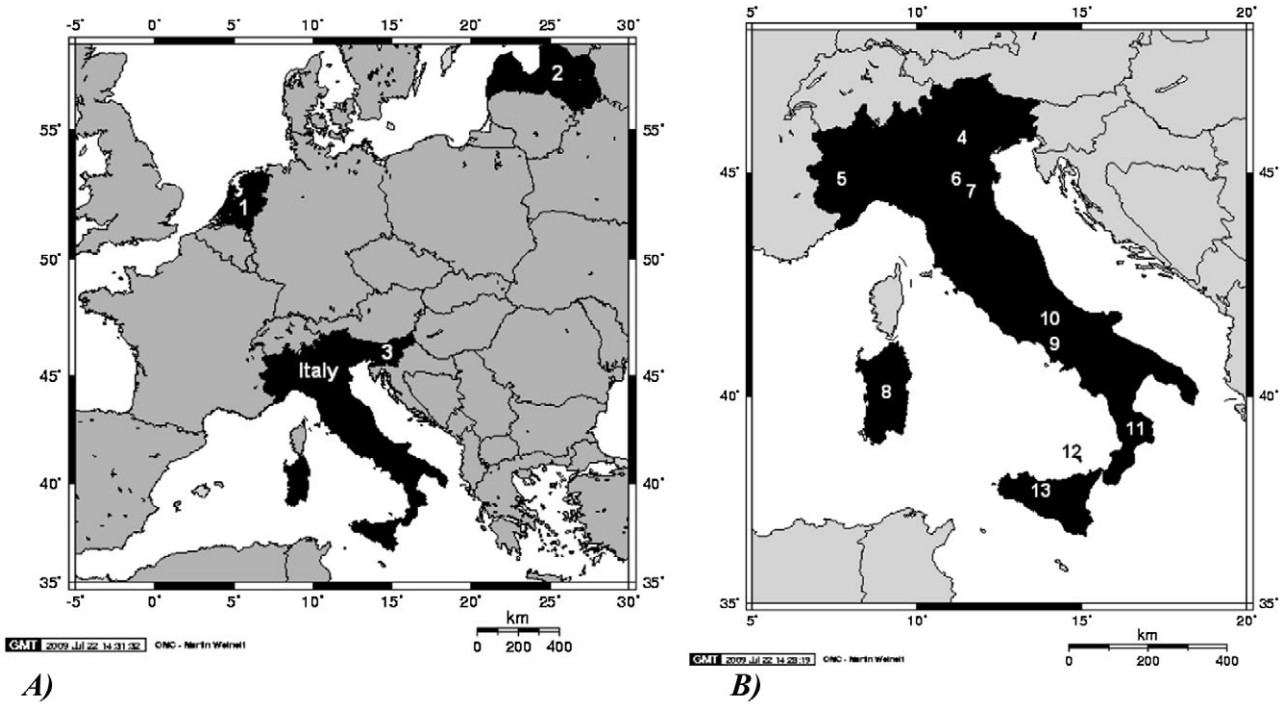

Figure 1. Sampling localities. A, European countries are indicated in black: Latvia [1]; Holland [2]; Slovenia [3]; and Italy. B, Italian localities: Asiago [4]; Torino [5]; Modena [6]; Bologna [7]; Sardegna [8]; Campania [9]; Molise [10]; Calabria [11]; Salina Island [12]; Madonie mountains [13] in Sicily.

Schneider, 2005), DNASP, version 4.10 (Rozas *et al*., 2003), and MEGA, version 4.0 (Tamura *et al*., 2007). Haplotype (h) and nucleotide (π) diversities (Nei, 1987) evaluated genetic variability.

The best-fit model (GTR+G) selected by the Akaike information criterion (AIC) in MRMODELTEST, version 2.3 (Nylander, 2004) was the same for the two single mitochondrial regions and for the concatenated fragment [cyt b + COI].

A partition homogeneity test [i.e. the incongruence length difference test (ILD; Farris *et al*., 1994)] implemented in PAUP 4.0b10 (Swofford, 2002) was calculated with heuristic search (1000 replicates) on the different partitions, cyt b and COI sequences, to assess congruence among data for the two different genes. Thus, phylogenetic analyses, conducted with bayesian analysis, using MrBayes 3.1 (Ronquist & Huelsenbeck, 2003), and the distance method, calculated by the Tajima & Nei model (1984), were performed considering both genes concatenated. A *Mus musculus domesticus* sequence, retrieved from GenBank (A.N. NC006914), was used as outgroup.

Furthermore, haplotype networks were constructed using TCS, version 1.21 (Clement, Posada & Crandall, 2000), using the statistical parsimony approach of Templeton, Crandall & Sing (1992).

The partitioning of genetic variability among populations was tested using a hierarchical analysis of molecular variance (AMOVA), performed in the ARLEQUIN software on the concatenated fragment [cyt b + COI]. Two alternative groupings were tested to obtain the best fit for the data. Because the samples were not representative of a population-level survey and they could not be used to assess a population structure, data were clustered according to the two main clades, and according to a finer structuring of the Italian branch. Φ-statistics, Φ_{ST}, analogous to Wright's (1951) F_{ST} statistic, based on molecular divergence was used to assess this type of group subdivision; the grouping of localities that maximizes Φ_{ST} was assumed to be the most probable geographical subdivision.

Divergence time inference

MDIV (Nielsen, 2002) software estimates the divergence time between two populations. Using an 'isolation by migration' model, the program jointly estimates theta or the effective population size scaled by the neutral mutation rate ($\theta = 2N_{ef}\mu$, where N_{ef} is the effective female population size and μ is the mutation rate per nucleotide per year), symmetric gene flow ($M = N_{ef}\ m$, where m is the fraction of effective migrants per generation), divergence time ($T = t_1/N_{ef}$, where t_1 is the population divergence time in years before present), and time to the most recent

common ancestor (TMRCA = t_2/N_{ef}, where t_2 is the gene coalescence time in years before present). We ran each simulation 5×10^6 times with a 10% burn-in period, *sensu* Nielsen (2002). Likelihood values were calculated and the values with the highest posterior probability accepted as the best estimate.

To convert population divergence time (t_1) and the coalescence time (t_2) in years before present (Y_{BP}), the equation according to Brito (2005) was used:

$$Y_{BP} = [(T \times \theta)/2L]1/\mu$$

where L is the sequence length (400 bp for cyt *b*, 673 bp for COI) and μ is the mutation rate per site per generation. The equation was applied for a wide range of plausible mutation rates for mitochondrial (mtDNA) (i.e. 2% to 10% per Myr) (Taberlet *et al*., 1998; Michaux *et al*., 2005; Nabholz, Glémin & Galtier, 2008; Trucchi & Sbordoni, 2009; Hürner *et al*., 2010) and, successively, multiplied by a generation time in the range 2–5 years.

RESULTS

Sequence variation

The cyt *b* sequences produced five different haplotypes. Two of the eleven variable sites displayed non-synonymous substitutions: Met→Thr at the 127th base pair, corresponding to the second codon positon, in the EU-CYTB-2 haplotype; and Ala→Thr at the 54th base pair, corresponding to the first codon position, in both the IT-CYTB-2 and IT-CYTB-3 haplotypes (Fig. 2A). The average haplotype (h) and nucleotide diversity (π) were 0.711 and 0.008, respectively.

The COI sequences revealed ten haplotypes, showing most of the genetic variation observed from specimens found all along the Italian Peninsula, including Sicily. The extent of variation of the COI sequences, estimated as the average haplotype (h) and nucleotide diversity (π), was 0.824 and 0.015 respectively.

The TCS parsimony network (Fig. 2) shows the partition of haplotypes in the geographical area under study.

Phylogenetic analysis

The partition homogeneity test (ILD) indicated no significant conflicts between the COI and cyt *b* datasets ($P = 1.00$), supporting the congruence of the phylogenetic signal from both cyt *b* and COI genes. The test also indicated that combining the congruent partitions phylogenetic accuracy increases (Cunningham, 1997).

All the phylogenetic analyses rendered the same topology, with high bootstrap support for the Neighbour-joining tree constructed on the Tajima–Nei model (data not showb) and high Bayesian posterior probabilities values (Fig. 3).

Phylogenetic analysis gave a dichotomy topology in which two lineages were evident. The two clades can be considered consistent with a model proposed by Bilton *et al*. (1998), where a well-defined monophyletic group occurred widespread in the central European, including samples from a huge geographical area (up to 7000 km distance) and named 'the major branch', whereas clades from the peninsulas were differentiated from each other and, mainly, formed peninsular branches. Because the data of the present study fitted such a pattern scored in small mammals by Bilton *et al*. (1998), the two lineages have been named here European major branch and Italian branch, referring to 'major' solely regarding the width of the geographical distribution.

The pairwise nucleotide distance according to the Tajima–Nei model was in the range 0.000–0.018 for the cyt *b* gene and 0.000–0.029 for the COI gene (Table 1). The COI dataset proved to be more informative than the cyt *b*. Consequently, an average divergence from the COI sequences of approximately 0.0261 separated the two lineages, whereas the average calculation from the cyt *b* gene produced a three-fold lower divergence (0.0082) between the two major clades.

Regarding the genetic distances scored in the cyt *b* gene, it should be noted that the Central and South Italy subgroup (Table 1), clustered within the Italian clade, diverged from the rest of the Italian lineage by the same order of magnitude (with a value equal to 0.017) (Table 1) as from the European major branch lineage (0.018), revealing a high degree of differentiation, which was clearly drawn by the position of the IT-CYTB-1 in the TCS parsimony network (Fig. 2B).

The European major branch is represented by samples originating from north European countries and northern Italy, whereas the Italian branch exclusively represented Italian samples (Fig. 3); however, the two clades were not allopatric. Indeed, they were characterized by an overlapping zone corresponding to the area of the Italian north Apennines, where specimens from the same area (i.e. Bologna locality) belonged to different lineages (Fig. 3). Such a result is demonstrated in the consensus tree (Fig. 3). It is noteworthy that the Italian clade is more substructured than the European clade, with most of the divergence present in the central and southern areas of the Italian Peninsula. The Italian branch resulted in three subgroups, namely 'Central and South Italy', 'South Italy' and 'Sicily', as shown in Table 1, in accordance with the geographical position of the sampled populations.

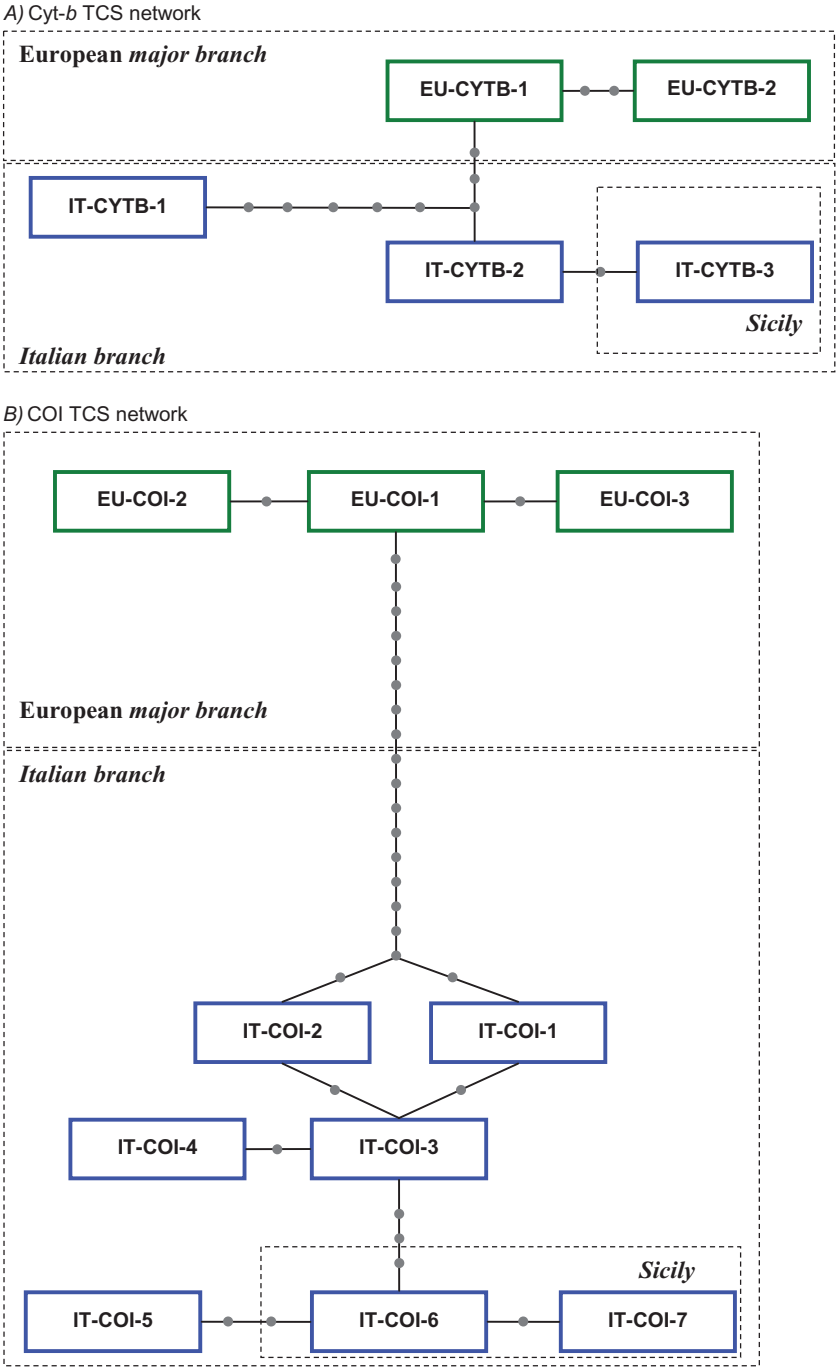

Figure 2. Parsimony network carried out using TCS, version 1.21, showing relationships found in cytochrome *b* (cyt *b*) (A) and cytochrome *c* oxidase subunit I (COI) (B) genes. Sequences were collapsed into haplotypes, represented by rectangles that are not proportional to the frequency of each haplotype. Haplotypes are drawn connected by estimates of genealogical information, on the basis of the mutational steps (small grey circles). Underlined is the separation of the two mitochondrial lineages as independent nets, European major branch (green rectangles) and Italian branch (blue rectangles); and the private haplotypes detected in Sicily. Cyt *b* haplotypes including the following localities/specimens. EU-CYTB-1: Slovenia, Latvia, Holland, Torino, Asiago, Modena; EU-CYTB-2: Bologna; IT-CYTB-1: Sardegna, Campania, Molise, Bologna, Calabria; IT-CYTB-2: Calabria; IT-CYTB-3: Madonie, Salina. COI haplotypes including the following localities/specimens. EU-COI-1: Holland, Slovenia, Latvia, Torino, Asiago, Modena, Bologna; EU-COI-2: Holland, Latvia; EU-COI-3: Modena; IT-COI-1: Bologna; IT-COI-2: Sardinia_1; IT-COI-3: Molise, Campania, Sardinia_5; IT-COI-4: Calabria_6; IT-COI-5: Calabria; IT-COI-6: Madonie, Salina; IT-COI-7: Madonie.

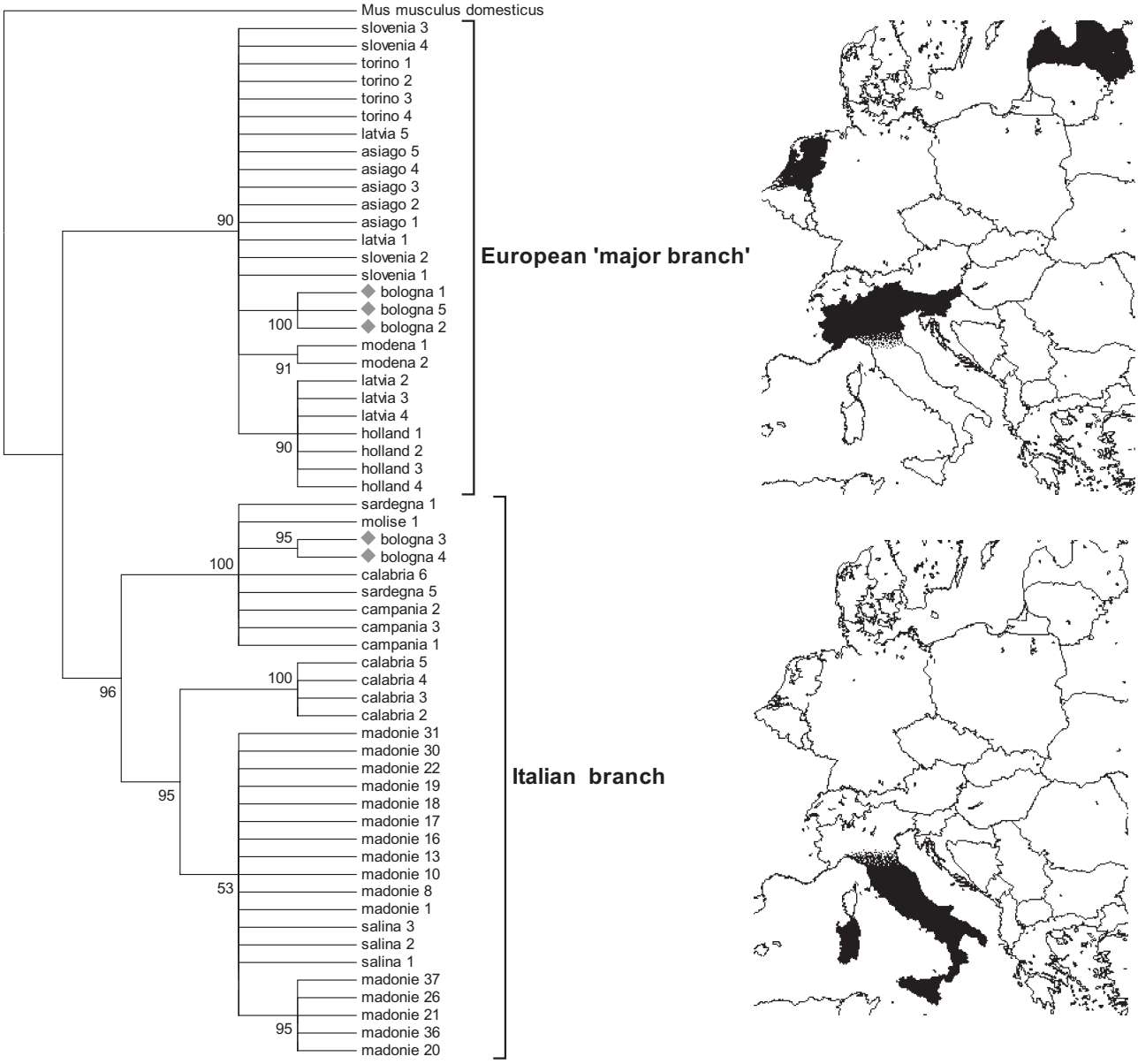

Figure 3. Consensus tree from the Bayesian analysis using the GTR+G model from the 1073 bp [cytochrome *b* (cyt *b*) and cytochrome *c* oxidase subunit I (COI)] concatenated sequences. Numbers at nodes indicate supports from Bayesian posterior probabilities in percentage. The dichotomy in the European major branch and Italian branch is demonstrated, together with their geographical separation represented in the maps; the putative transition zone between European clade and Italian clade, at the latitude corresponding to the North-Apennines Italian region, is shown. The two clades partially overlap and their boundary is not well marked. Indeed, the specimens from Bologna sample [sample-ID (7) in Fig. 1] belong at the same time to the two major clades; they are indicated by a symbol (grey rhombus). The structuring in subgroups within Italian clade is shown, such as the Sicily subgroup to whom the Salina [12] and Madonie [13] samples belong.

AMOVA was performed testing two alternative groupings according to the differentiation scored in the phylogenetic tree: I. [major branch] versus [Italian branch], and II. [major branch] versus [central and south Italy] versus [Sicily], the last according to a finer structuring of the Italian branch. AMOVA returned significant levels of genetic structure among the groups (I: $\Phi_{ST} = 0.854$, $P = 0.000$; II: $\Phi_{ST} = 0.908$, $P = 0.000$) indicating the existence of a geographical structure where the highest degree of discrimination was found when partitioning populations into three groups.

Table 1. Pairwise Tajima–Nei distances between localities

	European major branch							Italian branch						
								Central and South Italy			South Italy	Sicily		
Locality	(1) Latvia $N = 5$	(2) Holland $N = 4$	(3) Slovenia $N = 4$	(4) Asiago $N = 5$	(5) Torino $N = 4$	(6) Modena $N = 2$	(7) Bologna $N = 5$	(8) Sardegna $N = 2$	(9) Campania $N = 3$	(10) Molise $N = 1$	(11) Calabria $N = 5$	(12) Salina $N = 3$	(13) Madonie $N = 16$	
(1) Latvia		0.000	0.000	0.000	0.000	0.000	0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(2) Holland	0.000		0.000	0.000	0.000	0.000	0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(3) Slovenia	0.000	0.001		0.000	0.000	0.000	0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(4) Asiago	0.000	0.001	0.000		0.000	0.000	0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(5) Torino	0.000	0.001	0.000	0.000		0.000	0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(6) Modena	0.002	0.003	0.001	0.001	0.001		0.003	0.018	0.018	0.018	0.007	0.010	0.010	
(7) Bologna	0.003	0.004	0.003	0.003	0.003	0.004		0.007	0.007	0.007	0.004	0.009	0.009	
(8) Sardegna	0.026	0.027	0.026	0.026	0.026	0.027	0.008		0.000	0.000	0.009	0.018	0.018	
(9) Campania	0.028	0.029	0.027	0.027	0.027	0.029	0.009	0.000		0.000	0.009	0.018	0.018	
(10) Molise	0.028	0.029	0.027	0.027	0.027	0.029	0.009	0.000	0.000		0.009	0.018	0.018	
(11) Calabria	0.029	0.030	0.029	0.029	0.029	0.028	0.012	0.006	0.006	0.006		0.003	0.003	
(12) Salina	0.028	0.029	0.027	0.027	0.027	0.029	0.012	0.006	0.006	0.006	0.002		0.000	
(13) Madonie	0.028	0.029	0.028	0.028	0.028	0.029	0.012	0.006	0.006	0.006	0.003	0.000		

Above cytochrome *b*; below cytochrome *c* oxidase subunit I analysis. Clades and subgroups are shown.

Estimating divergence

Coalescent calculations, implemented in the software MDIV, of the two major clades divergence resulted in a value of $\theta = 2.293$, $M = 0.020$ and $T = 4.940$, and an estimate of TMRCA = 7.807 for COI; whereas $\theta = 0.866$, $M = 0.020$ and $T = 4.900$, and an estimate of TMRCA = 5.878 for the cyt *b* gene.

Estimated population divergence time (t_1) between the European major branch and the Italian line was 0.841–5.103 and 0.168–0.421 Myr BP, respectively for a 2% Myr^{-1} and a 10% Myr^{-1} mutation rate within a 2–5-year range of generation time. Estimates of the coalescent COI gene divergence from most recent common ancestor (t_2) were 1.330–3.325 Myr BP, assuming a substitution rate of 2% Myr^{-1} and 0.266–0.665 Myr assuming the faster substitution rate of 10% Myr^{-1}, within a 2–5-year range of generation time.

Regarding cyt *b* gene, estimated population divergence time (t_1) between the two major clades was 0.265–0.663 and 0.053–0.133 Myr BP, respectively for a 2% and a 10% mutation rate within a 2–5-year range of generation time. Estimates of the coalescent cyt *b* gene divergence from MRCA (t_2) were 0.318–0.795 Myr BP, assuming a substitution rate of 2% Myr^{-1} and 0.063–0.159 Myr BP assuming the faster substitution rate of 10% Myr^{-1}, within a 2–5-year range of generation time.

Because the existence of a lineage-specific nucleotide substitution rate in mammalian mtDNA has been demonstrated, even among species of the same order, (Gissi *et al*., 2000; Nabholz *et al*., 2008), the absence of a specific calibrated tree for *G. glis* limits the validation of a robust inference of divergence time.

Rodentia appears to be a fast-evolving mammal taxon (Nabholz *et al*., 2008) and, in the light of such observation, the divergence times to be considered more realistic should be the lowest. Such estimates should determine a temporal divergence consistent with a Pleistocene timescale. However, the time of the intraspecific differentiation is based on assumption of a not strongly defined mutation rate in the two genes. As argued by Nabholz *et al*. (2008), molecular dating studies relying on mtDNA require a cautious calibration step; thus, a fossil calibration within *Glis* genus requires further investigation.

DISCUSSION

Mitochondrial analysis divided the samples into two subsets, corresponding to two major clades, namely a European major branch and an Italian branch, which diverged with values of 0.8% and 2.6% for the cyt *b* and COI gene respectively. The split into two main genetic lineages is in accordance with a previous partition of the two different subspecies, *G. g. glis* and *G. g. italicus* (Capizzi & Angelici, 2008), thus supporting the current taxonomic rank and the status of species-complex of *G. glis* L. This latter should be further investigated especially in Sardinia and Sicily.

In accordance with the model proposed by Bilton *et al*. (1998) and the general pattern of 'northern purity versus southern richness' observed in most Palaearctic taxa (Hewitt, 2004; Stewart *et al*., 2010), the south Italian Peninsula appears as an area of high genetic diversity and evident differentiation among populations; this contrasts with the reduced diversity of the more extensive northern area of Europe. The Italian populations have probably experienced several events of isolation and expansion during Quaternary climatic fluctuations, which may have allowed the insurgence and survival of local haplotypes in different refugia, scattered throughout forest habitats along the Italian Peninsula and the island of Sicily.

The Italian Peninsula is traditionally classified as a Glacial Southern Refugium (i.e. a refugium for temperate species during glacial phases) (Stewart *et al*., 2010). In particular, two common phylogeographical patterns regarding the Italian Peninsula were scored in previous studies (Gippoliti & Amori, 2002; Hewitt, 2004): (1) Italian populations were found to be associated with populations coming from the Balkan peninsula (Michaux *et al*., 2005) because the frequent variation in the level of the Adriatic Sea during the Quaternary ice period allowed the Italian and Balkan populations to remain in contact with each other (Strahler & Strahler, 1989) and (2) other results obtained from genetic studies have identified the Alps as a biogeographical barrier because they limited the expansion of fauna (Stewart *et al*., 2010). Up to now, none of the phylogeographical patterns so far found in Italy and Europe have fitted the available data. The genetic structure of *G. glis* seems peculiar because the only Balkan sample (Slovenia) falls within the European clade rather than grouping with the Italian as is often reported (Taberlet *et al*., 1998; Hewitt, 2004; Michaux *et al*., 2005; Vega *et al*., 2010) and, in addition, because the boundary between the two main lineages is not marked by the Alps but is at a lower latitude of the Italian north Apennines region (Fig. 3). Obviously, the absence of relationship between the Balkan and Italian Peninsula is not sufficiently resolved with such data, and further investigations are needed. However, it is clear that the two European and Italian clades partially overlap at the borders of their geographical distribution in the northern Apennines region, as demonstrated by the co-presence of two mitochondrial forms in specimens from the same area (i.e. Bologna locality) (Fig. 3).

Hewitt (2004) and Stewart *et al*. (2010) quote cases of slow expansion, which depend on particular ecological conditions and the choice of a shorter route of dispersal. This matches our study case because vegetation also changed with the climate (Médail & Diadema, 2009) and species such as the fat dormouse, with a specialized niche in mixed deciduous forest (Milazzo *et al*., 2003; Koprowski, 2005), low dispersal capabilities, and a gregarious life history (Morris, 1997; Koprowski, 2005; Mortelliti, Santulli Sanzo & Boitani, 2009), can have expanded by passively tracking their habitat.

Accordingly, combined palaeobotanical and genetic data from the European beech *Fagus sylvatica* have revealed a scattered area of glacial refugia in central–southern Italy (Magri *et al*., 2006; Magri, 2008); during the postglacial period, southern Italian beech populations expanded at a slower rate and later than European beech populations and the former did not move much to the north to reach the lower reaches of the north Apennines (Magri *et al*., 2006; Magri, 2008).

Given the likely slow rate of colonization from the southern Italian refuge, European populations could have had the time to move across the Alps and to colonize empty niches in northern Italy. Another scenario hypothesizes a refugium in northern Italy, from where a propagule could have moved northwards. However, the genetic cohesion in the European major branch is so strong that a further extensive investigation needs to support such contrasting hypothesis.

The recognition of refugia is not easy (Stewart *et al*. 2010). A rigorous time-space scale should be considered; the climate cycles differed in amplitude and duration. Consequently long and short fluctuations can have had different effects on refugial isolation and range expansion of temperate species (Stewart *et al*., 2010). Furthermore, knowledge of the past geographical species distribution and the duration of isolated populations are pre-requisite to define a refugium, which represents the species' maximum contraction in the geographical range during the period of a glacial/interglacial cycle (Stewart *et al*., 2010).

Within the Italian clade, the TCS network (Fig. 2) and the Bayesian tree (Fig. 3) highlighted diverse subgroups associated with quite clear geographical patterns and, consequently, were named 'Central and South Italy', 'South Italy', and 'Sicily', for the approximate accordance to the location of the samples where the haplotypes were scored (Table 1). The geomorphology of the Italian Peninsula seems to lead us to suggesting a 'stepping stones' model, whereby adjacent populations exchange migrants along a one-dimensional direction, leaving the southern populations (i.e. the 'South Italy' and 'Sicily' subgroups) more isolated (Figs 2, 3). The interplay between complex historical processes and heterogeneous environmental conditions has given rise to considerable diversity and endemism along the Italian Peninsula, as highlighted by the diversity in such region.

If we compare the phylogeographical patterns in multiple co-distributed species, we observe that a structured geographical distribution among genetic lineages within different species in Italy is not rare. A genetic substructuring in Italy has been reported for the European roe deer *Capreolus capreolus* (Vernesi *et al*., 2002), the Italian hare *Lepus corsicanus* (Pierpaoli *et al*., 1999), the red squirrel *Sciurus vulgaris* (Grill *et al*., 2009), the Italian tree frog *Hyla intermedia* (Canestrelli, Cimmaruta & Nascetti, 2007), and, recently, the pygmy shrew *Sorex minutus* (Vega *et al*., 2010). However, these cases do not always fit the assumption of 'genealogical concordance' (i.e. concordance with the geographical positions of significant genetic partitions across multiple co-distributed species) (third aspect *sensu* Avise, 1998); especially, regarding the isolation of Sicilian populations from the rest of Italy.

The substructuring within the Italian lineage points to the role of Sicily as a putative refugium and this has been demonstrated to exert a powerful influence on current patterns of genetic diversity for other temperate European species (Pierpaoli *et al*., 1999; Canestrelli *et al*., 2007; Stöck *et al*., 2008). Evidently, the Quaternary features of Sicily, which have alternately increased and reduced the isolation from peripheral areas, have endowed the island with the capability of trapping populations of several species; so leading some of them to run independent evolutionary routes.

During glacial periods, Sicily was a refuge for temperate animals escaping the tough climatic conditions in Europe and northern Italy, and, contemporaneously, Sicily was an area to be colonized by north-African species that could move northwards thanks to the emergence of wide areas of the Sicilian Channel (Stöck *et al*., 2008). Conversely, during interglacial periods, Sicily became warmer and more isolated from surrounding regions, forcing temperate taxa into high altitude refuges and, possibly, facilitating the new arrival of southern fauna on the island. The current and past features of Sicily have endowed the island with the capability of trapping populations and leading them to their own independent evolutionary routes. This was the result of a marked degree of diversity of the landscape of the island and complex Sicilian palaeogeographical history, which has alternately increased and reduced the isolation of Sicily from peripheral areas.

The central position of Sicily, and contextually of Italy, in the Mediterranean basin implies that the study of phylogeography of species inhabiting this area is particularly appropriate for gathering exten-

sive information regarding the biological history of the entire Mediterranean region.

ACKNOWLEDGEMENTS

The authors are grateful to all of those who contributed by providing tissue or specimens, to Federico Marrone for the insightful comments during the preparation of the manuscript, and to the anonymous reviewers who provided valuable suggestions and comments during revision of the manuscript. The present study was supported by the University of Palermo (Fondi d'Ateneo, ex 60%).

REFERENCES

Avise JC. 1998. The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7:** 371–379.

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB. 1998. Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London Series B, Biological Sciences* **265:** 1219–1226.

Brito P. 2005. The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in Western Europe. *Molecular Ecology* **14:** 3077–3094.

Canestrelli D, Cimmaruta R, Nascetti G. 2007. Phylogeography and historical demography of the Italian treefrog, *Hyla intermedia*, reveals multiple refugia, population expansions and secondary contacts within peninsular Italy. *Molecular Ecology* **16:** 4808–4821.

Capizzi D, Angelici FM. 2008. Glis glis. In: Amori G, Contoli L, Nappi A, eds. *Fauna d'italia. Mammalia ii.* Erinaceomorpha – Soricomorpha – Lagomorpha – Rodentia, ed. Bologna: Calderini.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1659.

Cunningham CW. 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? *Systematic Biology* **46:** 464–478.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Farris JS, Kallersjo M, Kluge AG, Bult C. 1994. Testing the significance of incongruence. *Cladistics* **10:** 315–319.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3:** 294–299.

Gippoliti S, Amori G. 2002. Mammal diversity and taxonomy in Italy: implications for conservation. *Journal of Nature Conservation* **10:** 133–143.

Gissi C, Reyes A, Pesole G, Saccone C. 2000. Lineage-specific evolutionary rate in mammalian mtDNA. *Molecular Biology and Evolution* **17:** 1022–1031.

Graham RW, Lundelius EL, Graham MA Jr, Schroeder EK, Toomey RS III, Anderson E, Barnosky AD, Burns JA, Churcher CS, Grayson DK, Guthrie RD, Harington CR, Jefferson GT, Martin LD, McDonald HG, Morlan RE, Semken HA, Webb SD Jr, Werdelin L, Wilson MC. 1996. Spatial Response of mammals to late quaternary environmental fluctuations. *Science* **272:** 1601–1606.

Grill A, Amori G, Aloise G, Lisi I, Tosi G, Wauters LA, Randi E. 2009. Molecular phylogeography of European *Sciurus vulgaris*: refuge within refugia? *Molecular Ecology* **18:** 2687–2699.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **359:** 183–195.

Hürner H, Krystufek B, Sarà M, Ribas A, Ruch T, Sommer R, Ivashkina V, Michaux JR. 2010. Mitochondrial phylogeography of the edible dormouse (*Glis glis*) in the Western Palearctic Region. *Journal of Mammalogy* **91:** 233–242.

Irwin DM, Kocher TD, Wilson AC. 1991. Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution* **32:** 128–144.

Koprowski JL. 2005. The response of tree squirrels to fragmentation a review and synthesis. *Animal Conservation* **8:** 369–376.

Lister AM. 2004. The impact of Quaternary Ice Ages on mammalian evolution. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **359:** 221–241.

Lo Brutto S, Arculeo M, Sarà M. 2004. Mitochondrial simple sequence repeats and 12S-rRNA gene reveal two distinct lineages of *Crocidura russula* (Mammalia, Soricidae). *Heredity* **92:** 527–533.

Magri D. 2008. Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography* **35:** 450–463.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latalowa M, Litt T, Paule L, Roure JM, Tantau I, van der Knaap WO, Petit RJ, de Beaulieu JL. 2006. A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist* **171:** 199–221.

Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography* **36:** 1333–1345.

Michaux JR, Libois R, Filippucci MG. 2005. So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicolli*s) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Heredity* **94:** 52–63.

Milazzo A, Falletta W, Sarà M. 2003. Habitat selection of fat dormouse (*Glis glis italicus*) in deciduous woodlands of Sicily. *Acta Zoologica Academiae Scientiarum Hungaricae* **49** (Suppl. 1)**:** 117–124.

Morris P. 1997. *The edible dormouse* Glis glis. London: An occasional publication of the Mammal Society.

Mortelliti A, Santulli Sanzo G, Boitani L. 2009. Species' surrogacy for conservation planning: caveats from comparing the response of three arboreal rodents to habitat loss and fragmentation. *Biodiversity and Conservation* **18:** 1131–1145.

Nabholz B, Glémin S, Galtier N. 2008. Strong variations of mitochondrial mutation rate across mammals – the longevity hypothesis. *Molecular Biology and Evolution* **25:** 120–130.

Nei M. 1987. *Molecular evolutionary genetics*. New York, NY: Columbia University Press.

Nielsen R. 2002. *MDIV software*. Available at: http://ser-loopp.tc.cornell.edu/cbsu/mdiv.htm

Niethammer J. 1990. *Dormouse-group rodents. Grzimek's encyclopedia of mammals*, Vol. 3, New York, NY: McGraw-Hill, 276–291.

Nylander JAA. 2004. *MrModeltest 2.3*. Software distributed by the author. Evolutionary Biology Centre, Uppsala University.

Pierpaoli M, Riga F, Trocchi V, Randi E. 1999. Species distinction and evolutionary relationships of the Italian hare (*Lepus corsicanus*) as described by mitochondrial DNA sequencing. *Molecular Ecology* **8:** 1805–1817.

Pilastro A, Tavecchia G, Marin G. 2003. Long living and reproduction skipping in the fat dormouse. *Ecology* **84:** 1784–1792.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

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

Schlund W, Scharfe F, Ganzhorn JU. 2002. Long-term comparison of food availability and reproduction in the edible dormouse (*Glis glis*). *Mammalian Biology* **67:** 219–232.

Sciński M, Borowski Z. 2008. Spatial organization of the fat dormouse (*Glis glis*) in an oak-hornbeam forest during the mating and post-mating season. *Mammalian Biology* **73:** 119–127.

Stewart JR, Lister AM, Barnes I, Dalén L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society of London Series B, Biological Sciences* **277:** 661–671.

Stöck M, Sicilia A, Belfiore NM, Buckley D, Lo Brutto S, Lo Valvo M, Arculeo M. 2008. Post-Messinian evolutionary relationships across the Sicilian channel: mitochondrial and nuclear markers link a new green toad from Sicily to African relatives. *BMC Evolutionary Biology* **8:** 56.

Strahler AH, Strahler AN. 1989. *Elements of physical geography*, 3rd edn. New York, NY: John Wiley & Sons.

Swofford DL. 2002. *PAUP*: phylogenetic analysis using parsimony (*and other methods)*, Version 4. Sunderland, MA: Sinauer Associates.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453–464.

Tajima F, Nei M. 1984. Estimation of evolutionary distance between nucleotide sequences. *Molecular Biology and Evolution* **1:** 269–285.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24:** 1596–1599.

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132:** 619–633.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22:** 4673–4680.

Trucchi E, Sbordoni V. 2009. Unveiling an ancient biological invasion: molecular analysis of an old European alien, the crested porcupine (*Hystrix cristata*). *BMC Evolutionary Biology* **9:** 109.

Vega R, Amori G, Aloise G, Cellini S, Loy A, Searle JB. 2010. Genetic and morphological variation in a Mediterranean glacial refugium: evidence from Italian pygmy shrews, Sorex minutus (Mammalia: Soricomorpha). *Biological Journal of the Linnean Society* **100:** 774–787.

Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G. 2002. The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. *Molecular Ecology* **11:** 1285–1297.

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