

Research Space

Journal article

Phylogeographic structure of the pygmy shrew: revisiting the roles of southern and northern refugia in Europe

Vega, Rodrigo, McDevitt, Allan D, Stojak, Joanna, Mishta, Alina, Wójcik, Jan M, Kryštufek, Boris and Searle, Jeremy B

1 **Word count:**

2 10403 words (main body of text including references).

3

4 **Type of article:**

5 Original Article.

6

7 **Title:**

8 **Phylogeographic structure of the pygmy shrew: revisiting the roles of southern and**
9 **northern refugia in Europe.**

10

11 **Authors:**

12 RODRIGO VEGA^{1*}, ALLAN D. MCDEVITT², JOANNA STOJAK³, ALINA MISHTA⁴, JAN M.
13 WÓJCIK³, BORIS KRYŠTUFEK⁵ and JEREMY B. SEARLE⁶

14

15 **Affiliations:**

16 ¹*Ecology Research Group, Section of Natural and Applied Sciences, School of Human and*
17 *Life Sciences, Canterbury Christ Church University, Becket Bf23, North Holmes Road,*
18 *Canterbury, Kent, CT1 1QU, UK.*

19 ²*Ecosystems and Environment Research Centre, School of Science, Engineering and*
20 *Environment, University of Salford, Salford, M5 4WT, UK.*

21 ³*Mammal Research Institute, Polish Academy of Sciences, 17-230 Białowieża, Poland.*

22 ⁴*Schmalhausen Institute of Zoology, National Academy of Sciences of Ukraine, 01601 Kiev,*
23 *Ukraine.*

24 ⁵*Slovenian Museum of Natural History, Prešernova 20, SI-1000 Ljubljana, Slovenia.*

25 ⁶*Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York,*
26 *14853-2701, USA.*

27

28 *Corresponding author. E-mail: rodrigo.vega@canterbury.ac.uk

1

2 **Running title:**

3 Phylogeography of the pygmy shrew.

4

1 **ABSTRACT**

2 Southern and northern glacial refugia are considered paradigms that explain the complex
3 phylogeographic patterns and processes of European biota. Here, we provide a revisited
4 statistical phylogeographic analysis of the pygmy shrew *Sorex minutus* Linnaeus, 1766
5 (Eulipotyphla, Soricidae) examining the genetic diversity, genetic differentiation and
6 demographic history in the Mediterranean peninsulas and in Western and Central Europe.
7 The results showed support for genetically distinct and diverse phylogeographic groups
8 consistent with southern and northern glacial refugia, as expected from previous studies, but
9 also identified geographical barriers concordant with glaciated mountain ranges during the
10 Last Glacial Maximum (LGM), early diversification events dated between the Upper
11 Pleistocene and Lower Holocene for the main phylogeographic groups, and recent (post-
12 LGM) patterns of demographic expansions. This study is the most comprehensive
13 investigation of this species to date, and the results have implications for the conservation of
14 intraspecific diversity and the preservation of the evolutionary potential of *S. minutus*.

15

16 **KEYWORDS:** mitochondrially encoded cytochrome *b* – glacial refugia – historical
17 demography – Last Glacial Maximum – mammals – postglacial colonisation.

18

1 INTRODUCTION

2 During the Quaternary glaciations, species in Europe were restricted to glacial refugia at
3 glacial maxima (Bilton *et al.*, 1998; Taberlet *et al.*, 1998; Hewitt, 2000; Stewart & Lister, 2001;
4 Pazonyi, 2004; Sommer & Nadachowski, 2006). As glaciers retreated, a broad range of
5 recolonisation patterns emerged, as evidenced by palaeontological, biogeographic and
6 phylogeographic studies on various taxa, resulting in the complex contemporary patterns of
7 endemism, species richness and biodiversity hotspots observed across Europe. While
8 population contraction and lineage diversification within southern glacial refugia in the
9 Mediterranean peninsulas during the Last Glacial Maximum [LGM; 19-26.5 thousand years
10 ago (KYA) (Clark *et al.*, 2009)], and subsequent northward postglacial recolonisation of
11 Europe have been accepted and recognised since the 1990s (Bilton *et al.*, 1998; Taberlet *et*
12 *al.*, 1998; Hewitt 2000), the concept of northern glacial refugia also became a paradigm to
13 explain the complex phylogeographic patterns and processes of European biota (Stewart &
14 Lister, 2001; Pazonyi 2004; Sommer & Nadachowski, 2006). Fossil records and phylogenetic
15 analyses revealed that many species of flora and fauna could have survived during the LGM
16 in the Carpathian Basin (Stewart & Lister, 2001; Pazonyi, 2004; Sommer & Nadachowski,
17 2006; Stojak *et al.*, 2015) and in the Dordogne region (Steward *et al.*, 2010), and glacial
18 refugia could also be located in Crimea (Marková, 2011) or in the Russian Plain (Banaszek
19 *et al.*, 2012). Nowadays, locations of southern and northern glacial refugia during the LGM
20 are hotspots of genetic diversity (Petit *et al.*, 2003; Stojak *et al.*, 2016).

21 The Eurasian pygmy shrew *Sorex minutus* Linnaeus, 1766 (Eulipotyphla, Soricidae)
22 (Hutterer, 1990) has been used as a phylogeographic model species for understanding the
23 effects of the glaciations in Europe and the colonisation history during the Pleistocene and
24 postglacial times (Bilton *et al.*, 1998; McDevitt *et al.*, 2010; Vega *et al.*, 2010a, b). However,
25 little is still known about the phylogeographic structure, genetic diversity and structure, and
26 demographic history of this small mammal within these regions due to the limited number of
27 samples from Mediterranean peninsulas. An expanded phylogeographic study of the pygmy
28 shrew is therefore important for the understanding and further development of biogeographic

1 models of glacial refugia and postglacial recolonization, for depicting areas with high
2 intraspecific genetic diversity, for establishing conservation measures of rear-edge
3 populations, and for the preservation of the evolutionary potential of species, particularly in
4 the face of climate and anthropogenic change (Deffontaine *et al.*, 2005; Provan & Bennett,
5 2008; Stojak *et al.*, 2019; Stojak & Tarnowska, 2019).

6 In this study, we explored the evolutionary history and phylogeographic structure of
7 *Sorex minutus* using a statistical phylogeography approach (Knowles & Maddison, 2002;
8 Knowles, 2009). Here, we emphasise the genetic diversity and structure within and among
9 refugia, the inference of geographical barriers and the demographic history of *S. minutus*,
10 which are aspects that have not been studied in detail previously. Specifically, we asked the
11 following questions: 1) What are the geographical distribution and genetic diversity patterns
12 of the genealogical lineages of *S. minutus*? 2) Is there significant population genetic
13 structure across the geographic range of *S. minutus*? 3) What is the historical demography
14 of *S. minutus* in Europe? Our results showed support for distinct and genetically diverse
15 lineages, geographical barriers concordant with glaciated mountain ranges during the LGM,
16 and recent (post-LGM) population expansions with contemporary contact areas. The results
17 presented here have implications for the long-term conservation of intraspecific diversity and
18 the preservation of the evolutionary potential of *S. minutus* in the face of modern climate
19 change.

20

21 MATERIALS AND METHODS

22 *Study species*

23 *Sorex minutus* is common over most of its distribution but is rarely dominant and it occurs in
24 a wide range of terrestrial habitats with adequate ground cover and in relatively damp areas,
25 including swamps, grasslands, heaths, sand dunes, woodland edge, rocky areas, shrubland
26 and montane forests (Hutterer, 1990, 2016; Churchfield, 1990; Churchfield & Searle, 2008).
27 It is found from southern and western Europe to much of central and northern Europe,
28 Ireland and the British Isles, and Siberia to Lake Baikal in the east (Hutterer, 1990, 2016). It

1 is found from sea level up to 2260 m (in the Alps), but its distribution becomes patchy and
2 limited to higher altitudes in southern Europe where it occurs with some degree of
3 geographical isolation and differentiation, while in central and northern parts of Europe and
4 in Siberia it is more abundant and populations are more connected and widespread
5 (Hutterer, 1990, 2016).

6

7 *Samples and molecular methods*

8 A total of 671 mitochondrially encoded cytochrome *b* (*MT-CYB*) DNA sequences of *S.*
9 *minutus* from Europe and Siberia were used for this study (Fig. 1B; see Supplementary
10 information Table S1). DNA sequences were obtained from samples collected from the wild
11 following ethical guidelines (Sikes, Gannon & the Animal Care and Use Committee of the
12 American Society of Mammalogists, 2011), or from museums, and from published GenBank
13 data (including AB175132: Ohdachi *et al.*, 2006; AJ535393 – AJ535457: Mascheretti *et al.*,
14 2003; GQ272492 – GQ272518: Vega *et al.*, 2010a; GQ494305 – GQ494350: Vega *et al.*,
15 2010b; and JF510376 – JF510321: McDevitt *et al.*, 2011). In addition, four *MT-CYB*
16 sequences of *S. volnuchini*, which was used as an outgroup (Fumagalli *et al.*, 1999), were
17 incorporated into the analysis (including AJ535458: Mascheretti *et al.*, 2003).

18 Genomic DNA from wild and museum samples was extracted using a commercial kit
19 (Qiagen). Partial (1110 bp) *MT-CYB* sequences were obtained by PCR using two primer
20 pairs that amplified approximately 700 bp of overlapping fragments, or using five primer pairs
21 (for museum samples with highly degraded DNA) that amplified approximately 250 bp of
22 overlapping fragments (Vega *et al.*, 2010a). PCR amplification was performed in a 50 µl final
23 volume: 1X Buffer, 1 µM each primer, 1 µM dNTP's, 3 mM MgCl₂ and 0.5 U Platinum Taq
24 Polymerase (Invitrogen), with cycling conditions: 94°C for 4 min, 40 cycles at 94°C for 30 s,
25 55°C for 30 s and 72°C for 45 s, and a final elongation step at 72°C for 7 min. Purification of
26 PCR products was done with a commercial kit (Qiagen) and sequenced (Macrogen and
27 Cornell University Core Laboratories Center).

28

1 *Phylogenetic analysis*

2 DNA sequences were edited by eye in BioEdit version 7.0.9.0 (Hall, 1999), contigs were
3 made from forward and reverse sequences also in BioEdit, and sequences were aligned
4 using ClustalX version 2.0 (Larkin *et al.*, 2007). A haplotype data file was obtained using
5 DnaSP version 5.10.1 (Librado & Rozas, 2009). Newly obtained haplotypes were deposited
6 in GenBank (Accession Numbers: MN840358 - MN840484, Supplementary information
7 Table S1).

8 The model of evolution that best fitted the molecular data (haplotypes) was searched
9 using jModelTest version 2.1.10 (Darriba *et al.*, 2012) using the Bayesian Information
10 Criterion value. The substitution model supported was the General Time Reversible (GTR)
11 with specified substitution types (A–C=0.4250, A–G=23.5124, A–T=1.6091, C–G=1.8671, C–
12 T=17.2314, G–T=1.0000), proportion of invariable sites (0.6044), gamma shape parameter
13 (0.2816) and nucleotide frequencies (A=0.2777, C=0.3076, G=0.1416, T=0.2731).

14 The phylogenetic relationships among *MT-CYB* haplotypes of *S. minutus* were
15 inferred by Bayesian analysis and by generating a parsimony phylogenetic network. The
16 Bayesian analysis was done using MrBayes version 3.2.7 (Ronquist *et al.*, 2012) with two
17 independent runs (10 million generations and 5 chains each), a sampling frequency every
18 1000 generations and temperature of 0.1 for the heated chain, and checking for
19 convergence in Tracer version 1.7.1 (Rambaut *et al.*, 2018). Trees were summarized after a
20 burn-in value of 2500 to obtain the posterior probabilities of each phylogenetic branch. The
21 main phylogenetic groups (phylogroups) were identified based on monophyly of the
22 haplotypes, and were named based on the geographical origin of the samples. The
23 phylogenetic network was done using PopART version 1.7 (<http://popart.otago.ac.nz>)
24 implementing a median-joining algorithm.

25 Sequence polymorphism indices and diversity values, including the number of
26 haplotypes (H), polymorphic (segregating) sites (S) and parsimony informative sites (P),
27 haplotype diversity (Hd), nucleotide diversity (π), and average number of nucleotide
28 differences (k), were estimated using DnaSP. This was done for the whole data set (ingroup),

1 for the main phylogroups, and also for other relevant geographic groups, including island
2 populations and continental samples.

3

4 *Population genetic structure*

5 Pairwise genetic differentiation values (F_{ST}) between all pairs of phylogroups and other
6 relevant geographic groups, and an Analysis of Molecular Variance (AMOVA) were
7 calculated using Arlequin version 3.11 (Excoffier *et al.*, 2005). Ten thousand nonparametric
8 permutations were performed to generate a random distribution to test the significance of the
9 pairwise F_{ST} values and covariance components of the AMOVA, and $\alpha = 0.05$ was set as the
10 threshold for statistical significance.

11 With samples assigned to phylogroups and with the samples' geographical
12 coordinates, the geographic midpoints (i.e. mid-geographic location between two or more
13 coordinates) of the phylogroups were calculated using the Geographic Midpoint Calculator
14 (available at <http://www.geomidpoint.com/>). The geographic midpoints were used to obtain
15 the pairwise geographic distances among phylogroups with the Geographic Distance Matrix
16 Calculator version 1.2.3 (by PJ Ersts, available at
17 http://biodiversityinformatics.amnh.org/open_source/gdmg). A Mantel test was used to
18 evaluate the relationship between matrices of pairwise geographic distances and genetic
19 differentiation values (Slatkin's linearised pairwise F_{ST} as $D = F_{ST}/(1-F_{ST})$; Slatkin, 1995).
20 Despite criticisms, the Mantel test is still a widely used and can be a powerful statistical
21 approach to analyse sequence data to test evolutionary hypotheses (Diniz-Filho *et al.*,
22 2013). Due to the very low (or absence of) genetic variation in the Orkney islands, DNA
23 sequences originating from there were pooled to avoid issues with pairwise F_{ST} calculations.

24 Geographic barriers were computed using Barrier version 2.2 (Manni *et al.*, 2004).
25 This approach implements Monmonier's maximum difference algorithm to find edges
26 (boundaries) on a Voronoi tessellation associated with the highest rate of change in genetic
27 distances among samples interconnected by a geometric network (i.e. Delaunay
28 triangulation) (Manni *et al.*, 2004). A barrier highlights the geographic areas where a genetic

1 discontinuity is found, and where samples on each side of the barrier are genetically more
2 similar than samples taken on different sides of the boundary. Pairwise genetic distances
3 were estimated using continental samples only, limiting the data set in the geometric network
4 calculation to one sample per locality, and computing a maximum of 10 barriers.

5

6 *Historical demography*

7 A strict molecular clock was compared to the uncorrelated lognormal relaxed molecular clock
8 (Drummond *et al.*, 2006). Coalescent constant population size and Bayesian skyline
9 demographic models (Drummond *et al.*, 2005) were compared to identify the best-fitting
10 pattern of changes in the pygmy shrew population. For model selection, path sampling and
11 stepping-stone sampling (Baele *et al.*, 2013), based on four independent MCMC chains
12 (1000 steps of 100,000 generations each, following a 10 million generations burn-in period),
13 were used for calculating the log Marginal Likelihoods Estimates (MLEs) for each model.
14 MLEs were used to calculate Bayes Factors (BFs) for each comparison between tested
15 models to determine the best-fitting one (Kass & Raftery, 1995). The best-fitting models were
16 then used to estimate the Time of divergence from the Most Recent Common Ancestor
17 (TMRCA) and Bayesian Skyline Plots (BSP) (see below). The 95% Highest Posterior
18 Density (HPD) was included in the TMRCA and BSP estimations.

19 TMRCA for the ingroup (all *S. minutus* samples) and the phylogroups were
20 estimated using BEAST version 2.5.2 (Bouckaert *et al.*, 2014). The following prior
21 assumptions were: random starting tree, monophyletic groups (for the ingroup and the Irish
22 phylogroups) (Drummond *et al.*, 2006) to calculate the evolutionary rate, and the GTR
23 substitution model with four categories, gamma = 0.9680 and proportion of invariable sites =
24 0.4680 (from jModelTest using the full data set). The oldest record of *S. minutus* has been
25 found in Podlesice and Mała Cave, Poland dated between 5.3 and 3.6 MYA (Early Pliocene;
26 Mammal Neogene 14) (Rzebik-Kowalska, 1998). Using this fossil information, a calibration
27 point for the ingroup was set at 4.45 MYA (SD = 0.5 MY; 5.27 – 3.63 MYA) with a normal
28 prior distribution. Due to the absence of dated fossils of pygmy shrews that can be assigned

1 specifically to the main phylogroups, a second calibration was set for the node age of the
2 Irish lineage at 0.006 MYA (SD = 0.0005 MYA; 0.00682 – 0.00518 MYA). This secondary
3 calibration point, derived from a previous analysis and applied to our data set, was based on
4 the inferred colonisation time of Ireland by *S. minutus* in the Neolithic using multiple genetic
5 markers and fossil data (McDevitt *et al.*, 2009, 2011). The trace files were analysed in
6 Tracer, the tree information from the four runs were combined and resampled at a lower
7 frequency (for a total of 10,000 trees) using LogCombiner, and the information was
8 summarized using TreeAnnotator selecting Maximum clade credibility tree and median
9 heights. The phylogenetic tree showing the TMRCA was created using FigTree version
10 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) with median and 95% HPD values based on
11 those 10,000 trees.

12 Genetic evidence of population expansion for the phylogroups, island populations
13 and continental samples was investigated using the R_2 test of neutrality (Ramos-Onsins &
14 Rozas, 2002), based on the difference of the number of singleton mutations and the average
15 number of nucleotide differences, and Fu's F_s (Fu, 1997), a statistic based on the infinite-site
16 model without recombination that shows large negative F_s values when there has been a
17 demographic population expansion. Both population expansion tests were carried out in
18 DnaSP using coalescent simulations for testing significance (10,000 replicates).

19 Mismatch distributions (i.e. the distribution of the number of differences between
20 pairs of haplotypes) were estimated for the phylogroups (and where $N \geq 10$) to compare the
21 demography of the populations with the expectations of a sudden population expansion
22 model (Rogers & Harpending, 1992). For the phylogroups and continental samples that
23 showed a unimodal mismatch distribution and significant population expansion, the time
24 since the population expansion (t) was calculated as $t = \tau/2u$, where τ (tau) is the mode for
25 the unimodal mismatch distribution, and u is the cumulative (across the sequence)
26 probability of substitution (Schenekar & Weiss, 2011). The calculations were done using the
27 MS Excel Mismatch Calculator (Schenekar & Weiss, 2011) with sequence length = 1110 bp,
28 generation time = 1 year (Hutterer *et al.*, 2016), substitutions per site per million years

1 (subst/Site/MY) = 0.551 (based on the average substitution rate across all sites 'clock.rate'
2 results from BEAST) and cumulative substitutions/generation = 0.00062.

3 BSPs were calculated using BEAST based on the posterior distribution of effective
4 population size through time from a sample of gene sequences. This was done for the
5 phylogroups showing a unimodal mismatch distribution and significant signatures of recent
6 population expansion (where $N \geq 10$). The analysis was run for 100 million generations,
7 sampled every 1000, using the best-fitting model.

8

9 RESULTS

10 *Phylogenetic analysis*

11 For the complete *S. minutus* data set ($N = 671$) (Fig. 1B), there were 424 haplotypes with
12 390 polymorphic sites of which 277 were parsimony informative (Table 1). We report 160
13 newly sequenced specimens of *S. minutus* from the Iberian (4) and Balkan (19) peninsulas
14 and from Central and Northern Europe (137) from which 127 were new haplotypes. Also,
15 there were three new sequences and haplotypes of *S. volnuchini*, from which two were from
16 Turkey and one from the Crimean Peninsula.

17 The Bayesian phylogenetic analysis showed *S. minutus* as a monophyletic group and
18 revealed six distinct lineages corresponding to their geographical origin (i.e. phylogroups)
19 supported by high posterior probabilities (Fig. 2A). Samples from the Mediterranean
20 peninsulas clustered in three distinct phylogroups, namely the Iberian, Italian and Balkan
21 phylogroups. The Iberian group was represented with few DNA sequences ($N = 6$). It was
22 geographically restricted to the Iberian Peninsula and included samples from Rascafría,
23 Central Spain (Sierra de Guadarrama) and Picos de Europa, Northern Spain. The Italian
24 phylogroup ($N = 26$) was mostly restricted to the north-central regions of the Italian
25 peninsula; it included samples from the Apennines and the Alps in Italy, but also from
26 Switzerland, Slovenia, Southern and Eastern France near the border with Italy, Czech
27 Republic and Germany. The Balkan phylogroup ($N = 22$) included samples mostly from the
28 Balkan Peninsula and a few from further north in Central Europe. This phylogroup showed a

1 weak north/south subdivision, with one clade containing samples from Switzerland, Austria,
2 Slovakia, Czech Republic, Hungary and Montenegro, another clade containing samples from
3 Serbia, Bosnia and Herzegovina and North Macedonia, plus other ungrouped basal samples
4 from Montenegro, North Macedonia, Serbia and Turkey (East Thrace, Southeast Europe).

5 There was also a well-supported and geographically widespread Western phylogroup
6 ($N = 283$), which included samples from northern Spain (Cantabrian Mountain Range),
7 Southern France and Andorra (i.e. the Pyrenees), western and central France (including
8 Belle-Île), Ireland, the Orkney Islands, and western mainland Britain and offshore islands on
9 the western coast of mainland Britain. Samples from Ireland formed an internal monophyletic
10 lineage (i.e. the Irish phylogroup, $N = 94$) within the Western phylogroup. Notably, two
11 samples from Navarra in northern Spain (ESNa0861 and ESNa1131; Accession Number
12 JF510331) shared haplotypes with samples from Ireland (Hap_64). A monophyletic South
13 Italian phylogroup ($N = 4$) was most closely related to the Western phylogroup than to the
14 Italian phylogroup, and was geographically restricted to La Sila, Calabria in Southern Italy.

15 Samples from northern and central Europe and Siberia, namely the Northern
16 phylogroup ($N = 330$), formed the most geographically widespread lineage and included
17 samples ranging from Central France and Britain (excluding those within the Western
18 phylogroup), across Central and Northern Europe to Lake Baikal in Siberia, but did not
19 include samples from Southern Europe. Samples from mainland Britain belonging to the
20 Northern phylogroup did not form an internal monophyletic cluster.

21 The phylogenetic network had a complex structure (Fig. 2B), but the haplotypes
22 clustered into the same phylogroups detected with Bayesian phylogenetics and were
23 distantly related from each other (> 10 mutational steps). The Western phylogroup had a
24 star-like pattern and showed three most internal haplotypes; notably, Hap_61 was found in
25 the Pyrenees with other Western haplotypes directly connected to it, Hap_94 was found on
26 islands of the western coast of Scotland (Arran and Mull) with other Scottish and continental
27 Western haplotypes directly connected to it, and Hap_64 included samples from Northern
28 Spain and Ireland with other Irish haplotypes connected to it. The Northern phylogroup

1 showed a star-like pattern with many reticulations and three most internal haplotypes
2 separated from each other by few mutational steps. There was a weak geographical
3 subdivision within the Northern phylogroup, where samples from Siberia, Eastern and
4 Northern Europe were derived or most closely connected to samples from Central Ukraine
5 (Hap_287), samples from Central Europe were derived or most closely connected to
6 samples from The Netherlands (Hap_274), and all samples from Britain were derived or
7 most closely connected to other samples from The Netherlands than to the other central
8 haplotypes (Hap_90); however, the highly reticulated pattern of the inner haplotypes of the
9 Northern phylogroup indicated that this geographical subdivision was weak.

10 Sequence polymorphism indices and diversity values for the phylogroups and other
11 geographic groups are shown in Table 1. For the phylogroups, the haplotype diversity values
12 were high (>90%), and the nucleotide diversity values were either half or almost half as
13 much as the ingroup. Notably, the Northern phylogroup had the highest haplotype diversity
14 values, followed by the Balkan phylogroup; however, the Balkan phylogroup had the highest
15 nucleotide diversity values. The Irish phylogroup, which clustered within the Western
16 phylogroup, showed slightly lower haplotype diversity than any other phylogroups.

17 The continental groups (Northern continental and Western continental) showed
18 equivalent DNA polymorphism values as the main phylogroups, but the island groups
19 showed different levels of DNA polymorphism (Table 1). There was low DNA polymorphism
20 in islands of the Orkney Archipelago, with only 11 haplotypes in all Orkney Islands combined
21 ($N = 119$), but all haplotypes were unique to these islands. There were eight haplotypes in
22 Orkney Mainland ($N = 44$), from which seven were unique to this island (the largest island of
23 the archipelago), there were two unique haplotypes in Orkney South Ronaldsay ($N = 40$),
24 and there was only one haplotype in Orkney Westray ($N = 33$) also present in Orkney Hoy
25 ($N = 2$) and Orkney Mainland. There were five haplotypes in Belle-Île ($N = 5$), and only one
26 was present in the continent also belonging to the Western phylogroup. The British group (N
27 = 91) showed high haplotype diversity but moderate nucleotide diversity values and had 80
28 haplotypes from which 77 were unique haplotypes not found elsewhere.

1

2 *Population genetic structure*

3 The highest pairwise differentiation values were found between some southern phylogroups
4 and island groups, while the lowest values were between phylogroups and islands groups
5 that clustered within them (Supplementary information Table S2). There was higher
6 percentage of variation among (73.5 %) than within (26.5 %) groups, and there was a
7 significant population differentiation ($F_{ST} = 0.7349$, $P < 0.0001$). The Mantel test showed a
8 nonsignificant relationship between pairwise geographic and genetic distances based on
9 Slatkin's linearised F_{ST} ($R_2 = 0.0095$, $P = 0.2935$) (Supplementary information Fig. S1).

10 The barriers identified using the computational geometry approach reflected the
11 genetic differentiation between *S. minutus* and *S. volnuchini*, and among the phylogroups
12 within *S. minutus* (Fig. 1C). The first barrier separated *S. minutus* from *S. volunichini*. The
13 nine following barriers coincided with the location of mountain ranges, including a barrier
14 located in the north of the Balkan Peninsula, in the Alps and in the Pyrenees, which reflected
15 the genetic subdivisions and lineages in *S. minutus*.

16

17 *Historical demography*

18 Comparison of BFs for each model indicated the Bayesian skyline demographic model as
19 the best-fitting one (BF = 391), and the strict molecular clock was better than the
20 uncorrelated lognormal relaxed molecular clock (BF = 23). The MLEs for the constant
21 population size and Bayesian skyline demographic models using the strict molecular clock
22 were -10960 and -10569, while using the uncorrelated lognormal relaxed molecular clock
23 were -10907 and -10592, respectively. Therefore, the strict clock and Bayesian skyline
24 demographic model were selected as the best-fitting according to BFs. The effective sample
25 size (ESS) for all values was higher than 200.

26 All branches of the Bayesian genealogy (Fig. 3, Table 2) were well-supported
27 (posterior probabilities $PP \geq 0.97$), except for the clade containing all phylogroups excluding
28 Iberian ($PP = 0.05$). Molecular dating analysis revealed that the ingroup and outgroup

1 separated approximately 83.4 KYA, with lower and upper 95% HPD limits between 59.7 and
2 110.2 KYA (Fig. 3, Table 2). The diversification of *S. minutus* occurred approximately 31.8
3 KYA (95% HPD: 21.8 – 40.5 KYA) with the formation of the Iberian phylogroup, followed by
4 the formation of the Balkan group 29.6 KYA (95% HPD: 21.8 – 40.5 KYA), while in Western,
5 Central and Northern Europe, *S. minutus* continued its diversification with the Northern
6 phylogroup forming 24.1 KYA (95% HPD: 16.4 – 33.1 KYA), the Italian phylogroup forming
7 15.3 KYA (95% HPD: 10.7 – 21.5 KYA), the South Italian phylogroup forming 12.8 KYA (95%
8 HPD: 8.5 – 17.8 KYA), and the Western phylogroup forming 9.3 KYA (95% HPD: 6.7 – 12.6
9 KYA) (Fig. 3, Table 2). The TMRCA of the Balkan phylogroup was the earliest, dated back to
10 15.5 KYA (95% HPD: 9.7-22.7 KYA), followed by the Northern phylogroup, dated back to
11 11.8 KYA (95% HPD = 7.7-16.8 KYA), while the rest of the main phylogroups had TMRCA
12 dated approximately to about 6 and 9 KYA (Fig. 3, Table 2). Within the Western phylogroup,
13 the TMRCA for the Irish clade dated back to 5.9 KYA (95% HPD: 4.9 – 6.9 KYA).

14 The population expansion tests (R_2 and Fu's F_s) showed significant departures from
15 neutrality for the ingroup and several other phylogroups, except for the Balkan, Iberian and
16 South Italian (Table 2). The population expansions were not an effect of the island samples
17 belonging to these phylogroups, and continental samples analysed separately also
18 demonstrated a similar pattern (Table 2). For the island groups, only the Irish and British
19 groups showed signatures of recent population expansions (Table 2).

20 The mismatch distributions varied significantly among the phylogroups (Fig. 4A;
21 Supplementary information Fig. S2). The ingroup showed a bimodal mismatch distribution,
22 which reflected the pairwise comparisons within and among phylogroups in *S. minutus*. The
23 Northern (and Northern continental), Italian, Western (and Western continental) and Irish
24 phylogroups all had distinctly unimodal distributions with an almost perfect fit between
25 observed and expected pairwise differences of a sudden population expansion model. All
26 population expansions for the phylogroups were dated to the Holocene; the Italian and
27 Northern phylogroups had the oldest times of expansion (>8.0 KYA), while the Irish showed
28 a relatively recent population expansion dated to 1.6 KYA.

1 The BSP obtained for three phylogroups (Northern, Western and Irish) suggested
2 that demographic expansions of these populations started approximately 5.0 KYA (Fig. 4B).
3 BSP calculation for the Italian phylogroup indicated an even earlier demographic expansion
4 (approximately 5.5 KYA) (Fig. 4B).

5

6 DISCUSSION

7 Quaternary refugia represent the geographical regions that species inhabit during periods of
8 glacial or interglacial cycles when there is the maximum contraction in geographical range
9 (Stewart *et al.*, 2009). There is support for both southern (Taberlet *et al.*, 1998; Hewitt, 2000)
10 and northern glacial European refugia (Bilton *et al.*, 1998; Stewart & Lister, 2001; Kotlík *et*
11 *al.*, 2006; Provan & Bennett 2008; Fløjgaard *et al.*, 2009; Vega *et al.*, 2010a, b). Rather than
12 polarising the biogeographic patterns into southern and northern refugia (Tzedakis *et al.*,
13 2013), the paradigms of postglacial colonisation in Europe (Hewitt, 2000) can be improved
14 with the acceptance of southern hotspots of diversification without northward colonisation
15 (Bilton *et al.*, 1998) and the concept of refugia-within-refugia (Gómez & Lunt, 2007), as well
16 as with the findings of northern glacial refugia (Stewart & Lister 2001; Provan & Bennett,
17 2008; Stewart *et al.*, 2009), to reflect the evolutionary processes across varied topographical
18 areas that have shaped genetic diversity. The statistical phylogeographic results obtained
19 here contribute to the understanding of the phylogeographic patterns and processes during
20 the Quaternary glaciations that shaped the European biota, and provide further evidence to
21 the emerging pattern of complex biogeographical histories in Europe (Pedreschi *et al.*,
22 2019).

23

24 *Sorex minutus* phylogeography

25 The significant genetic structure among phylogroups defined in this study illustrate the
26 complex history of European colonisation, isolation and diversification of *S. minutus* during
27 the Pleistocene and Holocene, and is not a simple case of isolation by distance and
28 colonisation of Northern and Central Europe from expanding populations from the south.

1 While the southern phylogroups, including the Iberian, Balkan, Italian and South Italian, were
2 mostly restricted to the Southern European peninsulas (consistent with the traditional
3 southern glacial refugia), the Northern and Western phylogroups were widespread
4 geographically and were found north of the Mediterranean peninsulas, consistent with
5 previous studies with fewer samples (Bilton *et al.*, 1998; Mascheretti *et al.*, 2003; Vega *et al.*
6 2010a, b) and with different molecular markers (McDevitt *et al.*, 2010).

7 The hypothesis of northern refugia is further supported by palaeontological and
8 palynological evidence for other temperate and boreal species (Willis *et al.*, 2000; Willis &
9 van Andel, 2004; Magri *et al.*, 2006; Sommer & Nadachowski, 2006), as well as many
10 phylogeographic studies in small mammals, including the field vole *M. agrestis* (Jaarola &
11 Searle, 2002; Herman *et al.*, 2019), bank vole *M. glareolus* (Deffontaine *et al.*, 2005; Kotlík *et*
12 *al.*, 2006; Wójcik *et al.*, 2010), root vole *M. oeconomus* (Brunhoff *et al.*, 2003), common vole
13 *M. arvalis* (Heckel *et al.*, 2005; Stojak *et al.*, 2016), common shrew *S. araneus* (Bilton *et al.*,
14 1998; Yannic *et al.*, 2008) and weasels *Mustela nivalis* (McDevitt *et al.*, 2012). For several
15 small mammals, including *S. minutus*, suitable climatic conditions at the LGM could have
16 been widespread across Central and Eastern Europe (Fløjgaard *et al.*, 2009; Vega *et al.*,
17 2010b; McDevitt *et al.* 2012; Stojak *et al.*, 2019).

18 Until recently, it was unclear which species of *Sorex* inhabit Crimea. According to
19 Zagorodniuk (1996) it could be *S. (minutus) dahli* [mentioned in Hutterer (2005) as a
20 synonym of *Sorex volnuchini (dahli)*], and Zaitsev *et al.* (2014) and Hutterer *et al.* (2016)
21 indicated the presence of *S. minutus* in mainland Ukraine and in Crimea. Hutterer (2005)
22 mentioned that *S. volnuchini* might be present in Crimea, but in Hutterer *et al.* (2016) *S.*
23 *volnuchini* is only found in southern Russia and Caucasus States, Turkey and northern Iran.
24 Our research demonstrated that *S. volnuchini* may be present in the southern region of
25 Crimea (based on one *MT-CYB* sequence), while *S. minutus* is present in mainland Ukraine
26 and in the northern region of Crimea, but further sampling in this region is needed.

27 The finding of two phylogroups in the Iberian peninsula (i.e. Iberian and Western
28 phylogroups) and two in the Italian peninsula (i.e. Italian and South Italian phylogroups),

1 support the hypothesis of microevolutionary processes shaping the genetic diversity and
2 structure within the Mediterranean peninsulas. In the Iberian Peninsula, the topography of
3 the region with east-west mountain ranges and other high ground (over 1000 m a.s.l.), large
4 rivers (which could act as barriers to dispersal), and the distinct seasonal precipitation and
5 vegetation types (O'Regan, 2008), must have played an important role in the colonisation of
6 the region and the genetic differentiation of populations. McDevitt *et al.* (2010) proposed that
7 the Western phylogroup could have originated in the Dordogne region based on a limited
8 number of samples from France, but the presence of this phylogroup in northern Iberia and
9 the central position of Hap_61 (Pyrenees) could mean that an Iberian origin is possible
10 instead. A similar process could explain the presence of the two phylogroups in the Italian
11 peninsula (i.e. Italian and South Italian). The genetic differentiation of the South Italian
12 phylogroup, further supported by morphological data (Vega *et al.*, 2010a, 2016), could be
13 due to the unique geography of Southern Italy consisting of mountain massifs of Pollino, La
14 Sila and Aspromonte separated by lowland areas, which from the Pliocene to the end of the
15 Middle Pleistocene, at times of high sea level, were islands in a chain (Malatesta, 1985;
16 Caloi *et al.*, 1989; Bonardi *et al.*, 2001; Bonfiglio *et al.*, 2002). The patterns of differentiation
17 within refugial areas were concordant with the 'refugia-within-refugia' concept widely
18 recognized for the Iberian Peninsula (Gómez & Lunt, 2007; Abellán & Svenning, 2019) and
19 similar to microrefugia in the Balkans (Kryštufek *et al.*, 2007). For the Italian peninsula, a
20 comparable 'refugia-within-refugia' pattern was found in several species (Amori *et al.*, 2008;
21 Canestrelli *et al.*, 2008; Castiglia *et al.*, 2008; Vega *et al.*, 2010a, 2016; Senczuk *et al.*, 2017;
22 Bisconti *et al.*, 2018).

23 The genetic similarity between the Western and South Italian phylogroups indicates a
24 common history and it can be hypothesised that their common ancestor was more
25 widespread throughout the Italian peninsula, probably displaced later by the Italian lineage in
26 the Apennines and Western Alps. A similar scenario has been proposed for the water shrew
27 *Neomys fodiens* (Castiglia *et al.*, 2007), Alpine salamander *Salamandra salamandra*
28 (Steinfartz *et al.*, 2000), black pine *Pinus nigra* (Afzal-Rafii & Dodd, 2007) and green lizard

1 *Lacerta bilineata bilineata* (Böhme *et al.*, 2007), which showed closely related South Italian
2 and Western phylogroups most closely related to each other than to a North-Central Italian
3 lineage.

4 The phylogeographic patterns found here were further supported by the
5 determination of barriers that coincided with mountain ranges located on the north of the
6 Iberian, Italian and Balkan peninsulas. Contact zones among phylogroups (i.e. localities
7 where at least two *MT-CYB* phylogroups were present) were detected at the northern
8 extremes of the southern peninsulas. During the LGM, glaciers covered most of the Alpine
9 (Buoncristiani & Campy, 2004) and Pyrenean mountain ranges (Calvet, 2004), while glaciers
10 in the Carpathians (Reuther *et al.*, 2007) and in the Balkan Peninsula (Hughes *et al.*, 2006)
11 were found > 1,000 m a.s.l. When climate ameliorated and suitable habitat became
12 available, pygmy shrew populations belonging to different phylogroups on different sides of
13 the mountain ranges could have expanded and colonised previously glaciated areas thus
14 forming the observed contact zones. Moreover, the widespread distribution and absence of
15 phylogeographic structure of the Northern phylogroup could be explained by the apparent
16 absence of major geographical barriers across Central and Northern Europe, and
17 recolonization from northern refugia. Similarly, pygmy shrews belonging to the Western and
18 Northern phylogroups could have quickly colonised mainland Britain across a land
19 connection to continental Europe (i.e. Doggerland; Gaffney *et al.*, 2007), resulting in the
20 genetic similarities observed between the British Isles and continental Europe.

21

22 *Sorex minutus* demography

23 The oldest fossils assigned to *S. minutus* were found in Podlesice and Mała Cave, Poland
24 dated to the Early Pliocene between 4 and 5.3 MYA (Rzebik-Kowalska, 1998). An early
25 widespread colonisation of Europe by *S. minutus* might have been possible because it was
26 probably one of the first species of the genus *Sorex* in the continent (Rzebik-Kowalska,
27 1998, 2008). The Bayesian analysis revealed, however, more recent diversification events,
28 with TMRCA for the ingroup and the phylogroups in continental Europe between the Upper

1 Pleistocene and Lower Holocene. This is consistent with recent studies on field vole
2 (*Microtus agrestis*) phylogeography in Europe (Herman & Searle, 2011; Herman *et al.*, 2014)
3 which demonstrated the importance of the Younger Dryas (11.7-12.9 KYA) glacial re-
4 advance for the diversification within this species. Similar colonisation scenarios and
5 divergence before the LGM from Eastern to Western Europe have been proposed for other
6 species, including the common vole *Microtus arvalis* (Heckel *et al.*, 2005; Stojak *et al.*,
7 2016), the bank vole *Clethrionomys glareolus* (Deffontaine *et al.*, 2005; Kotlík *et al.*, 2006;
8 Wójcik *et al.*, 2010), and the root vole *M. oeconomus* (Brunhoff *et al.*, 2003).

9 The population expansion signatures for the Northern and Western phylogroups,
10 star-like patterns in phylogenetic networks and population expansion times support recent
11 and quick colonisation events of central and northern Europe, and appear to reflect
12 responses to postglacial climate warming. The Western lineage was restricted to Central,
13 Western and South-Eastern France and North-Western Spain in continental Europe, but it
14 was the only lineage found in Ireland and several islands off the west and north coasts of
15 Britain. The region of the Dordogne in South-Western France was situated outside the LGM
16 permafrost area and has temperate mammal fossil records dated to the end of the LGM;
17 therefore, it has been suggested as another likely northern refugium (Sommer &
18 Nadachowski, 2006; McDevitt *et al.*, 2010) where the Western lineage could have persisted
19 and recolonised Western and Central France after the LGM. But as stated above, an Iberian
20 origin for this phylogroup is also possible. However, SDM studies showed that suitable
21 climatic conditions during the LGM for *S. minutus* and other temperate small mammal
22 species could have been more continuous and present further north (Fløjgaard *et al.*, 2009;
23 Vega *et al.*, 2010b), which could explain its widespread distribution in Western Europe and
24 its presence in Britain. According to BSP results, it is plausible that Northern and Western
25 phylogroups spread across Europe after the Younger Dryas. The British (island) group,
26 belonging to the Northern phylogroup, showed a significant signature of population
27 expansion. This expansion could have selectively displaced pygmy shrew populations of the

1 Western lineage, which still remain in uplands and islands in the periphery to the north, west
2 and south of Britain forming a 'Celtic fringe' (Searle *et al.*, 2009).

3 The widespread Italian lineage may be presumed to derive from a glacial refugium
4 located somewhere within the vicinity of the Apennine mountain chain. A significant
5 population expansion signature demonstrates that the Italian phylogroup went through a
6 recent expansion phase, calculated in BSP for about 5.5 KYA. Contrastingly, the lack of a
7 population expansion signature, the high nucleotide and haplotype diversities, and the highly
8 divergent sequences showing a weak north/south subdivision of the Balkan phylogroup
9 warrants further attention. The Balkans is a European hotspot for biodiversity given its
10 environmental stability, topographic and climatic diversity and occasional connectedness
11 with Asia Minor (Kryštufek & Reed, 2004; Kryštufek *et al.*, 2007, 2009; Bužan *et al.*, 2010),
12 and it could be expected that some of these factors shaped the genetic diversity of the
13 Balkan lineage there. Similarly, the lack of significant population expansion values for the
14 Iberian lineage may relate to historical stable population sizes; however, the sample size
15 was low and this result should be taken with caution.

16

17 *Further considerations and implications*

18 The comparison of the results obtained here with those elsewhere shows an emerging
19 pattern of glacial refugia in Mediterranean peninsulas and further north in Central Europe for
20 several species.

21 Although *S. minutus* is considered as a least concern species by the IUCN (Hutterer
22 *et al.*, 2016), the distinct phylogroups deserve more attention than this implies. Genetic
23 diversity is considered an important aspect of global biodiversity (McNeely *et al.*, 1990), and
24 local and/or country-based conservation efforts are highly valued (for example, in Britain and
25 Ireland the pygmy shrew is protected by law). The refugial areas in Southern Europe are
26 often found in mountain ranges at the low-latitude margins of the present-day distribution
27 ranges of species and are most likely to contain rear-edge populations where selection for
28 local adaptations could have resulted in the evolution of distinct ecotypes (Cook, 1961;

1 Hampe & Petit, 2005). Rear-edge populations, including the southern lineages of *S. minutus*,
2 deserve further investigation and should be regarded for conservation because they are
3 important to determine the responses of species to modern climate change (Petit *et al.*,
4 2003; Hampe & Petit, 2005).

5 In conclusion, the Eurasian pygmy shrew *Sorex minutus* is a good model for
6 understanding biological diversity, colonisation patterns and the effects of past climate
7 change on biological diversity. There is a mosaic of genetic lineages across continental
8 Europe, characterised by different demographic histories and natural colonisation patterns,
9 while island populations are characterised by recent natural and human-mediated
10 colonisations. This study has notably expanded previous findings on *S. minutus*, with a more
11 precise statistical phylogeographic analysis of the genetic variability and structure,
12 colonisation routes, geographical barriers and historical demography across Europe.
13 Specifically, we provided new data from the Iberian and Balkan peninsulas, and from Central
14 and Eastern Europe (Poland, Ukraine and Russia), important for understanding postglacial
15 events. *Sorex minutus* is not an easy species to obtain in large numbers, and the sampling
16 described here represents a very substantial effort. However, it is a species that is unusually
17 widespread and genetically subdivided and therefore can inform better than almost any other
18 about the relative importance of southern and northern glacial refugia.

19

20 ACKNOWLEDGEMENTS

21 Specimens and species records of *Sorex minutus* were made available by several museums
22 and we acknowledge the help of the curators from the following institutions: Département
23 d'écologie et évolution (Université de Lausanne, Switzerland), Univerza na Primorskem and
24 Research Centre of Koper (Slovenia), Natuurhistorisch Museum (Rotterdam, Netherlands),
25 Dipartimento di Ecologia (Università della Calabria, Italy), Museo di Anatomia Comparata
26 and Museo di Zoologia "La Sapienza" (Università di Roma, Italy), Natuurmuseum Brabant
27 (Tilburg, Netherlands) and Mammal Research Institute PAS (Białowieża, Poland). We are
28 very grateful for the tissue samples provided by Glenn Yannic, Jacques Hausser, Jan Zima,

1 Fríða Jóhannesdóttir, Holger Bruns, Peter Borkenhagen, Natália Martínková, Elena
2 Gladilina, Pavel Goldin and Petr Kotlík. We thank Ricardo Castiglia and two anonymous
3 reviewers for their helpful comments.

4

5 REFERENCES

6 **Abellán P, Svenning J-C. 2014.** Refugia within refugia – patterns in endemism and genetic
7 divergence are linked to Late Quaternary climate stability in the Iberian Peninsula.
8 *Biological Journal of the Linnean Society* **113**: 13-28.

9 **Afzal-Rafii Z, Dodd RS. 2007.** Chloroplast DNA supports a hypothesis of glacial refugia
10 over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in
11 Western Europe. *Molecular Ecology* **16**: 723-736.

12 **Baele G, Lemey P, Vansteelandt S. 2013.** Make the most of your samples: Bayes factor
13 estimators for high-dimensional models of sequence evolution. *BMC Bioinformatics*
14 **14**: 85.

15 **Banaszek A, Ziomek J, Jadwiszczak KA, Kaczyńska E, Mirski P. 2012.** Identification of
16 the barrier to gene flow between phylogeographic lineages of the common hamster
17 *Cricetus cricetus*. *Acta Theriologica* **57**: 195-204.

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

22 **Bisconti R, Aloise G, Siclari A, Fava V, Provenzano M, Arduino P, Chiocchio A,**
23 **Nascetti G, Canestrelli D. 2018.** Forest dormouse (*Dryomys nitedula*) populations in
24 southern Italy belong to a deeply divergent evolutionary lineage: implications for
25 taxonomy and conservation. *Hystrix* **29**: 75-79.

26 **Böhme MU, Fritz U, Kotenko T, Džukić G, Ljubisavljević K, Tzankov N, Berendonk TU.**
27 **2007.** Phylogeography and cryptic variation within the *Lacerta viridis* complex
28 (Lacertidae, Reptilia). *Zoologica Scripta* **36**: 119-131.

- 1 **Bonardi G, Cavazza W, Perrone V, Rossi S. 2001.** Calabria-Peloritani terrane and
2 Northern Ionian Sea. In: Vai GB, Martini IP, eds. *Anatomy of an orogen: The*
3 *Apennines and adjacent Mediterranean basins*. Dordrecht: Kluwer, 287-306.
- 4 **Bonfiglio L, Mangano G, Marra AC, Masini F, Pavia M, Petruso D. 2002.** Pleistocene
5 Calabrian and Sicilian bioprovinces. *Geobios* **35**: 29-39.
- 6 **Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A,**
7 **Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary
8 analysis. *PLOS Computational Biology* **10**: e1003537.
- 9 **Brunhoff C, Galbreath KE, Fedorov VB, Cook JA, Jaarola M. 2003.** Holarctic
10 phylogeography of the root vole (*Microtus oeconomus*): implications for late
11 Quaternary biogeography of high latitudes. *Molecular Ecology* **12**: 957-968.
- 12 **Buoncristiani J-F, Campy M. 2004.** Palaeogeography of the last two glacial episodes in
13 France: the Alps and Jura. In: Ehlers J, Gibbard PL, eds. *Quaternary glaciations –*
14 *Extent and chronology Part I: Europe*. London: Developments in Quaternary Science
15 **2**: 101-110.
- 16 **Bužan EV, Förster DW, Searle JB, Kryštufek B. 2010.** A new cytochrome b phylogroup of
17 the common vole (*Microtus arvalis*) endemic to the Balkans and its implications for
18 the evolutionary history of the species. *Biological Journal of the Linnean Society* **100**:
19 788-796.
- 20 **Caloi L, Malatesta A, Palombo RM. 1989.** Biogeografia della Calabria meridionale durante
21 il Quaternario. *Atti Accademia Peloritana dei Pericolanti Classe I, Scienze*
22 *Matematiche Fisiche e Naturali* **67**: 261-278.
- 23 **Calvet M. 2004.** The Quaternary glaciation of the Pyrenees. In: Ehlers J, Gibbard PL, eds.
24 *Quaternary glaciations – Extent and chronology Part I: Europe*. London:
25 Developments in Quaternary Science **2**: 119-128.
- 26 **Canestrelli D, Cimmaruta R, Nascetti G. 2008.** Population genetic structure and diversity
27 of the Apennine endemic stream frog, *Rana italica*; insights on the Pleistocene
28 evolutionary history of the Italian peninsular biota. *Molecular Ecology* **17**: 3856-3872.

- 1 **Castiglia R, Annesi F, Aloise G, Amori G. 2007.** Mitochondrial DNA reveals different
2 phylogeographic structures in the water shrews *Neomys anomalus* and *N. fodiens*
3 (Insectivora: Soricidae) in Europe. *Journal of Zoological Systematics and*
4 *Evolutionary Research* **45**: 255-262.
- 5 **Churchfield S. 1990.** *The natural history of shrews*. London: Christopher Helm and A&C
6 Black.
- 7 **Churchfield S, Searle JB. 2008.** Pygmy shrew *Sorex minutus*. In: Harris S, Yalden DW,
8 eds. *Mammals of the British Isles: handbook*. London: The Mammal Society, 267-
9 271.
- 10 **Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX,**
11 **Hostetler SW, McCabe AM. 2009.** The Last Glacial Maximum. *Science* **325**: 710-
12 714.
- 13 **Cook LM. 1961.** The edge effect in population genetics. *American Naturalist* **95**: 295-307.
- 14 **Darriba D, Taboada GL, Doallo R, Posada D. 2012.** jModelTest 2: more models, new
15 heuristics and parallel computing. *Nature Methods* **9**: 772.
- 16 **Deffontaine V, Libois R, Kotlík P, Sommer R, Nieberding C, Paradis E, Searle JB,**
17 **Michaux JR. 2005.** Beyond the Mediterranean peninsulas: evidence of central
18 European glacial refugia for a temperate forest mammal species, the bank vole
19 (*Clethrionomys glareolus*). *Molecular Ecology* **14**: 1727-1739.
- 20 **Diniz-Filho JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, de Campos Telles**
21 **MP, Rangel TF, Bini LM. 2013.** Mantel test in population genetics. *Genetics and*
22 *Molecular Biology* **36**: 475-485.
- 23 **Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005.** Bayesian coalescent inference
24 of past population dynamics from molecular sequences. *Molecular Biology and*
25 *Evolution* **22**: 1185-1192.
- 26 **Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006.** Relaxed phylogenetics and
27 dating with confidence. *PLOS Biology* **4**: e88.

- 1 **Excoffier L, Laval G, Schneider S. 2005.** Arlequin ver. 3.0: An integrated software package
2 for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47-50.
- 3 **Fløjgaard C, Normand S, Skov F, Svenning J-C. 2009.** Ice age distributions of European
4 small mammals: insights from species distribution modelling. *Journal of*
5 *Biogeography* **36**: 1152-1163.
- 6 **Fu YX. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking
7 and background selection. *Genetics* **147**: 915-925.
- 8 **Fumagalli L, Taberlet P, Stewart DT, Gielly L, Hausser J, Vogel P. 1999.** Molecular
9 phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from
10 mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution* **11**: 222-
11 235.
- 12 **Gaffney V, Thomson K, Fitch S. 2007.** *Mapping Doggerland: the Mesolithic landscapes of*
13 *the southern North Sea*. Oxford: Archaeopress.
- 14 **Gómez A, Lunt DH. 2007.** Refugia within refugia: patterns of phylogeographic concordance
15 in the Iberian Peninsula. In: Weiss S, Ferrand N, eds. *Phylogeography of southern*
16 *European refugia*. Dordrecht: Springer, 155-188.
- 17 **Hall TA. 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis
18 program for Windows 95/98/NT/XP. *Nucleic Acids Symposium Series* **41**: 95-98.
- 19 **Hampe A, Petit RJ. 2005.** Conserving biodiversity under climate change: the rear edge
20 matters. *Ecology Letters* **8**: 461-467.
- 21 **Heckel G, Burri R, Fink S, Desmet J-F, Excoffier L. 2005.** Genetic structure and
22 colonization processes in European populations of the common vole, *Microtus*
23 *arvalis*. *Evolution* **59**: 2231-2242.
- 24 **Herman JS, Searle JB. 2011.** Post-glacial partitioning of mitochondrial genetic variation in
25 the field vole. *Proceedings of the Royal Society. Series B. Biological Sciences* **278**:
26 3601-3607.

- 1 **Herman JS, McDevitt AD, Kawałko A, Jaarola M, Wójcik JM, Searle JB. 2014.** Land-
2 bridge calibration of molecular clocks and the post-glacial colonization of Scandinavia
3 by the Eurasian field vole *Microtus agrestis*. *PLoS ONE* **9**: e103949.
- 4 **Herman JS, Stojak J, Paupério J, Jaarola M, Wójcik JM, Searle JB. 2019.** Genetic
5 variation in field voles *Microtus agrestis* from the British Isles: selective sweeps or
6 population bottlenecks. *Biological Journal of the Linnean Society* **126**: 852-865.
- 7 **Hewitt GM. 2000.** The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907-913.
- 8 **Hughes PD, Woodward JC, Gibbard PL. 2006.** Quaternary glacial history of the
9 Mediterranean mountains. *Progress in Physical Geography* **30**: 334-364.
- 10 **Hutterer R. 1990.** *Sorex minutus* Linnaeus, 1766. Zwergspitzmaus. In: Niethammer J, Krapp
11 F, eds. *Handbuch der Säugetiere Europas, band 3*. Wiesbaden: Aula-Verlag, 183-
12 206.
- 13 **Hutterer R. 2005.** Order Soricomorpha. In: Wilson DE, Reeder DM, eds. *Mammal species of*
14 *the world*, 3rd ed. Baltimore: Johns Hopkins University Press, 220-311.
- 15 **Hutterer R, Kryštufek B, Fernandes M, Meinig H. 2016.** *Sorex minutus* (errata version
16 published in 2017). The IUCN Red List of Threatened Species 2016:
17 e.T29667A115171222.[http://dx.doi.org/10.2305/IUCN.UK.2016-](http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T29667A22316362.en)
18 [3.RLTS.T29667A22316362.en](http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T29667A22316362.en). Downloaded on 04 February 2019.
- 19 **Jaarola M, Searle JB. 2002.** Phylogeography of field voles (*Microtus agrestis*) in Eurasia
20 inferred from mitochondrial DNA sequences. *Molecular Ecology* **11**: 2613-2621.
- 21 **Kass RE, Raftery AE. 1995.** Bayes factors. *Journal of the American Statistical Association*
22 **90**: 773-795.
- 23 **Knowles LL. 2009.** Statistical phylogeography. *Annual Review of Ecology, Evolution, and*
24 *Systematics* **40**: 593-612.
- 25 **Knowles LL, Maddison WP. 2002.** Statistical phylogeography. *Molecular Ecology* **11**: 2623-
26 2635.

- 1 **Kotlík P, Deffontaine V, Mascheretti S, Zima J, Michaux JR, Searle JB. 2006.** A northern
2 glacial refugium for bank voles (*Clethrionomys glareolus*). *Proceedings of the*
3 *National Academy of Sciences of the United States of America* **103**: 14860-14864.
- 4 **Kryštufek B, Bužan EV, Hutchinson WF, Hänfling B. 2007.** Phylogeography of the rare
5 Balkan endemic Martino's vole, *Dinaromys bogdanovi*, reveals strong differentiation
6 within the Western Balkan Peninsula. *Molecular Ecology* **16**: 1221-1232.
- 7 **Kryštufek B, Bryja J, Bužan EV. 2009.** Mitochondrial phylogeography of the European
8 ground squirrel, *Spermophilus citellus*, yields evidence on refugia for steppic taxa in
9 the southern Balkans. *Heredity* **103**: 129-135.
- 10 **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,**
11 **Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.**
12 **2007.** Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.
- 13 **Librado P, Rozas J. 2009.** DnaSP v5: A software for comprehensive analysis of DNA
14 polymorphism data. *Bioinformatics* **25**: 1451-1452.
- 15 **Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latałowa M,**
16 **Litt T, Paule L, Roure JM, Tantau I, Van Der Knaap WO, Petit RJ, De Beaulieu J.**
17 **2006.** A new scenario for the Quaternary history of European beech populations:
18 palaeobotanical evidence and genetic consequences. *New Phytologist* **171**: 199-221.
- 19 **Malatesta A. 1985.** *Geologia e paleobiologia dell'era glaciale*. Roma: La Nuova Italia
20 Scientifica. [In Italian]
- 21 **Manni F, Guerard E, Heyer E. 2004.** Geographic patterns of (genetic, morphologic,
22 linguistic) variation: how barriers can be detected by using Monmonier's algorithm.
23 *Human Biology* **76**: 173-190.
- 24 **Marková AK. 2011.** Small mammals from Palaeolithic sites of the Crimea. *Quaternary*
25 *International* **231**: 22-27.
- 26 **McDevitt AD, Rambau RV, O'Brien J, McDevitt CD, Hayden TJ, Searle JB. 2009.** Genetic
27 variation in Irish pygmy shrews *Sorex minutus* (Soricomorpha: Soricidae):

- 1 implications for colonization history. *Biological Journal of the Linnean Society* **97**:
2 918-927.
- 3 **McDevitt AD, Yannic G, Rambau RV, Hayden TJ, Searle JB. 2010.** Postglacial
4 recolonization of continental Europe by the pygmy shrew (*Sorex minutus*) inferred
5 from mitochondrial and Y chromosomal DNA sequences. In: Habel JC, Assman T,
6 eds. *Relict species: phylogeography and conservation biology*. Berlin: Springer
7 Verlag, 217-236.
- 8 **McDevitt AD, Vega R, Rambau RV, Yannic G, Herman JS, Hayden TJ, Searle JB. 2011.**
9 Colonization of Ireland: revisiting 'the pygmy shrew syndrome' using mitochondrial, Y
10 chromosomal and microsatellite markers. *Heredity* **107**: 548-557.
- 11 **McDevitt AD, Zub K, Kawałko A, Olivier MK, Herman JS, Wójcik JM. 2012.** Climate and
12 refugial origin influence mitochondrial lineage distribution of weasels (*Mustela nivalis*)
13 in a phylogeographic suture zone. *Biological Journal of the Linnean Society* **106**: 57-
14 69.
- 15 **McNeely JA, Miller KR, Reid WV, Mittermeier RA, Werner TB. 1990.** *Conserving the*
16 *world's biological diversity*. Washington, DC: World Conservation Union, World
17 Resources Institute, Conservation International, World Wildlife Fund–US, and the
18 World Bank.
- 19 **Ohdachi SD, Hasegawa M, Iwasa MA, Vogel P, Oshida T, Lin L-K, Abe H. 2006.**
20 Molecular phylogenetics of soricid shrews (Mammalia) based on mitochondrial
21 cytochrome b gene sequences: with special reference to the Soricinae. *Journal of*
22 *Zoology* **270**: 199-200.
- 23 **O'Regan HJ. 2008.** The Iberian Peninsula – corridor or cul-de-sac? Mammalian faunal
24 change and possible routes of dispersal in the last 2 million years. *Quaternary*
25 *Science Reviews* **27**: 2136-2144.
- 26 **Pazonyi P. 2004.** Mammalian ecosystem dynamics in the Carpathian Basin during the last
27 27,000 years. *Palaeogeography, Palaeoclimatology, Palaeoecology* **212**: 295-314.

- 1 **Pedreschi D, García-Rodríguez O, Yannic G, Cantarello E, Diaz A, Golicher D,**
2 **Korstjens AH, Heckel G, Searle JB, Gillingham P, Hardouin EA, Stewart JR.**
3 **2019.** Challenging the European southern refugium hypothesis: species-specific
4 structures versus general patterns of genetic diversity and differentiation among
5 small mammals. *Global Ecology and Biogeography* **28**: 262-274.
- 6 **Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R,**
7 **Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch**
8 **B, Palmé A, Martín JP, Rendell S, Vendramin GG. 2003.** Glacial refugia: hotspots
9 but not melting pots of genetic diversity. *Science* **300**: 1563-1565.
- 10 **Provan J, Bennett KD. 2008.** Phylogeographic insights into cryptic glacial refugia. *Trends in*
11 *Ecology and Evolution* **23**: 564-571.
- 12 **Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018.** Posterior summarisation
13 in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901-904.
- 14 **Ramos-Onsins SE, Rozas J. 2002.** Statistical properties of new neutrality tests against
15 population growth. *Molecular Biology and Evolution* **19**: 2092-2100.
- 16 **Reuther AU, Urdea P, Geiger C, Ivy-Ochs S, Niller H-P, Kubik PW, Heine K. 2007.** Late
17 Pleistocene glacial chronology of the Pietrele Valley, Retezat Mountains, Southern
18 Carpathians constrained by ¹⁰Be exposure ages and pedological investigations.
19 *Quaternary International* **164-165**: 151-169.
- 20 **Rogers AR, Harpending H. 1992.** Population growth makes waves in the distribution of
21 pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-569.
- 22 **Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, Larget B, Liu L,**
23 **Suchard MA, Huelsenbeck JP. 2012.** MrBayes 3.2: efficient Bayesian phylogenetic
24 inference and model choice across a large model space. *Systematic Biology* **61**: 539-
25 542.
- 26 **Rzebik-Kowalska B. 1998.** Fossil history of shrews in Europe. In: Wójcik J, Wolsan M, eds.
27 *Evolution of shrews*. Białowieża: Mammal Research Institute, Polish Academy of
28 Sciences, 23-71.

- 1 **Schenekar T, Weiss S. 2011.** High rate of calculation errors in mismatch distribution
2 analysis results in numerous false inferences of biological importance. *Heredity* **107:**
3 511-512.
- 4 **Searle JB, Kotlík P, Rambau RV, Marková S, Herman JS, McDevitt AD. 2009.** The Celtic
5 fringe of Britain: insights from small mammal phylogeography. *Proceedings of the*
6 *Royal Society Series B* **276:** 4287-4294.
- 7 **Senczuk G, Colangelo P, De Simone E, Aloise G, Castiglia R. 2017.** A combination of
8 long term fragmentation and glacial persistence drove the evolutionary history of the
9 Italian wall lizard *Podarcis siculus*. *BMC Evolutionary Biology* **17:** 6.
- 10 **Sikes RS, Gannon WL; the Animal Care and Use Committee of the American Society**
11 **of Mammalogists. 2011.** Guidelines of the American Society of Mammalogists for
12 the use of wild mammals in research. *Journal of Mammalogy* **92:** 235–253.
- 13 **Slatkin M. 1995.** A measure of population subdivision based on microsatellite allele
14 frequencies. *Genetics* **139:** 457-462.
- 15 **Sommer RS, Nadachowski A. 2006.** Glacial refugia of mammals in Europe: evidence from
16 fossil records. *Mammal Review* **36:** 251-265.
- 17 **Steinfartz S, Veith M, Tautz D. 2000.** Mitochondrial sequence analysis of *Salamandra* taxa
18 suggests old splits of major lineages and postglacial recolonizations of Central
19 Europe from distinct source populations of *Salamandra salamandra*. *Molecular*
20 *Ecology* **9:** 397-410.
- 21 **Stewart JR, Lister AM. 2001.** Cryptic northern refugia and the origins of the modern biota.
22 *Trends in Ecology and Evolution* **16:** 608-613.
- 23 **Stewart JR, Lister AM, Barnes I, Dalén L. 2009.** Refugia revisited: individualistic
24 responses of species in space and time. *Proceedings of the Royal Society Series B*
25 **277:** 661-671.
- 26 **Stojak J, McDevitt AD, Herman JS, Searle JB, Wójcik JM. 2015.** Post-glacial colonization
27 of eastern Europe from the Carpathian refugium: evidence from mitochondrial DNA of

1 the common vole *Microtus arvalis*. *Biological Journal of the Linnean Society* **115**:
2 927-939.

3 **Stojak J, McDevitt AD, Herman JS, Kryštufek B, Uhlíková J, Purger JJ, Lavrenchenko**
4 **LA, Searle JB, Wójcik JM. 2016.** Between the Balkans and the Baltic:
5 phylogeography of a common vole mitochondrial DNA lineage limited to Central
6 Europe. *PLoS ONE* **11**: e0168621.

7 **Stojak J, Borowik T, Górny M, McDevitt AD, Wójcik JM. 2019.** Climatic influences on the
8 genetic structure and distribution of the common vole and field vole in Europe.
9 *Mammal Research* **64**: 19-29.

10 **Stojak J, Tarnowska E. 2019.** Polish suture zone as the goblet of truth in post-glacial
11 history of mammals in Europe. *Mammal Research* **64**: 463-475.

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

14 **Tzedakis PC, Emerson BC, Hewitt GM. 2013.** Cryptic or mystic? Glacial tree refugia in
15 northern Europe. *Trends in Ecology and Evolution* **28**: 696-704.

16 **Vega R, Amori G, Aloise G, Cellini S, Loy A, Searle JB. 2010.** Genetic and morphological
17 variation in a Mediterranean glacial refugium: evidence from Italian pygmy shrews,
18 *Sorex minutus* (Mammalia: Soricomorpha). *Biological Journal of the Linnean Society*
19 **100**: 774-787.

20 **Vega R, Fløjgaard C, Lira-Noriega C, Nakazawa Y, Svenning J-C, Searle JB. 2010.**
21 Northern glacial refugia for the pygmy shrew *Sorex minutus* in Europe revealed by
22 phylogeographic analyses and species distribution modelling. *Ecography* **33**: 260-
23 271.

24 **Vega R, McDevitt A, Kryštufek B, Searle JB. 2016.** Ecogeographical patterns of
25 morphological variation in pygmy shrews *Sorex minutus* (Soricomorpha: Soricinae)
26 within a phylogeographical and continental-and-island framework. *Biological Journal*
27 *of the Linnean Society* **119**: 799-815.

- 1 **Willis KJ, van Andel TH. 2004.** Trees or no trees? The environments of central and eastern
2 Europe during the Last Glaciation. *Quaternary Science Reviews* **23**: 2369-2387.
- 3 **Willis KJ, Rudner E, Sümegi P. 2000.** The full-glacial forests of central and southeastern
4 Europe. *Quaternary Research* **53**: 203-213.
- 5 **Wójcik JM, Kawałko A, Marková S, Searle JB, Kotlík P. 2010.** Phylogeographic signatures
6 of northward post-glacial colonization from high-latitude refugia: a case study of bank
7 voles using museum specimens. *Journal of Zoology* **281**: 249-262.
- 8 **Yannic G, Basset P, Hausser J. 2008.** A new perspective on the evolutionary history of
9 western European *Sorex araneus* group revealed by paternal and maternal
10 molecular markers. *Molecular Phylogenetics and Evolution* **47**: 237-250.
- 11 **Zagorodniuk I. 1996.** Rare shrew species in the territory of Ukraine: legends, facts and
12 diagnostics. *Zoology Bulletin* **30**: 53-69. [In Russian]
- 13 **Zaitsev MV, Voyta LL, Sheftel BI. 2014.** *The mammals of Russia and adjacent territories.*
14 *Lipotylans.* Saint Petersburg: Nauka. [In Russian]
- 15

1 SUPPORTING INFORMATION

2

3 **Table S1.** *Sorex minutus* dataset and sample information

4 **Table S2.** Pairwise geographic distances (in Km, below diagonal) and genetic differentiation

5 (Slatkin's F_{ST} , above diagonal) among *Sorex minutus* phylogroups and other geographic

6 groups

7 **Figure S1.** Correlogram of pairwise geographic and genetic distances among *Sorex minutus*

8 phylogroups and other geographic groups.

9 **Figure S2.** Mismatch distributions of *Sorex minutus* phylogroups and other geographic

10 groups.

11

12

1 TABLES

2

3

Table 1. DNA sequence polymorphism in *Sorex minutus* phylogroups and other geographic groups

Phylogroups	<i>N</i>	<i>S</i>	<i>P</i>	<i>H</i>	<i>Hd</i>	<i>Hd</i> (SD)	π	π (SD)	<i>k</i>
Ingroup	671	390	277	424	0.9899	0.0015	0.0143	0.0000	15.8670
Italian	26	51	19	18	0.9600	0.0230	0.0061	0.0004	6.7720
South Italian	4	16	0	4	1.0000	0.1770	0.0072	0.0020	8.0000
Balkan	22	55	28	17	0.9610	0.0290	0.0097	0.0009	10.7970
Iberian	6	15	6	5	0.9330	0.1220	0.0058	0.0013	6.4000
Western	283	147	83	102	0.9458	0.0067	0.0049	0.0002	5.4400
Irish	94	53	21	42	0.8920	0.0270	0.0020	0.0002	2.2180
Northern	330	311	197	278	0.9984	0.0005	0.0065	0.0002	7.1840
Continental groups									
Western (Continental)	15	28	11	13	0.9810	0.0310	0.0050	0.0006	5.5430
Northern (Continental)	226	241	142	188	0.9978	0.0007	0.0062	0.0002	6.9300
Other island groups									
Orkney Islands (All)	119	17	13	11	0.7720	0.0210	0.0027	0.0001	3.0140
Orkney Mainland	44	9	7	8	0.7550	0.0550	0.0013	0.0002	1.4790
Orkney South Ronaldsay	40	1	1	2	0.1420	0.0710	0.0001	0.0001	0.1420
Orkney Westray	33	0	0	1	0.0000	0.0000	0.0000	0.0000	0.0000
Orkney Hoy	2	2	0	2	1.0000	0.5000	0.0018	0.0009	2.0000
Belle Île	5	9	3	5	1.0000	0.1260	0.0038	0.0010	4.2000
British	91	146	61	80	0.9960	0.0030	0.0055	0.0003	6.1210

N = Sample size; *S* = Number of polymorphic (segregating) sites; *P* = Parsimony informative sites; *H* = Number of haplotypes; *Hd* = Haplotype diversity; SD = Standard Deviation; π = Nucleotide diversity; *k* = Average number of nucleotide differences.

4

1

Table 2. Population expansion tests for *Sorex minutus* phylogroups and other geographic groups

Phylogroups	R_2	P -value	F_s	P -value	τ	t (in years)	TMRCA (in KYA)	95% HPD (in KYA)
Ingroup	0.0198	0.0004	-741.2620	***	7.8590	6425	31.8	22.0-43.1
Italian	0.0521	0.0000	-5.8766	0.0152	6.7720	5536	7.2	4.8-10.2
South Italian	0.1822	0.1658	0.0687	0.2975	5.6340	-	7.7	4.2-12.1
Balkan	0.0830	0.0542	-3.6701	0.0768	7.1500	-	15.5	9.7-22.7
Iberian	0.1462	0.0888	0.0731	0.4290	4.0100	-	6.2	3.9-10.0
Western	0.0175	0.0004	-114.6990	***	3.6660	2997	9.3	6.7-12.6
Irish	0.0187	0.0000	-52.5664	***	1.3040	1066	5.9	4.9-6.9
Northern	0.0105	0.0000	-663.4730	***	6.5390	5346	11.8	7.7-16.8
Continental groups								
Western (Continental)	0.0793	0.0045	-6.0342	0.0035	5.5430	4532	-	-
Northern (Continental)	0.0128	0.0000	-386.4520	***	5.8010	4742	-	-
Other island groups								
Orkney Islands (All)	0.0880	0.5209	0.6044	0.6437	1.1740	-	-	-
Orkney Mainland	0.0839	0.2301	-1.6879	0.1892	1.4790	-	-	-
Orkney Hoy	0.5000	1.0000	NC	NC	2.0000	-	-	-
Orkney South Ronaldsay	0.0712	0.1770	-0.2182	0.4420	0.1420	-	-	-
Orkney Westray	NC	NC	NC	NC	NC	-	-	-
Belle Île	0.1915	0.2467	-1.6330	0.0732	3.5500	-	-	-
British	0.0161	0.0000	-122.8550	***	6.1210	5004	-	-

R_2 = Ramos-Onsins and Rozas (2002) test of neutrality; P -value = P -values of expansion tests expected under neutrality (*** = $P < 0.001$); F_s = Demographic population expansion test (Fu 1997); $\tau = (2ut)$ The mode of a mismatch distribution; t = Time of population expansion (for phylogroups with bi- or unimodal mismatch distributions); TMRCA = Time of divergence from the Most Recent Common Ancestor; 95% HPD = 95% Highest Posterior Density; KYA = Thousand Years Ago; NC = Not computable (not enough variation or samples)

2

3

1 FIGURE LEGENDS

2

3 **Figure 1.** A) Map of Eurasia showing the geographical distribution of the Eurasian pygmy
4 shrew *Sorex minutus* (Hutterer *et al.*, 2016). B) Sample localities of *S. minutus* used for this
5 study and divided into mitochondrially encoded cytochrome b (*MT-CYB*) phylogroups
6 (symbols with a dot represent samples used for inferring geographic barriers). C)
7 Geographic barriers (red lines) for *S. minutus*; the barriers (up to a maximum of 10) were
8 inferred using Monmonier's maximum difference algorithm which finds edges (boundaries)
9 on the Voronoi tessellation (blue polygons) associated with the highest rate of change in
10 genetic distances among a subset of continental samples (dots) interconnected with a
11 Delaunay triangulation (green lines).

12

13 **Figure 2.** Phylogenetic reconstructions of the Eurasian pygmy shrew *Sorex minutus* using
14 *MT-CYB* sequences. A) Bayesian phylogenetic tree (with posterior probabilities on branches)
15 showing the phylogroups. B) Haplotype phylogenetic network with haplotypes represented
16 as nodes and their evolutionary relationships represented by edges; relevant haplotypes
17 named at the centre of star-like patterns.

18

19 **Figure 3.** Maximum Clade Credibility tree for 671 sequences of *Sorex minutus* from Europe
20 and Siberia, annotated from 10,000 Bayesian genealogy sampling. Posterior probabilities of
21 basal nodes indicate support for each lineage. Horizontal bars represent 95% HPD intervals
22 for Time to Most Recent Common Ancestor (TMRCA) of each lineage (in thousand years
23 ago, KYA).

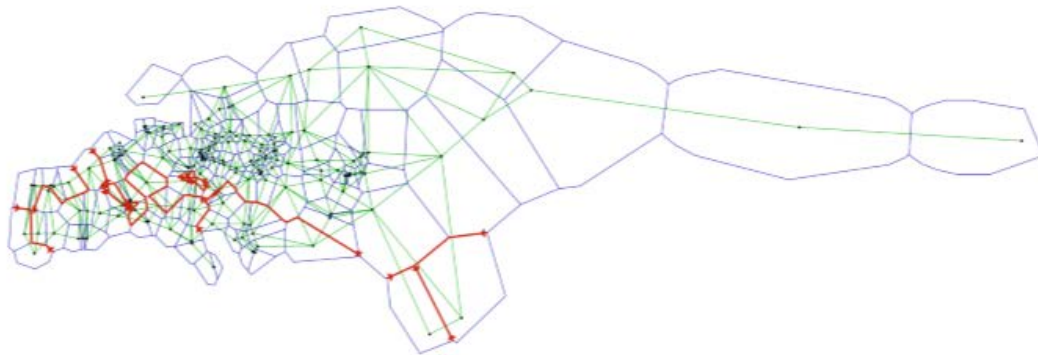
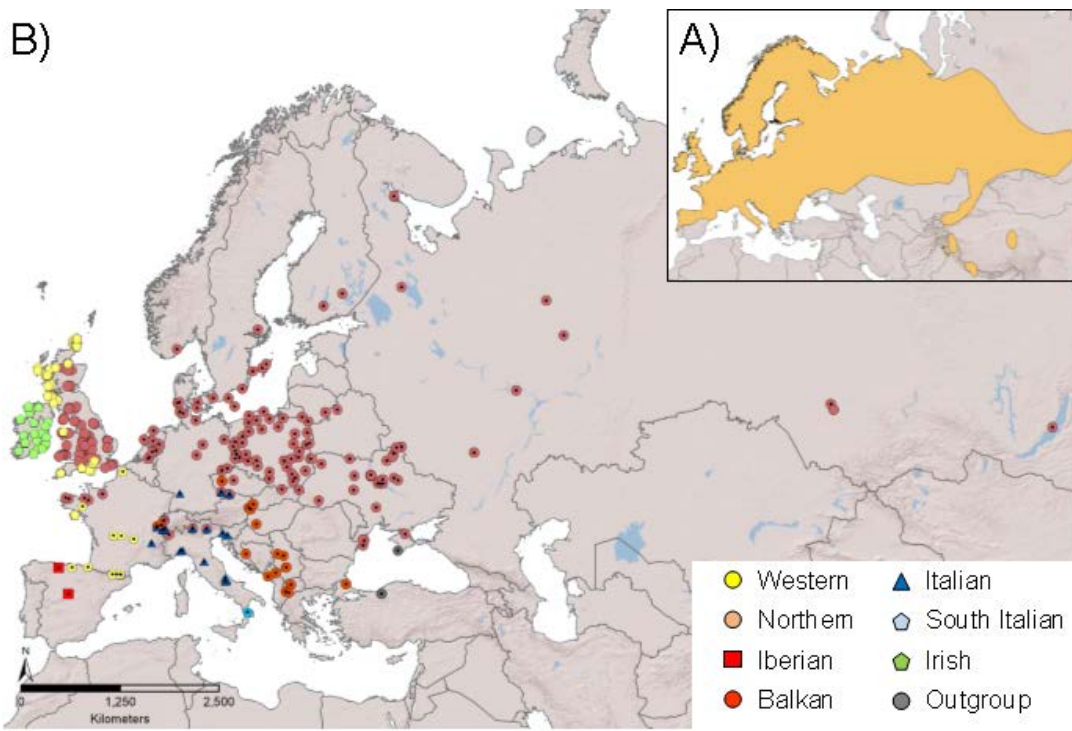
24

25 **Figure 4.** Historical demography of the Eurasian pygmy shrew *Sorex minutus*. A) Mismatch
26 distributions of groups with significant signatures of population expansion. B) Bayesian
27 Skyline Plots (BSP) of phylogroups with significant signatures of population expansion. The

- 1 solid lines in BSP are median estimates and the shaded areas represent 95% Highest
- 2 Probability Densities (confidence intervals).
- 3
- 4
- 5

1 FIGURES

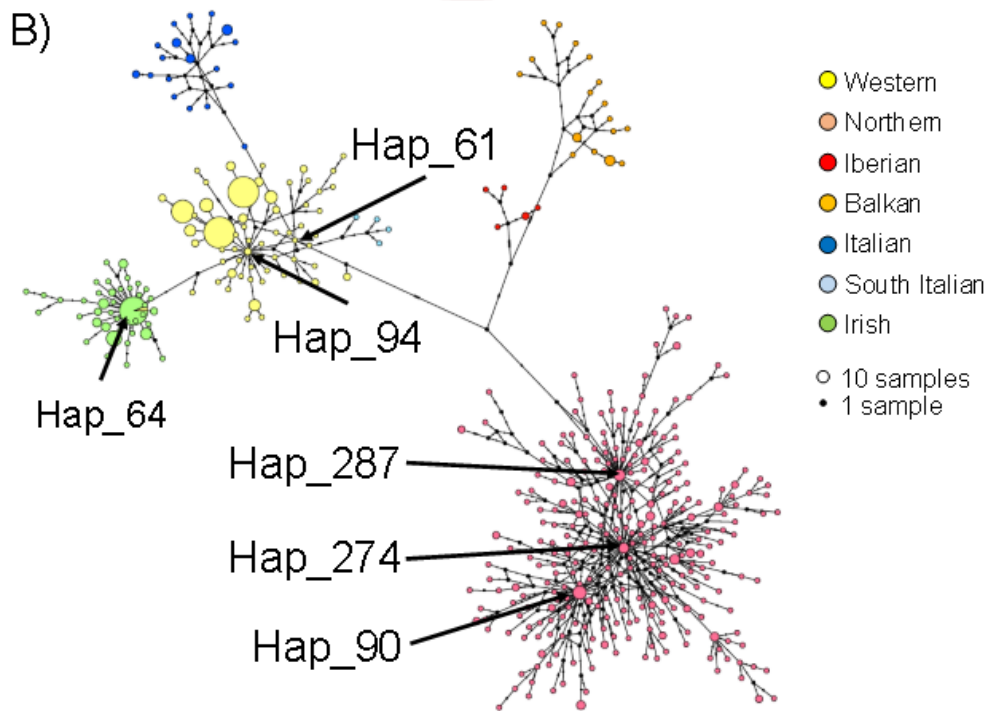
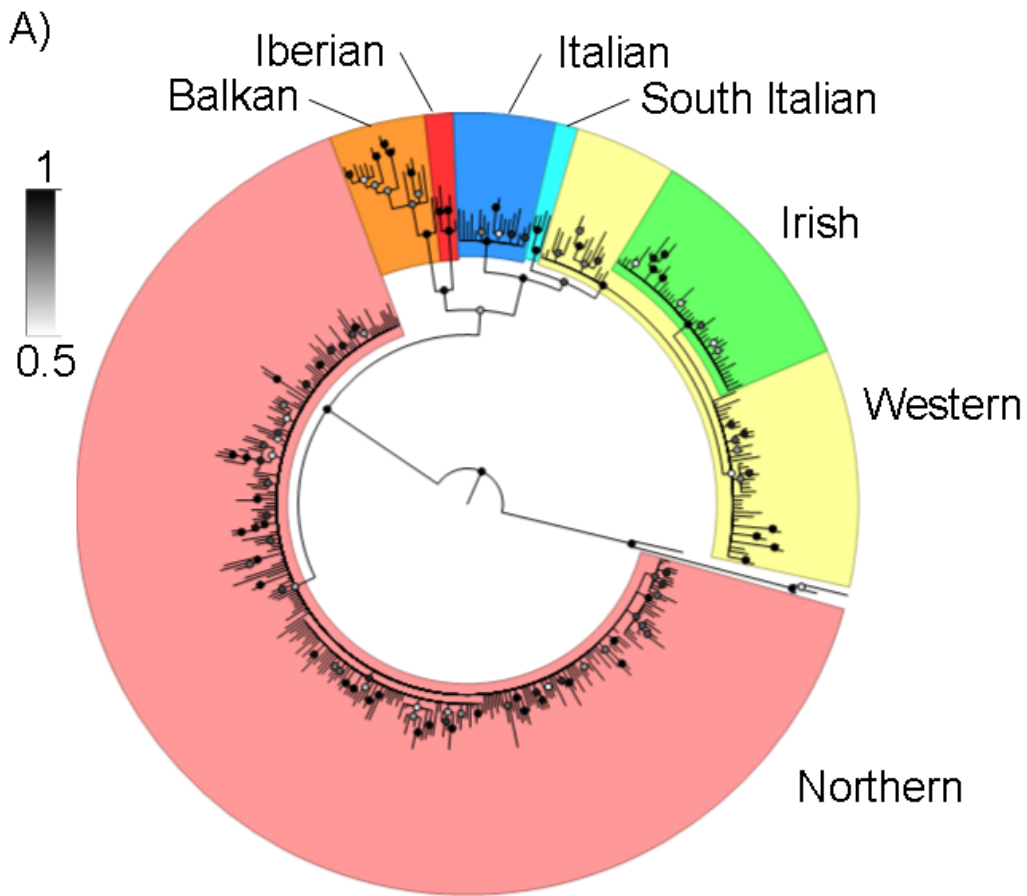
2 Figure 1.



3

4

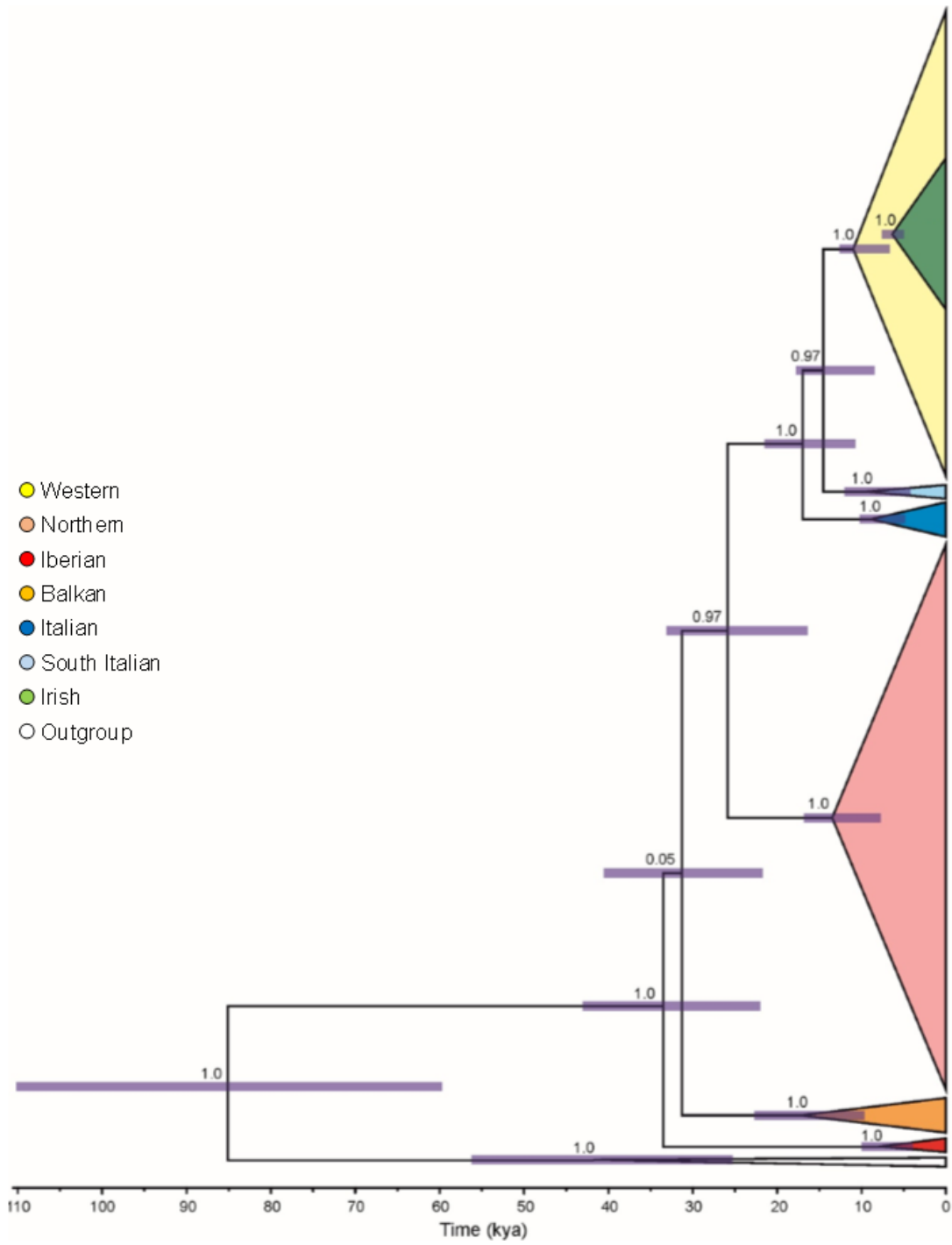
1 Figure 2.



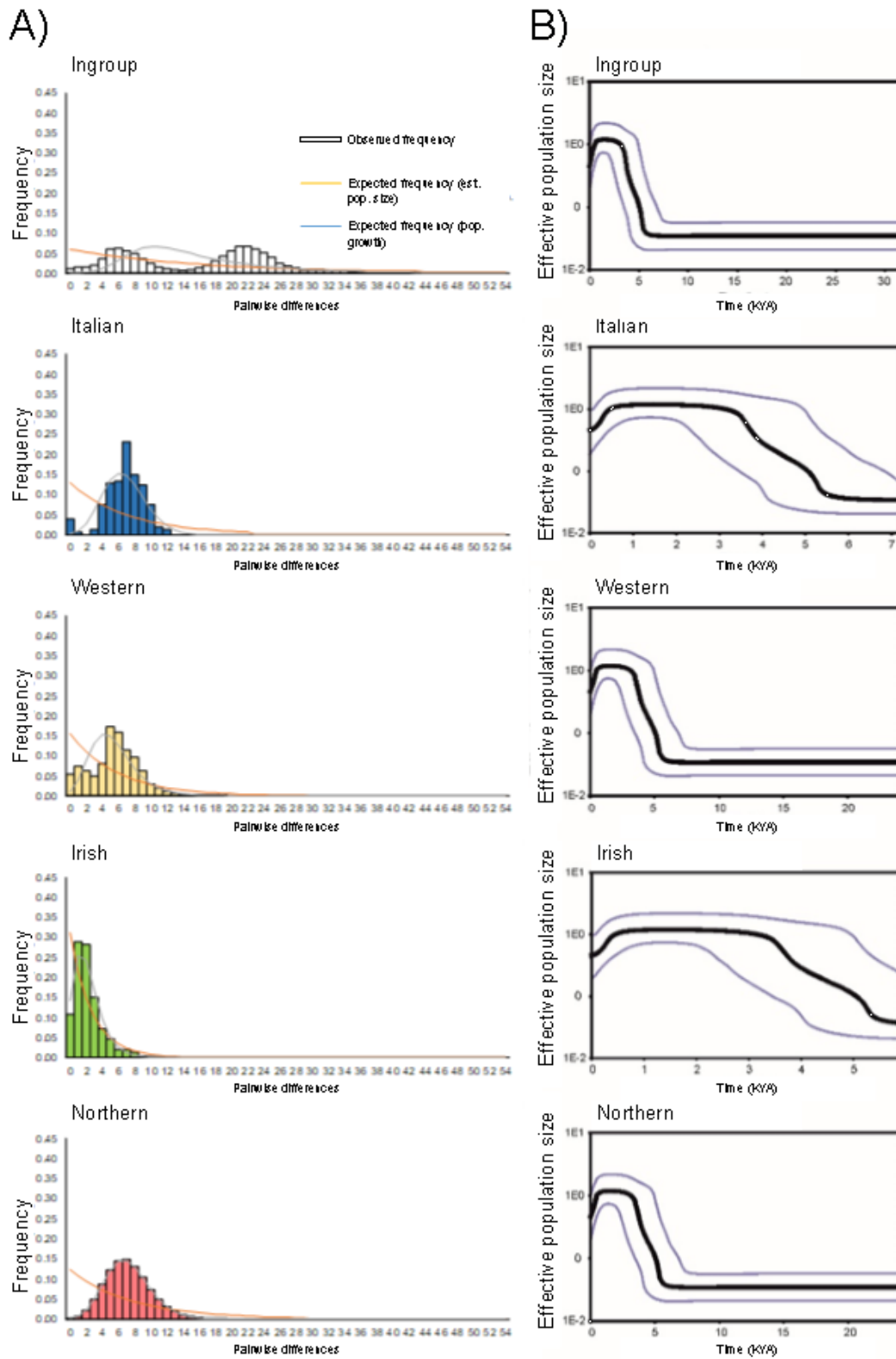
2

3

1 **Figure 3.**



1 **Figure 4.**

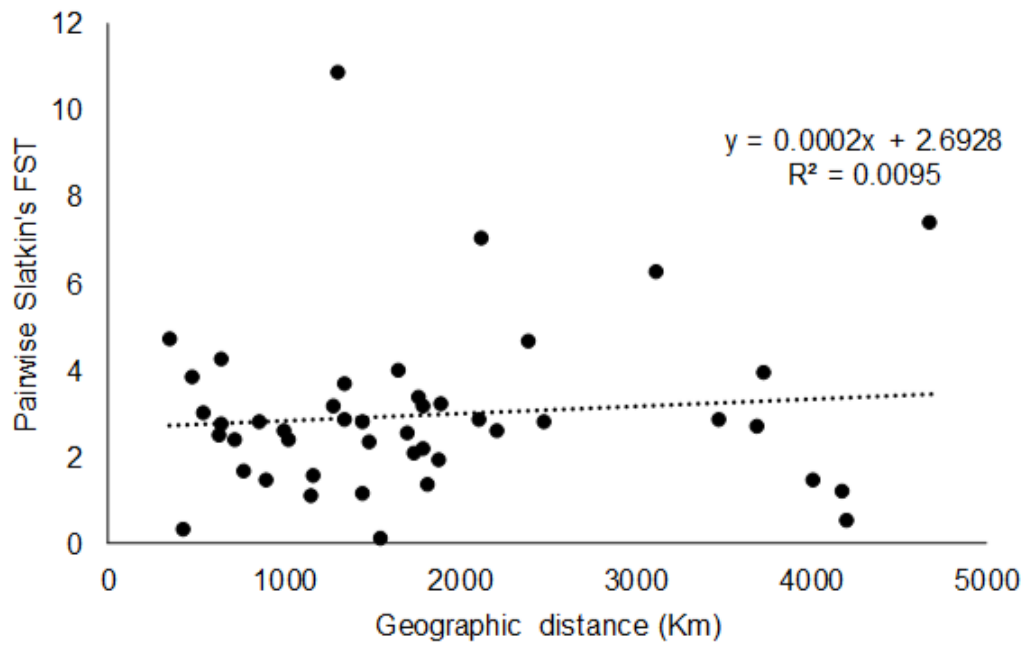


2

3

1 SUPPLEMENTARY FIGURES

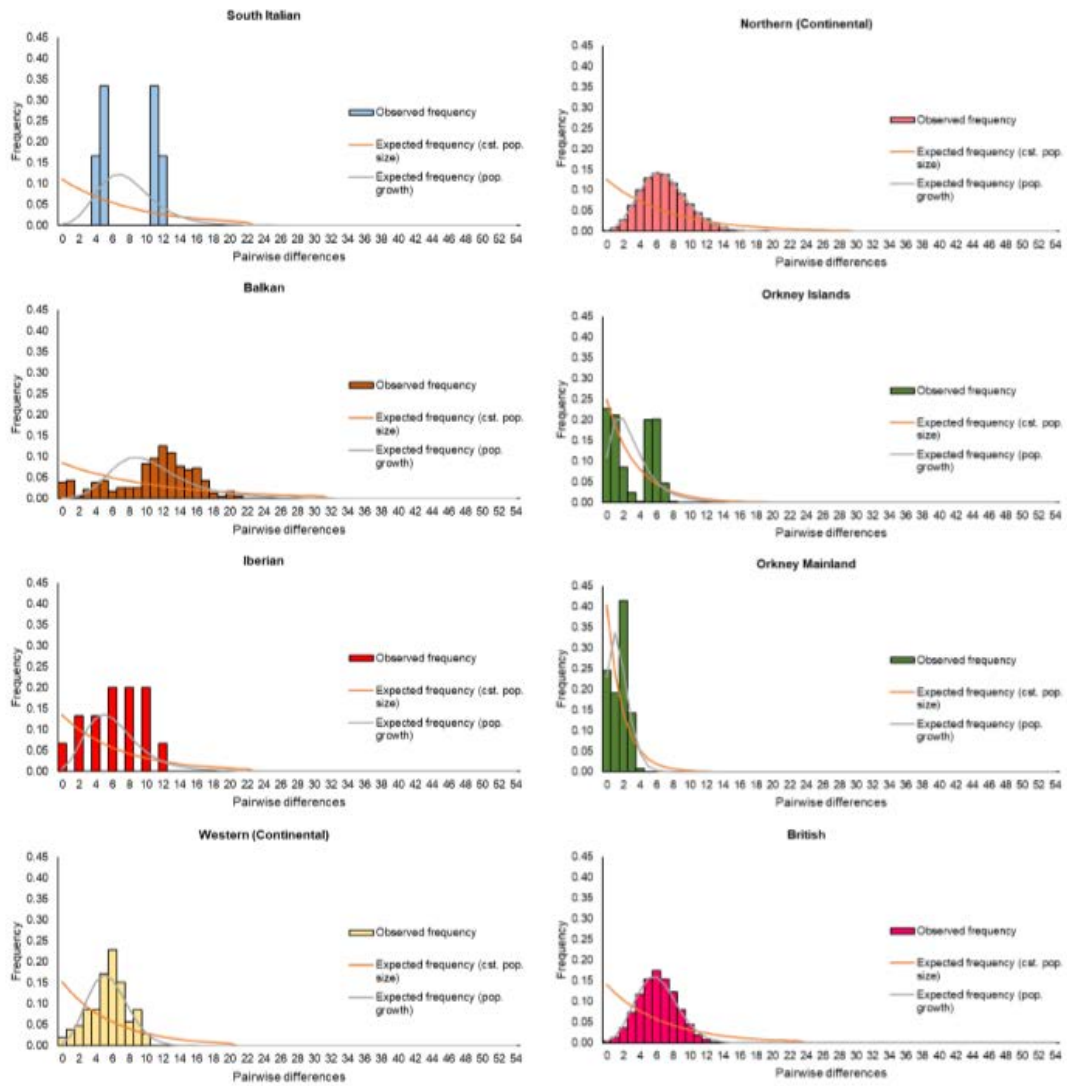
2 Figure S1.



3

4

1 Figure S2.



2

3

4

1 **SUPPLEMENTARY TABLES**

2

3 **Table S1.** *Sorex minutus* dataset and sample information. *Contact corresponding author.*

1

Table S2. Pairwise geographic distances (in Km, below diagonal) and genetic differentiation (Slatkin's F_{ST} , above diagonal) among *Sorex minutus* phylogroups and other geographic groups

	Italian	South Italian	Balkan	Iberian	Belle Île	Britain	Northern (Continental)	Western (Continental)	Orkney Islands	Irish
Italian	-	1.6558	3.0387	3.6919	1.5940	2.8852	2.3798	1.4534	2.8673	4.0090
South Italian	773.14	-	2.5113	3.3869	1.3569	2.8562	2.3204	1.1713	2.7079	4.6820
Balkan	547.27	628.96	-	1.9234	2.5617	3.1494	2.8093	2.7975	6.2456	7.0533
Iberian	1349.26	1768.56	1880.98	-	4.2650	3.1850	2.6191	3.8498	7.3804	10.8797
Belle Ile	1162.82	1815.47	1701.44	640.58	-	2.7595	2.1790	0.3345	1.2148	2.4003
Britain	1347.34	2108.66	1795.12	1286.79	647.11	-	0.1449	2.6083	3.9225	4.7265
Northern (Continental)	1022.36	1488.84	863.37	2218.78	1788.28	1554.91	-	2.0767	2.7910	3.2035
Western (Continental)	903.90	1444.78	1448.34	476.18	434.81	1006.85	1742.75	-	0.5436	1.1193
Orkney Islands	3476.34	3693.42	3127.67	4679.14	4175.54	3739.02	2477.61	4206.50	-	1.4635
Irish	1652.35	2396.73	2124.06	1311.98	726.86	346.14	1897.29	1151.96	4016.56	-

2

3