The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2015, **114**, 495–512. With 6 figures

Tracing the evolutionary history of the mole, *Talpa europaea*, through mitochondrial DNA phylogeography and species distribution modelling

ROBERTO FEUDA[1,2,*], ANNA A. BANNIKOVA[3], ELENA D. ZEMLEMEROVA[3], MIRKO DI FEBBRARO[4], ANNA LOY[4], RAINER HUTTERER[5], GAETANO ALOISE[6], ALEXANDER E. ZYKOV[7], FLAVIA ANNESI[1] and PAOLO COLANGELO[1,8,*]

[1]*Department of Biology and Biotechnology 'Charles Darwin', University 'La Sapienza', Via Borelli 50, 00161 Roma, Italy*
[2]*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena CA 91125*
[3]*Lomonosov Moscow State University, Department of Vertebrate Zoology, Leninskiye Gory 1/12, 119234 Moscow, Russia*
[4]*Environmetrics Lab, Department Bioscience and Territory, University of Molise, Contrada Fonte Lappone s.n.c, I-86090 Pesche, Italy*
[5]*Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany*
[6]*Museo di Storia Naturale della Calabria e Orto Botanico, University of Calabria, Via Savinio – Edificio Polifunzionale, I-87036 Rende, Italy*
[7]*Educational and Scientific Centre 'Institute of Biology' Taras Shevchenko National University of Kyiv, Build. 12, Academician Glushkov Ave., 03022 Kyiv, Ukraine*
[8]*National Research Council, Institute of Ecosystem Study, Largo Tonolli 50, 28922 Verbania Pallanza, Italy*

Received 12 August 2014; revised 3 November 2014; accepted for publication 3 November 2014

Our understanding of the effect of Pleistocene climatic changes on the biodiversity of European mammals mostly comes from phylogeographical studies of non-subterranean mammals, whereas the influence of glaciation cycles on subterranean mammals has received little attention. The lack of data raises the question of how and to what extent the current amount and distribution of genetic variation in subterranean mammals is the result of Pleistocene range contractions/expansions. The common mole (*Talpa europaea*) is a strictly subterranean mammal, widespread across Europe, and represents one of the best candidates for studying the influence of Quaternary climatic oscillation on subterranean mammals. Cytochrome *b* sequences, as obtained from a sampling covering the majority of the distribution area, were used to evaluate whether Pleistocene climate change influenced the evolution of *T. europaea* and left a trace in the genetic diversity comparable to that observed in non-subterranean small mammals. Subsequently, we investigated the occurrence of glacial refugia by comparing the results of phylogeographical analysis with species distribution modelling. We found three differentiated mitochondrial DNA lineages: two restricted to Spain and Italy and a third that was widespread across Europe. Phylogenetic inferences and the molecular clock suggest that the Spanish moles represent a highly divergent and ancient lineage, highlighting for the first time the paraphyly of *T. europaea*. Furthermore, our analyses suggest that the genetic break between the Italian and the European lineages predates the last glacial phase. Historical demography and spatial principal component analysis further suggest that the Last Glacial Maximum left a signature both in the Italian and in the European lineages. Genetic data combined with species distribution models support the presence of at least three putative glacial refugia in southern Europe (France, Balkan Peninsula and Black Sea) during the

*Corresponding authors. E-mail: paolo.colangelo@uniroma1.it; rfeuda@caltech.edu

last glacial maximum that likely contributed to post-glacial recolonization of Europe. By contrast, the Italian lineage remained trapped in the Italian peninsula and, according to the pattern observed in other subterranean mammals, did not contribute to the recolonization of northern latitudes. *Biological Journal of the Linnean Society*, 2015, **114**, 495–512.

ADDITIONAL KEYWORDS: cytochrome *b* – Europe – glacial refugia – historical demography – Last Glacial Maximum – paraphyly – phylogenetics – SDM – sPCA.

INTRODUCTION

The influence of Quaternary climatic oscillations on the current distribution and genetic diversity of European mammals is well established (Hewitt, 2004; Hofreiter & Stewart, 2009). Traditionally, phylogeographical analyses have suggested a contraction of species ranges during the glacial phases and a subsequent expansion in the interglacial periods (Hewitt, 2000). The effects of these movements affected the current pattern of genetic variability. During glaciation, many species survived in relatively small areas of refuge, usually in the southern part of the European continent, such as the Italian, Iberian, and Balkan peninsulas (Randi, 2007). However, recent studies have suggested that these refugia could have been much more common and widespread than previously considered (Provan & Bennett, 2008; Hampe & Jump, 2011). Our understanding of the effect of Pleistocene climatic changes mostly comes from phylogeographical studies of non-subterranean mammals, whereas the influence of the glaciation cycles on subterranean mammals has received little attention.

Moles are strictly subterranean mammals occurring in Europe, with six species belonging to the monophyletic genus *Talpa* (Wilson & Reeder, 2005). With the exception of sand dunes, permanently waterlogged and acid soils (Mitchell-Jones *et al*., 1999), moles are ubiquitous from lowland meadows up to mountains. Subterranean mammals are characterized by low vagility, and they live in a fairly constant environment that is characterized by the absence of light and small fluctuations of temperature and humidity (Lacey, Patton & Cameron, 2000). Therefore, they are expected to be less subject to seasonal climatic oscillations unless these lead to extreme drought or freezing, which dramatically increase of hardness of soils and lower the availability of feeding resources represented by the soil fauna.

The present study aimed to investigate whether and how Pleistocene climatic changes have shaped the current genetic variability of the widespread common mole *Talpa europaea*. By contrast to other mole species, *T. europaea* is widely distributed across Europe (Mitchell-Jones *et al*., 1999; Wilson & Reeder, 2005). Therefore, it represents the best candidate for studying the influence of Quaternary climatic oscillation on subterranean mammals. The common mole ranges from the Ebro River in Spain to the Ob and Irtish Rivers in Russia (Wilson & Reeder, 2005). According to the morphology of the mandibles, Loy & Corti (1996) suggested the existence of four distinct clades (i.e. those in Italy, the Balkan peninsula, Spain, and Central Europe). The distribution of these four clades and the existence of step clines of morphological variation have been related to interglacial secondary contact zones between populations expanding from the glacial refugia in the Balkans, Italy, and Spain (Loy & Corti, 1996).

To identify the historical and ecological factors that could have shaped the actual genetic variability of *T. europaea*, we combined phylogeographical inference and species distribution modelling. First, we evaluated the phylogeographical structure of *T. europaea* across its range using mitochondrial cytochrome *b* sequences from specimens covering a large portion of the species range. To validate the molecular results and to provide insights concerning the locations of putative glacial refugia, we developed species distribution models (SDMs) at three time intervals: present-day, Last Glacial Maximum (LGM) and Last Interglacial (LIG).

MATERIAL AND METHODS

Sample collection and laboratory procedures

We analyzed 85 samples of *T. europaea* from 46 localities covering most of the species distribution range (Fig. 1, Table 1). Tissues were gathered from various sites (Table 1) and preserved in ethyl alcohol 70%. We also retrieved 16 sequences available from GenBank (Table 1), two of which (FN640556 and FN640557) were incorrectly assigned by Colangelo *et al*. (2010) to *Talpa occidentalis*. The re-evaluation of the taxonomic status of these two specimens was performed on the basis of morphological characters (body size, humerus, and skull), which unambiguously show that these two Spanish specimens belong to *T. europaea* (for further details, see Discussion).

DNA was extracted as described in Aljanabi & Martinez (1997). Cytochrome *b* sequences (1140 bp)

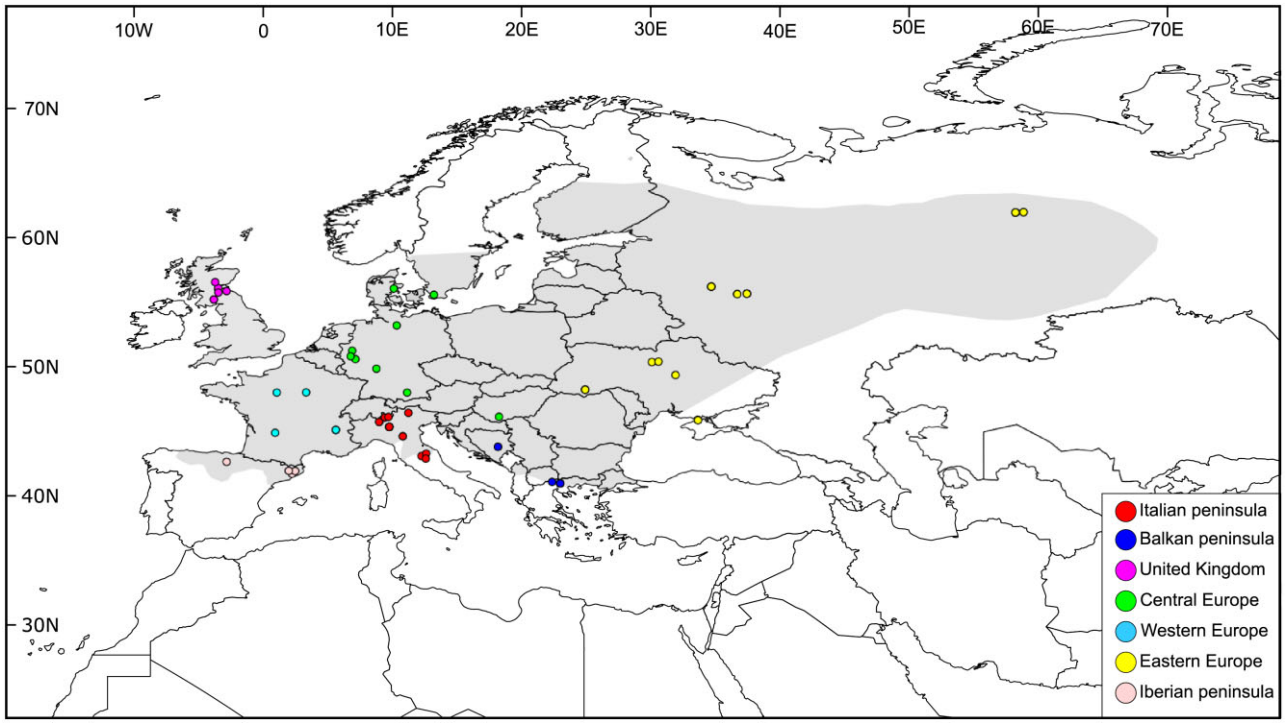

Figure 1. Map of sampling localities for mitochondrial DNA analysis. The current distribution range of *Talpa europaea*, obtained from IUCN (Amori *et al*., 2008), is reported as shaded area. Localities were coloured according to their geographical origin (for details, see Table 1).

were obtained in accordance with the amplification protocol and using primers described in Colangelo *et al*. (2010). The polymerase chain reaction product was purified with SureClean (BioLine) and sequencing was carried out by Macrogen Inc. (http://www.macrogen.com). The chromatograms were edited using FINCHTV, version 1.3.1 (Geospiza Inc.; http://www.geospiza.com/), and the alignment was performed using SEAVIEW, version 4.0 (Gouy, Guindon & Gascuel, 2010).

Phylogeographical structure

Phylogenetic trees were obtained using maximum likelihood (ML) and Bayesian inference (BA). The trees were rooted with *Talpa romana* and *Talpa caeca* (Accession numbers: FN640563-4 and FN640558-9). ML and BA were performed under a general time reversible (GTR) model (Tavaré, 1986) and a gamma distribution of rates across sites (Γ_4) that was the best fitting model according to the Akaike information criterion as implemented in JMODELTEST, version 2 (Darriba *et al*., 2012) ML analysis was performed using PHYML, version 3.0 (Guindon *et al*., 2010) with 100 bootstrap replicates. Bayesian inference was performed using MrBayes, version 3.2.1 (Ronquist & Huelsenbeck, 2003) using 10 million (burn-in = 25%) generations and sampling every 1000 steps. Convergence between chains was verified by manual inspection using the sump command and by using the mean SD of split frequencies, which was below 0.01 in our runs, as an indicator of convergence.

The relationships among haplotypes were also investigated by using a statistical parsimony (SP) network as implemented in TCS, version 1.2.1 (Clement, Posada & Crandall, 2000). Phylogenetic networks have the advantage over tree-based phylogenetic methods in that they allow multifurcation and, thus, are well suited when genetic divergence is low, as in intraspecific data. Before calculating the SP network, sites with missing data were excluded to reduce the ambiguities introduced in mutational connections.

To further investigate the spatial distribution of the genetic diversity within each lineage, we used a spatial principal component analysis (sPCA) as implemented in the R package ADEGENET, version 1.3-8 (Jombart, 2008). sPCA is a tool used to investigate cryptic spatial patterns of genetic variability using georeferenced genetic data. This method relies on modifications of the PCA such that not only the variance between the studied entities (individuals or populations), but also their spatial autocorrelation is taken into account (Jombart *et al*., 2008). sPCA

Table 1. Country, locality, geographical origin (EE, Eastern Europe; WE, Western Europe; CE, Central Europe; IT, Italian peninsula; BK, Balkan peninsula; SP, Spain), coordinates (provided in decimal degrees), haplotype number, and accession number of the specimens included in the analyses

ID	Country	Locality	Geographical origin	Latitude	Longitude	Haplotype	Accession number
RO3	Ukraine	Ishun	EE	45.92	33.82	H10	KF801516
RO2	Ukraine	Ishun	EE	45.92	33.82	H11	KF801517
SAR2	Bosnia-Herzegovina	Saraievo	BK	43.85	18.30	H28	KF801542
SAR1	Bosnia-Herzegovina	Sarajevo	BK	43.85	18.30	H28	KF801541
AB076829*	Denmark	Arhus	CE	56.15	10.22	H1	AB076829
4238	France	Courtalain	WE	48.07	1.14	H4	KF801511
6868	France	Dixmont	WE	48.08	3.41	H5	KF801512
11165	France	Eyzies-de-Tayac	WE	44.94	1.01	H6	KF801513
FN640550*	France	Grenoble	WE	45.17	5.72	H7	FN640550
FRA2	France	Grenoble	WE	45.17	5.72	H8	KF801514
FRA3	France	Grenoble	WE	45.17	5.72	H9	KF801515
ZFMK2004-056	Germany	Bonn	CE	50.73	7.10	H30	KF801545
ZFMK2005-254	Germany	Cologne	CE	50.93	6.95	H29	KF801543
AB037601*	Germany	Leubechk	CE	51.32	6.98	H2	AB037601
ZFMK2007-005	Germany	Munster	CE	49.92	8.87	H1	KF801544
ZFMK2005-055	Germany	Wesseling-Adendorf	CE	53.28	10.44	H31	KF801546
ZFMK2002-138	Germany	Wesseling-Urfeld	CE	48.07	11.25	H32	KF801547
FJ688080*	Greece	Eidomeni	BK	41.12	22.50	H36	FJ688080
FJ688081*	Greece	NeaMalgara	BK	41.01	23.13	H33	FJ688081
FJ688083*	Greece	NeaMalgara	BK	41.01	23.13	H34	FJ688083
FJ688085*	Greece	NeaMalgara	BK	41.01	23.13	H35	FJ688085
FJ688082*	Greece	NeaMalgara	BK	41.01	23.13	H37	FJ688082
FJ688084*	Greece	NeaMalgara	BK	41.01	23.13	H38	FJ688084
FJ715340*	Hungary	Nova	BK	46.18	18.38	H39	FJ715340
FN640551*	Italy	Belvedere	IT	43.15	12.60	H15	FN640551
BZ1	Italy	Bolzano	IT	46.49	11.35	H12	KF801518
Como1	Italy	Como	IT	45.78	9.08	H22	KF801533
Como2	Italy	Como	IT	45.78	9.08	H23	KF801534
T2-1134	Italy	CosioValtellino	IT	46.13	9.52	H17	KF801525
CDM3	Italy	Crocedimora	IT	43.13	12.37	H13	KF801522
FN640549*	Italy	Modena	IT	44.67	10.92	H16	FN640549
MO3	Italy	Modena	IT	44.67	10.92	H16	KF801524
MV11	Italy	Montevillano	IT	43.15	12.60	H13	KF801519
MV14	Italy	Montevillano	IT	43.15	12.60	H14	KF801520
MV15	Italy	Montevillano	IT	43.15	12.60	H14	KF801521
FN640552*	Italy	Pune Alto	IT	43.07	12.62	H13	FN640552
PNA5	Italy	Pune Alto	IT	43.07	12.62	H13	KF801523
VAL2	Italy	Sondalo	IT	46.33	10.32	H19	KF801527
VAL3	Italy	Sondalo	IT	46.33	10.32	H19	KF801528
VAL4	Italy	Sondalo	IT	46.33	10.32	H19	KF801529
VAL7	Italy	Sondalo	IT	46.33	10.32	H19	KF801532
VAL5	Italy	Sondalo	IT	46.33	10.32	H20	KF801530
VAL6	Italy	Sondalo	IT	46.33	10.32	H21	KF801531
T3-660	Italy	Talamona	IT	46.14	9.61	H18	KF801526
PI 08-17-2	Russia	Garevka	EE	62.06	58.47	H42	KF801562
PI 08-20-5	Russia	Garevka	EE	62.06	58.47	H42	KF801565
PI 08-21-6	Russia	Garevka	EE	62.06	58.47	H42	KF801561
PI 08-18-3	Russia	Garevka	EE	62.06	58.47	H43	KF801563
PI 08-19-4	Russia	Garevka	EE	62.06	58.47	H44	KF801564
Tver 2004-4	Russia	Krutitsy	EE	56.30	34.86	H1	KF801549
Tver 2000-2	Russia	Krutitsy	EE	56.30	34.86	H1	KF801548
Tver 2000-3	Russia	Krutitsy	EE	56.30	34.86	H1	KF801554
Tver 2001	Russia	Krutitsy	EE	56.30	34.86	H1	KF801550
Tver 2001-10	Russia	Krutitsy	EE	56.30	34.86	H1	KF801552

Table 1. *Continued*

ID	Country	Locality	Geographical origin	Latitude	Longitude	Haplotype	Accession number
Tver 2001-11	Russia	Krutitsy	EE	56.30	34.86	H1	KF801553
Tver 2001-9	Russia	Krutitsy	EE	56.30	34.86	H1	KF801551
M 11-6	Russia	Moscow	EE	55.75	37.62	H1	KF801556
M 09xx	Russia	Moscow	EE	55.75	37.62	H40	KF801557
M 11-5	Russia	Moscow	EE	55.75	37.62	H40	KF801555
M 09	Russia	Moscow	EE	55.75	37.62	H41	KF801558
PI 08–12	Russia	Mt.Yanapupner	EE	62.08	59.09	H1	KF801567
PI 08–13	Russia	Mt.Yanapupner	EE	62.08	59.09	H45	KF801566
Zven 11-1	Russia	Zvenigorod	EE	55.72	36.87	H1	KF801568
Zven 11-4	Russia	Zvenigorod	EE	55.72	36.87	H1	KF801560
Y19192*	Sweden	Dalby	CE	55.67	13.33	H3	Y19192
Brovari 09-2	Ukraine	Brovary	EE	50.47	30.76	H50	KF801575
Cherkassy	Ukraine	Cherkassy	EE	49.43	32.08	H49	KF801572
Kiev	Ukraine	Kiev	EE	50.44	30.52	H48	KF801571
Kiev 08-1	Ukraine	Kiev	EE	50.44	30.52	H48	KF801573
Kosovo 11-1	Ukraine	Kosovo	EE	48.29	25.07	H46	KF801568
Kosovo 11-2	Ukraine	Kosovo	EE	48.29	25.07	H47	KF801569
Kosovo 11-3	Ukraine	Kosovo	EE	48.29	25.07	H47	KF801570
Kosovo 09-1	Ukraine	Kosovo	EE	48.29	25.07	H48	KF801574
UK6	UK	Ballinluig	UK	56.65	–3.65	H27	KF801540
UK18	UK	Cleish	UK	56.20	–3.42	H24	KF801535
UK1	UK	Durisdeer	UK	55.30	–3.75	H24	KF801536
UK2	UK	Durisdeer	UK	55.30	–3.75	H26	KF801538
UK5	UK	Hopetoun	UK	55.98	–3.40	H25	KF801537
UK3	UK	Monkrigg	UK	55.95	–2.76	H24	KF801539
95070507	Spain	Fogars de Montclús	SP	41.74	2.43		KF801510
FN640556*	Spain	Haro	SP	42.87	–2.85		FN640556
FN640557*	Spain	Haro	SP	42.87	–2.85		FN640557
94042301	Spain	Montseny	SP	41.75	2.38		KF801509
95000001	Spain	Montseny	SP	41.75	2.38		KF801508
93111701	Spain	Sant Pere De Vilamajor	SP	41.7	2.39		KF801507

Samples were collected from different museums: MNHN Paris, France; Museo di Anatomia Comparata, University of Rome 'La Sapienza', Italy; Museu de Granollers, Spain, ZFMK Bonn, Germany; Department of Vertebrate Zoology of Lomonosov Moscow State University, Russia; Zoological Museum of National Museum of Natural History, Kiev, Ukraine; National Museums of Scotland, UK. Sequences gathered from GenBank are denoted by an asterisk.

incorporates Moran's *I* (Moran, 1948, Moran, 1950) to detect spatial features in the data. A Gabriel graph was constructed based upon individual sample locations to define neighbours for the calculation of Moran's *I*. We used the global and local tests based on Monte Carlo permutations ($N = 9999$) to interpret global and local components of sPCA. The presence of a significant 'global structure' might be related to patterns of spatial genetic structure, such as patches, clines, isolation-by-distance, whereas 'local structure' refers to strong differences between local neighbourhoods (Jombart *et al.*, 2008).

Time of divergence

The fossil record of *T. europaea* is not well understood (van Cleef-Roders & van den Hoek Ostende, 2001) and is not suitable for a precise calibration of the molecular clock. To overcome this issue, we tentatively used two different methods to estimate major clade divergences.

First, we calibrated the molecular clock using a geological event (i.e. the time elapsed between the end of the LGM and the disappearance of the Doggerland, a land bridge that was connecting Great Britain to the European continent). Paleontological evidence suggests that, during the main Glaciation, Great Britain was an inhospitable environment for mammals (Stuart, 1995). However, the reappearance of a diverse fauna in the late glacial could be explained by the existence of a land connection between Great Britain and Europe (Stuart, 1995). The Doggerland remained exposed from the end of the LGM, from approximately 12 000 years BP to approximately 6000 years BP (Coles, 2000). Fossil records for *T. europaea* from the

UK extend as far back as the Early Middle Pleistocene, (Stuart, 1995). However, fossils of *T. europaea* from late Pleistocene were not found in Great Britain (Currant, 1989), suggesting that this species became extinct in Great Britain in the late Pleistocene. This scenario would suggest that the common moles currently found in Great Britain are derived from a late colonization that occurred during the Holocene (Montgomery *et al*., 2014), when the Doggerland was exposed. The time of divergence was estimated with BEAST, version 1.8 (Drummond & Rambaut, 2007) using a coalescent tree prior. The period of exposition of the Doggerland was used to calibrate the time of the most common ancestor of Britain specimens using a log-normal distribution (mean = –6.0, SD = 0.07, offset = 0.005) with a median of 0.00747 and a 95% highest posterior density (HPD) ranging from 0.00562 to 0.01477 million years ago. The molecular clock prior (i.e. strict clock versus log-normal relaxed clock) was selected using the Bayes factor (BF). The marginal likelihood of the two hypotheses was estimated using stepping-stone sampling as implemented in MrBayes using 50 million of generations. According to Kass & Raftery (1995), there is strong evidence (BF = 53.03) in favour of the strict clock model that was used for further analysis. Finally, the choice of the prior for the substitution rate prior was based on data in the literature. In talpids, the substitution rate (substitutions per million years) may range from 0.014, as observed in the genus *Talpa* (Colangelo *et al*., 2010), to 0.0224, as observed in *Galemys pyrenaicus* (Igea *et al*., 2013). Accordingly, we defined the substitution rate in advance by using a normal distribution with its mean set to 0.018 and SD set to 0.0024.

Second, because the BF strongly supports a clock-like behaviour of the dataset, we obtained an approximate estimation of the divergence among the main *T. europaea* lineages according to the formula $T = d/2\mu$; where T is the time of divergence, d is the genetic distance between groups, and μ is the substitution rate in millions of year. We calculated the uncorrected distance and used a substitution rate (μ) for cytochrome *b* of 0.01407 per lineage per million of years, as estimated by Colangelo *et al*. (2010).

Molecular diversity and historical demography

Uncorrected (p) and Kimura 2-parameter (K2P) genetic distances among major lineages of *T. europaea* and closely-related species (*T. occidentalis*, *T. romana* and *T. caeca*) were calculated using MEGA, version 5.0 (Tamura *et al*., 2011). The number of haplotypes, nucleotide diversity (π), and haplotype diversity (H_D) were calculated using the software DNASP, version 5.10.1 (Librado & Rozas, 2009). To investigate the level of diversity in different areas of the range of *T. europaea*, specimens were grouped according to their geographical origin (Fig. 1, Table 1). The demographic history of the main lineages was reconstructed using a coalescent-based Gaussian Markov random field (GMRF) Bayesian skyride plot (Minin, Bloomquist & Suchard, 2008) with a uniform smoothing as implemented in BEAST, version 1.8 (single run of 50 million generations). The effective sample size was evaluated using TRACER, version 1.6 (Rambaut *et al*., 2014). The GMRF skyride plot was reconstructed assuming a strict clock with a substitution rate defined according to the results of molecular clock analysis.

Species distribution models

To evaluate current and past habitat suitability for *T. europaea*, we produced distribution models under different climatic conditions using a species distribution modelling approach. The majority of *T. europaea* occurrences were not only derived from the online database 'Global Biodiversity Information Facility' (http://www.gbif.org), but also came from museum collections and the literature (see Supporting information, Appendix S1). The download dataset was trimmed to the spatial resolution of the environmental layer used (2.5 arc min, approximately 4 km) to avoid duplicate coordinates in each cell. The final dataset includes 1259 distribution records (see Supporting information, Appendices S1, S2) covering almost the entire extent of occurrence accounted by the IUCN for *T. europaea* (Amori *et al*., 2008). The model is subject to spatial bias because the western part of the species range is oversampled relative to the eastern range. To reduce this bias, species records were treated with a subsampling procedure (details and R script available in the Supporting information, Appendix S3). The subsampling procedure was replicated 10 times, obtaining 10 different datasets of 354 points that were used for the calibration of the model.

As environmental layers, we used available climatic data from the Worldclim database (Hijmans *et al*., 2005), which contains interpolated surfaces for 19 climatic variables, to infer the habitat suitability for *T. europaea*. Although the subterranean niche is generally stable, changing climate alters annual cycles of temperature and humidity. These factors can affect the survival of moles in a variety of ways, especially extreme climatic events, such as drought, freezing, and flooding. More specifically, drought driven by high temperatures increases soil hardness and limits the availability of feeding resources represented by soil invertebrates (Gorman & Stone, 1990; Loy, 2008). Extreme rainfall events may cause flooding of the soil,

which may affect mole survival by increasing the risk of death as a result of thermoregulation problems (Gorman & Stone, 1990). Accordingly, Worldclim bioclimatic variables can provide a good proxy for subterranean habitat suitability, and we chose 11 variables on the basis of the current knowledge of the ecology and habitat preferences of moles: annual mean temperature, mean diurnal range in temperature, isothermality, annual range in temperature, mean temperature of the driest quarter of the year, mean temperature of the warmest quarter of the year, precipitation of the driest month, precipitation seasonality, precipitation of the wettest quarter of the year, precipitation of the warmest quarter of the year, and precipitation of the coldest quarter of the year. We successively calculated Pearson correlation coefficients for pairwise comparisons among the set of 11 variables to evaluate the level of autocorrelation. According to the Pearson correlation coefficient, our set of variables shows a general low r value in all of the pairwise comparisons. We found an r value higher than 0.75 for only three pairwise comparisons (BIO1/BIO9, BIO19/BIO16 and BIO19/BIO14) (see Supporting information, Appendix S4), suggesting a low bias as a result of autocorrelation in our dataset.

We predicted the European common mole distribution using an ensemble forecasting approach, as implemented in the R package 'biomod2' (Thuiller *et al*., 2009). SDMs for *T. europaea* were generated using four different algorithms: Maxent (Phillips, Anderson & Schapire, 2006; Phillips & Dudík, 2008), Gradient Boosting Machines (GBM; Ridgeway, 1999), Generalized Linear Model (GLM; McCullagh & Nelder, 1989), and General Additive Model (GAM; Hastie & Tibshirani, 1990). All of the methods estimate species distributions using environmental predictors together with species occurrences. Here, the four algorithms were used to predict the species occurrence under present-day, last glacial maximum (LGM, ~26–20 kya), and LIG (~120–140 kya) conditions (Otto-Bliesner *et al*., 2006; Braconnot *et al*., 2007).

To obtain a reliable evaluation of the models, each of the 10 occurrence datasets derived from the subsampling procedure was randomly split into two subsets: 70% of records to calibrate the models and the remaining 30% for evaluation (Thuiller *et al*., 2009). This procedure was replicated 10 times. Each time, we randomly selected different 70–30% portions of the occurrence dataset, obtaining 100 different occurrence datasets for the calibration of the models. For each of the 10 replicated occurrence datasets, a random selection of 10 000 background points was performed (drawn in a large area of the Palaearctic region, including the entire extent of occurrence of *T. europaea*) to characterize the climate of the study area and to represent pseudo-absences. Predictive performance of the model was assessed by measuring the area under the receiver operating characteristic curve (AUC; Hanley & McNeil, 1982) and the true skill statistic (TSS; Allouche *et al*., 2006). This data-splitting procedure was repeated 10 times, obtaining 100 final occurrence datasets, and the evaluation values were averaged. To avoid using poorly calibrated models, only projections from models with AUC ≥ 0.7 and TSS ≥ 0.4 were considered in all subsequent analyses. Model averaging was performed by weighting the individual model projections by their AUC or TSS scores and averaging the result because this method was shown to be particularly robust (Marmion *et al*., 2009).

For LGM prediction, we used data from both the Community Climate System Model (CCSM) and the Model for Interdisciplinary Research on Climate (MIROC) available in Worldclim database. To counteract the extrapolation effect of the model projections over the past (LGM and LIG) climates, we calculated the multivariate environmental similarity surface (MESS) index (Elith, Kearney & Phillips, 2010), estimating similarities between the climate values used in training and the past climate projections across all of the *T. europaea* distribution range. The index was then plotted on a map where positive values and those near zero indicated a good extrapolation.

RESULTS

Phylogeographical structure and the spatial distribution of genetic diversity

According to phylogenetic analyses, *T. europaea* includes three main lineages, European, Italian and Spanish, and is paraphyletic because the six specimens from Northern Spain are more closely-related to *T. occidentalis* than to *T. europaea* from other parts of Europe (Fig. 2). The genetic divergence between Spain and the other two lineages (Table 2) is quite large ($p > 0.07$; K2P > 0.08). The times of divergence of the Spanish lineage estimated using two different approaches were comparable, suggesting an ancient separation (2.85 Mya; 95% HPD: 2.03–3.71 Mya using the Bayesian methods and approximately 2.5 Mya using the distance methods).

The other two *T. europaea* lineages (European and Italian) are reciprocally monophyletic (Fig. 2). The Italian lineage includes the entire Italian sample, with the exception of a specimen from Bolzano, whereas the European lineage includes all of the remaining specimens (Fig. 2). The monophyly of the European lineage is well supported in both of the phylogenetic trees, whereas the monophyly of the Italian lineage is not strongly supported (Fig. 2). The genetic divergence between the Italian and European lineages is lower with respect to that observed for

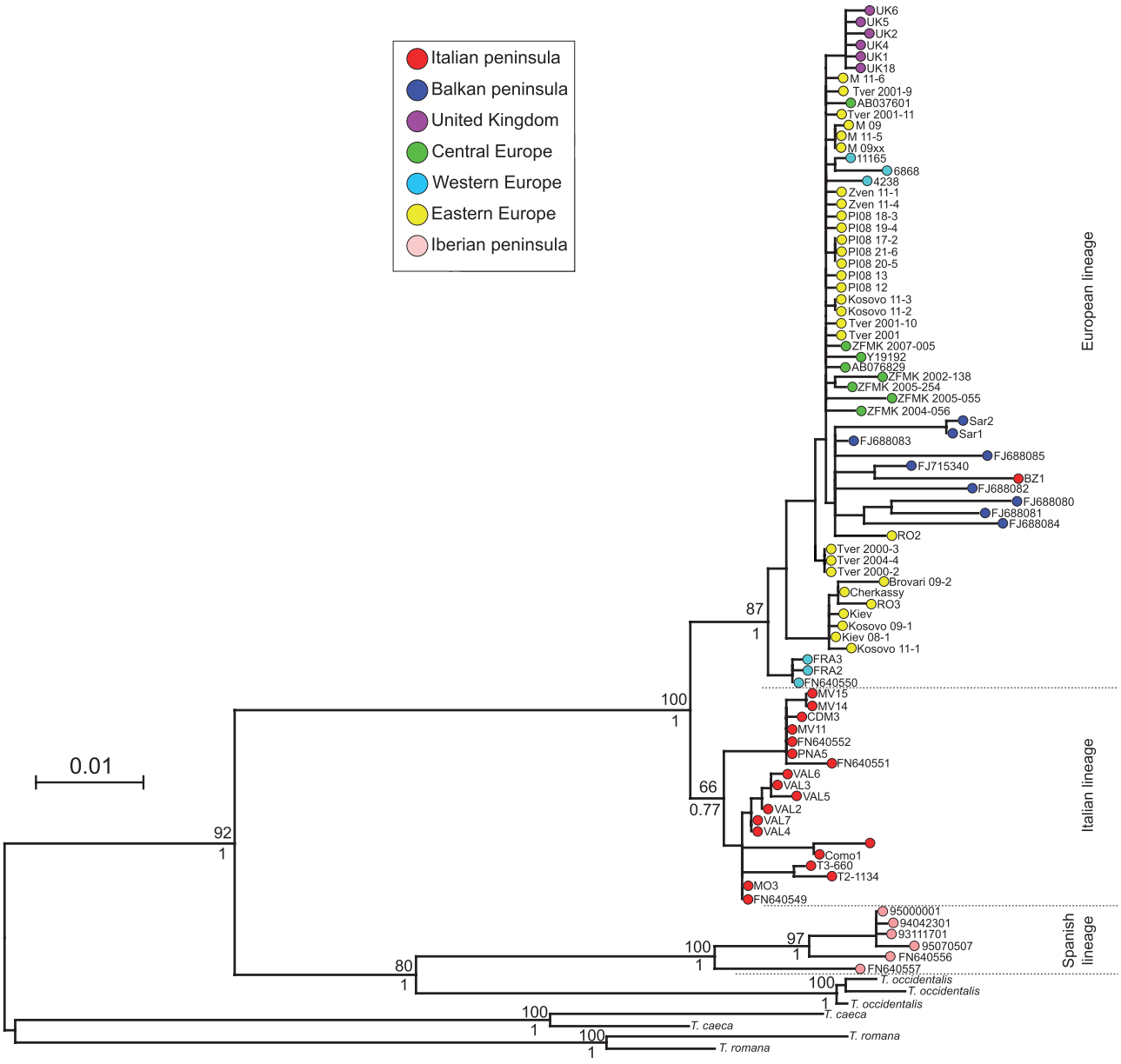

Figure 2. Phylogenetic tree based on 94 sequences of cytochrome *b*. Branch length and the topology were obtained from the maximum likelihood analysis under GTR+Γ_4. Bootstrap support and posterior probability are reported only for the key nodes. Sequences are labelled based on their geographical origin (coloured circles) according to Table 1 and Figure 1. The three main phylogroups (Europe, Italy, and Spain) are separated by dashed lines.

the Spanish lineage (Table 2). We estimated a separation of the two lineages at approximately 860 kya (95% HPD = 0.58–1.14 Mya) by using the relaxed molecular clock. This estimate is consistent with the estimation that we made using genetic distance (approximately 700 kya).

The SP network corroborates the results of phylogenetic analyses (Fig. 3), providing insights concerning the intralineage relationships. Specimens from Spain, separated by a large genetic distance, were not included in this analysis. Among the 79 remaining specimens, we found a total of 50 haplotypes (Table 1). Haplotype H1 is the most common and shared by 13 specimens, 11 of which were from eastern Europe (Russia) and two from central Europe (Denmark and Germany). Other haplotypes from central and eastern Europe are connected to H1 by few mutations producing a starlike topology. However, haplotypes from south-eastern Ukraine (H10, H46, H48, H49, and H50) show a certain degree

Table 2. Mean genetic distances among the three main lineages of *Talpa europaea*, identified by phylogenetic analyses (Fig. 2), and the more closely-related species

	Italian lineage ($N = 19$)	European lineage ($N = 60$)	Spanish lineage ($N = 6$)	*Talpa occidentalis*	*Talpa caeca*	*Talpa romana*
Italian lineage		0.021	0.081	0.088	0.098	0.107
European lineage	0.021		0.083	0.094	0.101	0.110
Spanish lineage	0.075	0.077		0.072	0.098	0.110
Talpa occidentalis	0.081	0.086	0.067		0.098	0.112
Talpa caeca	0.090	0.093	0.090	0.090		0.103
Talpa romana	0.097	0.100	0.100	0.101	0.093	

The values below the diagonal are *p*-distances and the numbers above the diagonal are K2P distances.

of differentiation. Three western European haplotypes from France (H4, H5, and H6) and those from UK appear closely related to the central-eastern European haplogroup. Three specimens from western Europe (H7, H8, and H9 from southern France) are those showing the highest divergence from the eastern-central haplogroup and are connected with a haplotype from south-eastern Europe (H48) by three mutations. Haplotypes from the Balkan peninsula are connected with the central-eastern European haplogroup but are separated by a number of mutations, ranging from 1 to 10. Furthermore Balkan haplotypes show a high intragroup diversity.

According to phylogenetic analyses, we found one haplotype in the Italian group from Bolzano (H12) that was more related to the other European haplotypes and directly connected with a haplotype from Eastern Europe (H11). The remaining Italian specimens form a distinct haplogroup separated by the closest haplotype from southern France by at least six mutations. Furthermore, the Italian haplogroup shows large intragroup diversity, with specimens included in two distinct clusters separated by seven mutations. These two clusters reflect the geographical origin of the haplotypes within the Italian peninsula from the north and the centre of Italy, respectively. According to the results of phylogenetic and network analyses, the specimen from Bolzano (BZ1) is not part of the Italian genetic pool and, in all further analysis (sPCA, molecular diversity and historical demography), it was considered as part of the European lineage.

sPCA was performed separately on the Italian and European lineages to highlight the intralineage spatial distribution of mtDNA diversity. The result of the sPCA supports the separation of the Italian lineage in two distinct haplogroups (Fig. 4A). Concerning the European lineages (Fig. 4B), the sPCA showed a clear partitioning of genetic diversity between northern and southern latitudes, with all the southern European specimens having positive values of sPC1. For both the Italian and the European lineages, the local test rejected a local population structure ($P = 0.99$ and $P = 0.54$, respectively), whereas the global test was significant (both with $P < 0.0001$).

Molecular diversity and historical demography

Molecular diversity indices suggest a low cytochrome *b* diversity of *T. europaea* across Europe, with few exceptions. According to our results (Table 3), western, central, eastern Europe, and UK have low nucleotide diversity and high haplotype diversity. Within eastern Europe, the southern localities from Ukraine harbour a higher diversity ($N = 10$, $H = 9$, $H_D = 0.978 \pm 0.054$, $\pi = 0.0070 \pm 0.0014$) compared to northern localities from Russia ($N = 20$, $H = 8$, $H_D = 0.816 \pm 0.071$, $\pi = 0.0011 \pm 0.0002$). The Balkan peninsula shows both highest nucleotide and haplotype diversity in Europe. The Italian peninsula shows high haplotype diversity and a nucleotide diversity lower than that observed in the Balkans but still higher than that observed in the rest of Europe. A different framework emerges when we analyze Italian specimens according to the clustering highlighted by the SP network (Fig. 3) and sPCA (Fig. 4). Although the low diversity in northern Italy ($N = 12$, $H = 9$, $H_D = 0.955 \pm 0.047$, $\pi = 0.0047 \pm 0.00111$) was similar to that observed in other regions of Europe, central Italy showed a high nucleotide diversity ($N = 7$, $H = 3$, $H_D = 0.667 \pm 0.16$, $\pi = 0.0184 \pm 0.0007$) comparable to that observed in the Balkan peninsula (Table 3).

The GMRF skyride plots (Fig. 5) suggest that both the Italian and European lineages experienced a population shrinking followed by successive demographic expansions. Specifically, in both lineages, we observed a bottleneck corresponding to the last glaciation (between 100 and 150 kya) and a subsequent increase of the population size of approximately

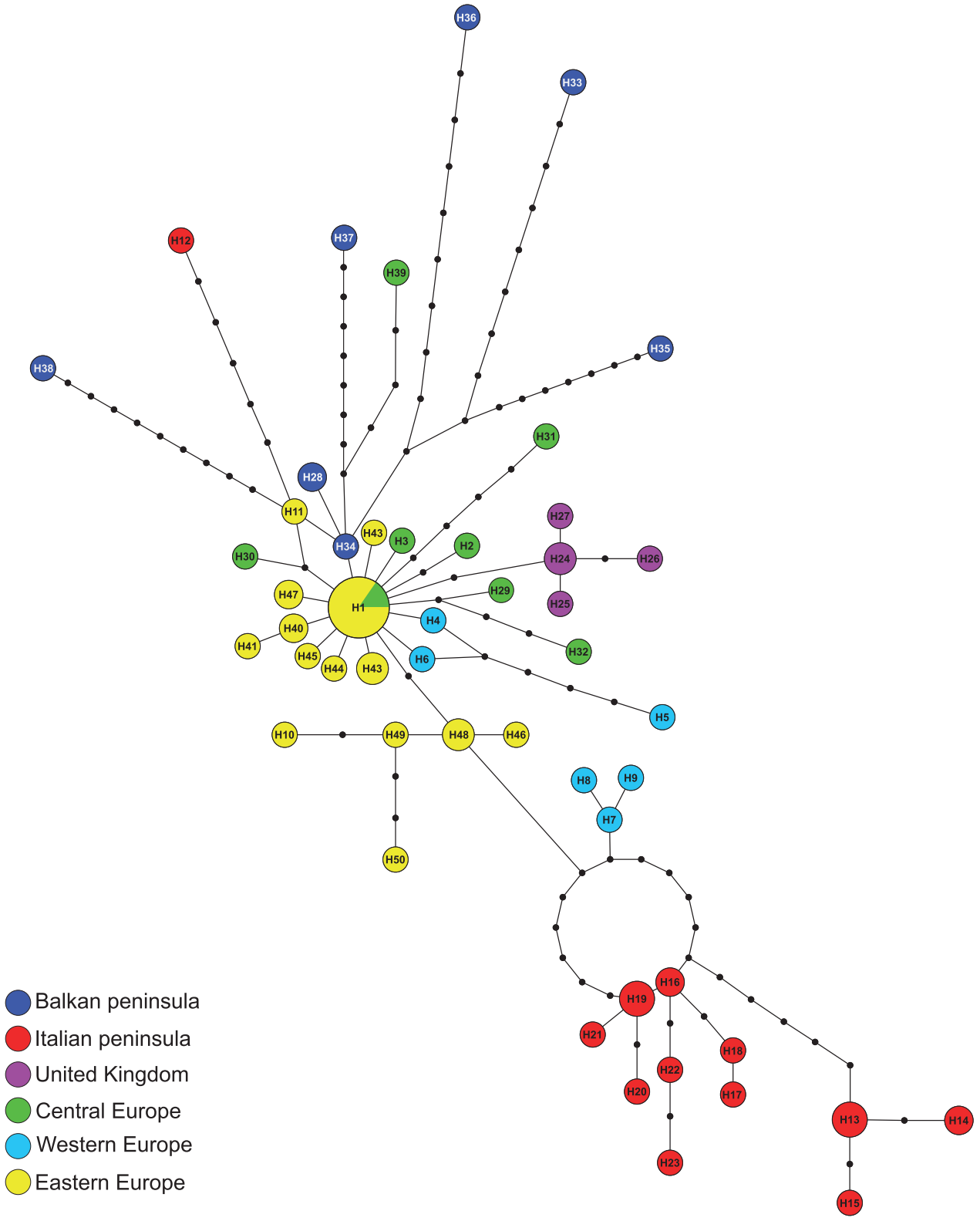

Figure 3. Parsimony network of *Talpa europaea* haplotypes. Circles represent haplotypes and with size proportional to the number of individuals sharing the same haplotype. Small filled circles are unsampled intermediate haplotypes; lines indicate single sequence differences (mutations) joining haplotypes. Circles are coloured according to the geographical origin of the haplotypes (Fig. 1 and Table 1).

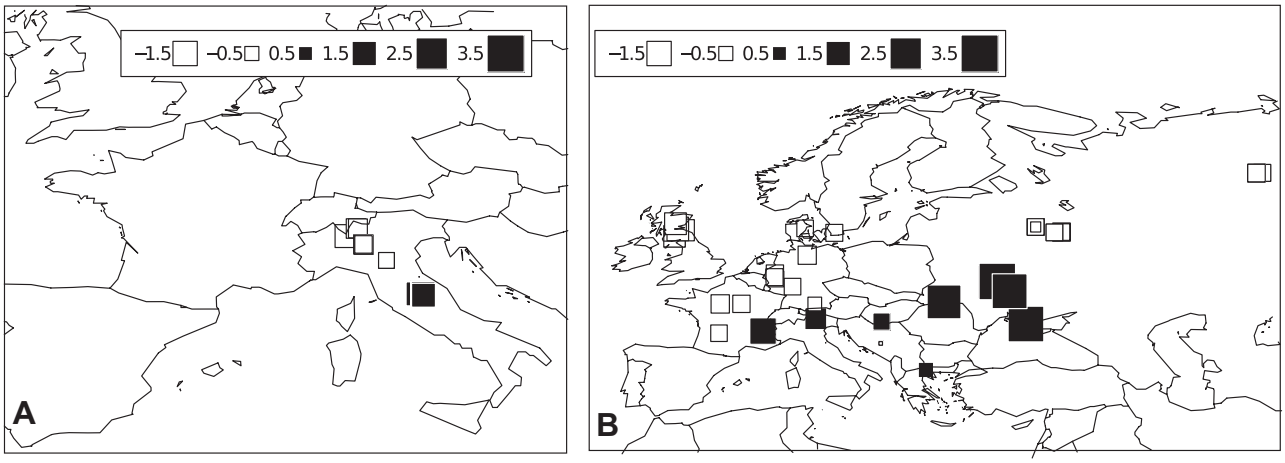

Figure 4. Spatial distribution of the scores of the first principal component obtained from the spatial principal component analysis (sPCA) for the European (A; $N = 60$) and for the Italian (B; $N = 19$) lineages. The values of scores component of the sPCA are represented using black (positive values) and white (negative values) symbols. The two maps reveal that cytochrome *b* diversity is partitioned in a northern and a southern component both in the Italian peninsula and in Europe.

Table 3. Molecular diversity within *Talpa europaea*

	N	H	Hd	π
Western Europe	6	6	1 ± 0.096	0.0078 ± 0.0015
Central Europe	9	8	0.972 ± 0.064	0.0053 ± 0.0012
Eastern Europe	30	17	0.917 ± 0.038	0.0045 ± 0.0008
United Kingdom	6	5	0.933 ± 0.122	0.0018 ± 0.0004
Balkan peninsula	8	7	0.964 ± 0.077	0.0162 ± 0.0028
Italian peninsula	20	13	0.947 ± 0.032	0.0094 ± 0.0019

Specimens were grouped according to their geographical origin (Fig. 1 and Table 1). For each group, the number of specimens (*N*), number of haplotypes (*H*), haplotypic diversity (H_D), and nucleotide diversity (π) are reported.

50 kya. The European lineage also shows a demographic decrease of approximately 350 kya, probably associated with the Mindel glaciation (ranging from 450 to 300 kya), and an increase of approximately 200 kya.

SDM

The SDMs had very good predictive abilities (mean AUC > 0.9 and mean TSS > 0.7 for all of the models), indicating the good reliability and accuracy of the models. The four models report very similar habitat suitability, as predicted on the basis of the four bioclimatic datasets (current, LIG, and two for LGM). The MESS index (map shown in the Supporting information, Appendix S4) shows a generally good similarity between the reference set of points and the set of predictor variables.

The ensemble model (obtained combining the GAM, GBM, GLM, and Maxent models) revealed high habitat suitability for *T. europaea* across Europe under present-day bioclimatic conditions (Fig. 6A), in agreement with the current known distribution of *T. europaea*. The projection of the model over the LIG (~120–140 kya) returned a lower habitat suitability in Europe with respect to the present-day, more evident in the eastern part of the continent and in a small area of present-day Poland, although still good over almost all the range (Fig. 6B), and comparable to that observed under present-day bioclimatic conditions. The habitat suitability drastically decreased during the LGM irrespective of the data used [i.e. Community Climate System Model (CCSM) versus MIROC]. Indeed, both models support southern European patches of suitable habitat for *T. europaea* during the LGM (Fig. 6C, D). Suitable areas were found in the three southern European peninsulas (Italian, Balkan, and Iberian), around the Black Sea, and a north central area corresponding to portions of the France and UK. In both models, a large portion of the suitable areas in the northern latitudes fall in areas that are currently under the sea level. Differences between

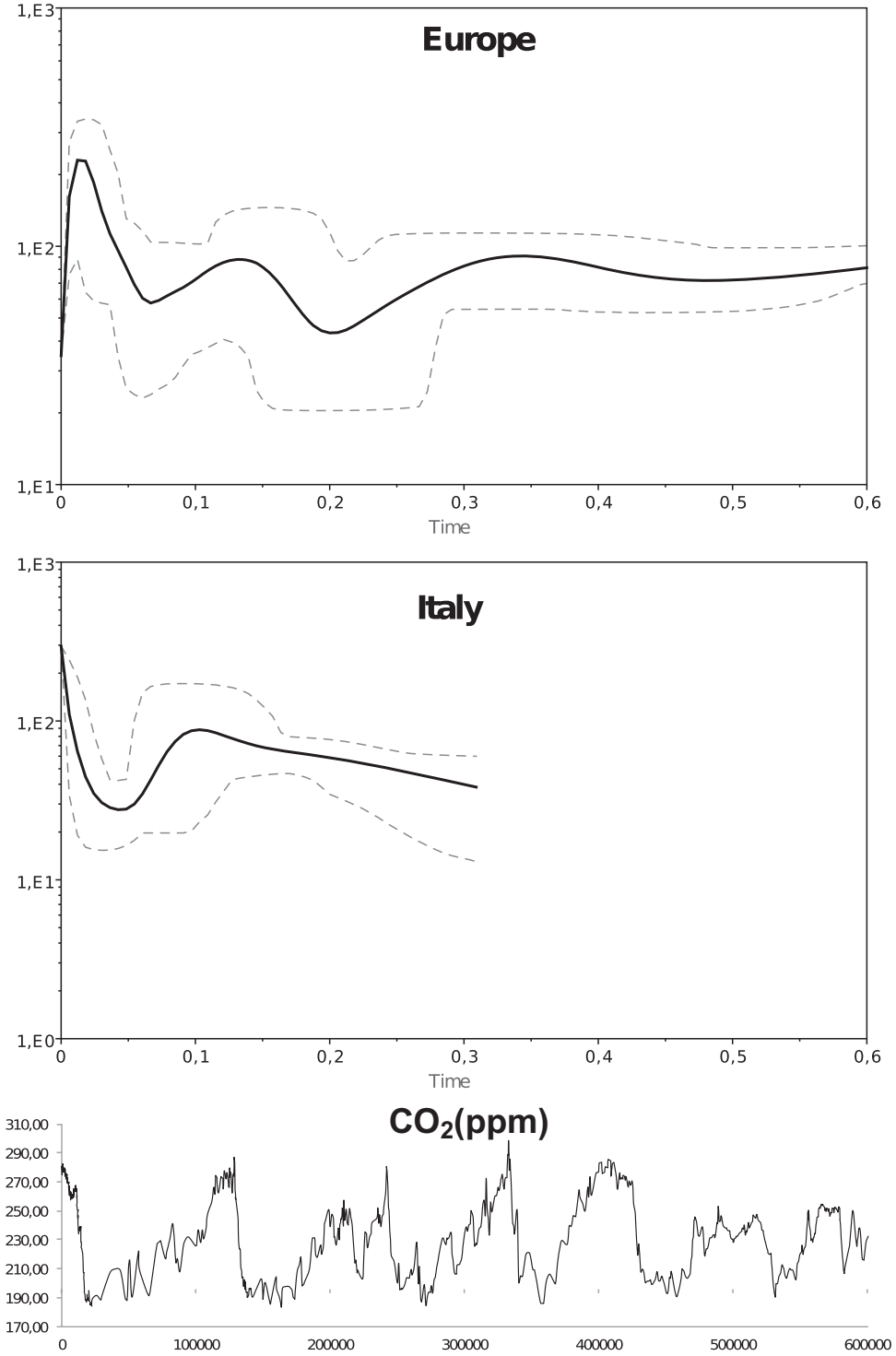

Figure 5. Gaussian Markov random field Bayesian skyride plots based on cytochrome *b* sequences. Analyses were reported for each of the two major lineages (i.e. Italian peninsula and Europe). On the *y*-axis, the effective population size is reported. The solid line is the median estimate, and the dotted lines show the 95% highest posterior density limits. Time of expansion (in million of years, on the *x*-axis) was calculated using the mean clock rate obtained for *Talpa europaea* in molecular clock analysis (0.0196). Bottom: glacial and interglacial cycles as represented by atmospheric CO_2 concentration (parts per million by volume), as measured by the composition of air trapped in ice cores from Antarctica (Lüthi *et al*., 2008).

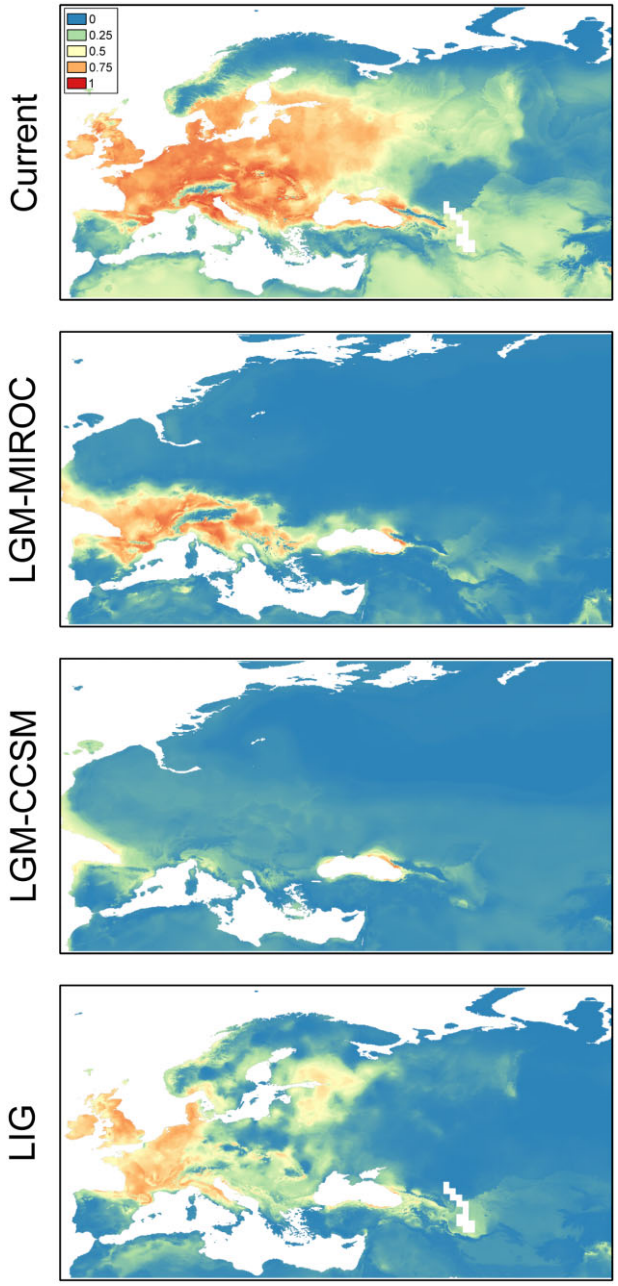

Figure 6. Ensemble species distribution model for *Talpa europaea* developed for present-day conditions, for the Last Glacial Maximum (LGM) based on the Model for Interdisciplinary Research on Climate (MIROC) and on Community Climate System Model (CCSM) palaeo-climatic models and for the Last Interglacial (LIG). Warmer colours show areas with better-predicted conditions.

the two models are mainly based on the degree of suitability. The model based on MIROC data showed higher levels of suitability and also supported the expansion of suitable areas to central-western Europe. By contrast, the model based on CCSM also suggests a suitable area for *T. europaea* in the same south western region of Europe, although to a lesser extent than that proposed by MIROC. The Italian peninsula shows an area of high suitability during the LGM, although it is not continuously connected with other areas in France and the Balkans.

DISCUSSION

Paraphyly of *T. europaea*

The analysis of the phylogeographical structure of *T. europaea* indicates the presence of three different lineages within the species range: the first includes only the Spanish haplotypes, the second includes specimens from the Italian peninsula except one from Bolzano (BZ1 from northern Italy close to the border with Austria), and a third includes the moles from central, northern, and eastern Europe. The Spanish lineage forms a monophyletic clade with *T. occidentalis*, suggesting the paraphyly of *T. europaea* (but see below). Levels of genetic differentiation and the inferred time of divergence support an early separation of this lineage, dated at the Pliocene–Pleistocene boundary.

The European mole has been recorded from north-east Spain subsequent to the first report by Miller (1912) and is currently regarded as part of the Spanish fauna; its range slightly overlaps with the Iberian mole *T. occidentalis* (Palomo & Gisbert, 2002). We examined the morphology of a subsample of 10 *T. europaea* and 17 *T. occidentalis* from Spain (skins and skulls in ZFMK Bonn). The Spanish *T. europaea* shows the general features of the species and differs from *T. occidentalis* by the larger sizes of its body, humerus, and skull. The condylobasal length ranges from 32.4 to 37.6 mm ($N = 8$) in the *T. europaea* sample, with no overlap with *T. occidentalis* (29.4 to 31.2 mm; $N = 15$). Miller (1912) and subsequent studies found similar size differences. Our samples clearly belong to the *T. europaea* clade. However, the Spanish common moles, despite showing the general morphological features of *T. europaea*, have a mean genetic distance from other *T. europaea* lineages more than three times greater than the distance observed between the European and the Italian lineages. It is interesting to note that the K2P distances of the Spanish lineage from *T. occidentalis* and from other *T. europaea* lineages are both very high (0.072 and 0.08, respectively) and very close to those reported by Colangelo *et al*. (2010) for pairs of sister species within the genus *Talpa*.

According to our results, two different scenarios can be suggested. The first scenario claims an introgression of mtDNA from *T. occidentalis* to *T. europaea*. Introgression of DNA among insectivore mammals has been reported along parapatric contact zones for

moles (Loy *et al*., 2001), shrews (Yannic, Basset & Hausser, 2008; Bannikova & Lebedev, 2010), and hedgehogs (Bogdanov *et al*., 2009), as well as for other small mammals such as voles (Abramson, Rodchenkova & Kostygov, 2009). According to the levels of genetic divergence between the Spanish *T. europaea* and *T. occidentalis*, we can exclude a recent introgression of mtDNA, although we cannot a priori rule out an ancient introgression during some phases of the late Pliocene, similar to what was hypothesized for the introgression events between *T. europaea* and *T. romana* in the parapatric contact zone in central Italy (Loy *et al*., 2001).

A second scenario considers the possible existence of a differentiated taxonomic entity in northern Spain. This scenario would be congruent with the observed level of genetic divergence among the Spanish lineage, the other *T. europaea* lineages and *T. occidentalis*. Moles are quite common in Europe, although few studies have focused on the description of taxonomic diversity using molecular methods. It is interesting to note that the genus *Talpa* currently includes nine species and, with the exception of *T. europaea*, most species are endemic to restricted geographical regions (Mitchell-Jones *et al*., 1999; Wilson & Reeder, 2005). It should also be stressed that, within the tribe Talpini, the possible occurrence of new cryptic species has been recently claimed for the genera Mogera and Euroscaptor (Zemlemerova *et al*., 2013; He *et al*., 2014). Based on the genetic divergence and the fact that the Iberian peninsula is rich in endemism, it is thus plausible that *T. europaea*, as currently defined, could include more than one species.

Ancient Pleistocene genetic break in the common mole

The other two European lineages show a large genetic gap, congruent with a long-term separation of the Italian lineage from the rest of Europe. The split of these two lineages occurred much later than the separation of the Spanish lineage, approximately 700–800 kya (Fig. 3), in agreement with the data reported by Colangelo *et al*. (2010). Most importantly, the estimated time of divergence largely predated the LGM, leading us to hypothesize that the causes of the deep genetic divergence observed between the two lineages must be found far back in the past, during the middle Pleistocene, which is a pattern found for several others mammals in Europe (Hewitt, 2000). After colonization of Italy, the common mole apparently remained trapped in the peninsula. The transition between the Early and Middle Pleistocene, approximately 0.78 Mya, has been characterized by a change in the cyclicity of the Earth's climate (Head, Pillans & Farquhar, 2008). This change had a strong effect on the composition of the ecological communities. For example, in Italy, the mammalian record of the late early Pleistocene is in accordance with open arid conditions followed by a wetter and substantially cooler climate in the Early-Middle Pleistocene (Head *et al*., 2008). Moreover, in the Alpine region, the Pleistocene glacial stages starting at approximately 0.87 Mya were characterized by enhanced glacier development and glacio-eustatic lowstands, as well as by a substantial forest withdrawal with long phases of persistent forest cover at low altitudes (Muttoni, Scardia & Kent, 2010). These factors might have created unsuitable conditions for *T. europaea* both during the glacial and the interglacial phases. It is interesting to note that the morphology supports an independent evolution of the Italian populations with respect to other European common moles. Loy & Corti (1996) found a sudden change in mandible morphology over a short geographical range, highlighting a significant morphological difference of the Italian *T. europaea* versus the other populations from central Europe. Specifically, an abrupt change was found between the Italian populations and the nearby populations from Switzerland, Germany, and Austria (Loy & Corti, 1996).

We found only one haplotype in northeastern Italy (BZ1 from Bolzano) that is more closely related to the Balkan haplotypes than to the Italian ones. The origin of this haplotype is unclear because of a lack of good sampling in that area. However, it is possible to speculate that it is related to a recent migration from the Balkan area that may have occurred during the late Pleistocene.

The influence of the last glaciation on the distribution of genetic variability

For the recent evolutionary history of *T. europaea*, our results support a phylogeographical pattern congruent with that reported for non-subterranean small mammals. The sPCA suggests that cytochrome *b* diversity is partitioned into a northern and a southern component. Southern Europe presents a higher molecular diversity, fitting a pattern of 'southern richness and northern purity' that is usually attributed to the prolonged population stability at southern latitudes, and recent recolonization events at northern latitudes (Hewitt, 2004). This hypothesis is in agreement with the palaeo-ecological models, indicating that, during the LGM, central and eastern Europe were largely unsuitable for the common mole. Extremely low temperatures for extended periods likely favoured the establishment of permafrost, where soil hardness and scarcity of edaphic resources could have represented the main limiting factors for moles. Indeed, moles are missing from latitudes higher than 63°N (Mitchell-Jones *et al*., 1999).

The results of the historical demographic analysis (Fig. 6A) for the European lineage are coherent with a model of expansion after a bottleneck. Central Europe showed the lowest molecular diversity, as expected for an area recently colonized by few founders. In the eastern European populations, the low molecular diversity is comparable to that recorded for central Europe, suggesting a recent origin of moles in this area. However, south-eastern localities from Ukraine showed a certain degree of differentiation, as suggested by SP network (Fig. 4), sPCA (Fig. 5A), and diversity indices. According to the palaeo-ecological models, during the LGM, suitable areas were still available around the Black Sea. It is thus possible that the southern latitudes were the first to be recolonized and/or were subject to a less severe bottleneck during the LGM. The same could have occurred in Western Europe, where two different haplogroups were found. The one found in north and central France is closely related to the eastern and the central European haplotypes and can thus be considered of recent post-glacial origin. The haplogroup found is southern France is clearly differentiated from the other European haplotypes (Fig. 4), and it is interesting to note that SDM models indicate a good habitat availability in this area during the LGM, suggesting a possible role of putative refugia for southern France, as was also hypothesized for the semi fossorial rodent *Microtus arvalis* (Tougard *et al*., 2008). Finally, the highest level of molecular diversity and the high suitability of the Balkan peninsula during the LGM suggest that this area could have acted both as a refuge and a source population for the recolonization of Europe during the interglacial phases.

A final point to address is the isolation of the Italian lineage and the lack of a contribution of this genetic pool to the post-glacial recolonization of Europe, a pattern shared with other non-subterranean mammals (Grill *et al*., 2009; Hewitt, 2011; Colangelo *et al*., 2012). The Italian lineage showed a level of genetic diversity higher than the rest of Europe, with the exception of the Balkans, supporting its role as a long-term Pleistocene refuge. The GMRF skyride plot suggests that the Italian haplogroup underwent a decrease of the effective population size during the last glaciation, which, according to molecular diversity indices (Table 2), principally affected northern Italy. The successive demographic expansion did not coincide to an expansion north of the Alps that acted as barrier to post-glacial recolonization of Europe. Thus, the Alps acted as an initial force of isolation and continued to prevent the northern expansion of Italian moles, driving a significant increase in genetic differentiation.

Conclusions

According to cytochrome *b* phylogeny, the common mole *T. europaea*, as currently defined, is a paraphyletic taxon. To date, we cannot exclude the possibility that the Spanish lineage of *T. europaea* represents a new cryptic species, nor that it could be the result of past introgression of mtDNA from *T. occidentalis*. The causes of the paraphyly and of the genetic differentiation observed between the Spanish and the other two *T. europaea* lineages deserve further investigation using nuclear markers and morphological analyses.

According to our results, the phylogeographical structure of the common mole in Europe is in line with predictions made for non-subterranean mammals. We found an ancient genetic break between the Italian and the European lineages that can be dated back to middle Pleistocene. Thus, the cause of this genetic break must be sought in events that predated the LGM and probably comprised a combination of ecological factors driven by climate shifts. By contrast, the phylogeographical structure observed within the Italian and the European lineages is a clear by-product of the last glaciation. According to our findings from both genetic and SDM, *T. europaea* might have retreated to different refugia during the unsuitable glacial phases. These refugia were located not only in the Balkan and the Italian peninsulas, but also possibly all along a southern European belt from southern France up to the shores of the Black Sea.

ACKNOWLEDGEMENTS

We widh to thank A. V. Bobretzov, J. Herman (Department of Natural Sciences, National Museums of Scotland), Lídia Freixas Mora (Museu de Granollers-Ciències Naturals), Christiane Denys (MNHN Paris), and Holger Meinig (Werther), who kindly provided tissues of moles, respectively, from North Ural mountains (Komi Republic), UK, Spain, and France. We also thank anonymous reviewers for their helpful comments. The research of A. A. Bannikova and E. D. Zemlemerova was financially supported by Russian Foundation for Basic Research, grant N11-04-00020.

REFERENCES

Abramson NI, Rodchenkova EN, Kostygov A. 2009. Genetic variation and phylogeography of the bank vole (*Clethrionomys glareolus*, Arvicolinae, Rodentia) in Russia with special reference to the introgression of the mtDNA of a closely related species, red-backed vole (*C. rutilus*). *Genetika* **45:** 610–623.

Aljanabi SM, Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25:** 4692–4693.

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* **43:** 1223–1232.

Amori G, Hutterer R, Mitsain G, Yigit N, Kryštufek B, Palomo LJ. 2008. Talpa europaea. *The IUCN Red List of Threatened Species*, Version 2014.2. Available at: http://www.iucnredlist.org.

Bannikova A, Lebedev V. 2010. Genetic heterogeneity of the Caucasian shrew *Sorex satununi* (Mammalia, lipotyphla, soricidae) inferred from the mtDNA markers as a potential consequence of ancient hybridization. *Molecular Biology* **44:** 658–662.

Bogdanov A, Bannikova A, Pirusskii YM, Formozov N. 2009. The first genetic evidence of hybridization between West European and Northern white-breasted hedgehogs (*Erinaceus europaeus* and *E. roumanicus*) in Moscow region. *Biology Bulletin* **36:** 647–651.

Braconnot P, Otto-Bliesner B, Harrison S, Joussaume S, Peterchmitt J-Y, Abe-Ouchi A, Crucifix M, Driesschaert E, Fichefet T, Hewitt C. 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – part 1: experiments and large-scale features. *Climate of the Past* **3:** 261–277.

van Cleef-Roders J, van den Hoek Ostende L. 2001. Dental morphology of *Talpa europaea* and *Talpa occidentalis* (Mammalia: Insectivora) with a discussion of fossil *Talpa* in the Pleistocene of Europe. *Zoologische Mededelingen* **75:** 51–68.

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

Colangelo P, Aloise G, Franchini P, Annesi F, Amori G. 2012. Mitochondrial DNA reveals hidden diversity and an ancestral lineage of the bank vole in the Italian peninsula. *Journal of Zoology* **287:** 41–52.

Colangelo P, Bannikova AA, Krystufek B, Lebedev VS, Annesi F, Capanna E, Loy A. 2010. Molecular systematics and evolutionary biogeography of the genus Talpa (Soricomorpha: Talpidae). *Molecular Phylogenetics and Evolution* **55:** 372–380.

Coles BJ. 2000. Doggerland: the cultural dynamics of a shifting coastline. *Geological Society, London, Special Publications* **175:** 393–401.

Currant A. 1989. The Quaternary origins of the modern British mammal fauna. *Biological Journal of the Linnean Society* **38:** 23–30.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Elith J, Kearney M, Phillips S. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* **1:** 330–342.

Gorman ML, Stone RD. 1990. *The natural history of moles*. Ithaca, NY: Cornell University Press.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27:** 221–224.

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.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–321.

Hampe A, Jump AS. 2011. Climate relicts: past, present, future. *Annual Review of Ecology, Evolution and Systematics* **42:** 313–333.

Hanley JA, McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143:** 29–36.

Hastie TJ, Tibshirani RJ. 1990. *Generalized additive models*. London: Chapman & Hall.

He K, Shinohara A, Jiang XL, Campbell KL. 2014. Multilocus phylogeny of talpine moles (Talpini, Talpidae, Eulipotyphla) and its implications for systematics. *Molecular Phylogenetics and Evolution* **70:** 513–521.

Head MJ, Pillans B, Farquhar SA. 2008. The Early–Middle Pleistocene transition: characterization and proposed guide for the defining boundary. *Episodes* **31:** 255.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

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.

Hewitt GM. 2011. Mediterranean peninsulas: the evolution of hotspots. In: Zachos FE, Habel JC, eds. *Biodiversity hotspots*. Amsterdam: Springer, 123–147.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Hofreiter M, Stewart J. 2009. Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology* **19:** 584–594.

Igea J, Aymerich P, Fernandez-Gonzalez A, Gonzalez-Esteban J, Gomez A, Alonso R, Gosalbez J, Castresana J. 2013. Phylogeography and postglacial expansion of the endangered semi-aquatic mammal *Galemys pyrenaicus*. *BMC Evolutionary Biology* **13:** 115.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24:** 1403–1405.

Jombart T, Devillard S, Dufour AB, Pontier D. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* **101:** 92–103.

Kass R, Raftery A. 1995. Bayes factors. *Journal of the American Statistical Association* **90:** 773–795.

Lacey EA, Patton JL, Cameron GN. 2000. *Life underground: the biology of subterranean rodents*. Chicago, IL: University of Chicago Press.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451–1452.

Loy A. 2008. Famiglia Talpidae. Mammalia II. Erinaceomorpha, Soricomorpha, Lagormorpha, Rodentia. In: Amori G, Contoli L, Nappi A, eds. Fauna d'Italia*, Vol. XLIV*. Milano: Calderini de Il Sole 24 Ore, 91–93.

Loy A, Capula M, Palombi A, Capanna E. 2001. Genetic and morphometric evidence of introgression between two species of moles (Insectivora: *Talpa europaea* and *Talpa romana*) in central Italy. *Journal of Zoology* **254:** 229–238.

Loy A, Corti A. 1996. Distribution of *Talpa europea* (Mammalia, Insectivora, Talpidae) in Europe: a biogeographic hypothesis based on morphometric data. *Italian Journal of Zoology* **63:** 277–284.

Lüthi D, Le Floch M, Bereiter B, Blunier T, Barnola J-M, Siegenthaler U, Raynaud D, Jouzel J, Fischer H, Kawamura K, Stocker TF. 2008. High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* **453:** 379–382.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W. 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distribution* **15:** 59–69.

McCullagh P, Nelder JA. 1989. *Generalized linear models*. London: Chapman and Hall.

Miller GS. 1912. *Catalogue of the mammals of Western Europe (Europe exclusive of Russia) in the collection of the British Museum*. London: Trustees of the British Museum.

Minin VN, Bloomquist EW, Suchard MA. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular Biology and Evolution* **25:** 1459–1471.

Mitchell-Jones AJ, Amori G, Bogdanowicz W, Krystufek B, Reijnders P, Spitzenberger F, Stubbe M, Thissen J, Vohralik V, Zima J. 1999. *The atlas of European mammals*. London: Academic Press.

Montgomery WI, Provan J, McCabe AM, Yalden DW. 2014. Origin of British and Irish mammals: disparate post-glacial colonisation and species introductions. *Quaternary Science Reviews* **98:** 144–165.

Moran PA. 1948. The interpretation of statistical maps. *Journal of the Royal Statistical Society, B* **10:** 243–251.

Moran PA. 1950. Notes on continuous stochastic phenomena. *Biometrika* **37:** 17–23.

Muttoni G, Scardia G, Kent DV. 2010. Human migration into Europe during the late Early Pleistocene climate transition. *Palaeogeography, Palaeoclimatology, Palaeoecology* **296:** 79–93.

Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A. 2006. Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* **311:** 1751–1753.

Palomo LJ, Gisbert J. 2002. *Atlas de los mamíferos terrestres de Espana*. Madrid: Dirección General de Conservación de la Naturaleza-SECEM-SECEMU.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31:** 161–175.

Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution* **23:** 564–571.

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. *Tracer*, Version 1.6. Available at: http://beast.bio.ed.ac.uk/Tracer.

Randi E. 2007. Phylogeography of south European mammals. In: Weiss S, Ferrand N, eds. *Phylogeography of southern European refugia*. Berlin: Springer, 101–126.

Ridgeway G. 1999. The state of boosting. *Computing Science and Statistics* **31:** 172–181.

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

Stuart AJ. 1995. Insularity and Quaternary vertebrate faunas in Britain and Ireland. *Geological Society, London, Special Publications* **96:** 111–125.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28:** 2731–2739.

Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* **17:** 57–86.

Thuiller W, Lafourcade B, Engler R, Araújo MB. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography* **32:** 369–373.

Tougard C, Renvoise E, Petitjean A, Quere JP. 2008. New insight into the colonization processes of common voles: inferences from molecular and fossil evidence. *PLoS One* **3:** e3532.

Wilson DE, Reeder DM. 2005. *Mammal species of the world: a taxonomic and geographic reference*, 3rd edn. Baltimore, MD: Johns Hopkins University Press.

Yannic G, Basset P, Hausser J. 2008. A new perspective on the evolutionary history of western European *Sorex araneus* group revealed by paternal and maternal molecular markers. *Molecular Phylogenetics and Evolution* **47:** 237–250.

Zemlemerova ED, Bannikova AA, Abramov AV, Lebedev VS, Rozhnov VV. 2013. New data on molecular phylogeny of the East Asian moles. *Doklady Biological Sciences* **451:** 257–260.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Map of localities and sources of data used to train the Maxent species distribution model (SDM).
Appendix S2. Coordinates of *Talpa europaea* specimens used to obtain species distribution models.
Appendix S3. Subsampling procedure and R script.
Appendix S4. Pearson correlation coefficient for Worldclim bioclimatic variables pairwise comparisons and results of the multivariate environmental similarity surface (MESS) index.