The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2016, **119**, 943–959. With 5 figures.

Pleistocene diversification in Morocco and recent demographic expansion in the Mediterranean pond turtle *Mauremys leprosa*

JOANA VERÍSSIMO[1,2], MOHAMED ZNARI[3], HEIKO STUCKAS[4], UWE FRITZ[4], PAULO PEREIRA[1,2], JOSÉ TEIXEIRA[1,5], MARCO ARCULEO[6], FEDERICO MARRONE[6], FRANCESCO SACCO[6], MOHAMED NAIMI[3], CHRISTIAN KEHLMAIER[4] and GUILLERMO VELO-ANTÓN[1]*

[1]*CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, Instituto de Ciências Agrárias de Vairão, R. Padre Armando Quintas, 4485-661, Vairão, Portugal*
[2]*Departamento de Biologia da Faculdade de Ciências da Universidade do Porto, Rua Campo Alegre, 4169-007, Porto, Portugal*
[3]*Laboratory 'Biodiversity and Ecosystem Dynamics', Department of Biology, Faculty of Science–Semlalia, Cadi Ayyad University, Avenue Prince Moulay Abdellah, 40000, Marrakech, Morocco*
[4]*Museum of Zoology, Senckenberg Dresden, A. B. Meyer Building, 01109, Dresden, Germany*
[5]*CIIMAR – Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Rua dos Bragas, n.289, 4050-123, Porto, Portugal*
[6]*Dipartimento di Scienze e Tecnologie Biologiche, Chimiche e Farmaceutiche (STEBICEF), Università di Palermo, Via Archirafi18, 90123, Palermo, Italy*

Received 27 April 2016; revised 30 May 2016; accepted for publication 30 May 2016

Quaternary climatic oscillations and geographic barriers have strongly influenced the distribution and diversification of thermophilic species occurring in the Mediterranean Basin. The Western Mediterranean pond turtle, *Mauremys leprosa*, is widely distributed throughout the Iberian Peninsula, southern France and most of the Maghreb region, with two subspecies currently recognized. In this work, we used 566 samples, including 259 new individuals, across the species range, and sequenced two mitochondrial markers (cytochrome *b* gene and control region; 163 samples in a concatenated mtDNA dataset) and one nuclear intron (R35; 23 samples representing all identified sublineages) to study the evolutionary history of *M. leprosa*. We combined phylogenetic methods and phylogeographic continuous diffusion models with spatial analysis. Our results (1) show a high level of genetic structure in Morocco originated during the Pleistocene; (2) reveal two independent population expansion waves from Morocco to Tunisia and to southern Europe, which later expanded throughout the Iberian Peninsula, and (3) identify several secondary contact zones in Morocco. Our study also sheds new light on the role of geographical features (Moroccan mountains ranges and the Strait of Gibraltar) and Pleistocene climatic oscillations in shaping genetic diversity and structure of *M. leprosa*, and underlines the importance of the Maghreb as a differentiation centre harbouring distinct glacial refugia. © 2016 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2016, **119**, 943–959.

KEYWORDS: climatic oscillations – colonization – Moroccan mountains – Strait of Gibraltar.

INTRODUCTION

The Mediterranean Basin harbours a high degree of species richness and endemism (Myers *et al.*, 2000), which has been mostly associated with a combination of geological and climatic events. In the Western Mediterranean, diversification events in different taxonomic groups can be matched with tectonic movements that led to the separation of the Balearic Islands, Sardinia and Corsica and the split of the Rif-Betic mountain range (e.g. Martínez-Solano *et al.*, 2004; Magri *et al.*, 2007; Bidegaray-Batista &

*Corresponding author. E-mail: guillermo.velo@cibio.up.pt

Arnedo, 2011; Miraldo *et al.*, 2011). Furthermore, the late Miocene closure of the Strait of Gibraltar resulted in the desiccation of the Mediterranean Basin (Krijgsman *et al.*, 1999; Duggen *et al.*, 2003), which in turn produced land bridges that connected the European and African continents acting as migration corridors for terrestrial biota. The refilling of the Mediterranean Sea *c.* 5.3 Mya reinstated a separation between many European and North African taxa (e.g. Veith *et al.*, 2004; Sousa *et al.*, 2012; Velo-Antón *et al.*, 2012). Then, diversification events continued during the Plio-Pleistocene when the effects of the Milankovitch climatic oscillations became more frequent and intense, especially during the Pleistocene (Dynesius & Jansson, 2000; Hewitt, 2000).

Low temperatures like those experienced during the glacial periods have a high impact on a species thermoregulation processes, easily inducing population extinction and range shifts. Many species likely survived glacial cold periods in climatically suited refugia, usually located in the southern European peninsulas (Taberlet *et al.*, 1998; Hewitt, 1999, 2000) and North Africa (Husemann *et al.*, 2014). In Pleistocene warm periods (interglacials and interstadials), range expansion occurred and led to a genetic exchange between previously isolated conspecific populations (Taberlet *et al.*, 1998; Hewitt, 1999, 2004). This combination and sequence of palaeogeographic and climatic events have rendered distinct phylogeographic patterns for Mediterranean species whose response to the above events could differ due to their ecological constraints, such as dispersal abilities and ectothermal physiology.

The Western Mediterranean pond turtle *Mauremys leprosa* (Schweigger, 1812) is one of most widely distributed and abundant reptile species in the Iberian Peninsula (with a few scattered populations in the southern France) and in the Maghreb (from western Morocco to Libya; Fig. 1). The species occurrence in the Iberian Peninsula dates back to the Middle Pleistocene, or perhaps even to the Pliocene, based on fossil records (de Lapparent de Broin & Antunes, 2000; de Lapparent de Broin, 2001; Fèlix *et al.*, 2006; de Soler *et al.*, 2012), while the oldest North African records originate from the Ruscinian (5.4–3.4 Mya; de Lapparent de Broin, 2000).

Two genetic main lineages have been identified: *Mauremys leprosa leprosa* inhabits the Iberian Peninsula and northern Morocco, and *M. l. saharica* occurs in southern Morocco, easternmost Algeria and Tunisia (Fritz *et al.*, 2006). The wide and continuous distribution of *M. leprosa*, together with its ectothermic condition, makes it a good model to assess the effects of climatic oscillations and geographic barriers in shaping the evolutionary history in Western Mediterranean taxa.

In this work, we assess current patterns of genetic diversity and structure in *M. leprosa*. Using a comprehensive sampling covering most of the species' range, we combine information from mitochondrial DNA sequences (cytochrome b gene and control region fragment) and from a nuclear intron (R35) to examine its evolutionary history within an integrative framework of phylogenetic methods, phylogeographic continuous diffusion models and spatial analyses. We aim to: (1) identify the geographic diversification centre of the species; (2) determine the genetic structure within *M. leprosa* and estimate the origin of major lineages; (3) evaluate the effect of major barriers, such as the Atlas Mountains and the Strait of Gibraltar, in shaping current genetic diversity; and (4) identify contact zones of mitochondrial genetic lineages.

MATERIAL AND METHODS

Sampling and sequencing

Blood or tissue (tail tips) samples of 259 *Mauremys leprosa* were collected across the Iberian Peninsula and the Maghreb region and preserved in absolute ethanol (Fig. 1, Supporting Information, Table S1). Genomic DNA was extracted with commercial kits (Easyspin; innuPREP DNA Mini Kit; innuPREP Blood DNA Mini Kit), following the manufacturer's protocol, and the lysis period was extended to overnight to enhance the extraction.

Two mitochondrial DNA fragments, the cytochrome b gene (Cytb) and the control region (D-loop), and one nuclear intron (R35), were targeted in this study. The former fragment was selected in order to include the available sequences from previous studies (Fritz *et al.*, 2006; Palacios *et al.*, 2015). For a better resolution of phylogenetic relationships, a subset of the collected samples were also sequenced for D-loop and R35 fragments. To accomplish the study goals, four datasets were used in subsequent analyses: (A) a concatenated mtDNA dataset including 163 samples covering the species distribution (with the exception of the French populations), was used to identify phylogenetic clades and their relationships, to assess genetic diversity and divergence between lineages, as well as for demographic and phylogeographic reconstruction analyses; (B) a comprehensive dataset of 566 Cytb sequences (including all available GenBank sequences, Supporting Information, Table S1) was used to identify the mtDNA sublineage of all available sequences; (C) a subset of the latter ($N = 497$, excluding samples from Algeria and Tunisia) was used to visualize geographic patterns of

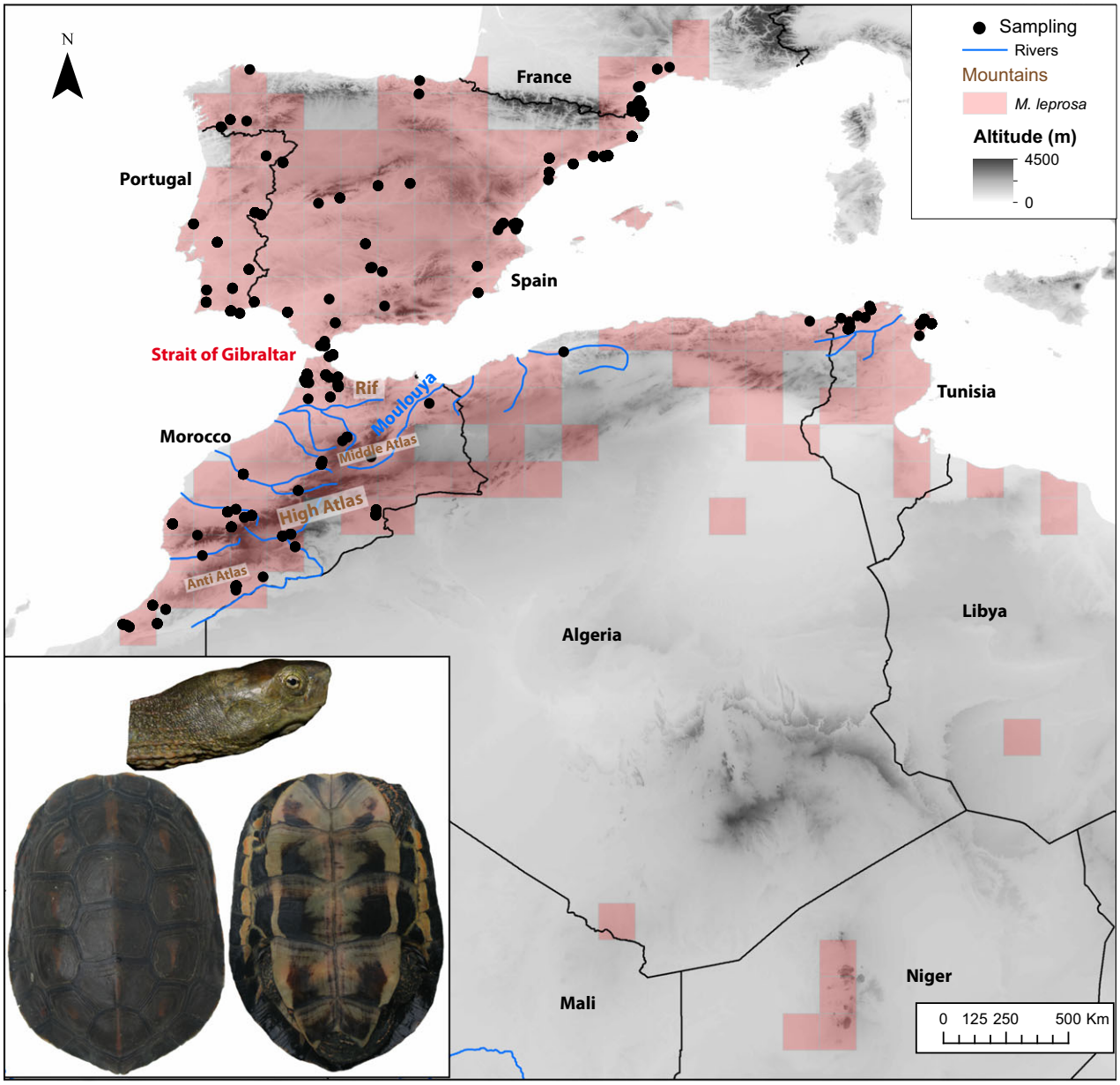

Figure 1. Distribution of the Mediterranean pond turtle, *Mauremys leprosa* (coarse grid size, 50 × 50 km) based on herpetological atlases from France, Morocco, Portugal and Spain. The distribution range from the remaining North African countries is based on Iverson (1992). Black circles indicate sampling locations. Note that distributional records from the Aïr Mountains (Niger), southern Algeria and southern Libya are old and perhaps a misidentification with *Pelomedusa* spp. Main North African rivers and mountains are also depicted. The inset figure shows the head, carapace and plastron of *M. leprosa*.

genetic diversity, and (D) a small representation of each mtDNA lineage (eight and 15 samples per lineage, representing all sublineages) was obtained for R35 to evaluate previous phylogenetic relationships obtained with the concatenated mtDNA dataset.

Cytb was amplified using the primers mt-a-neu (Lenk & Wink, 1997) and H-15909 (Lenk *et al.*, 1999). The D-loop fragment was amplified with specific primers designed with OLIGOEXPLORER 1.2 (http://www.genelink.com/tools/gl-oe.asp): MauMut_tThr.for (forward, 5′-ACTCTAGTAGCTTAACCCAT-3′) and MauMut_Dloop_2.rev (reverse, 5′-TCAGTTTAGTTGC TCTCGGA-3′). PCR reactions were conducted in a final volume of 10 μL from which 5 μL (1 U μL^{-1})

corresponded to MyTaq Mix (Bioline), 0.4 μM for each primer, 3.2 μL of ultra-pure water, and 1 μL of DNA (50–100 ng μL^{-1}). PCRs were carried out in a BioRad T100 Thermal Cycler with the following procedures: initial denaturation of 95 °C for 10 min; 10 cycles at 95 °C for 30 s, 55–50 °C (for Cytb) and 59–54 °C (for D-loop) decreasing 0.5 °C per cycle for 20 s (for Cytb) and 30 s (for D-loop); 72 °C for 1 min; 30 cycles of 95 °C for 30 s, 50 °C for 20 s (for Cytb) and 54 °C for 30 s (for D-loop); 72 °C for 1 min; and a final elongation step at 72 °C for 10 min. R35 was amplified with primers R35 Ex1 and R35 Ex2 (Fujita *et al.*, 2004) and reactions were conducted in a final volume of 10 μL from which 5 μL (1 U μL^{-1}) corresponded to MyTaq Mix (Bioline), 0.4 μM of each primer, 3.2 μL of ultra-pure water, and 1 μL of DNA (50–100 ng μL^{-1}). PCR was carried out in a BioRad T100 Thermal Cycler with the following procedure: initial denaturation at 95 °C for 10 min; 35 cycles of 95 °C for 30 s, 60°C for 90 s; 72 °C for 2 min; and a final elongation step of 72 °C for 10 min.

Remaining primers and nucleotides were removed from PCR products using ExoSAP (USB ExoSAP-IT PCR Product Cleanup, Affymetrix) following the manufacturer's instructions. All amplified fragments were sequenced in both directions using the forward and reverse PCR primers in a BioRad T100 Thermal Cycler with the BigDye Terminator v3.1 Cycle Sequencing Kits (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. Cycle sequencing-PCR products were purified using Sephadex (GE Healthcare, Munich, Germany). Finally, all sequencing reactions were run on an ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA, USA). All the obtained chromatograms were verified, aligned, and corrected by eye using Geneious Pro v4.8.5 (http://www.geneious.com/). Due to an indel presence and heterozygous positions in the R35 intron, sequences were phased in DnaSP v5.10 (Librado & Rozas, 2009) to reconstruct haplotypes. The MUSCLE plug-in implemented in Geneious Pro v4.8.5 (http://www.geneious.com/) was used for the alignments that were later manually checked.

Phylogenetic analyses of mtDNA

A phylogenetic tree was constructed using samples for which both mtDNA markers were available (A). jMODELTEST v.2.1.4 (Darriba *et al.*, 2012) was used to test for the best fitting model of nucleotide substitution for our dataset, under Akaike information criteria correction (AICc; TIM + I + G for both markers). Bayesian gene tree reconstructions were conducted in BEAST v1.7.5 (Drummond *et al.*, 2012). The dataset was partitioned by fragment and run under the corresponding evolutionary model. Markov Chain Monte Carlo (MCMC) analyses were run in three independent runs of 10^6 generations with four chains, with a sampling frequency of 1000 generations. Parameter convergence was verified by examining the effective sample sizes (ESSs) using TRACER v1.6. After discarding 10% trees as burn-in, the remaining trees were used to obtain the subsequent maximum clade credibility summary tree with posterior probabilities for each node using TREEANNOTATOR. FIGTREE v. 1.4.1 was used for tree visualization.

Estimation of divergence times

A mean substitution rate of 0.00626 substitutions/site/million years suggested for the mitochondrial DNA in turtles (Lourenço *et al.*, 2013) was used to estimate the time to the most recent common ancestor (TMRCA) of mtDNA lineages, with a standard deviation of 0.0002. The GTR + I + G nucleotide substitution model was included in the analysis, and a uncorrelated lognormal clock and a coalescence constant size model were used as tree priors. Divergence time estimates were calculated in BEAST using the concatenated mtDNA dataset (dataset A) but the substitution rate was only applied to the Cytb partition. Three independent runs of 10^6 generations were performed, sampling every 1000 generations and 10% of the trees were discarded as burn-in. Convergence was confirmed examining the likelihood and posterior trace plots of the independent runs with TRACER, and effective sample sizes of the parameters were above 1000.

Spatial diffusion modelling using continuous phylogeography

Bayesian implementation of the spatial diffusion approach was used to infer geographic origins of main mitochondrial clades in *M. leprosa*. This approach uses the geospatial coordinates of DNA sequences to reconstruct the spatial diffusion of organisms across a continuous landscape through time, and it has been proven useful to identify the location of lineages' most recent common ancestors (MRCA). A Cauchy Relaxed Random Walk (RRW) diffusion model (Lemey *et al.*, 2010) was used in BEAST v1.7.5 including unique haplotypes of the concatenated mtDNA dataset and their corresponding geographic coordinates. To create unique coordinates for individuals from identical localities (duplicate coordinates can confound the RRW model) a random jitter was applied. A coalescent tree with constant population size was used as the demographic prior. MCMC chains were run for 100

million generations, sampling every 10 000 generations, with the first 10% removed as burn-in. TRACER was used to evaluate the parameter convergence by examining the ESSs values and TREEANNOTATOR was used to summarize the posterior sample of trees and to create a maximum clade credibility tree. The surfaces output representing uncertainty for the diffusion process was formatted as KML using SPREAD (Bielejec *et al.*, 2011).

Genetic diversity and demographic analyses

DnaSP v5.10 was used to assess the number of segregating sites (S), and the nucleotide (π) and haplotype diversity (H_d) in the concatenated mtDNA dataset (A). Three tests of selective neutrality (Tajima's D, R_2, and Fu's Fs) were performed in DnaSP v5.10 to infer signatures of demographic expansion in each lineage, using 10 000 bootstrap replicates.

Three haplotype networks were constructed to visualize haplotypes relationships within *M. leprosa*: (1) a haplotype network based on the concatenated mtDNA dataset (dataset A) using statistical parsimony implemented in TCS v1.21 (Clement, Posada & Crandall, 2000); (2) a neighbour-net network based on concatenated mtDNA dataset (dataset A) using uncorrected patristic distance (p-distance) and bootstrap analysis with 1000 replicates in SPLITSTREE v4.6 (Huson & Bryant, 2006); and (3) a haplotype network based on R35 sequences using statistical parsimony implemented in TCS (dataset D).

Uncorrected p-distances were calculated for each mitochondrial fragment in MEGA v5.1 (Tamura *et al.*, 2011) to estimate genetic divergence between the six sublineages (see Results).

Phylogeographic interpolation

To visualize geographic patterns of genetic structure, the phylogeographic data (concatenated mtDNA dataset) of *M. leprosa* were interpolated using a modified method of kriging implemented in PHYLIN R package (Tarroso, Velo-Antón & Carvalho, 2015). As spatial interpolations were done with concatenated mtDNA dataset, samples from France, central Algeria and other disperse sampling points throughout the Iberian Peninsula and Morocco were not included in this analyses. The spatial dependence of the data (i.e. distance from where the samples can be considered independent) was first investigated using a total variogram with default values and fitted with a spherical model. Then, the occurrence, following the 0.95 probability, of the six sublineages obtained within the *M. leprosa* phylogeny was predicted, and potential contact zones between the six sublineages were identified using a regular sampling and a single threshold (hs = 0.8).

To visualize geographic patterns of genetic diversity, nucleotide diversity (π) was spatially interpolated following a kriging interpolation method (Oliver & Webster, 1990). Only Cytb data (N = 497; dataset C) was used for these interpolation and samples from Tunisia and Algeria were excluded in order to avoid the sampling gap across Algeria that would induce artefacts in the analysis. To identify the spatial distribution of genetic diversity, π values were calculated by pooling samples contained in a buffer with a radius of 0.449 decimal degrees (~50 km), which represent the potential genetic diversity of the original point. Nucleotide diversity values were then interpolated by generating a continuous surface with a kriging interpolation method (Oliver & Webster, 1990), implemented in the 'Geostatistical Analyst' extension of GIS ArcMap 9.3 (ESRI, 2006). The resulting raster was then reclassified into five classes, using Natural Jenks as the division criteria.

RESULTS

Phylogenetic analyses

We obtained 75 unique haplotypes in our concatenated mtDNA data set (163 samples) from an aligned matrix of 1769 base pairs (bp) (933 bp for Cytb and 862 bp of D-loop). Bayesian inference based on mtDNA data showed a largely resolved phylogeny with two major lineages (BPP > 0.95): lineages A and B, which correspond to the nominal subspecies *Mauremys leprosa leprosa* and *M. l. saharica* respectively (Fritz *et al.*, 2006; Fig. 2). Due to the lack of phenotypic data and limited nuclear data needed to accurately evaluate the distribution of both subspecies, we only refer to lineage A (*M. l. leprosa*) and lineage B (*M. l. saharica*) hereafter. We found lineage A haplotypes distributed across the Iberian Peninsula, southern France and north of the Atlas Mountains in Morocco, while lineage B haplotypes were confined to North Africa (i.e. north and south of the High Atlas in Morocco, Algeria and northern Tunisia) and in southern France due to recent introductions (see Palacios *et al.*, 2015). Further, we identified three well supported sublineages within both main lineages (BPP > 0.95; Fig. 2). Within A: sublineage A1 occurs in southwestern Morocco, north of the High Atlas; A2 mostly across central-north Morocco, but also south of the High Atlas, while A3 occurs in the Middle Atlas and the Rif Mountains, and throughout the Iberian Peninsula. Two haplotypes (A3-9 and A3-15) occur on both sides of the Strait of Gibraltar (Supporting Information,

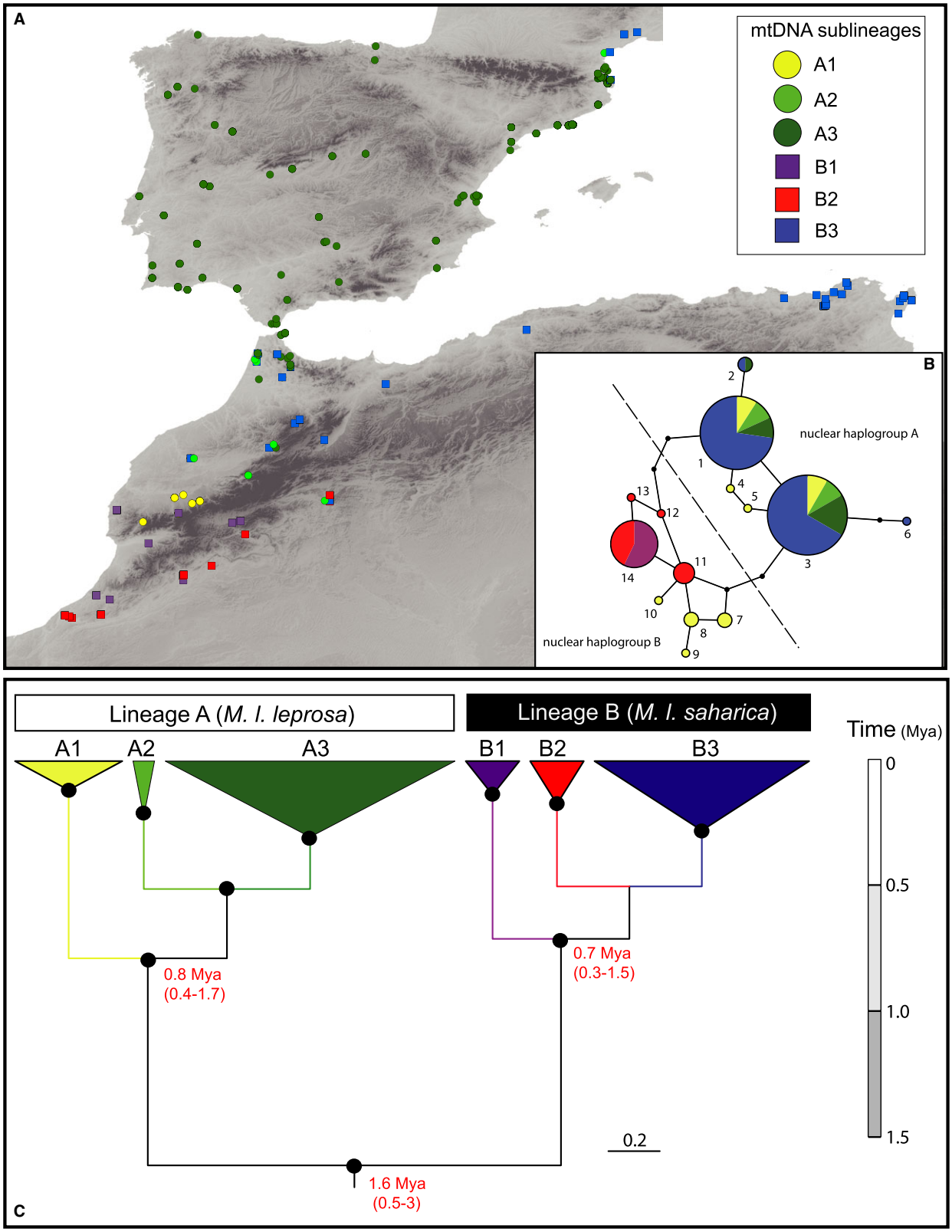
A
mtDNA sublineages
A1
A2
A3
B1
B2
B3
B
nuclear haplogroup A
nuclear haplogroup B
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Lineage A (*M. l. leprosa*)
Lineage B (*M. l. saharica*)
A1
A2
A3
B1
B2
B3
Time (Mya)
0
0.5
1.0
1.5
0.8 Mya
(0.4-1.7)
0.7 Mya
(0.3-1.5)
1.6 Mya
(0.5-3)
0.2
C

Figure 2. A, spatial distribution of the mtDNA lineages and sublineages. Circles and squares correspond to lineages A and B respectively. Note that sublineage B3 in southern France is the result of human introductions from Morocco (see Palacios *et al.*, 2015). B, haplotype network for R35 sequences inferred by TCS under the 95% parsimony threshold. The size of the 19 haplotypes are proportional to its frequency and lines represent mutational steps separating observed haplotypes. C, Bayesian consensus phylogram based on mtDNA data (Cytb and D-loop). Times to most recent common ancestor for the split between and within lineages are written in red (TMRCA; 95% highest posterior density interval). The 0.2 length corresponds to 0.2% of sequence divergence. Black circles indicate high nodal support (posterior probability > 0.95). Both symbol shapes and colours are concordant with those used for Figures 3 and 5.

Table S1). For lineage B, two sublineages (B1 and B2) are admixed across the south of the High Atlas, although B1 is mostly distributed between the High Atlas and Anti Atlas, as well north of the High Atlas. The third sublineage, B3, ranges from the Rif and Middle Atlas to Tunisia (Fig. 2), and is also found in the southeastern slope of the High Atlas. Sequence divergences (uncorrected *p*-distance for Cytb) between the two main lineages range from 0.9% to 1.2% for Cytb and 2.7% to 3.6% for D-loop, and high genetic divergence values are also found between lineages within both subspecies (Table 1).

Assessment of divergence times using BEAST estimates the time to the MRCA for lineages A and B to the Early Pleistocene (mean = 1.6; 95% HPD = 3–0.5 Myr), and the time of the MRCA for each lineage to the Early–Middle Pleistocene (lineage A, mean = 0.8; 95% HPD = 1.7–0.4 Myr; lineage B, mean = 0.7; 95% HPD = 1.5–0.3 Myr) (Fig. 2). According to this analysis, sublineages diverged during the Late Pleistocene (Fig. 2). However, we should bear in mind that these dates represent the coalescence time of the different mtDNA haplotypes, and thus the above lineages could have diverged at a much more recent time.

Spatial diffusion modelling

The continuous phylogeographic reconstructions show the root of both lineages in Morocco. The geographic origin of lineage A is located in the Western Rif while lineage B is distributed in the Atlas and Middle Atlas Mountains during Early Pleistocene (Fig. 4A). Subsequent climatic changes from glacial to interglacial stages likely allowed northwards expansions for both lineages, with the colonization of lineage A into the Iberian Peninsula, while lineage B expanded to the Rif during the Middle Pleistocene (Fig. 4B). More recently, in the Late Pleistocene or during the Holocene, lineage A expanded across the Iberian Peninsula while lineage B followed an eastward expansion, resulting in the colonization of the northeastern extreme of its range (Fig. 4C).

The R35 intron alignment contained 23 sequences with a maximum length of 969 bp that resulted in 14 haplotypes (Supporting Information, Table S1 and Fig. 2). There were 12 variable sites in all samples analyzed (six parsimony-informative sites and six insertions/deletions of 1–3 bp length). Two main nDNA haplogroups (A and B) separated by three mutations could be identified (Fig. 2, Supporting

Table 1. Genetic distances between mtDNA sublineages using the concatenated mtDNA dataset

Sublineage	A1	A2	A3	B1	B2	B3
A1	**0.3 (± 0.1)/ 0.1 (± 0.1)**	1.6 (± 0.4)	1.7 (± 0.4)	3.3 (± 0.6)	3.1 (± 0.6)	2.7 (± 0.5)
A2	0.4 (± 0.2)	**0.5 (± 0.2)/ 0.1 (± 0.1)**	0.7 (± 0.2)	3.6 (± 0.6)	3 (± 0.6)	2.8 (± 0.5)
A3	0.3 (± 0.1)	0.4 (± 0.2)	**0.1 (± 0.1)/ 0.1 (± 0)**	3.5 (± 0.6)	2.9 (± 0.6)	2.7 (± 0.5)
B1	0.9 (± 0.3)	1.1 (± 0.3)	1 (± 0.3)	**0.4 (± 0.1)/ 0.1 (± 0)**	2.3 (± 0.5)	1.7 (± 0.4)
B2	0.9 (± 0.3)	1.2 (± 0.3)	1.1 (± 0.3)	0.6 (± 0.2)	**0.3 (± 0.1)/ 0.1 (± 0)**	1.4 (± 0.4)
B3	0.9 (± 0.3)	1.2 (± 0.4)	1.1 (± 0.3)	0.6 (± 0.2)	0.4 (± 0.2)	**0.2 (± 0.1)/ 0.1 (± 0)**

Mean uncorrected *p*-distances and standard deviations for are shown above and below the diagonal respectively (mean and SE; in percentages). On the diagonal (shown in bold): mean uncorrected *p*-distance within each sublineage.

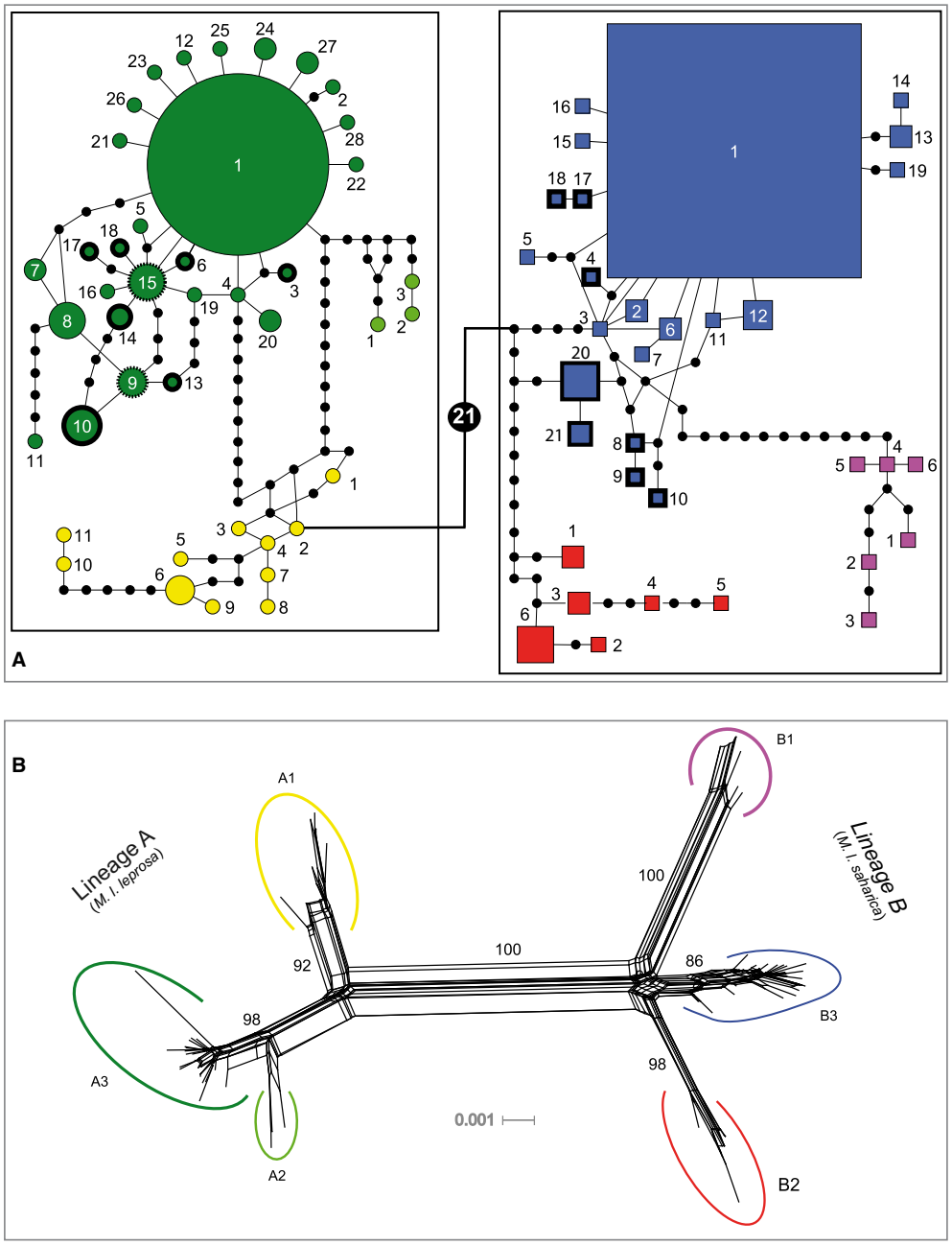

Figure 3. A, haplotype networks inferred for the concatenated mtDNA dataset in TCS under the 95% parsimony threshold. Circles and squares correspond to lineages A and B respectively. The size of each haplotype is proportional to its frequency and numbers identify the haplotype of each sublineage. Lines represent mutational steps separating observed haplotypes. Bold haplotype outline corresponds to haplotypes only found in Morocco, while dashed outline corresponds to haplotypes found in Morocco and Iberian Peninsula (for sublineages A3, in green, and B3, in blue). B, mitochondrial neighbour-net networking inferred in SplitsTree. Scale bar represents 1% sequence divergence while numbers correspond to bootstrap values. Both symbol shapes and colours are concordant with those used for Figures 3 and 5.

Information, Fig. S2). The first haplogroup includes a total of six haplotypes, with two main haplotypes, 1 and 4, shared among mtDNA sublineages A1, A2, A3 and B3. The second nDNA haplogroup contains eight haplotypes, with one haplotype shared by individuals from B2 and B3 mtDNA sublineages, four unique haplotypes to the sublineage A1, and three haplotypes from B2.

Genetic diversity and demographic analyses

Both lineages showed similar values of genetic diversity, although nucleotide diversity (π) was higher in lineage B (Table 2). Within this lineage, the two sublineages from south of the High Atlas (B1 and B2) showed higher genetic diversity (π and H_d) compared to the one distributed across the Middle Atlas, the Rif and northeastern Maghreb (B3). Within lineage A, the highest genetic diversity is found in the two sublineages endemic to Morocco, with a much less variation in the lineage distributed in both continents (A3). When this sublineage A3 was divided into two groups (North Africa and the Iberian Peninsula) we found higher genetic diversity in North Africa than in the Iberian Peninsula (results not shown), despite the high number of samples analyzed from the Iberian Peninsula. The most widely distributed sublineages, A3 and B3, showed negative and significant values of Tajima's, R_2 and Fu's statistics (Table 2).

Parsimony analyses, for the concatenated dataset, in TCS using the 95% connection limit yielded independent haplotype networks for the two mtDNA lineages, which can be manually connected by 26 mutational steps (Fig. 3). All sublineages are also well separated from each other by 7–15 mutational steps. Sublineages A3 and B3 clearly show a star-like network which is characteristic of demographic expansion scenarios, with A3-1 widely spread across the Iberian Peninsula and B3-1 widely distributed in the easternmost Algeria and Tunisia. The SPLITSTREE network shows identical relations to the above described for all sublineages.

Phylogeographic interpolation

The spherical model selected in PHYLIN had a good fit to the total variogram (Supporting Information, Fig. S3), which indicates a pattern of isolation-by distance, with small genetic differences at short distances and stabilization of the semi-variance at larger distances. The interpolated genetic distances surface using the regular sampling approach detected an abrupt change in the Atlas Mountains in North Africa (results not shown). The potential contact zones as represented by the average probability of the presence of multiple sublineages by using a single threshold on the mtDNA phylogenetic tree were identified at different regions in Morocco (Fig. 5A). In particular, the High Atlas is shown as both a partial barrier and a contact zone, while contact zones between lineages were found at both sides of this mountain chain, being the northern Morocco, the Middle Atlas and the southeastern slope of the High Atlas as the most evident contact zones where at least three sublineages occur in sympatry. The large contact zone area detected in Algeria is likely an artefact, resulting from a lack of sampling points in this country.

As for the genetic diversity, the Iberian Peninsula shows a homogeneous surface reflecting the lack of genetic divergence across these populations (Fig. 5B). The kriging interpolation produced a continuous surface of nucleotide diversity that clearly shows the highest genetic diversity in North Africa, particularly in the Rif, Middle Atlas and the High Atlas (π ranges from ~0.009 to ~0.0056; Fig. 5B). Within the Iberian Peninsula, the southern region presents moderate levels of genetic diversity that is reduced towards north of the Iberian Peninsula (π ranges from ~0.0001 to ~0.0004).

DISCUSSION

Overall, our study unveils a high level of genetic structure within *Mauremys leprosa* in Morocco originated during the Pleistocene, highlighting the importance of Western Maghreb as a differentiation centre and North African refugia. It also reveals two expansion waves from Morocco to Tunisia and to the Iberian Peninsula and identifies sympatric and syntopic areas between mtDNA lineages and sublineages in northern Morocco.

North African diversification

A wide distribution gap in the central Mediterranean region separates *M. leprosa* from its western Palaearctic congeners, *M. caspica* and *M. rivulata*, which are distributed in the southeastern Balkans and the Near and the Middle East (Fritz *et al.*, 2008). This distribution suggests an old and allopatric divergence between *M. leprosa* and their congeners. A previous phylogeographic study by Fritz *et al.* (2006) identified two clades of *M. leprosa* in Morocco, *M. l. leprosa* and *M. l. saharica*, of which the former was also found in Europe, suggesting a North African origin for the species. However, the sparse geographical sampling and the use of only a single mtDNA marker precluded a thorough evaluation of the evolutionary history of *M. leprosa*. Our study considerably increases our understanding with respect to the diversification patterns and biogeographic barriers of *M. leprosa*. In particular, our study reveals a pattern of genetic structuring and high levels of genetic diversity in Morocco.

The fact that Moroccan populations harbour all *M. leprosa* lineages and the highest diversity values, supports the hypothesis of a northwestern Maghreb origin for the species and a later colonization of the Iberian Peninsula and northeastern Maghreb. Contrary to the observed patterns reported for many

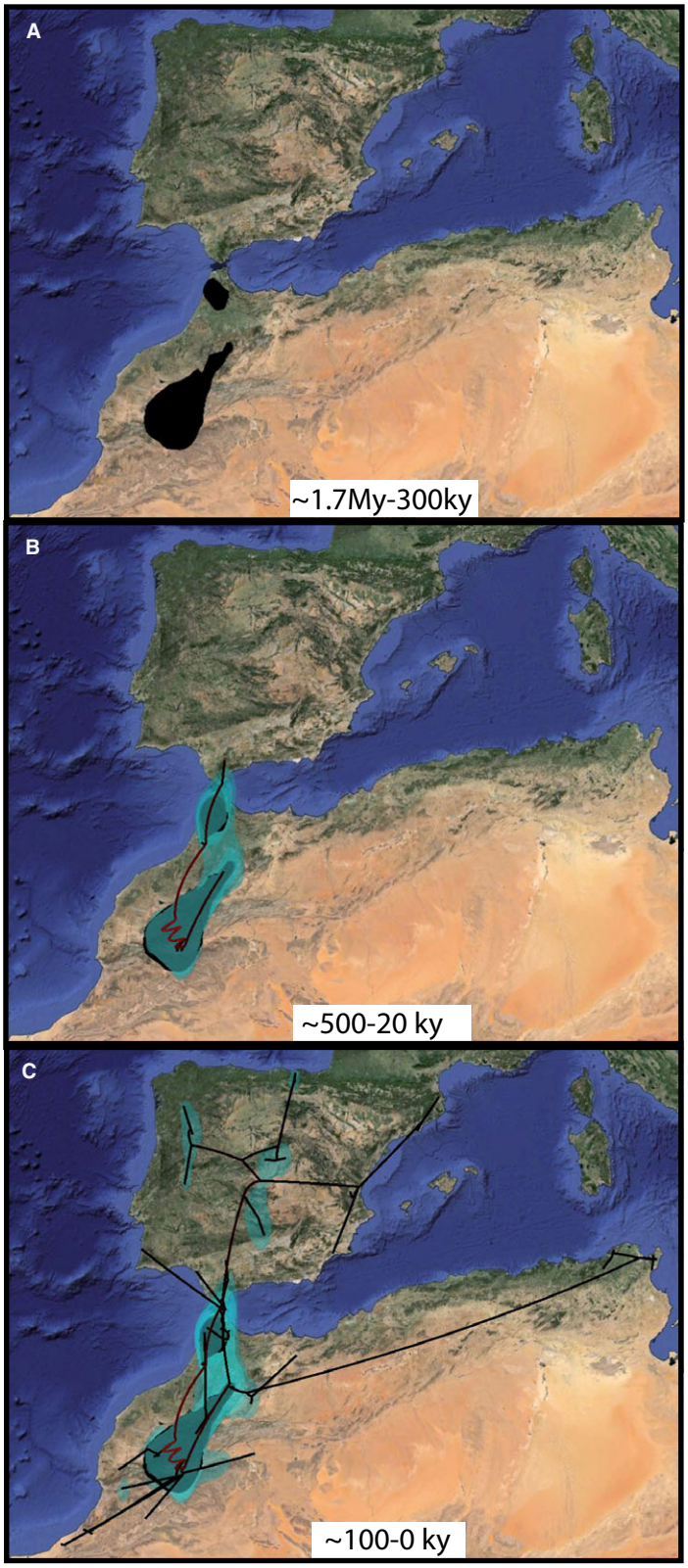
A
~1.7My-300ky
B
~500-20 ky
C
~100-0 ky

Figure 4. Continuous diffusion phylogeographic reconstruction in *M. leprosa* inferred from concatenated mtDNA dataset. The dispersal dynamics for the species through time (approximate estimated time) is represented for three important biogeographic events: (A) the split between the two main mtDNA lineages; (B) the expansion northwards and the colonization of the Iberian Peninsula from northern Morocco; and (C) the subsequent expansion of the species throughout the Iberian Peninsula and eastwards to Tunisia. The polygons in the map represent the 80% HPD intervals for the location of each node of the sampled genealogies, a measure of the uncertainty of the estimated location. The black–green gradient of the polygons represents the relative age of the dispersal events.

Table 2. Genetic diversity values and demographic estimates measured for the different lineages and sublineages of *Mauremys leprosa* using the concatenated mtDNA dataset

Group	N	π	S	H_n	H_d	R_2	D	F_S
Sublineage A1	13	0.0023	13	11	0.92	0.13	−0.13	**−3.81**
Sublineage A2	3	0.0030	8	3	1	NA	NA	NA
Sublineage A3	67	0.0010	29	21	0.75	**0.03**	**−21.44**	**−15.48**
Lineage A	83	0.0036	52	35	0.83	0.09	−11.36	**−9.77**
Sublineage B1	6	0.0022	9	6	1	0.14	−0.11	**−2.69**
Sublineage B2	11	0.0016	9	6	0.69	0.15	−0.28	1.34
Sublineage B3	63	0.0012	27	19	0.55	**0.03**	**−2.09**	**−10.76**
Lineage B	80	0.0043	58	33	0.71	0.06	−12.27	−4.80
ALL	163	0.0115	106	68	0.87			

Significant results for D and Fs shown in bold ($P < 0.01$). Lineages with low sampling size could not be used in demographic analyses and are represented as NA.

N, sample size; π, nucleotide diversity; S, number of polymorphic sites; H_n, number of haplotypes; H_d, haplotype diversity; R_2, Ramos-Osins and Rozas; D, Tajima's D; Fs, Fu's Fs; NA, not available.

taxa throughout the Maghreb region that show deep genetic divergence (e.g. Cosson *et al.*, 2005; Fritz *et al.*, 2009; Rato, Carranza & Harris, 2012; Velo-Antón *et al.*, 2012; Stuckas *et al.*, 2014; Lalis *et al.*, 2015; Nicolas *et al.*, 2015; Ben Hassine *et al.*, 2016), our study shows that diversification events in *M. leprosa* were restricted to Morocco, where all present sublineages occur. The genetic pattern observed in *M. leprosa* partially coincides with the phylogeographic pattern of *Buthus* scorpions (Sousa *et al.*, 2012), revealing cryptic diversity in the northwestern Maghreb region.

The role of climate and geographic barriers in North Africa

The high elevation (up to 4167 m) and large west–east extent (2500 km) of the High Atlas in Morocco has served as a major barrier to dispersal for North African species. This was previously evaluated in reptile species (e.g. Brown, Suárez & Pestano, 2002; Fritz *et al.*, 2006; Rato *et al.*, 2012). Our results support the hypothesis that the High Atlas functioned primarily as a north–south barrier to contact between the two main lineages (early Pleistocene), and later as a source of diversity and lineage divergence. Likewise most phylogeographic studies in this region, the combination of topographic heterogeneity and climatic oscillations during the Pleistocene are thought to be the main drivers promoting genetic differentiation among allopatric populations isolated in several climatic refugia (see Husemann *et al.*, 2014). During interglacial periods, population expansions of both mtDNA lineages allowed to circumvent these mountains and led to the establishment of several secondary contact zones and sympatric areas throughout Morocco. However, the nDNA dataset partially disagrees with the genetic structure inferred with mtDNA, as samples recovered in the mitochondrial sublineage B3 shared all nuclear haplotypes with mitochondrial lineage A, and nuclear haplotypes from mtDNA lineage A1 are closely related to the ones obtained from mtDNA sublineages (B1 and B2). With the data at hand, an scenario of mtDNA introgression north of the Atlas Mountains seem to be a more plausible explanation than incomplete lineage sorting in R35 to explain the observed geographic variation (Supporting Information, Fig. S2). The fact that B3 is mostly distributed in northern Atlas may facilitate a model of mtDNA capture mediated by male-biased dispersal following range shifts due to climatic fluctuations. Moreover, the nuclear network identifies samples of mtDNA sublineage A1 in both nuclear

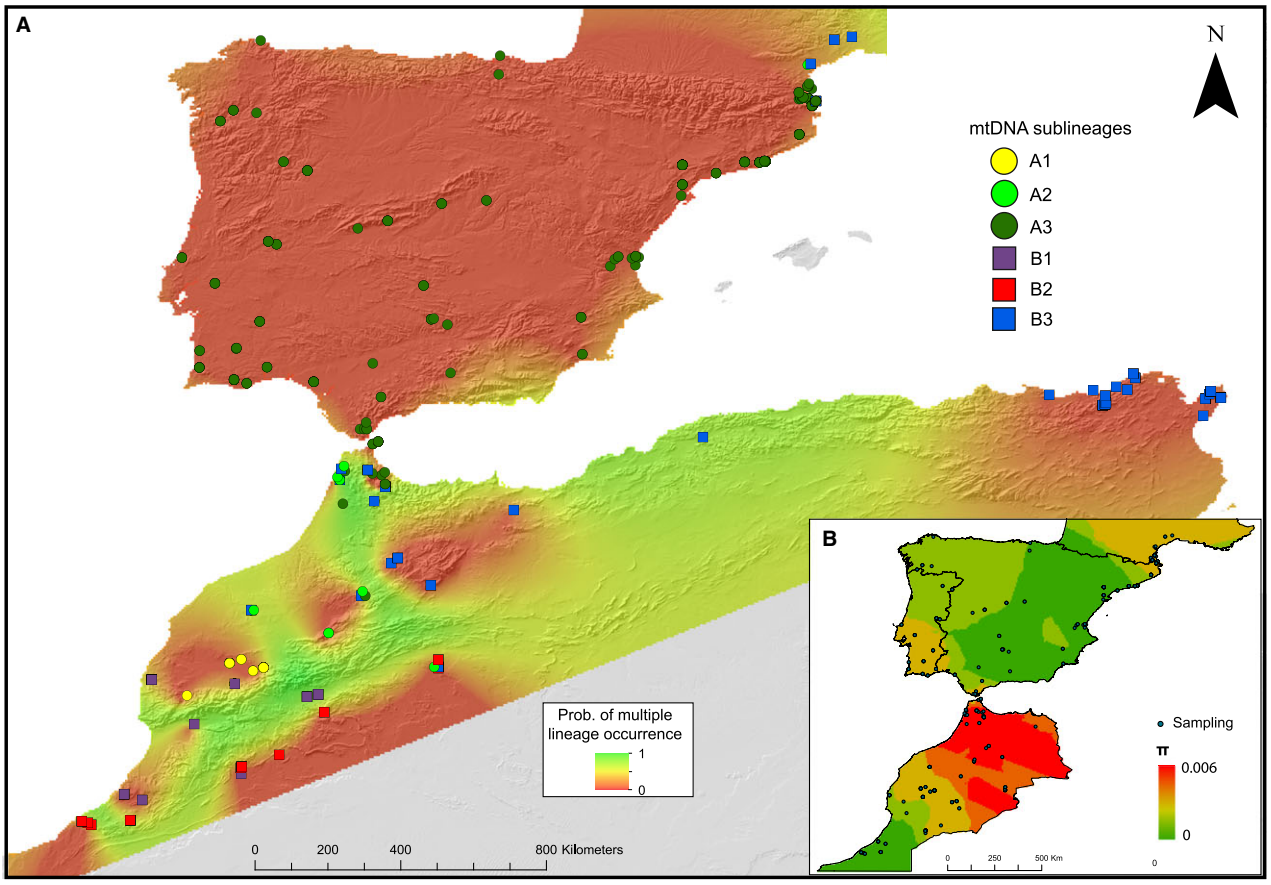

Figure 5. A, predicted occurrence of contact zones (green) between the six mtDNA sublineages of *Mauremys leprosa* using phylogeographic interpolation implemented in PHYLIN. Note that the large contact zone depicted in Algeria is an artifact of the sampling gap in this region. B, spatial interpolation of genetic diversity based on Cytb nucleotide diversity (π) and reclassified into five different classes. Tunisia and Algeria were excluded in order to avoid the sampling gap across this region that would induce artefacts on the analysis. Darker green and red colours correspond to areas with the lowest and highest genetic diversity respectively.

haplogroups, suggesting an introgression pattern through migrations across an eastern corridor in the Atlas Mountains. However, further samples from unstudied areas and further genetic data would be needed to scrutinize this biogeographic hypothesis.

Further genetic sub-structuring at mtDNA level is present north (sublineages A1, A2, A3, B3) and south of the High Atlas and the Middle Atlas Mountains (sublineages A2, B1, B2, B3). The spatial distribution of some of these sublineages are partially similar to the ones described in other taxa (e.g. *Buthus*, Sousa *et al.*, 2012; *Tarentola*, Rato *et al.*, 2012), which could have emerged during the Pleistocene through vicariance when river valleys were temporally transgressed by the sea or due to other orographic structures at both sides of the Atlas Mountains.

Despite the large sampling gap in Algeria, we suggest that specimens from Tunisia and Algeria are related to specimens from eastern Morocco (the Rif and Middle Atlas), as they group in the widespread sublineage B3. To the best of our knowledge, this phylogeographic pattern has never been observed in other phylogeographic studies on vertebrates with relatively continuous distribution throughout the Maghreb. Conversely, putative barriers, such as the arid valley of Moulouya, acted as an effective barrier to population dispersal of several taxa (e.g. Fritz *et al.*, 2009; Vences *et al.*, 2014; Lalis *et al.*, 2015; Nicolas *et al.*, 2015). Our results show a series of evidence that supports a demographic expansion of sublineage B3 from Morocco eastwards to Tunisia. Moroccan populations of B3 show higher diversity than Tunisian populations, and our demographic (both star-like network and Tajima's; R_2 and Fu's statistics), together with phylogeographic diffusion model analyses, suggest a west-east expansion from a glacial refugium located in an area between the Middle and High Atlas. However, the lack of a

comprehensive sampling in Algeria prevents from an accurate assessment of population expansion and directionality within this sublineage.

The role of the Strait of Gibraltar as a permeable biogeographic barrier

The Strait of Gibraltar is considered one of the major barriers to dispersal for the Mediterranean taxa (e.g. Paulo *et al.*, 2008; Velo-Antón *et al.*, 2012). However, in the past decades, several studies revealed that many species managed to cross it, regardless of the direction (see Husemann *et al.*, 2014 and references therein; Velo-Antón *et al.*, 2015), but only a few rapidly expanded across the Iberian Peninsula (e.g. Carranza, Arnold & Pleguezuelos, 2006; Recuero *et al.*, 2007; Velo-Antón, García-París & Cordero, 2008).

The lack of genetic differentiation between Iberian and North African populations of *M. leprosa* and the recent origin of sublineage A3 in the Iberian Peninsula suggest either a relatively very recent colonization in Europe or that constant gene flow across the sea strait prevented the evolution of distinct lineages. Several fossil records identified with *M. leprosa* dated to the Middle Pleistocene, or perhaps even to the Pliocene (de Lapparent de Broin & Antunes, 2000; de Lapparent de Broin, 2001; Fèlix *et al.*, 2006; de Soler *et al.*, 2012). However, the assignment of fossil records to present species is not always straightforward. When old fossil remains are commonly fragmented, an accurate identification between the present species and a lineage or species that went extinct becomes challenging (e.g. Pedall *et al.*, 2011). The old Pliocene fossils found in northeastern Iberian Peninsula (de Soler *et al.*, 2012) are an example of dubious records for *M. leprosa*, and thus we cannot unambiguously determine the presence of this species in the Iberian Peninsula. Conversely, the fossils of Middle Pleistocene age (de Lapparent de Broin, 2001) do not support a very recent arrival in the Iberian Peninsula but massive transoceanic gene flow has been reported for the eastern Mediterranean species *M. rivulata* (Vamberger *et al.*, 2014), making this also likely for *M. leprosa*. Yet, a series of evidences (i.e. lack of genetic differentiation at both sides of the strait and low genetic diversity in Iberian populations) supports the hypothesis that the Strait of Gibraltar is permeable to dispersal and gene flow between African and Iberian populations. This scenario matches with the lower sea level during Pleistocene glacials that facilitated gene flow across the Strait of Gibraltar. The separation between Africa and Europe could decrease in glacial periods to only 5 km (Brandt, Alpers & Backhaus, 1996; Zazo, 1999). Due to haplotype sharing, however, between continents (A3–15 and A3–9), we could not rule out the hypothesis of human-mediated translocations of *M. leprosa* between both sides of the strait. Human-mediated introductions might explain biogeographic patterns of chameleons (Paulo *et al.*, 2002) and hylids (Recuero *et al.*, 2007), and similar events can certainly influence turtle migrations in the recent past (Velo-Antón *et al.*, 2011b). However, dating very recent divergence events is challenging when trying to distinguish Late Pleistocene colonizations from anthropogenic introductions (e.g. Graciá *et al.*, 2013).

Rapid south–north colonization wave across the Iberian Peninsula

The most parsimonious scenario to explain the mismatch of old fossil records and observed genetic patterns, implies an ancient (Pliocene or earlier) invasion of Europe, followed by a massive extinction of the species in this region due to Pleistocene climatic oscillations, and a later re-colonization from North Africa and subsequent rapid population expansion throughout Iberia.

Interestingly, the co-distributed terrapin distributed in the Iberian Peninsula and northern Maghreb (*Emys orbicularis occidentalis*) shows a similar pattern of re-colonization from North Africa (Stuckas *et al.*, 2014; Velo-Antón *et al.*, 2015), and a rapid population expansion throughout the Iberian Peninsula that caused a reduction of genetic diversity from southern to northern populations (Velo-Antón *et al.*, 2008), and an increase of carapace scute anomalies through a series of bottleneck effects (Velo-Antón, Becker & Cordero-Rivera, 2011a). Thus, it appears that climatic conditions occurred in the Iberian Peninsula during glacial phases were too harsh for both terrapins, which likely caused large extinctions of *Emys* and *Mauremys* populations. However, the palaeoecological similarities of southern Europe and North Africa would point to other unidentified factors that would better explain the vanishing of these thermophilic species in the Iberian Peninsula, where many other reptiles remained in suitable refugia during Pleistocene climatic fluctuations.

Contact zones within *Mauremys leprosa*

Our study shows several secondary contact zones between the two mtDNA lineages and syntopic localities where mitochondrial haplotypes of both lineages currently co-occur, with a large contact zone in the Rif and the Middle Atlas. In a previous study (Fritz *et al.*, 2006), the presence of a single specimen of lineage B in northern Morocco led these authors to consider anthropogenic introduction as reptile trade, which commonly occurred in terrapins (e.g.

Velo-Antón *et al.*, 2011b). However, the high number of terrapins with haplotypes of lineage B found in this region points to a natural occurrence in the Rif and the Middle Atlas, probably as the result of range expansions during the Pleistocene. We found haplotypes of both lineages in syntopy in several Moroccan localities of the Rif (Tazia, Tetouan, Fifi, Zoumi) and the Middle Atlas (Sidi Mimoun). These populations should be re-examined using nuclear markers to find out whether extensive gene flow occurs between lineages or whether they evolved isolation mechanisms (prezygotic or postzygotic barriers).

Other contact and sympatric zones between sublineages occur throughout Morocco, mostly north of the High Atlas. Although the ancestor of lineage B likely originated south of the High Atlas, it remains unclear where the widespread sublineage B3 differentiated from its sister clades. Our results point to an eastern Moroccan origin for this lineage that could take place either on the southeastern High Atlas or north of this massive mountain chain, in the Middle Atlas or the Rif Mountains, facilitating the posterior expansion eastwards Maghreb, and also to the northwestern Rif.

Overall, our study unveils a remarkable diversification within *Mauremys leprosa* and underlines the importance of the Maghreb as a differentiation centre and North African refugia. The combination of the high topographical complexity in Morocco, i.e. mountains acting as geographic barriers, and climatic oscillations occurred during the Pleistocene might lead to distinct glacial refugia at both sides of the Atlas Mountains, resulting in high levels of genetic structure due to allopatric differentiation. Subsequent population expansions during interglacial periods allowed two main population expansion waves from Morocco, with one sublineage colonizing southern Europe and later expanded throughout the Iberian Peninsula, and another sublineage that expanded eastwards to Tunisia. Sympatric and syntopic areas between the two most divergent lineages are also common in northern Morocco but further studies combining phenotypic and genetic data are needed to evaluate the distribution of both subspecies and to detect admixture patterns across the species range.

ACKNOWLEDGEMENTS

We especially thank A. Cordero-Rivera, C. Ayres and M. Casal for their help and support during some field work campaigns in the Iberian Peninsula. We also thank P. Segurado, M, Franch, G. Alarcos, A. Álvarez, V. Sancho, A. Acuña, J. Hidalgo, Matati, A. Portheault, A. Andreu, X. Buenetxea, J. Pleguezuelos, L. García-Cardenete, I. Martínez-Solano, E. Recuero, A. Egea, B. El Marnisi, S. Fahd and BIOSESERTS members for helping during field work or providing samples. S. Lopes and P. Tarroso provided technical support in the laboratory and using PHYLIN package, respectively. We also thank three anonymous reviewers for their helpful comments. This work was funded by Instituto de Estudios Ceutíes in 2012 (to GVA) and by LIFE-Trachemys (LIFE09/NAT/ES/000529; to JT). Additional sampling was funded by grants PGIDT01MAM37101PR and PGIDTI05RFO37101PR from Xunta de Galicia. GVA and JT were supported by Fundação para a Ciência e Tecnologia (FCT, Portugal: IF/01425/2014 and SFRH/BPD/27173/2006). Sampling permits were obtained from regional administrations in both the Iberian Peninsula and Morocco. The authors have declared that no competing interests exist.

REFERENCES

Ben Hassine J, Gutiérrez-Rodríguez J, Escoriza D, Martínez-Solano I. 2016. Inferring the roles of vicariance, climate and topography in population differentiation in *Salamandra algira* (Caudata, Salamandridae). *Journal of Zoological Systematics and Evolutionary Research* **54:** 116–126.

Bidegaray-Batista L, Arnedo MA. 2011. Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology* **11:** 317.

Bielejec F, Rambaut A, Suchard MA, Lemey P. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27:** 2910–2912.

Brandt P, Alpers W, Backhaus JO. 1996. Study of the generation and propagation of internal waves in the Strait of Gibraltar using a numerical model and synthetic aperture radar images of the European ERS 1 satellite. *Journal of Geophysical Research* **101:** 237–252.

Brown RP, Suárez NM, Pestano J. 2002. The Atlas mountains as a biogeographical divide in North-West Africa: evidence from mtDNA evolution in the Agamid lizard *Agama impalearis*. *Molecular Phylogenetics and Evolution* **24:** 324–332.

Carranza S, Arnold EN, Pleguezuelos JM. 2006. Phylogeny, biogeography, and evolution of two Mediterranean snakes, *Malpolon monspessulanus* and *Hemorrhois hippocrepis*(Squamata, Colubridae), using mtDNA sequences. *Molecular Phylogenetics and Evolution* **40:** 532–546.

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

Cosson JF, Hutterer R, Libois R, Sara M, Taberlet P, Vogel P. 2005. Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology* **14:** 1151–1162.

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

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Duggen S, Hoernle K, van den Bogaard P, Rüpke L, Morgan JP. 2003. Deep roots of the Messinian salinity crisis. *Nature* **422:** 602–605.

Dynesius M, Jansson R. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 9115–9120.

ESRI. 2006. *ArcGIS 9.2*. Environmental Systems Research Inst., Redlands, CA.

Fèlix J, Budó J, Capalleras X, Mascort R. 2006. The Fossil Register of the Genera *Testudo*, *Emys* and *Mauremys* of the Quaternary in Catalonia. *Chelonii* **4:** 47–51.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R. 2006. Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, *Mauremys leprosa* (Reptilia, Testudines, Geoemydidae). *Zoologica Scripta* **35:** 97–108.

Fritz U, Ayaz D, Buschbom J, Kami HG, Mazanaeva LF, Aloufi AA, Auer M, Rifai L, Šilić T, Hundsdörfer AK. 2008. Go east: phylogeographies of *Mauremys caspica* and *M. rivulata*–discordance of morphology, mitochondrial and nuclear genomic markers and rare hybridization. *Journal of Evolutionary Biology* **21:** 527–540.

Fritz U, Harris DJ, Fahd S, Rouag R, Martínez EG, Casalduero AG, Široký P, Kalboussi M, Jdeidi TB, Hundsdörfer AK. 2009. Mitochondrial phylogeography of *Testudo graeca* in the Western Mediterranean: old complex divergence in North Africa and recent arrival in Europe. *Amphibia-Reptilia* **30:** 63–80.

Fujita MK, Engstrom TN, Starkey DE, Shaffer HB. 2004. Turtle phylogeny: insights from a novel nuclear intron. *Molecular Phylogenetics and Evolution* **31:** 1031–1040.

Graciá E, Giménez A, Anadón JD, Harris DJ, Fritz U, Botella F. 2013. The uncertainty of Late Pleistocene range expansions in the western Mediterranean: a case study of the colonization of south-eastern Spain by the spur-thighed tortoise, *Testudo graeca*. *Journal of Biogeography* **40:** 323–334.

Hewitt G. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68:** 87–112.

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

Hewitt G. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London* **359:** 183–195.

Husemann M, Schmitt T, Zachos FE, Ulrich W, Habel JC. 2014. Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. *Journal of Biogeography* **41:** 81–94.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23:** 254–267.

Iverson JB. 1992. A revised checklist with distribution maps of the turtles of the world. Richmond, IN, USA: Iverson JP.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ. 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature* **400:** 652–655.

Lalis A, Leblois R, Stoetzel E, Benazzou T, Souttou K, Denys C, Nicolas V. 2015. Phylogeography and demographic history of Shaw's Jird (*Meriones shawii* complex) in North Africa. *Biological Journal of the Linnean Society* **118:** 262–279.

de Lapparent de Broin F. 2000. African chelonians from the Jurassic to the present: phases of development and preliminary catalogue of the fossil record. *Palaeontologia Africana* **36:** 43–82.

de Lapparent de Broin F. 2001. The European turtle fauna from the Triassic to the present. *Dumerilia* **4:** 155–217.

de Lapparent de Broin F, Antunes MT. 2000. Pleistocene chelonians from Gruta da Figueira Brava (Arrábida, Portugal). *Memórias da Academia das Ciências de Lisboa* **38:** 105–116.

Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution* **27:** 1877–1885.

Lenk P, Wink M. 1997. A RNA/RNA heteroduplex cleavage analysis to detect rare mutations in populations. *Molecular Ecology* **6:** 687–690.

Lenk P, Fritz U, Joger U, Wink M. 1999. Mitochondrial phylogeography of the European pond turtle, *Emys orbicularis* (Linnaeus 1758). *Molecular Ecology* **8:** 1911–1922.

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

Lourenço JM, Glémin S, Chiari Y, Galtier N. 2013. The determinants of the molecular substitution process in turtles. *Journal of Evolutionary Biology* **26:** 38–50.

Magri D, Fineschi S, Bellarosa R, Buonamici A, Sebastiani F, Schirone B, Simeone MC, Vendramin GG. 2007. The distribution of *Quercus suber* chloroplast haplotypes matches the palaeogeographical history of the western Mediterranean. *Molecular Ecology* **16:** 5259–5266.

Martínez-Solano I, Gonçalves HA, Arntzen JW, García-París M. 2004. Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography* **31:** 603–618.

Miraldo A, Hewitt GM, Paulo OS, Emerson BC. 2011. Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. *BMC Evolutionary Biology* **11:** 170.

Myers N, Mittermeier R, Mittermeier CG, da Fonseca G, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403:** 853–858.

Nicolas V, Mataame A, Crochet PA, Geniez P, Ohler A. 2015. Phylogeographic patterns in north African water frog *Pelophylax saharicus* (Anura: Ranidae). *Journal of Zoological Systematics and Evolutionary Research* **53:** 239–248.

Oliver MA, Webster R. 1990. Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems* **4:** 313–332.

Palacios C, Urrutia C, Knapp N, Quintana MF, Bertolero A, Simon G, Du Preez L, Verneau O. 2015. Demographic structure and genetic diversity of *Mauremys leprosa* in its northern range reveal new populations and a mixed origin. *Salamandra* **51:** 221–230.

Paulo OS, Pinto I, Bruford MW, Jordan WC, Nichols RA. 2002. The double origin of Iberian peninsular chameleons. *Biological Journal of the Linnean Society* **75:** 1–7.

Paulo OS, Pinheiro J, Miraldo A, Bruford MW, Jordan WC, Nichols RA. 2008. The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. *Molecular Ecology* **17:** 1535–1551.

Pedall I, Fritz U, Stuckas H, Valdeón A, Wink M. 2011. Gene flow across secondary contact zones of the *Emys orbicularis* complex in the Western Mediterranean and evidence for extinction and re-introduction of pond turtles on Corsica and Sardinia (Testudines: Emydidae). *Journal of Zoological Systematics and Evolutionary Research* **49:** 44–57.

Rato C, Carranza S, Harris DJ. 2012. Evolutionary history of the genus *Tarentola* (Gekkota: Phyllodactylidae) from the Mediterranean Basin, estimated using multilocus sequence data. *BMC Evolutionary Biology* **12:** 14.

Recuero E, Iraola A, Rubio X, Machordom A, García-París M. 2007. Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern. *Journal of Biogeography* **34:** 1207–1219.

de Soler BG, Vall-Llosera GC, Van der Made J, Oms O, Agustí J, Sala R, Blain H-A, Burjachs F, Claude J, García Catalán S, Riba D, Rosillo R. 2012. A new key locality for the Pliocene vertebrate record of Europe: the Camp dels Ninots maar (NE Spain). *Geologica Acta* **10:** 1–17.

Sousa P, Harris DJ, Froufe E, Meijden A. 2012. Phylogeographic patterns of Buthus scorpions (Scorpiones: Buthidae) in the Maghreb and South-Western Europe based on CO1 mtDNA sequences. *Journal of Zoology* **288:** 66–75.

Stuckas H, Velo-Antón G, Fahd S, Kalboussi M, Rouag R, Arculeo M, Marrone F, Sacco F, Vamberger M, Fritz U. 2014. Where are you from, stranger? The enigmatic biogeography of North African pond turtles (*Emys orbicularis*). *Organisms Diversity & Evolution* **14:** 295–306.

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

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.

Tarroso P, Velo-Antón G, Carvalho SB. 2015. PHYLIN: an R package for phylogeographic interpolation. *Molecular Ecology Resources* **15:** 349–357.

Vamberger M, Stuckas H, Ayaz D, Lymberakis P, Široký P, Fritz U. 2014. Massive transoceanic gene flow in a freshwater turtle (Testudines: Geoemydidae: *Mauremys rivulata*). *Zoologica Scripta* **43:** 313–322.

Veith M, Mayer C, Samraoui B, Barroso D, Bogaerts S. 2004. From Europe to Africa and vice versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography* **31:** 159–171.

Velo-Antón G, García-París M, Cordero RA. 2008. Patterns of nuclear and mitochondrial DNA variation in Iberian populations of *Emys orbicularis* (Emydidae): conservation implications. *Conservation Genetics* **9:** 1263–1274.

Velo-Antón G, Becker CG, Cordero-Rivera A. 2011a. Turtle carapace anomalies: the roles of genetic diversity and environment. *PLoS ONE* **6:** e18714.

Velo-Antón G, Wink M, Schneeweiß N, Fritz U. 2011b. Native or not? Tracing the origin of wild-caught and captive freshwater turtles in a threatened and widely distributed species (*Emys orbicularis*). *Conservation Genetics* **12:** 583–588.

Velo-Antón G, Godinho R, Harris DJ, Santos X, Martínez-Freiria F, Fahd S, Larbes S, Pleguezuelos JM, Brito JC. 2012. Deep evolutionary lineages in a Western Mediterranean snake (*Vipera latastei/monticola* group) and high genetic structuring in Southern Iberian populations. *Molecular Phylogenetics and Evolution* **65:** 965–973.

Velo-Antón G, Pereira P, Fahd S, Teixeira J, Fritz U. 2015. Out of Africa: did *Emys orbicularis occidentalis* cross the Strait of Gibraltar twice? *Amphibia-Reptilia* **36:** 133–140.

Vences M, de Pous P, Nicolas V, Díaz-Rodríguez J, Donaire D, Hugemann K, Hauswaldt JS, Amat F, Barnestein JA, Bogaerts S, Bouazza A. 2014. New insights on phylogeography and distribution of painted frogs (*Discoglossus*) in northern Africa and the Iberian Peninsula. *Amphibia-Reptilia* **35:** 305–320.

Zazo C. 1999. Interglacial sea levels. *Quaternary International* **55:** 101–113.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Distribution of the two nuclear haplogroups identified with R35 intron in a subset of samples from Iberian Peninsula and Morocco.
Figure S2. Total variogram with fitted model.
Table S1. Origin and genetic information for all samples used in this study.