

Genealogy of the nuclear β -fibrinogen intron 7 in *Lissotriton boscai* (Caudata, Salamandridae): concordance with mtDNA and implications for phylogeography and speciation

José Teixeira^{1,3}, Iñigo Martínez-Solano^{1,4,7}, David Buckley⁵, Pedro Tarroso¹, Mario García-París⁵, Nuno Ferrand^{1,2,6}

¹ CIBIO-InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos. Campus Agrário de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

² Departamento de Biologia, Faculdade de Ciências do Porto, 4099-002 Porto, Portugal.

³ CIIMAR - Centro Interdisciplinar de Investigação Marinha e Ambiental. Rua dos Bragas. 4169-007 Porto, Portugal

⁴ Ecology, Evolution and Development Group, Department of Wetland Ecology, Estación Biológica de Doñana (EBD-CSIC), Avenida Américo Vespucio, s/n, 41092 Sevilla, Spain

⁵ Museo Nacional de Ciencias Naturales (MNCN-CSIC), c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain

⁶ Department of Zoology, University of Johannesburg, Kingsway Campus, Johannesburg, 2006 South Africa

⁷ E-mail: inigomsolano@gmail.com

Key words: β -fibint7, genetic landscape shape, Iberian Peninsula, mitochondrial DNA, nuclear DNA

Abstract

The power of phylogeographic inference resides in its ability to integrate information from multiple sources in an iterative hypothesis-testing framework. In this paper, we build upon previous mtDNA-based hypotheses about the evolutionary history of the Iberian newt *Lissotriton boscai* using sequences of the highly variable nuclear β -fibrinogen intron 7. In addition to the nuclear sequences, we produced new mtDNA data across the species range to delineate contact zones and test the congruence between nuclear and mitochondrial datasets at the same level of spatial organization. Through a combination of phylogenetic, phylogeographic continuous diffusion, and genetic landscape modelling analyses, we infer the evolutionary history of the species. We found notable congruence between nuclear and mtDNA datasets, which confirms deep and consistent differentiation between two major lineages that originated in the Miocene. Additionally, we found a new nuclear haplogroup with no mtDNA counterpart, roughly circumscribed to the Iberian Sistema Central mountains, and extensive areas of nuclear admixture across mtDNA lineages. We describe potential historical dispersal routes from an ancestral hypothetical refugium in the western end of the Sistema Central in central Portugal and highlight how deep phylogeographic breaks do not necessarily indicate cryptic speciation events.

Contents

Introduction	193
Material and methods	195
<i>Sampling design</i>	195
<i>PCR amplification and sequencing</i>	195
<i>Genetic diversity</i>	196
<i>Phylogenetic, phylogeographic, and demographic analyses</i>	196

<i>Geographical analyses</i>	197
<i>Comparative analyses between mitochondrial and nuclear DNA lineage distributions</i>	199
Results	199
<i>Genetic diversity</i>	199
<i>Phylogenetic and phylogeographic analyses</i>	201
<i>Geographical analyses</i>	203
<i>Comparative analyses between mitochondrial and nuclear DNA lineage distributions</i>	203
Discussion	204
<i>Nuclear DNA phylogeography of <i>Lissotriton boscai</i></i>	204
<i>Implications for species formation</i>	206
Acknowledgements	207
References	207
Appendices	212

Introduction

The field of phylogeography, originally defined as a bridge between phylogenetics and population genetics (Avice *et al.*, 1987), has rapidly grown to become one of the most successful disciplines in evolutionary biology (Hickerson *et al.*, 2010), with a prominent role in our understanding of the origin and conservation of biodiversity (Dufresnes *et al.*, 2013). Currently, some of the major trends in the study of phylogeographic patterns include: 1) analysis of multilocus datasets (with a rapid shift from the early reliance on mtDNA to high-throughput next-generation sequencing, McCormack *et al.*, 2013; Reitzel *et al.*, 2013), 2) strong emphasis on hypothesis testing (Forester *et al.*, 2013), and 3) use of spatially-explicit methods of data visualisation and

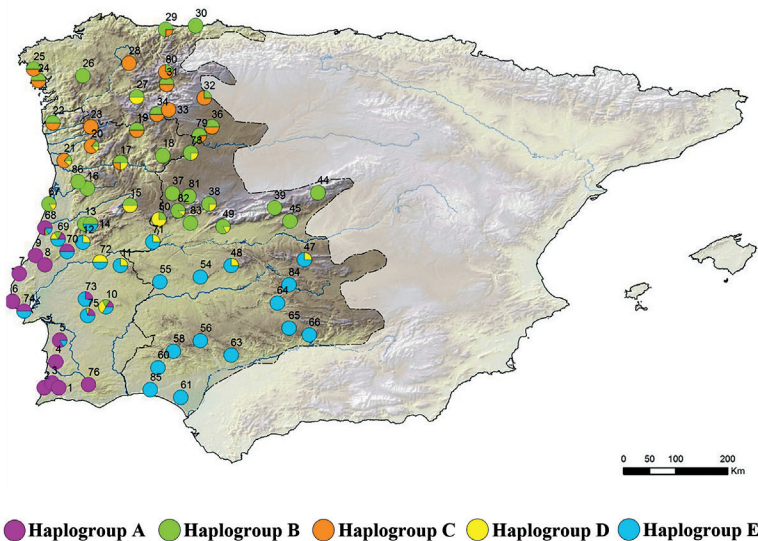
analysis (Alvarado-Serrano and Knowles, 2014).

One of the keys to the field's success is proper consideration of the iterative nature of phylogeographic research: the continuous feedback between data from related fields that allows new hypotheses to be successively tested and refined (Buckley, 2009). For instance, nuclear DNA data are used to test and revise hypotheses based on mtDNA, and GIS data are used to visualise and compare genetic patterns across markers, populations, species and regions and to formulate new hypotheses. There are many possibilities for the integration of genetic and geographic data, yet the use of geo-

referenced data in phylogeography is still relatively limited (Chan *et al.*, 2011) and methods are still not unified.

Amphibians are suitable model systems in phylogeography because their low vagility and philopatry usually result in strong genetic differentiation, even at small geographic scales (Martínez-Solano and Lawson, 2009; Milá *et al.*, 2010; Zhao *et al.*, 2013). Many studies have documented varying levels of divergence of inter- and intraspecific phylogroups, sometimes (especially in temperate regions) showing genetic signatures of past population fluctuations related to major

a) β -fibint7



b) mtDNA

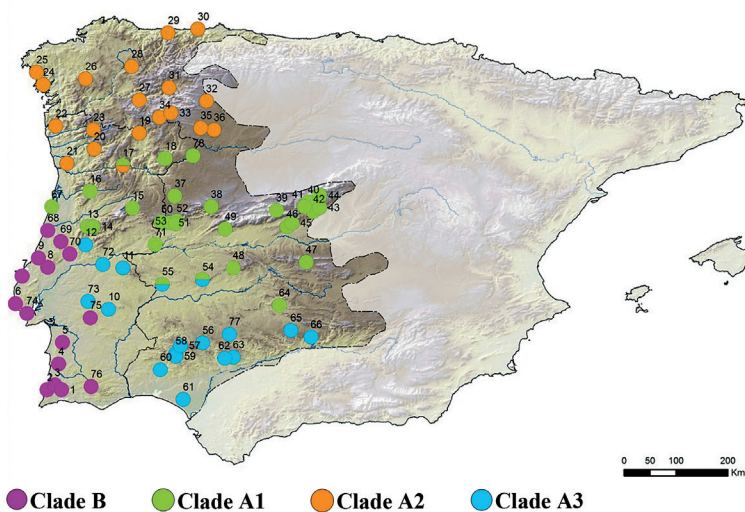


Fig. 1. Range of *Lissotriton boscai* in the Iberian Peninsula (adapted from Díaz-Paniagua, 2002 and Teixeira, 2008), showing the distribution of β -fibint7 haplogroups (a) and mtDNA clades (b). Numbers represent localities in Appendix I. Colour codes as in Fig. 2.

climatic changes in the past, mostly during the Pleistocene glaciations (Gonçalves *et al.*, 2009; Nuñez *et al.*, 2011; Wielstra *et al.*, 2013).

Southern European peninsulas are well known as hotspots for specific and subspecific diversity across taxonomic groups, including amphibians (Gómez and Lunt, 2007; Habel *et al.*, 2010). In particular, the Iberian Peninsula is rich in amphibian endemics, both at the specific and intraspecific levels. For instance, strong phylogeographical structure has been described for the morphologically homogeneous Iberian endemic Bosca's newt, *Lissotriton boscai* (Lataste in Tourneville, 1879), in which mtDNA lineages diverged during the Miocene (Martínez-Solano *et al.*, 2006). Previous analyses have focused on the relative timing of events and historical demography, but some major questions about its evolutionary history are still pending. For instance: are patterns in mitochondrial DNA mirrored in the nuclear genome? What were the most important historical barriers to gene flow and potential dispersal routes for the different haplogroups/lineages? Where were the main refugia located during climatically unfavourable periods? To address some of these issues, we analyse new intron data (the highly variable nuclear beta-fibrinogen intron 7, β -fibint7 (Johnson and Clayton, 2000; Pritchitko and Moore, 2003; Godinho *et al.*, 2006; Sequeira *et al.*, 2006), plus an increased mtDNA dataset, with several new populations from the vicinity of inferred contact zones between the previously identified mtDNA lineages, in a geographically comprehensive sample (86 populations). We investigate the evolutionary history of *L. boscai* in an integrative framework, combining phylogenetic methods and phylogeographic continuous diffusion models with GIS-based data analysis and visualisation tools (Genetic Landscape Shape and Monmonier's Maximum Difference Algorithm). Specifically, we assess the congruence of mtDNA and nuclear DNA genealogies in an explicit spatial context and advance new hypotheses regarding the process of diversification in *L. boscai*. We identify a new refugial area in the mountains of the Sistema Central, and evaluate the possibility that one or more cryptic genetic entities are present within the species.

Material and methods

Sampling design

We screened β -fibint7 haplotypic variation in 149 individuals of *Lissotriton boscai* from 73 localities across

Western Iberia (Fig. 1a, Appendix I). We used tissue samples from 98 individuals from 54 localities available from the study of Martínez-Solano *et al.* (2006). In order to improve the geographical coverage of the species range at a finer scale, especially along the previously identified contact zones between the main mtDNA lineages, we sampled 51 additional individuals from 19 new localities (populations 68 to 86, Fig. 1a, Appendix I). The sampling was designed to obtain a balanced and thorough coverage of the species range (Fig. 1).

Tissue samples consisted of tail tips of larvae and toe or tail clips of adults. All individuals were released in the wild after tissue collection. The number of samples per location varied between one and seven (Appendix I).

PCR amplification and sequencing

DNA was extracted from ethanol-preserved tissues using a standard proteinase K protocol (Sambrook *et al.*, 1989). The entire intron 7 and incomplete exon 8 from the β -fibrinogen nuclear gene were amplified using a combination of two pairs of primers: FIBX7 (5' – GGA GAN AAC AGN ACN ATG ACA ATN CAC – 3') and FIBX8 (5' – ATC TNC CAT TAG GNT TGG CTG CAT GGC – 3'), followed by the internal primers BFXF (5' – CAG YAC TTT YGA YAG AGA CAA YGA TGG – 3') and BFXR (5' – TTG TAC CAC CAK CCA CCA CCR TCT TC – 3') (Sequeira *et al.*, 2006). For all specimens, a nested two-step amplification procedure was used following the conditions described in Sequeira *et al.* (2006).

Sequences were determined using one or both (eight individuals were sequenced for both) of the primers BFXF and BFXR in a 3100 Applied Biosystems (Carlsbad, CA, USA) DNA sequencing analyser. After removing the 5' flanking exon sequences, the 575 base pair (bp) intron 7 and 39 bp exon 8 sequences were aligned manually in BioEdit 5.0.9 (Hall, 1999). The identity of the nuclear intron was confirmed by comparing the flanking exon sequence with GenBank sequences. For several of the samples, sequencing revealed that the sample consisted of two alleles that differed at one or more sites. In these cases, we inferred haplotype phase using the software PHASE 1.1 (Stephens *et al.*, 2001) with haplotype probabilities assessed through 1,000 replicates.

We used the published mitochondrial sequences of the 103 individuals analysed in Martínez-Solano *et al.* (2006). Additional examination of mtDNA haplotypic variation was performed in 22 samples from 12 new

localities (populations 67 to 78, Appendix I, Fig. 1b) from the contact zones between the main previously identified mitochondrial lineages. Amplification and sequencing of the *nad4* and control region markers were performed as described in Martínez-Solano et al. (2006). All new sequences have been deposited in GenBank under the accession numbers KP265980–KP266321.

Genetic diversity

In the β -fibint7 dataset, the number of haplotypes (h), nucleotide diversity index (π , Nei 1987), haplotype diversity (Hd) and the average number of nucleotide differences (k) were used as estimates of the genetic diversity of the total dataset and within the main haplogroups detected. Genetic diversity calculations were performed excluding recombinant sequences for the global data set and within distinctive haplogroups using DnaSP v4.10.9 (Rozas et al., 2003). We used the software PHIPACK (Bruen et al., 2006) to evaluate the presence of recombination in β -fibint7 sequences using the pairwise homoplasy index. This method has been shown by simulation studies to be less sensitive than other available statistics to falsely infer recombination when levels of recurrent mutation are high (Bruen et al., 2006). We used a permutation test (1000 permutations) to estimate the p -values for the null hypothesis of no recombination.

Intraspecific median-joining haplotype networks were constructed for both β -fibint7 and the combined *nad4* and control region sequences in Network version 4.2 (Bandelt et al., 1999).

Phylogenetic, phylogeographic, and demographic analyses

Phylogenetic analyses of the β -fibint7 data were performed on a reduced dataset including single haplotypes (and excluding recombinant haplotypes). *Lissotriton italicus* and *L. helveticus* were used as outgroups, with sequences downloaded from GenBank (accession numbers: DQ823400, Sequeira et al., 2006; and JF812960 and JF812963, Recuero and García-París, 2011). Analyses were run in Mr. Bayes v3.2.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). We sampled across the GTR substitution model parameter space (Huelsenbeck et al., 2004) by using the 'nst=mixed rates=gamma' option. Here and hereafter, we used the gamma parameter (+G) to describe rate heterogeneity among sites. The proportion of invariable

sites parameter (I) was not included in the models, given the correlation of the two parameters and the difficulty of optimising them simultaneously (e.g. Yang, 2006; Stamatakis, 2014). Analyses were run for 100 million generations, with a sampling frequency of 10,000. Of the resulting 10,000 trees, the first 10% were discarded as burn in.

The evolutionary and demographic history of *L. boscai* was further investigated through the use of phylogeographic continuous diffusion models and skyline plots as implemented in BEAST 1.8.0 (Drummond and Rambaut, 2007; Lemey et al., 2010). These methods produce a spatially explicit visualisation of the evolutionary history of the species based on inference of gene genealogies and reconstruction of ancestral clade ranges using information from the sequences and the latitude and longitude coordinates of sampling locations, together with the reconstruction through time of the effective population sizes. First, we obtained an independent time estimate for the major population split in *L. boscai*. We selected one haplotype of each of the two major mtDNA lineages in *L. boscai* and added sequences of other *Lissotriton* species (*L. italicus* and *L. helveticus*), and the related genera *Triturus* (*T. marmoratus*, *T. pygmaeus*, *T. cristatus*, *T. dobrogicus*, *T. karelinii*), *Ichthyosaura* (*I. alpestris*), *Calotriton* (*C. asper*), *Ommatotriton* (*O. vittatus*), and *Neurergus* (*N. strauchii*) available in GenBank (Appendix II) (Zhang et al., 2008; Wiens et al., 2011; Marjanovic and Laurin, 2013). The mtDNA dataset was divided in four partitions (by gene and codon position), as indicated by results of PartitionFinder v1.1 (Lanfear et al., 2012). Substitution models for each partition were also selected based on results of this program: p1 (*nad4*, first positions+ tRNA-Hist): TrN+G; p2 (*nad4*, second positions): HKY+I+G, implemented as HKY+G, see above; p3 (*nad4*, third positions): TrN+G; and p4 (control region): HKY+G. Relaxed molecular clocks (uncorrelated LogNormal) were unlinked across partitions but tree topologies were linked. Substitution rates for each partition were estimated from the data, using a diffuse gamma prior for the parameter (shape=0.01, scale=100, offset=0). As calibration points for the relaxed molecular clocks, we specified divergence times for three nodes based on the fossil record and previously published estimates of divergence times. First, we specified a maximum age for the root of the trees of 41–61 mya [lognormal distribution; mean=51 mya, log(stdev)=0.1]. Although wide, this range includes previously published estimates based on two different calibration strategies (Zhang et al., 2008; Wiens et al., 2011). We

then added two calibration points based on the fossil record. The dates of the fossils were used as minimum ages for clades, incorporating these dates as priors in the analysis. We used lognormal distributions for these priors, with an offset equal to the age of the fossil, a mean of 5 myr, and a standard deviation of 1 (see Wiens *et al.*, 2011). In this way, the age of the clade is placed between 1 myr and 15 myr (mean roughly 3 myr) before the appearance of the fossil, a conservative approach that incorporates the uncertainty about the maximum age for the clade. We calibrated the origin of *Triturus* based on the oldest known fossil for the genus [offset: 23.8 mya; mean=5; stdev=1; resulting in a range between 24–40 mya] (Estes, 1981; Zhang *et al.*, 2008; Wiens *et al.*, 2011; Marjanovic and Laurin, 2013). We also calibrated the origin of *Lissotriton* based on some of the oldest known fossils for the genus (offset=17 mya; mean=5; stdev=1; resulting in a range between 17–32 mya) (Rage and Bailon, 2005; Böhme, 2010; Martín and Sanchiz, 2013). As a speciation model, we specified the Yule coalescent prior (Gernhard, 2008). The time estimates for the split between the two major lineages in *L. boscai* obtained in this analysis were then used in subsequent analyses as a prior for the root of the trees. Using multiple calibration points is sometimes problematic, since the information provided by the different priors may be antagonistic and hard to reconcile given the data, leading to unrealistic results. Therefore, we ran the same analysis but with an empty alignment, sampling only from the prior, to be sure that the priors for the calibrated nodes were working properly, and to analyse the relative contribution of priors and data to the parameter posterior estimates.

We then reconstructed the demographic history of *L. boscai* through a multilocus Extended Bayesian Skyline Plot (EBSP) as implemented in BEAST1.8.0 (Heled and Drummond, 2008). We ran the analysis on the full dataset (excluding recombinant haplotypes for the nuclear dataset), excluding outgroups, and the data partitioned in four partitions according to Partition-Finder results: p1 (*nad4*, first positions): K80+G; p2 (*nad4*, second positions, tRNA, D-Loop): K81uf+I, implemented as K81uf+G; p3 (*nad4*, third positions): TN93; p4 (β -fibint7): HKY+I, implemented as HKY+G. Models and clocks were unlinked across partitions, and the topologies for p1, p2, and p3 were linked (mitochondrial genealogy). We used a strict clock for all the partitions in this analysis, with a diffuse prior for the clock rates (gamma distribution, shape=0.01; scale=100), and 10 categories for the parameter +G. For the genealogical reconstruction, we

used the coalescent prior Extended Bayesian Skyline Plot (EBSP) and an uninformative gamma prior (shape=1; scale=500) for the parameter Demographic.popMean. The analysis was calibrated and rooted specifying a prior for the age of the root derived from the results of the previous analysis (lognormal distribution, mean=9 mya, st. dev.=0.17).

We also calculated Tajima's *D* (Tajima, 1989) and Fu's *F_s* (Fu, 1997) statistics for the β -fibint7 dataset in DnaSP, to test for evidence for population expansion in the total dataset and for each of the main haplogroups identified. Analyses were run for the complete dataset and in a reduced dataset including a single sequence per sampling locality in order to conform to the expectations of the unstructured coalescent (see Wakeley, 2004); since results were similar, we report only those from the analysis of the full dataset (see Table 1).

Finally, we ran independent phylogeographic continuous diffusion analyses for the two markers (Lemey *et al.*, 2010). We included the geographic coordinates of each sample in the analysis. For these geographic traits, we assumed gamma-distributed variation in geographic diffusion rates across branches (Lemey *et al.*, 2010). We also included a jitter module in the xml file to add random variation to locations for sequences with the exact same geographic coordinates (random noise uniformly drawn from a window size of 0.1). The analyses were again calibrated and rooted assigning a temporal prior to the root of the trees consistent with previous results (lognormal prior, mean=9 mya, st. dev.=0.17), estimating substitution rates (prior: gamma distribution, shape=0.01; scale=100; offset=0), and using the EBSP as coalescent prior for the tree topologies. The resulting annotated trees were analysed in SPREAD v1.0.5 (Bielejec *et al.*, 2011) using the option of the 'Time slicer' to generate 80% HPD polygons for the inferred geographic location through time of every node of the tree.

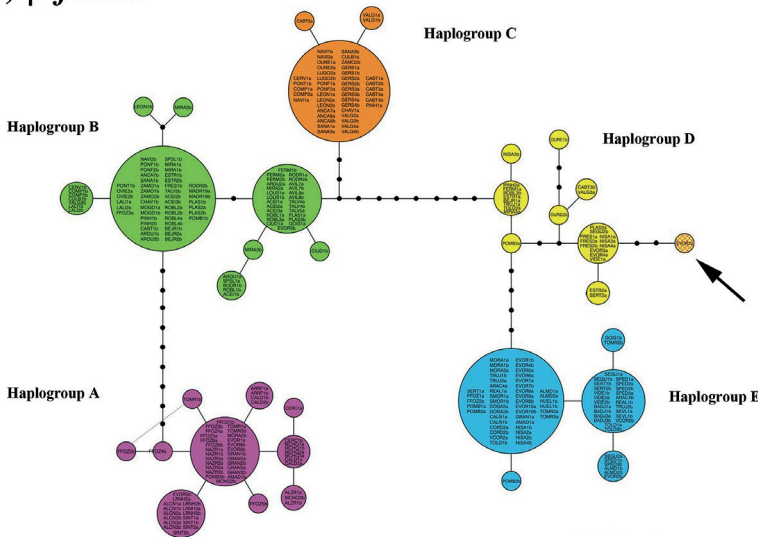
Geographical analyses

Mantel tests (Mantel, 1967) and analyses of spatial autocorrelation were performed for both nuclear and mitochondrial markers as implemented in Alleles in Space (AIS) v1.0 (Miller, 2005) to determine the correlation between genetic and geographical distances across the study area. The spatial autocorrelation analyses were conducted using 10, 20 and 50 geographic distance classes of equal size. The significance of both tests was evaluated through the use of a randomisation procedure based on 1000 permutation replicates.

To visualise the spatial patterns of genetic diversity at both nuclear and mitochondrial markers across the study area, we produced genetic distance synthetic maps, also named ‘Genetic Landscape Shape’ interpolations (Miller, 2005; Miller *et al.*, 2006). We used Alleles in Space v1.0 (Miller, 2005) to construct a connectivity network based on Delaunay triangulations (Delaunay, 1934) among all of the sampling locations and to calculate the genetic distances between observations connected in the network. The analyses were conducted using residual genetic distances derived from the linear regression of all pairwise genetic and geo-

graphical distances as recommended by Manni *et al.* (2004) to overcome the potential problems caused by correlations between genetic and geographical distances. The genetic distances and respective geographical coordinates of midpoints of each connection were then imported to the Geographical Information System Arc Gis 9.0 (ESRI, Redland, CA, USA). The ‘Inverse Distance Weighted’ (IDW) procedure with a power of one was used for the interpolation of the genetic distance surfaces as suggested by Miller (2005). IDW assumes that each input point has a local influence that diminishes with distance. The resulting genetic distances

a) *β-fibint7*



b) mtDNA

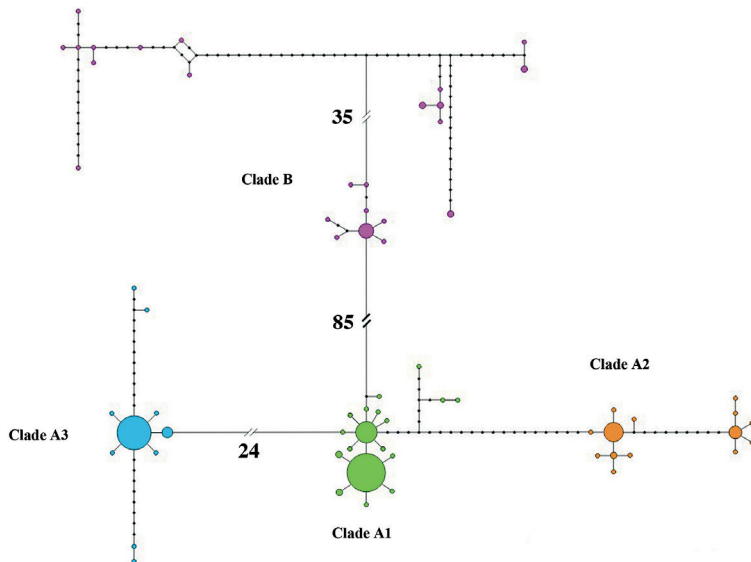


Fig. 2. Median-joining network of *β-fibint7* (a) and mtDNA haplotypes (b) in *L. boscai*. *β-fibint7* haplogroups are labelled A-E, and mtDNA clades are named as in Martínez-Solano *et al.* (2006). Colour codes as in Fig. 1. Circle size is proportional to haplotype frequencies. Codes in *β-fibint7* haplotypes represent localities as in Appendix I. Note the position of recombinant haplotype H30 (Code: EVOR2, hatched).

maps were plotted as 3-dimensional surfaces over the study area. The spatial representation of the ‘Genetic Landscape Shape’ interpolation areas was adjusted with a mask of the species’ range in order to exclude areas outside its known distribution.

We also used Monmonier’s Maximum Difference algorithm (Monmonier, 1973), as implemented in Alleles in Space v1.0 (Miller, 2005) to visualise in an explicit geographic context the main genetic barriers between the populations. This technique has been used extensively to depict the boundaries associated with the highest rate of change in genetic distances between sample locations (e.g. Manni *et al.*, 2004; Barbujani and Belle, 2006; Soltis *et al.*, 2006). The procedure begins by generating a connectivity network of all sampling locations using Delaunay triangulation (Delaunay, 1934). Next, a matrix of genetic distances between neighbouring samples is constructed and the pairwise values are associated to the respective edge of connection in the triangulation. Monmonier’s algorithm is then used to identify barriers by tracing the highest genetic distances between pairs of samples in the connection network (Patten and Smith-Patten, 2008). The analyses were conducted using residual genetic distances derived from the linear regression of the pairwise genetic and geographical distances. The first two boundaries identified at each marker by Monmonier’s algorithm were imported to the GIS program by using the geographical coordinates from the selected edges given by AIS and then represented over the ‘Genetic Landscape Shape’ interpolation map.

Comparative analyses between mitochondrial and nuclear DNA lineage distributions

We used Pearson and Spearman’s Bivariate Correlation coefficients in SPSS 14.0 (SPSS Inc, Chicago, IL, USA)

to compare the genetic distance maps of mitochondrial and nuclear DNA markers. Pearson’s correlation coefficient measures linear relationships between variables. It is based on the assumption that both X and Y values are sampled from populations that approximately follow a Gaussian distribution (with large samples, this assumption is less important). The nonparametric Spearman correlation is based on the ranking of the two variables, thus, making no assumption about the distribution of the values (Cann, 2002). We used the nuclear and mtDNA genetic distance values generated by Alleles in Space for a regular grid of 50×50 cells.

Results

Genetic diversity

A total of 35 β -fibint7 haplotypes and 49 variable (40 parsimony-informative) sites were detected in the 298 sequences analysed (Fig. 2a, Appendix III). The variable positions included three insertions and one deletion of 1bp. An A + T biased content of 59% was observed in β -fibint7 sequences, which is in agreement with values reported for this marker in other amphibian (Sequeira *et al.*, 2006; Gonçalves *et al.*, 2007) and reptile species (Godinho *et al.*, 2006). Five haplogroups separated by four to 10 mutations (labelled A to E, Fig. 2a) can be identified, exhibiting strong geographical correspondence (Fig. 1a). Estimates of genetic variation for each haplogroup are shown in Table 1. Haplotype and nucleotide diversity averaged respectively 0.907 and 0.016 in the total dataset.

Haplogroup A exhibited the highest values of genetic diversity ($h=10$, $Hd=0.740$) and differentiation from other groups. This haplogroup is exclusively present in

Table 1. Sample sizes (n), number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), average number of nucleotide differences (k), Tajima’s D and Fu’s F_s statistics for each β -fibint7 haplogroup in *Lissotriton boscai*. Note that sample sizes refer to number of alleles (two per sampled individual).

Group	n	h	Hd	π	k	Tajima’s D		Fu’s F_s	
						D value	P value	F_s	P value
Haplogroup A	60	10	0.740	0.0018	1.107	-0.936	P>0.10	-4.423	P>0.05
Haplogroup B	86	8	0.644	0.0014	1.443	-0.232	P>0.10	-0.861	P>0.10
Haplogroup C	45	3	0.129	0.0002	0.131	-1.306	P>0.10	-2.178	P>0.05
Haplogroup D	28	8	0.758	0.0036	2.382	0.460	P>0.10	-0.908	P>0.10
Haplogroup E	77	5	0.554	0.0011	0.659	0.1497	P>0.10	0.117	P>0.10
Total	296	34	0.907	0.0160	9.747	1.021	P>0.10	0.459	P>0.05

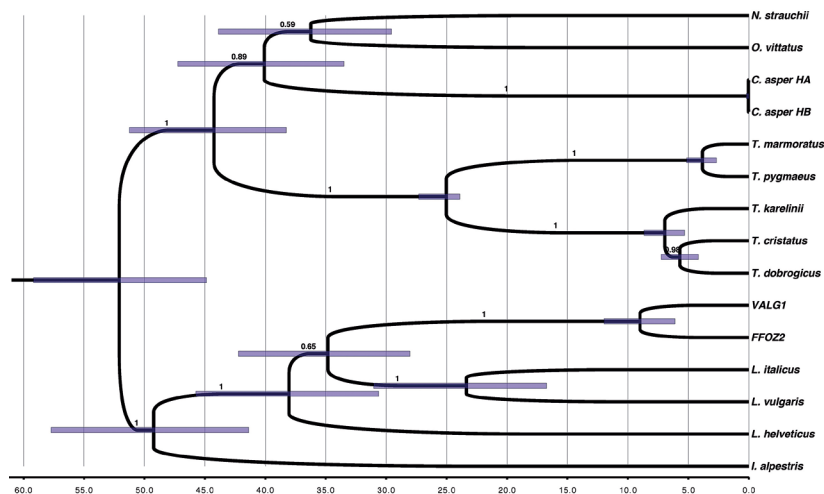


Fig. 3. Calibrated chronogram including mitochondrial haplotypes from the most divergent clades in *L. boscai* (FFOZ2 and VALG1) and haplotypes from related species and genera within Salamandridae. Mean divergence times are indicated for each node, together with their 95% Highest Posterior Density intervals (95% HPD, blue bars). Posterior probabilities are indicated above the branches. The temporal scale represents millions of years (mya)

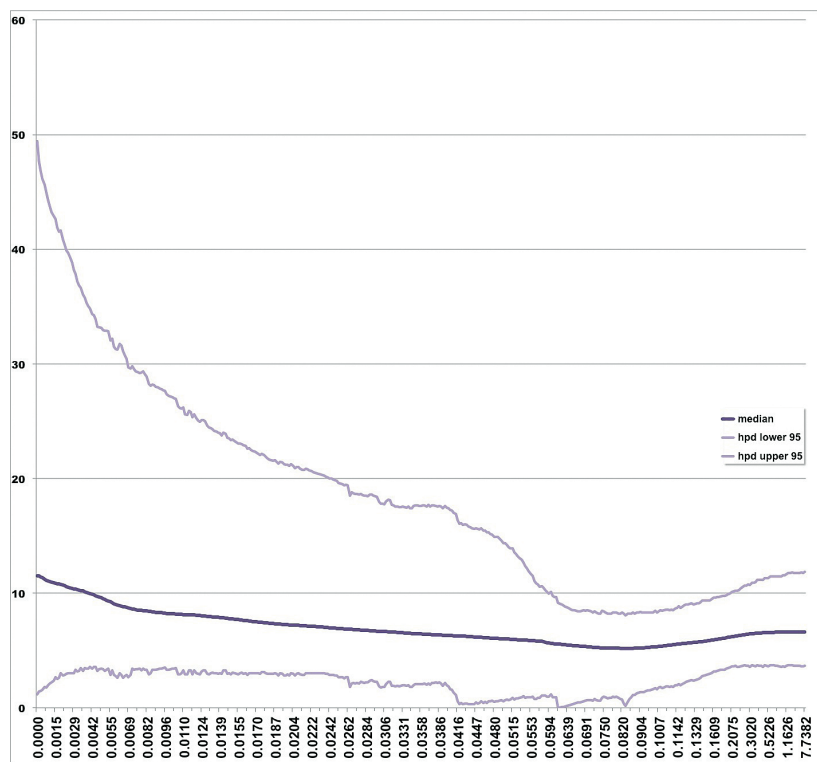


Fig. 4. Extended Bayesian Skyline Plot (EBSP), representing the effective population size through time in *L. boscai*. The mean population size (dark purple) is represented together with the lower and higher bounds of its 95% HPD. The x-axis represents the temporal scale in millions of years (mya). Population size is represented in the y-axis (units= population size x generation time). The plot does not show clear signals for expansions or bottlenecks, except a slight decrease in the effective population size between 200 and 70 kya, and an increase from 60 kya onwards, more evident in the last 20 ky.

populations from Southwestern Iberia (Fig. 1a). Haplogroup B is distributed mainly north of the Tejo River, with the exception of one sequence in the population of Évora. It is the most common haplogroup in central Portugal and the Spanish Sistema Central but it is also common throughout the northern limit of the species' range. Haplogroup C is restricted to the area north of the Douro River, where it co-occurs substantially with haplogroup

B. Haplogroup D shows a rare but widespread distribution over the central area of the species distribution, co-occurring with all other haplogroups. Although this haplogroup was present in 20 populations, only in one location it was the most frequent (Valverde del Fresno, population ID: 50). Haplogroup E was found mainly south of the Tejo River, being exclusive throughout the South-eastern region of the species' range (Fig. 1a).

Despite the lower levels of genetic divergence in the nuclear dataset, results of phylogenetic analyses of β -fibint7 data (excluding the recombinant haplotype H30, see below) support three main clades (Supplementary file S1). One of them corresponds to the most differentiated haplogroup (A), whereas the other two group together haplotypes from two different haplogroups each (B+C and D+E, respectively). Values of statistical support (posterior probability, PP) for these clades are 1.0 in Bayesian analyses. However, relationships between these clades are unresolved. Some additional subclades were recovered with significant support within the latter two clades, including haplogroups C and E (Bayesian Posterior Probabilities: 1.0 in both cases). Tajima's *D* and Fu's *F_s* statistics failed to detect a significant departure from neutral expectations in the β -fibint7 dataset or in any haplogroup (Table 1).

The recombination test using the pairwise homoplasy index identified the occurrence of recombination ($p < 0.05$). Visual inspection of haplotype variable positions suggests that H30 probably resulted from a recombination event between haplotypes from haplogroups A and D, since it shows a mosaic composition of the most frequent haplotypes from these haplogroups (haplotypes H6 and H28, Appendix III). Both parental lineages are found in sympatry in its population of origin: Azaruja, Évora (population ID: 10).

The 22 additional *nad4* and control region sequences from individuals from 12 populations unscreened in Martínez-Solano *et al.* (2006) clarified the distribution limits of mtDNA clades and helped to track down potential contact zones with a finer resolution (Fig. 1b). New samples grouped with the previously described mtDNA clades (Martínez-Solano *et al.*, 2006, Fig. 1b). We did not find haplotypes from different mtDNA clades coexisting in any of the new sampled locations. The mitochondrial haplotype network consists of two highly differentiated clades (A and B) separated by 85 mutations. Each of them is in turn subdivided into three clades (labelled A1 to A3 and B1 to B3 as in Martínez-Solano *et al.*, 2006, Fig. 2b). The geographical distribution of clades shows high congruence between nuclear haplogroups A, B, C, and E and mitochondrial clades B, A1, A2 and A3, respectively.

Phylogenetic and phylogeographic analyses

The phylogenetic tree resulting from the analysis of the reduced dataset is shown in Fig. 3. The topology is congruent with the most widely accepted phylogenetic hypotheses for the family except for the sister clade rela-

tionship between *Calotriton* and *Triturus*, not recovered in our analysis (Zhang *et al.*, 2008; Wiens *et al.*, 2011). Divergence time estimates among groups are also congruent with these and other studies, except slightly younger ages obtained for species within *Triturus* (e.g. Wielstra and Arntzen, 2011). Mean substitution rates and their associated 95% highest posterior density (HPD) intervals estimated for each partition were as follows (units: substitutions/site/my⁻¹): p1 (*nad4*, first positions+ tRNA-Hist): 0.0038 [0.0029-0.0047]; p2 (*nad4*, second positions): 0.0011 [0.0008-0.0015]; p3 (*nad4*, third positions): 0.020 [0.015-0.025]; and p4 (control region): 0.0028 [0.0023-0.0033]; total mean= 0.007. These values are within the range of previous estimates of mtDNA substitution rates in urodeles (Mueller, 2006; Schwartz and Mueller, 2010). The split between the main mtDNA lineages in *L. boscai* was estimated between 6 and 12 mya (mean=9 mya). This estimate was used as a temporal prior for the subsequent analyses. Running the analysis with an empty alignment, and therefore sampling only from the priors, confirmed the adequacy of the calibration points used in the study (Supplementary file S2). The temporal priors behaved as expected, with similar marginal likelihood distributions in the two analyses, but the non-calibrated nodes were dated differently, as it is expected when data is considered into the analysis.

We reconstructed the demographic history of *L. boscai* through an Extended Bayesian Skyline Plot (EBSP), which is shown in Fig. 4. The plot did not show clear signals for expansions or bottlenecks, except a slight decrease in the effective population size between 200 kya and 70 kya, and an increase from 60 kya onwards, more evident in the last 20 ky.

The continuous phylogeographic reconstructions for the two markers showed similar although not identical patterns, which are depicted in Fig. 5. The geographic location for the root in the two reconstructions is found in the same region, in central Portugal, at the western-most end of the Iberian Sistema Central (Fig. 5af), between 9 and 6 mya. The uncertainty about the exact geographic location is, however, notable along this time interval. From there, the species would have expanded latitudinally around 2.5-2 mya (Fig. 5bg), reaching the southernmost area of its distribution around 900-700 kya (Fig. 5ch). The expansion towards the east proceeded along two axes, yet the patterns are different between markers. The mtDNA reconstruction showed a main dispersal axis along the Iberian Sistema Central, starting around 600 kya (Fig. 5cd). From there, there was a southward expansion,

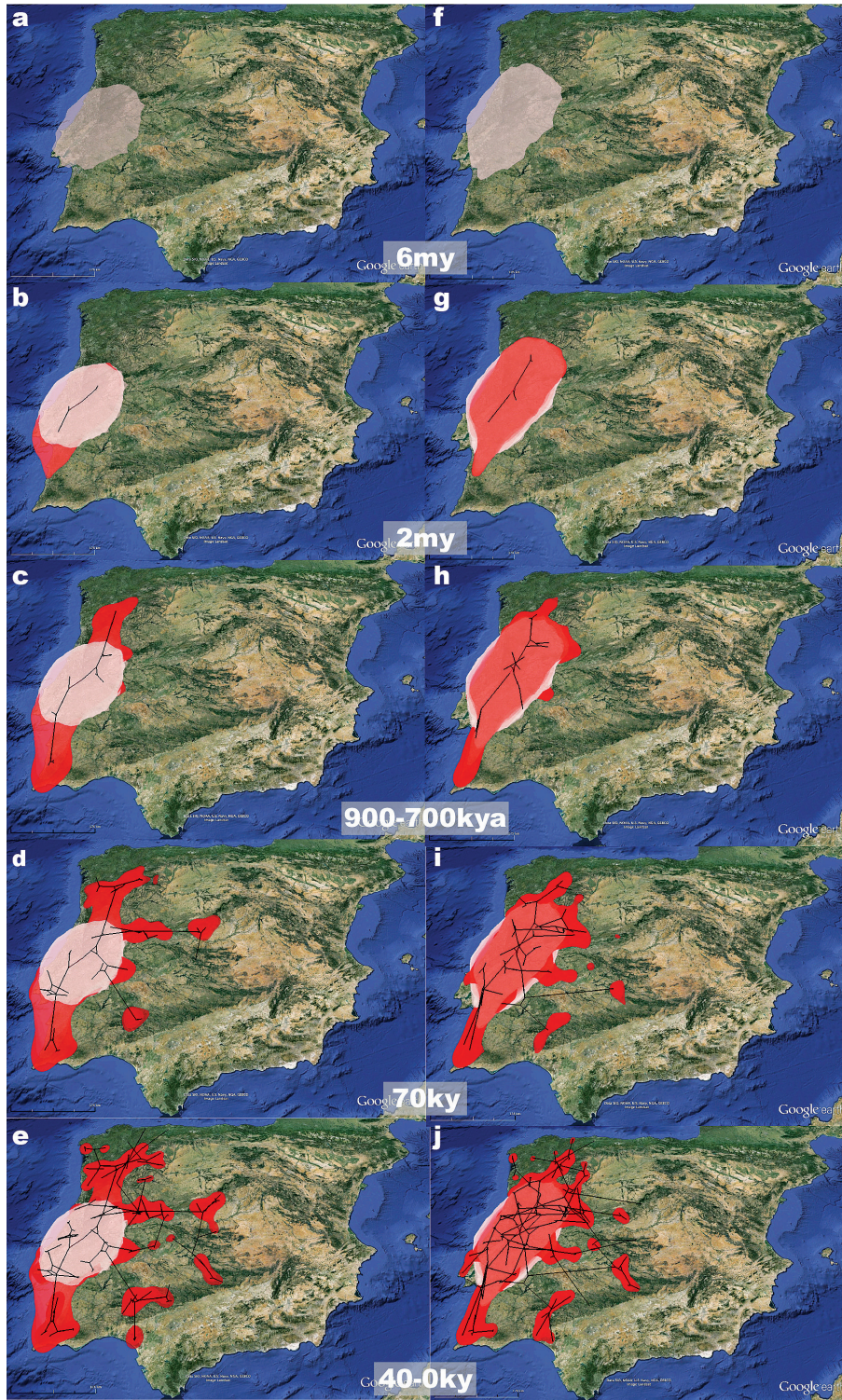


Fig. 5. Continuous diffusion phylogeographic reconstruction in *L. boscai*. We represent the dispersal dynamics through time for the species inferred from mtDNA (left panels, **a-e**) and the nuclear marker β -fibin7 (right panels, **f-j**). The polygons in the map represent the 80% HPD intervals for the location of each node of the sampled genealogies, a measure of the uncertainty of the estimated location. The white-red gradient of the polygons represents the relative age of the dispersal events, from white (old) to red (recent).

resulting in the colonization of the southeastern extreme of the range (200–70 kya) (Fig. 5d). A second axis of dispersal was established in a southeast direction, but dispersion along this axis was limited (Fig. 5cd). The easternmost limit of the distribution along the Iberian Sistema Central was reached 40 kya and the colonization of the northernmost area of the distribution occurred very recently, after the Last Glacial Maximum (LGM) (Fig. 5e).

In contrast, the reconstruction based on the nuclear marker showed the two eastern axes as equally important for dispersal and colonization of the east/southeast distribution areas. The south-eastern part of the distribution is not colonized from the Iberian Sistema Central as inferred in the mtDNA reconstruction, but from this second south-eastern axis, starting 600 kya, and reaching the south-eastern limits of the distribution 200 kya (Fig. 5hi). Similarly, the northernmost area of the distribution was reached very recently (20 kya) (Fig. 5j).

The diffusion rates for the species estimated in both analyses yielded similar values (mean around 176–193 km/myr, with 95% HPD intervals ranging between 115 and 280 km/myr). The continuous phylogeographic reconstructions are provided as Supplementary files S3 and S4, and can be visualized in GoogleEarth®.

Geographical analyses

Mantel tests detected significant correlations between genetic and geographic distances for both nuclear ($r=0.408$, $p<0.001$) and mitochondrial ($r=0.316$, $p<0.001$) data. Likewise, spatial autocorrelation analyses were also significant ($p<0.001$) for all grouping sets, indicating the existence of a clear relationship between genetic and geographic distances.

The ‘Genetic Landscape Shape’ interpolations produced a highly informative spatial illustration of genetic differentiation between populations (Fig. 6). For the nuclear β -fibint7, this technique detected a major ‘ridge’ (indicating greatest genetic distances) located along the centre of the distribution of the species, essentially separating haplogroup E from haplogroup B (Fig. 6a). A second evident ‘ridge’ is depicted in southwestern Iberia, separating haplogroup A from haplogroups E and B. For mtDNA, the major ‘ridge’ is detected in southwestern Iberia, separating populations from lineages A and B (Fig. 6b). Two other minor ‘ridges’ are apparent from Central Portugal to the southeastern limit of the species’ range and from northwestern Portugal to the Cantabrian Mountains in Northern

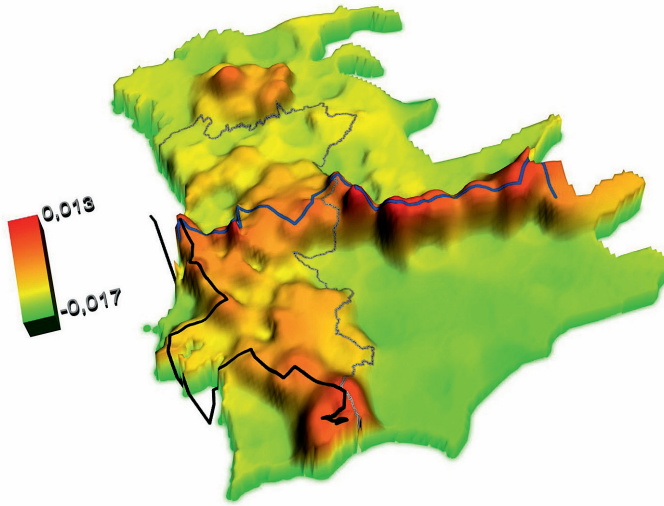
Spain, separating populations from mtDNA clades A1 and A3 and clades A1 and A2, respectively. These results are in close concordance with those reported in Martínez-Solano *et al.* (2006).

The use of Monmonier’s Maximum Difference algorithm (Monmonier, 1973) to detect the main genetic barriers for Iberian populations of *L. boscai* is in close agreement with the ‘Genetic Landscape Shape’ interpolations as illustrated in Fig. 6. For the nuclear dataset, the two major genetic barriers were identified along the centre of the species distribution and in southwestern Iberia, depicting the same geographical pattern as in the genetic distance synthetic map. For the mtDNA this method identified the main genetic barriers in southwestern Iberia. Although the method detected three barriers, the second and third are contiguous and interpreted as a single barrier, separating mtDNA clades A2 and A3 (Fig. 6b). Both of these barriers overlap closely with the main mtDNA ridges of the ‘Genetic Landscape Shape’ interpolations.

Comparative analyses between mitochondrial and nuclear DNA lineage distributions

The phylogeographical patterns obtained for the β -fibint7 nuclear marker are highly congruent with those obtained for the mtDNA markers, both in terms of number, phylogenetic relationships and distribution of haplogroups (Figs. 1, 2). Both Pearson and Spearman’s coefficients depicted significant correlations ($p<0.001$, $r=0.547$ and $r_s=0.463$, respectively) for the genetic distance synthetic maps between the analysed mitochondrial and nuclear markers, confirming the existence of similar spatial patterns between these markers.

Nonetheless, some differences are apparent between these markers including: 1) the existence of a fifth group in the nuclear β -fibint7 dataset (haplogroup D), which is not present in the mtDNA dataset (Fig. 1); 2) an overall lower genetic differentiation between the nuclear haplogroups, particularly visible within haplogroup A, between haplogroup A and the remaining haplogroups, and between haplogroups B and C (Fig. 2); 3) a proportionally higher genetic differentiation between the nuclear haplogroup E and the other clades in comparison with the corresponding mtDNA lineage A3 (Fig. 2); and 4) a broader distribution and admixture of all nuclear clades, especially of haplogroup B north of the Douro river, where it co-occurs widely with haplogroup C, and haplogroup E in the southeastern area of the species’ distribution, where based on mtDNA

a) β -*fibint7*

b) mtDNA

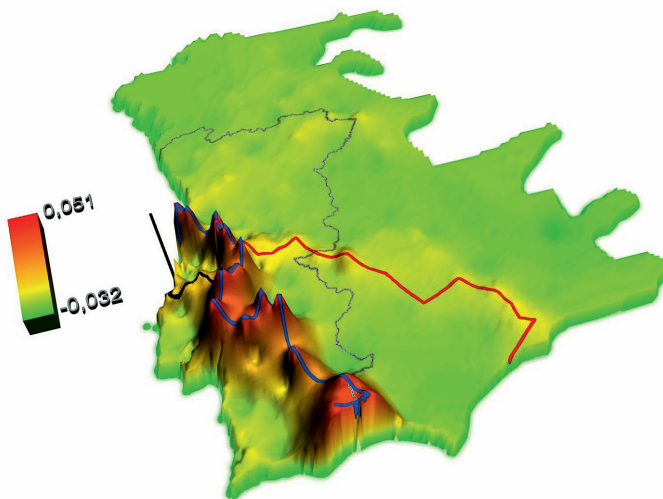


Fig. 6. Three-dimensional ‘Genetic Landscape Shape’ interpolations and graphical representation of Monmonier’s Maximum Difference algorithm showing areas characterised by major genetic breaks in the β -*fibint7* (a) and the mtDNA (b) datasets in *L. boscai*.

results we expected the presence of haplogroup B instead (Fig. 1). Most of these features are also reflected in the genetic distances synthetic map, where 1) the southwestern ‘ridge’ geographically overlaps for the two markers, although with a decrease of intensity in the nuclear data set, 2) the central ‘ridge’ shows conversely a higher intensity in the nuclear data and a geographical displacement to the north in comparison to the mitochondrial data, and 3) the northern ‘ridge’ depicted in the mitochondrial synthetic map, separating lineages A1 and A2, is not evident in the nuclear map,

probably because of geographical admixture of nuclear clades B and C (Fig. 6).

Discussion

Nuclear DNA phylogeography of Lissotriton boscai

Previous studies have revealed the utility of the nuclear intron β -*fibint7* for phylogenetic analyses (Johnson and Clayton, 2000; Prychitko and Moore, 2003; Sequeira et

al., 2006; Gonçalves *et al.*, 2007), but few have also employed this marker in intraspecific studies (*e.g.* Godinho *et al.*, 2006; Velo-Antón *et al.*, 2008; Gonçalves *et al.*, 2009). In this study, despite relatively shallow levels of genetic divergence among intraspecific clades due to expected low substitution rates for the nuclear marker, we found clear genetic structure and geographic patterns in the β -fibint7 sequences, suggesting its potential utility for phylogeographic studies in other taxa.

We identified five well-differentiated nuclear haplogroups (Fig. 2a), with localised distributions but wide areas of overlap (Fig. 1a), suggesting either extensive admixture across some mtDNA borders or incomplete lineage sorting. However, some areas are characterised by little admixture between nuclear haplogroups, like southwestern Portugal and Spain or the eastern Sistema Central mountains in Central Spain, suggesting restricted gene flow between adjacent haplogroups or founder effects, with both events potentially causing fixation of certain haplotype variants. Sampling issues, however, cannot be completely ruled out due to the low number of individuals analysed per location. Historical demographic analyses failed to detect signatures of recent demographic expansions in all β -fibint7 haplogroups, perhaps limited by the relatively low number of variable characters in the dataset. The inferred EBSP, however, is consistent with this inference of overall demographic stability, punctuated by modest population declines and expansions, the most recent dating back to the LGM.

An interesting peculiarity of the nuclear dataset is the presence of an additional haplogroup with no match in the mtDNA dataset (haplogroup D). This haplogroup has a wide distribution, but is more frequent in the Sistema Central mountains in central Spain, an area that has been characterised as a potential glacial refuge for other endemic amphibians and reptiles (*e.g.* *Salamanca salamandra almanzoris*, Martínez-Solano *et al.*, 2005; *Iberolacerta cyreni*, *I. martinezricai*, Arribas and Carranza, 2004; *Lacerta schreiberi*, Paulo *et al.*, 2001; *Podarcis carbonelli*, Pinho *et al.*, 2007). The low frequency of this haplogroup within its range (Fig. 1a) and the lack of a mitochondrial counterpart could be explained by drift as it is known that mtDNA is much more sensitive to bottleneck effects than nuclear genes because of its four times lower effective population size (Moore, 1995).

The overall geographic patterns for β -fibint7 and mtDNA haplotypes are notably congruent. The phylogenetic relationships between clades are also in accordance for the nuclear and the mitochondrial datasets,

although support values are lower in the nuclear dataset. The similarity between nuclear and mitochondrial phylogeographic patterns is highly relevant because this is not often the case (*e.g.* Ballard *et al.*, 2001; Gamache *et al.*, 2003; Mosen and Blouin, 2003; Schaschl *et al.*, 2003; Bensch *et al.*, 2006) and confirms the existence of deep and consistent differentiation across intraspecific lineages. In spite of the overall similarities, the phylogeographic diffusion models and geographical analyses reveal distinct primary dispersal routes and phylogeographic breaks for the nuclear and mtDNA (Fig. 6). The primary 'ridge' in the mtDNA dataset (Fig. 1b, Fig. 6b) reflects the phylogenetic break between the most differentiated mitochondrial lineage B from southwestern Iberia and the other Iberian mtDNA lineages (Martínez-Solano *et al.*, 2006). By contrast, the primary 'ridge' in the nuclear data set divides the range of the species into north and south groups, separating haplogroup E from adjacent haplogroup B (Fig. 1a, Fig. 6a). However, the second most important β -fibint7 'ridge', separating the southwestern haplogroup A from the remaining clades, closely overlaps with the primary mtDNA 'ridge'. Similarly, the second most important mtDNA 'ridge', which separates lineages A1 and A3, corresponds roughly to the primary β -fibint7 'ridge' (although with a displacement to the south in the southeastern limit of the species' range reflecting different nuclear and mtDNA clade assignment in this area). Most of the depicted 'ridges' correspond roughly to geographic areas where the species is less abundant due to ecological factors (as shown by the areas where the species was not detected within its range, Fig. 1a, b; Díaz-Paniagua, 2002; Teixeira, 2008). According to our temporal estimates, the two major intraspecific lineages in *L. boscai* would have originated during the Miocene, with subsequent splits around the Pliocene–Pleistocene boundary. These lineages would have survived in different allopatric refugia during the Pleistocene glacial cycles, expanding their ranges from these refugia to the current distribution of the species, during climatologically favourable periods (Fig. 5). According to this interpretation, supported by inference from phylogeographic diffusion models, both nuclear and mtDNA breaks would reveal areas of secondary contact between anciently diverged lineages.

Patterns of gene flow and admixture are not similar between mitochondrial and nuclear markers, nor are they similar in all the contact zones revealed. Historical and ecological factors, together with the explicit spatial context where the contact zones occur, contribute to the differences observed. However, little is

known about some basic biological traits of the species, such as patterns of dispersal, necessary to understand the processes that shaped the differences observed. It might be the case that some differences are due to male-biased dispersal in the species (as observed in *Ichthyosaura alpestris*, Joly and Grolet, 1996). It could be also possible that different haplogroup dispersal is affected by the local habitat quality (Denoël and Lehmann, 2006), by competitive exclusion (Dolmen, 1988), or is density-dependent (Jánosi and Scheuring, 1997; Aars and Ims 2000), in which case individuals may exhibit limited dispersal to geographical areas already occupied by conspecifics. As another plausible scenario, individuals might disperse but fail to reproduce or exhibit asymmetrical reproductive success in occupied areas due to sexual selection (Aragón *et al.*, 2000; Gabor *et al.*, 2000; Babik *et al.*, 2003) or genetic barriers (*e.g.* Herrero, 1991). These alternative but not mutually exclusive hypotheses have yet to be empirically tested with further field and laboratory experiments, explicitly taking into account the historical, ecological and geographical contexts in which the processes are taking place, *i.e.* the distinct secondary contact zones revealed in our analyses.

The implementation of Monmonier's algorithm and 'Genetic Landscape Shapes' in GIS programs produces overall similar results and provides a number of different approaches for portraying patterns of genetic diversity and divergence across landscapes. Both methods share the common goal of identifying spatial patterns associated with the largest genetic distances in a data set but the 'Genetic Landscape Shape' procedure potentially allows for inferences about the entire landscape as opposed to identifying sets of maximally differentiated individuals. The Monmonier's algorithm approach shows in some cases difficulties to consistently identify a main barrier (as in the case of barrier #2 in the mtDNA data set, Fig. 6b), being highly susceptible to isolated genetic distance peaks or high values located near a main barrier. Performing the 'Genetic Landscape Shapes' interpolations and the representation of the synthetic maps in a GIS program instead of in Alleles in Space (Miller, 2005) revealed three major advantages: 1) the interpolation surface is continuous and no grid cell size needs to be defined, 2) the graphical representation of genetic diversity patterns can be depicted either as 2 or 3-dimensional surface plot over a study area map instead of over a geographically unreferenced grid, and 3) the interpolation area does not need to be a regular grid and can be adjusted to the species' range. As suggested by Kidd and

Ritchie (2006), the applications of GIS-mediated analysis and spatial representation methods confirm to be very informative and highly promising tools for phylogeographical analysis.

Implications for species formation

Deep phylogeographic breaks in mtDNA can identify population groups of particular conservation value (ESUs *sensu* Moritz, 1994), but are in many cases insufficient to clarify species status (*e.g.* Zangari *et al.*, 2006). In this respect, our study provides a direct comparison of nuclear and mitochondrial DNA patterns and suggests admixture between parapatric and very divergent mtDNA lineages, illustrating how deep phylogeographic breaks do not necessarily indicate instances of cryptic speciation. Even in a case in which mitochondrial breaks are a consequence of ancient speciation processes, current population admixture, through secondary contact, might overcome the incipient reciprocal coalescence of isolated lineages, resulting in incomplete speciation. Our approach revealed broad areas of nuclear DNA admixture superimposed over the patterns observed with mtDNA, and thus suggested lack of reproductive isolation in otherwise genetically differentiated groups of populations (the mtDNA lineages A and B of *L. boscai*), through gene flow (as suggested by the finding of heterozygote individuals and evidence of recombination events). The disruption of the expected outcome of a long speciation process initiated in the Miocene emphasises, once more, the problems of single-marker species delineation.

It has long been recognised that delineating species in amphibians is a difficult task. One of the major problems relates to the widespread absence of effective reproductive isolation mechanisms between taxa that have evolved in isolation for millions of years (Hickerson *et al.*, 2006; Malone and Fontenot, 2008). This makes it hard to predict the outcome of evolutionary interactions between deeply divergent clades from tree topologies alone. This pattern-based approach should be accompanied by the study of the actual mechanisms driving lineage divergence (Wiens, 2004). For instance, our results allowed us to pinpoint potential contact zones of interest to study the evolutionary mechanisms related to lineage diversification in *L. boscai*. Future studies should therefore emphasise the analysis of variation in genetic, physiological, behavioural and ecological traits and their interrelations in these contact zones. The need for an 'integrative taxonomy' (Padiál *et al.*, 2010) is undisputed; however, this should not just

rely on data addition but, more importantly, on data integration, which would reflect the evolutionary processes and mechanisms leading to species formation.

Acknowledgements

We would like to acknowledge and thank E. Albert, M. Alcobendas, J.W. Arntzen, J. Barbadillo, R. Begonha, A. Cunha, J.P.G. de la Vega, C. Díaz-Paniagua, B. Docherty, P. Esteves, H. Gonçalves, C. Grande, M. Lapeña, R. Márquez, F. Queirós, R. Rebelo, E. Recuero, B. Ribeiro, R. Ribeiro, I. Sánchez, S. Schönhuth, P. Sá-Sousa, F. Sequeira, N. Sillero, C. Soares, N. Sousa and I. Urbán for their help in collecting samples or providing tissues. The comments and suggestions on preliminary drafts of this manuscript by G. Luikart and J.W. Arntzen were very useful, as well as the relevant references provided by E.G. Crespo, A. Salvador and B. Sanchiz. We also thank the comments provided by the editor and two anonymous reviewers, which have certainly improved the manuscript. The ‘Consejerías de Medio Ambiente’ of Galicia, Asturias, Castilla-León, Madrid, Extremadura, Castilla-La Mancha and Andalucía (Spain) and ICN (Portugal) provided the legal collecting permits. This study was funded by projects CGL2007-64621, CGL2010-15786 (PI: MGP), CGL2008-04271-C02-01/BOS and CGL2011-28300 (PI: IMS) (Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad, Spain, and FEDER), PIII10-0097-4200 (PI: IMS) (Junta de Comunidades de Castilla la Mancha, and FEDER), the Program Operacional Factores de Competitividade (COMPETE), and by national funds from Fundação para a Ciência e a Tecnologia (FCT), through the research Project PTDC/BIA-BEC/105083/2008. JT was supported by a PhD grant from Fundação da Ciência e Tecnologia (Ref.: BD/3376/2000). DB was supported by a JAE-DOC fellowship from the CSIC under the program ‘Junta para la Ampliación de Estudios’ co-financed by the European Social Fund (ESF). IMS was supported by Project ‘Biodiversity, Ecology and Global Change’, co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2–O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF) and is currently supported by funding from the Spanish Severo Ochoa Program (SEV-2012-0262).

References

Aars J, Ims RA. 2000. Population dynamic and genetic consequences of spatial density-dependent dispersal in patchy populations. *American Naturalist* 155: 252-265.

Alvarado-Serrano DF, Knowles LL. 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources* 14: 233-248.

Aragón P, López P, Martín J. 2000. Conspecific chemical cues influence pond selection by male newts *Triturus boscai*. *Copeia* 2000: 874-878.

Arribas O, Carranza S. 2004. Morphological and genetic evidence of the full species status of *Iberolacerta cyreni marinezricai* (Arribas, 1996). *Zootaxa* 634: 1-24.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.

Babik W, Szymura JM, Rafinski J. 2003. Nuclear markers, mitochondrial DNA and male secondary sexual traits variation in a newt hybrid zone (*Triturus vulgaris*; *T. montandoni*). *Molecular Ecology* 12: 1913-1930.

Babik W, Branicki W, Crnobrnja-Isailović J, Cogalniceanu D, Sas I, Olgun K, Poyarkov N, García-París M, Arntzen J. 2005. Phylogeography of two European newt species-discordance between mtDNA and morphology. *Molecular Ecology* 14: 2475-2491.

Ballard JWO, Chernoff B, James AC. 2001. Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in *Drosophila simulans*. *Evolution* 56: 527-545.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37-48.

Barbujani G, Belle EMS. 2006. Genomic boundaries between human populations. *Human Heredity* 61: 15-21.

Bensch S, Irwin DE, Irwin JH, Kvist L, Åkesson S. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology* 15: 161-171.

Bielejec F, Rambaut A, Suchard MA, Lemey P. 2011. SPREAD: Spatial Phylogenetic Reconstruction of Evolutionary Dynamics. *Bioinformatics* 27: 2910-2912.

Böhme M. 2010. Ectothermic vertebrates (Actinopterygii, Alcaudata, Urodela, Anura, Crocodylia, Squamata) from the Miocene of Sandelzhausen (Germany, Bavaria) and their implications for environment reconstruction and palaeoclimate. *Paläontologische Zeitschrift* 84: 3-41.

Bruen TC, Philippe H, Bryant D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172: 2665-2681.

Buckley D. 2009. Toward an organismal, integrative, and iterative phylogeography. *BioEssays* 31: 784-793.

Cann AJ. 2002. *Maths from Scratch for Biologists*. Oxford: Wiley.

Chan LM, Brown JL, Yoder AD. 2011. Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution* 59: 523-537.

Delaunay B. 1934. Sur la sphère vide. *Izvestia Akademii Nauk SSSR* 7: 793-800.

Denoël M, Lehmann A. 2006. Multi-scale effect of landscape processes and habitat quality on newt abundance: Implications for conservation. *Biological Conservation* 130: 495-504.

Díaz-Paniagua C. 2002. *Triturus boscai*. Pp. 61-63 in: Pleguezuelos JM, Márquez R, Lizana M, eds., *Atlas y libro rojo de los anfibios y reptiles de España*. Madrid: Dirección General de Conservación de la Naturaleza and Asociación Herpetológica Española.

Dolmen D. 1988. Coexistence and niche segregation in the newts *Triturus vulgaris* (L.) and *T. cristatus* (Laurenti). *Amphibia-Reptilia* 9: 365-374.

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

- Dufresnes C, Wassef J, Ghali K, Brelsford A, Stöck M, Lymbekakis P, Crnobrnja-Isailović J, Perrin N. 2013. Conservation phylogeography: does historical diversity contribute to regional vulnerability in European tree frogs (*Hyla arborea*)? *Molecular Ecology* 22: 5669-5684.
- Estes R. 1981. Gymnophiona, Caudata. In: Wellnhofer P., ed., *Handbuch der Paläoherpetologie*. Part 2. Stuttgart: Fischer Verlag.
- Forester BR, DeChaine EG, Bunn AG. 2013. Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Diversity and Distributions* 19: 1480-1495.
- Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.
- Gabor CR, Krenz J, Jaeger RG. 2000. Female choice, male interference, and sperm precedence in the red-spotted newt. *Behavioral Ecology* 11: 115-124.
- Gamache I, Jaramillo-Correa JP, Payette S, Bousquet J. 2003. Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with postglacial colonization. *Molecular Ecology* 12: 891-901.
- Gernhard T. 2008. The conditioned reconstructed process. *Journal of Theoretical Biology* 253: 769-778.
- Godinho R, Mendonça B, Crespo EG, Ferrand N. 2006. Genealogy of the nuclear β -fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. *Heredity* 96: 454-463.
- Gómez A, Lunt DH. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. Pp. 155-188 in: Weiss S, Ferrand N, eds., *Phylogeography of European Refugia*. Amsterdam: Springer.
- Gonçalves H, Martínez-Solano I, Ferrand N, García-París M. 2007. Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution* 44: 494-500.
- Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho BM, García-París M, Ferrand N. 2009. High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. *Molecular Ecology* 18: 5143-5160.
- Habel J, Drees C, Schmitt T, Assmann T. 2010. Review: Refugial Areas and Postglacial Colonizations in the Western Palearctic. Pp. 189-197 in: Habel JC, Assmann, eds., *Relict Species: Phylogeography and Conservation Biology*. Berlin: Springer.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95-98.
- Heled J, Drummond A. 2008. Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology* 8: 289.
- Herrero P. 1991. Polytypic chromosomal variation in *Triturus boscai* (Urodela: Salamandridae). *Génétique, Sélection et Évolution* 23: 263-272.
- Hickerson MJ, Meyer C, Moritz C. 2006. DNA barcoding will often fail to discover new animal species over broad parameter space. *Systematic Biology* 55: 729-739.
- Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* 54: 291-301.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754-755.
- Huelsenbeck JP, Larget B, Alfaro ME. 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution* 21: 1123-1133.
- Jánosi IM, Scheuring I. 1997. On the evolution of density dependent dispersal in a spatially structured population model. *Journal of Theoretical Biology* 187: 397-408.
- Johnson KP, Clayton DH. 2000. Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). *Molecular Phylogenetics and Evolution* 14: 141-151.
- Joly P, Grolet O. 1996. Colonization dynamics of newly-created ponds in the Alpine newt, *Triturus alpestris*: influence of landscape configuration and age structure of the colonizers. *Acta Oecologica* 17: 599-608.
- Kidd DM, Ritchie MG. 2006. Phylogeographic information systems: putting the geography into phylogeography. *Journal of Biogeography* 33: 1851-1865.
- Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29: 1695-1701.
- 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.
- Malone J, Fontenot B. 2008. Patterns of reproductive isolation in toads. *PLoS ONE* 3: e3900.
- Manni F, Guerard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* 76: 173-190.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209-220.
- Marjanović D, Laurin M. 2013. An updated paleontological timetree of lissamphibians, with comments on the anatomy of Jurassic crown-group salamanders (Urodela). *Historical Biology* 26: 535-550.
- Martín C, Sanchiz B. 2013. Lisanfos KMS. Version 1.2. Online reference accessible at <http://www.lisanfos.mncn.csic.es/>. Madrid: Museo Nacional de Ciencias Naturales, CSIC.
- Martínez-Solano I, Alcobendas M, Buckley D, García-París M. 2005. Molecular characterisation of the endangered *Salamandra salamandra almanzoris* (Caudata, Salamandridae). *Annales Zoologici Fennici* 42: 57-68.
- Martínez-Solano I, Teixeira J, Buckley D, García-París M. 2006. Mt-DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. *Molecular Ecology* 15: 3375-3388.
- Martínez-Solano I, Lawson R. 2009. Escape to Alcatraz: evolutionary history of slender salamanders (*Batrachoseps*) on the islands of San Francisco Bay. *BMC Evolutionary Biology* 9: 38.

- McCormack JEJ, Hird SMS, Zellmer AJA, Carstens BCB, Brumfield RTR. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* 66: 526-538.
- Milá B, Carranza S, Guillaume O, Clobert J. 2010. Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology* 19: 108-120.
- Miller MP. 2005. Alleles In Space (AIS): Computer software for the joint analysis of inter-individual spatial and genetic information. *Journal of Heredity* 96: 722-724.
- Miller MP, Bellinger RM, Forsman E, Haig SM. 2006. Effects of historical climate change, habitat connectivity, and vicariance on genetic structure and diversity across the range of the red tree vole (*Phenacomys longicaudus*) in the Pacific Northwestern United States. *Molecular Ecology* 15: 145-159.
- Monmonier MS. 1973. Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis* 5: 245-261.
- Monsen KJ, Blouin MS. 2003. Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology* 12: 3275-3286.
- Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49: 718-726.
- Moritz C. 1994. Defining "Evolutionarily Significant Units" for conservation. *Trends in Ecology and Evolution* 9: 373-375.
- Mueller R. 2006. Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology* 55: 289-300.
- Nei M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Núñez JJ, Wood NK, Rabanal FE, Fontanella FM, Sites JW. 2011. Amphibian phylogeography in the Antipodes: Refugia and postglacial colonization explain mitochondrial haplotype distribution in the Patagonian frog *Eupsophus calcarius* (Cycloramphidae). *Molecular Phylogenetics and Evolution* 58: 343-352.
- Padial JM, Miralles A, De la Riva I, Vences M. 2010. The integrative future of taxonomy. *Frontiers in Zoology* 7: 16.
- Patten MA, Smith-Patten BD. 2008. Biogeographical boundaries and Monmonier's algorithm: a case study in the northern Neotropics. *Journal of Biogeography* 35: 407-416.
- Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA. 2001. The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London B* 268: 1625-1630.
- Pinho C, Harris DJ, Ferrand N. 2007. Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. *Molecular Ecology* 16: 1191-1205.
- Prychitko T, Moore WS. 2003. Alignment and phylogenetic analysis of β -fibrinogen intron 7 sequences across avian orders reveal conserved nucleotides and strong phylogenetic signal. *Molecular Biology and Evolution* 20: 762-771.
- Rage J-C, Bailon S. 2005. Amphibians and squamate reptiles from the late early Miocene (MN 4) of Béon 1 (Montréal-du-Gers, southwestern France). *Geodiversitas* 27: 413-441.
- Recuero E, García-París M. 2011. Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. *Molecular Phylogenetics and Evolution* 60: 170-182.
- Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM. 2013. Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* 22: 2953-2970.
- Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
- Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning*. New York: Cold Spring Harbor Laboratory Press.
- Schaschl H, Kaulfuss D, Hammer S, Suchentrunk F. 2003. Spatial patterns of mitochondrial and nuclear gene pools in chamois (*Rupicapra r. rupicapra*) from the Eastern Alps. *Heredity* 91: 125-135.
- Schwartz RS, Mueller RL. 2010. Limited effects of among-lineage rate variation on the phylogenetic performance of molecular markers. *Molecular Phylogenetics and Evolution* 54: 849-856.
- Sequeira F, Ferrand N, Harris DJ. 2006. Assessing the phylogenetic signal of the nuclear β -fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia* 27: 409-418.
- Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* 15: 4261-4293.
- Stamatakis A. 2014. The RAxML v8.1.X Manual. Available at: <http://sco.h-its.org/exelixis/resource/download/NewManual.pdf>
- Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978-989.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.
- Teixeira J. 2008. *Lissotriton boscai*. Pp. 98-99 in: Loureiro AM, Carretero MA, Ferrand N, Paulo O., eds., *Atlas dos Anfíbios e Répteis de Portugal*. Lisboa: Instituto da Conservação da Natureza.
- Tourneville A. 1879. Description d'une nouvelle espèce de batracien urodèle d'Espagne (*Pelonectes boscai* Lataste). *Bulletin de la Société Zoologique de France* 4: 70-87.
- Velo-Antón G, Martínez-Solano I, García-París M. 2008. β -fibrinogen intron 7 variation in *Discoglossus* (Anura: Discoglossidae): implications for the taxonomic assessment of morphologically cryptic species. *Amphibia Reptilia* 29: 523-533.
- Wakeley, J. 2004. Metapopulation models for historical inference. *Molecular Ecology* 13: 865-875.
- Wielstra B, Arntzen JW. 2011. Unraveling the rapid radiation of crested newts (*Triturus cristatus* superspecies) using complete mitogenomic sequences. *BMC Evolutionary Biology* 11: 162.
- Wielstra B, Crnobrnja-Isailović J, Litvinchuk SN, Reijnen BT, Skidmore AK, Sotiropoulos K, Toxopeus AG, Tzankov N, Vukov T, Arntzen JW. 2013. Tracing glacial refugia of *Triturus* newts based on mitochondrial DNA phylogeography and species distribution modeling. *Frontiers in Zoology* 10: 13.

- Wiens JJ. 2004. What is speciation and how should we study it? *American Naturalist* 163: 914-923.
- Wiens JJ, Sparreboom M, Arntzen JW. 2011. Crest evolution in newts: implications for reconstruction methods, sexual selection, phenotypic plasticity and the origin of novelties. *Journal of Evolutionary Biology* 24: 2073-2086.
- Yang Z. 2006. *Computational Molecular Evolution*. Oxford, UK: Oxford University Press.
- Zangari F, Cimmaruta R, Nascetti G. 2006. Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae). *Biological Journal of the Linnean Society* 87: 515-536.
- Zhang P, Papenfuss TJ, Wake MH, Qu L, Wake DB. 2008. Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution* 49: 586-597.
- Zhao Y, Zhang Y, Li X. 2013. Molecular phylogeography and population genetic structure of an endangered species *Pachyhynobius shangchengensis* (hynobiid Salamander) in a fragmented habitat of southeastern China. *PLoS ONE* 8: p.e78064.

Received: 8 September 2014

Revised and accepted: 9 February 2015

Published online: 28 August 2015

Editor: B. Wielstra

Online supplementary material

S1. Bayesian 50% majority rule phylogram for the β -fibint7 haplotypes in *L. boscai*. Three main clades were recovered in the analysis, although the relationships between them were not fully resolved: the first one, **in purple**, corresponds to the most differentiated haplogroup (A); the second one groups together haplotypes from haplogroups B, **in green**, and C, **in orange**; the third included haplogroups D, **in yellow**, and E, **in blue**.

S2. Marginal likelihood distributions for (A) the root of the tree, (B) the Time to the Most Recent Common Ancestor (TMRCA) for *Triturus*, (C) the TMRCA for the split between *Lissotriton* and *Ichthyosaura*, and (D) the TMRCA for the two most divergent lineages within *L. boscai*. The distributions obtained when sampling only from the prior (**in black**) are compared to the marginal distributions obtained when data is considered into the analysis (**in blue**). Calibration points were added in (A) and (B), and both distributions are congruent, as expected when the information contained in the priors is not conflicting. When no prior calibration is included [nodes (C) and (D)], the resulting distributions are very different, meaning that the final posterior distributions will be shaped by the data, as expected. The x-axis is scaled in millions of years.

S3. .kml file showing the inferred continuous diffusion process in *Lissotriton boscai* based on a time-calibrated β -fibint7 genealogy reconstructed with BEAST. This file can be visualized with the software Google Earth®.

S4. .kml file showing the inferred continuous diffusion process in *Lissotriton boscai* based on a time-calibrated mtDNA genealogy reconstructed with BEAST. This file can be visualized with the software Google Earth®.

Appendix I
 Sampling localities, population identifiers (ID), sample sizes (number of individuals sequenced) and GenBank accession numbers for the sequences used in this study. In parentheses, locality codes (see also Fig. 2 and Appendix III).

Population ID	Locality	Nuclear Sample size	GenBank Accession numbers	MtDNA Sample size	GenBank Accession numbers	control region
1	Portugal: Monchique (MCHQ)	3	β-fibint7 KP265980-KP265985	4	na44 DQ491737-491740	DQ491840-491843
2	Portugal: Arrifana, Aljezur (ARRF)	1	KP265986-KP265987	1	DQ491741	DQ491844
3	Portugal: Zambujeira de Baixo, Aljezur (ALZR)	1	KP265988-KP265989	1	DQ491742	DQ491845
4	Portugal: Barqueiros, Odemira (CERC)	1	KP265990-KP265991	1	DQ491743	DQ491846
5	Portugal: Santa Margarida da Serra, Grândola (GRAN)	3	KP265992-KP265997	2	DQ491744-491745	DQ491847-491848
6	Portugal: Sintra (SINT)	2	KP265998-KP266001	1	DQ491746	DQ491849
7	Portugal: S. Bartolomeu dos Galegos, Lourinhã (LRNH)	2	KP266002-KP266005	2	DQ491747-491748	DQ491850-491851
8	Portugal: Alcanena, Alcobaca (ALCN)	3	KP266006-KP266011	2	DQ491749-491750	DQ491852-491853
9	Portugal: Valado dos Frades, Nazaré (NAZR)	3	KP266012-KP266017	4	DQ491751-491754	DQ491854-491857
10	Portugal: Azaruja, Évora (EVOR1-3)	3	KP266018-KP266023	3	DQ491755-491757	DQ491858-491860
11	Portugal: Castelo de Vide (VIDE)	2	KP266024-KP266027	2	DQ491758-491759	DQ491861-491862
12	Portugal: Cernache de Bonjardim, Sertã (SERT)	2	KP266028-KP266031	2	DQ491760-491761	DQ491863-491864
13	Portugal: Lousã (LOUS)	1	KP266032-KP266033	1	DQ491762	DQ491865
14	Portugal: Góis (GOIS)	1	KP266034-KP266035	1	DQ491763	DQ491866
15	Portugal: Farnalhão, Guarda (ESTR)	2	KP266036-KP266039	3	DQ491764-491766	DQ491867-491869
16	Portugal: São Pedro do Sul (SPSL)	1	KP266040-KP266041	1	DQ491767	DQ491870
17	Portugal: Pinhão (PINH)	2	KP266042-KP266045	2	DQ491768-491769	DQ491871-491872
18	Portugal: Mogadouro (MOGD)	1	KP266046-KP266047	1	DQ491770	DQ491873
19	Portugal: Chaves (CHAV)	1	KP266048-KP266049	1	DQ491771	DQ491874
20	Portugal: Cabeceiras de Basto (CABT)	4	KP266050-KP266057	2	DQ491772-491773	DQ491875-491876
21	Portugal: Alfena, Valongo (VALG)	4	KP266058-KP266065	4	DQ491774-491777	DQ491877-491880
22	Portugal: Vila Nova de Cerveira (CERV)	1	KP266066-KP266067	1	DQ491778	DQ491881
23	Portugal: Carris, Terras de Bouro (GERS)	4	KP266068-KP266075	1	DQ491779	DQ491882
24	Spain: A Pobra do Caramiñal, A Coruña (PONT/RIBE)	1	KP266076-KP266077	1	DQ491780	DQ491883
25	Spain: Abelleira, A Coruña (COMP)	2	KP266078-KP266081	2	DQ491781-491782	DQ491884-491885
26	Spain: Noceda, Pontevedra (LALJ)	2	KP266082-KP266085	1	DQ491783	DQ491886
27	Spain: Alto da Ermida, Ourense (OURE)	2	KP266086-KP266089	1	DQ491784	DQ491887
28	Spain: Alfonxe, Lugo (LUGO)	1	KP266090-KP266091	1	DQ491785	DQ491888
29	Spain: Villacandide, Asturias (NAVI)	2	KP266092-KP266095	1	DQ491786	DQ491889
30	Spain: Esqueiro, Asturias (OVIE)	1	KP266096-KP266097	1	DQ491787	DQ491890
31	Spain: Carracedelo, León (PONF)	2	KP266098-KP266101	2	DQ491788-491789	DQ491891-491892
32	Spain: Destriana, León (LEON)	2	KP266102-KP266105	1	DQ491790	DQ491893
33	Spain: Sanabria, Zamora (SANA3)	1	KP266106-KP266107	1	DQ491791	DQ491894
34	Spain: Lubián, Zamora (SANA1-2)	1	KP266108-KP266109	1	DQ491792	DQ491895
35	Spain: Sarracín de Aliste, Zamora (CULB2-3)	-		1	DQ491793	DQ491896
36	Spain: Tábara, Zamora (CULB1)	1	KP266110-KP266111	1	DQ491794	DQ491897
37	Spain: Saelices el Chico, Salamanca (RODR)	2	KP266112-KP266115	1	DQ491795	DQ491898
38	Spain: Miranda del Castañar, Salamanca (BEJR)	2	KP266116-KP266119	1	DQ491796	DQ491899

39	Spain: Navatalgordo, Ávila (AVIL7-8)	2	KP266120-KP266123	1	DQ491797	DQ491900
40	Spain: Navas del Marqués, Ávila (AVIL4-5)	-	-	2	DQ491798-491799	DQ491901-491902
41	Spain: Robledo de Chavela, Madrid (MADR12-13)	-	-	2	DQ491800-491801	DQ491903-491904
42	Spain: Navas del Rey, Madrid (MADR2)	-	-	1	DQ491802	DQ491905
43	Spain: Navalagamella, Madrid (MADR10)	-	-	1	DQ491803	DQ491906
44	Spain: Guadarrama, Madrid (MADR19-20)	1	KP266124-KP266125	2	DQ491804-491805	DQ491907-491908
45	Spain: Sierra de la Higuera, Madrid (TALV4-5)	2	KP266126-KP266129	1	DQ491806	DQ491909
46	Spain: El Real de San Vicente, Toledo (TALV1)	-	-	1	DQ491807	DQ491910
47	Spain: San Pablo de los Montes, Toledo (TOLD)	2	KP266130-KP266133	1	DQ491808	DQ491911
48	Spain: Garciaz, Cáceres (TRUJ)	2	KP266134-KP266137	1	DQ491809	DQ491912
49	Spain: Garganta la Olla, Cáceres (PLAS)	3	KP266138-KP266143	2	DQ491810-491811	DQ491913-491914
50	Spain: Valverde del Fresno, Cáceres (FRES)	2	KP266144-KP266147	1	DQ491812	DQ491915
51	Spain: Acebo, Cáceres (ACEB)	-	-	2	DQ491813-491814	DQ491916-491917
52	Spain: Gata, Cáceres (GATA)	-	-	2	DQ491815-491816	DQ491918-491919
53	Spain: Villasbuenas de Gata, Cáceres (VGAT)	-	-	2	DQ491817-491818	DQ491920-491921
54	Spain: Montánchez, Cáceres (SPEI)	3	KP266148-KP266153	2	DQ491819-491820	DQ491922-491923
55	Spain: Villar del Rey, Badajoz (BADJ)	2	KP266154-KP266157	2	DQ491821-491822	DQ491924-491925
56	Spain: Pañares, Badajoz (SMOR)	1	KP266158-KP266159	1	DQ491823	DQ491926
57	Spain: Fuentes de León, Badajoz (ARAC4)	1	KP266160-KP266161	1	DQ491824	DQ491927
58	Spain: Cortelazor, Huelva (ARAC1-2)	-	-	2	DQ491825-491826	DQ491928-491929
59	Spain: Linares de la Sierra, Huelva (ARAC3)	-	-	1	DQ491827	DQ491930
60	Spain: Calañas, Huelva (CALÑ)	1	KP266162-KP266163	1	DQ491828	DQ491931
61	Spain: Doñana, Huelva (DOÑA)	1	KP266164-KP266165	4	DQ491829-491832	DQ491932-491935
62	Spain: El Pedroso, Sevilla (SEVL1)	1	KP266166-KP266167	2	DQ491833-491834	DQ491936-491937
63	Spain: Constantina, Sevilla (SEVL2-3)	-	-	1	DQ491835	DQ491938
64	Spain: Almadén, Ciudad Real (ALMD)	2	KP266168-KP266171	2	DQ491836-491837	DQ491939-491940
65	Spain: Villanueva de Córdoba, Córdoba (VCOR)	1	KP266172-KP266173	1	DQ491838	DQ491941
66	Spain: Venta del Charco, Córdoba (CORD)	1	KP266174-KP266175	1	DQ491839	DQ491942
67	Spain: Azuaya, Badajoz (AZUA)	-	-	1	KP266300	KP266278
68	Portugal: Mira (MORA)	2	KP266176-KP266179	2	KP266301-KP266302	KP266279-KP266280
69	Portugal: Corte Figueira, Almodóvar (CALD)	7	KP266180-KP266193	2	KP266303-KP266304	KP266281-KP266282
70	Portugal: Escoural, Évora (EVOR4-10)	1	KP266194-KP266195	1	KP266305	KP266283
71	Portugal: Charneca da Caparica, Almada (AMAD)	1	KP266196-KP266199	2	KP266306-KP266307	KP266284-KP266285
72	Portugal: Gavião, Nisa (NISA)	4	KP266200-KP266207	1	KP266308	KP266286
73	Portugal: Segura, Idanha-a-Nova (SEGU)	2	KP266208-KP266211	1	KP266309	KP266287
74	Portugal: Água Boa, Tomar (TOMR)	3	KP266212-KP266217	3	KP266310-KP266312	KP266288-KP266290
75	Portugal: Santiais, Pombal (POMB)	3	KP266218-KP266223	2	KP266313-KP266314	KP266291-KP266292
76	Portugal: Carvalhais, Figueira da Foz (FFOZ)	6	KP266224-KP266235	3	KP266315-KP266317	KP266293-KP266295
77	Portugal: Mira (MIRA)	3	KP266236-KP266241	2	KP266318-KP266319	KP266296-KP266297
78	Portugal: Albergaria das Cabras, Arouca (AROU)	2	KP266242-KP266245	-	-	-
79	Spain: Formillos de Fermoselle, Zamora (FERM)	2	KP266246-KP266249	2	KP266320-KP266321	KP266298-KP266299
80	Spain: Samir de los Caños, Zamora (ZAMO)	2	KP266250-KP266253	-	-	-
81	Spain: Penoselo, León (ANCA)	2	KP266254-KP266257	-	-	-
82	Spain: Tenebrón, Salamanca (CIUD)	1	KP266258-KP266259	-	-	-
83	Spain: Robledillo de Gata, Salamanca (ROBL)	4	KP266260-KP266267	-	-	-
84	Spain: Aceituna, Cáceres (ACEI)	3	KP266268-KP266273	-	-	-
85	Spain: Puebla de Don Rodrigo, Ciudad Real (REAL)	1	KP266274-KP266275	-	-	-
86	Spain: Huelva (HUEL)	1	KP266276-KP266277	-	-	-
	Total number of individuals	149		125		

H14	..	G	.G	C	.A	..	T	C	.A	..	T	..	T	..	A	..	T	..	A	..	T	..	C	..	1
H15	..	G	.G	C	.A	..	T	.T	.A	..	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	26
H16	..	G	.G	C	.A	..	T	.T	.C	..	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	1
H17	..	G	.G	C	.A	..	T	.T	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	1
H18	..	G	.G	C	.A	..	T	.T	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	5
H19	..	T	.A	G	C	.A	..	T	.T	..	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	43
H20	..	T	.A	G	C	.A	..	T	.T	..	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	1
H21	..	T	.A	G	C	.A	..	T	.T	..	T	..	.A	..	.G	..	.A	..	.C	..	.A	..	.C	..	2
H22G	C	.A	8
H23G	C	.A	..	T	1
H24G	C	.A	..	T	1
H25G	C	.A	..	T	1
H26G	C	.A	..	T	1
H27G	C	.A	..	T	..	.C	1
H28G	C	.A	..	T	1
H29G	C	.A	..	T	1
H30G	C	.A	1
H31G	C	.A	..	T	..	.C	.A	45
H32G	C	.A	..	T	.G	..	.C	.A	1
H33	C	..	.G	C	.A	..	T	..	.C	.A	23
H34	C	..	.G	C	.A	..	T	..	.C	.A	6
H35	C	..	.G	C	.A	..	T	..	.C	.A	2

(b)

H6	26
H30	TA	.GA	.TA	A C I
H28G	C	.A	..	TA	.GA	.TA	A C II

