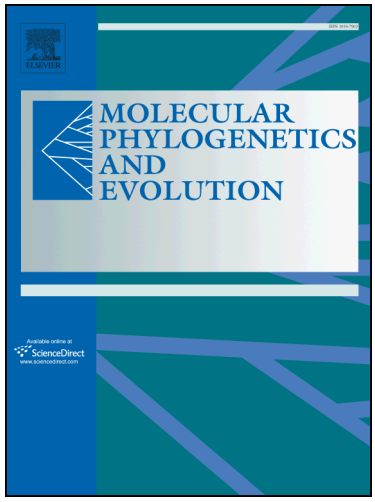


Integrative inference of population history in the Ibero-Maghrebian endemic *Pleurodeles waltl* (Salamandridae)

Jorge Gutiérrez-Rodríguez, A. Márcia Barbosa, Í ñigo Martínez-Solano

PII: S1055-7903(17)30291-9
DOI: http://dx.doi.org/10.1016/j.ympev.2017.04.022
Reference: YMPEV 5808

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 5 July 2016
Revised Date: 1 March 2017
Accepted Date: 4 April 2017

Please cite this article as: Gutiérrez-Rodríguez, J., Márcia Barbosa, A., Martínez-Solano, I., Integrative inference of population history in the Ibero-Maghrebian endemic *Pleurodeles waltl* (Salamandridae), *Molecular Phylogenetics and Evolution* (2017), doi: http://dx.doi.org/10.1016/j.ympev.2017.04.022

Integrative inference of population history in the Ibero-Maghrebian endemic *Pleurodeles waltl* (Salamandridae)

Jorge Gutiérrez-Rodríguez[1], A. Márcia Barbosa[2] & Íñigo Martínez-Solano[1,3,4,5*]

1 Museo Nacional de Ciencias Naturales, CSIC, c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain.

2 Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO/InBIO) – Universidade de Évora, 7004-516 Évora, Portugal

3 Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), Ronda de Toledo, s/n, 13071 Ciudad Real, Spain

4 CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, R. Padre Armando Quintas, 4485-661 Vairão, Portugal

5 Ecology, Evolution, and Development Group, Department of Wetland Ecology, Doñana Biological Station, CSIC, c/ Américo Vespucio, s/n, 41092 Seville, Spain

Keywords: mtDNA, microsatellites, species distribution models, Holocene, Pliocene, Pleistocene, Guadalquivir river

*corresponding author: Museo Nacional de Ciencias Naturales, CSIC, c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain. E-mail: inigomsolano@gmail.com, Phone: (+34) 91 411 13 28, Fax: (+34) 91 564 50 78

Running title: Integrative phylogeography of *P. waltl*

Abstract

Inference of population histories from the molecular signatures of past demographic processes is challenging, but recent methodological advances in species distribution models and their integration in time-calibrated phylogeographic studies allow detailed reconstruction of complex biogeographic scenarios. We apply an integrative approach to infer the evolutionary history of the Ribbed Newt (*Pleurodeles waltl*), an Ibero-Maghrebian endemic with populations north and south of the Strait of Gibraltar. We analyzed an extensive multilocus dataset (mitochondrial and nuclear DNA sequences and ten polymorphic microsatellite loci) and found a deep east-west phylogeographic break in Iberian populations dating back to the Plio-Pleistocene. This break is inferred to result from vicariance associated with the formation of the Guadalquivir river basin. In contrast with previous studies, North African populations showed exclusive mtDNA haplotypes, and formed a monophyletic clade within the Eastern Iberian lineage in the mtDNA genealogy. On the other hand, microsatellites failed to recover Moroccan populations as a differentiated genetic cluster. This is interpreted to result from post-divergence gene flow based on the results of IMa2 and Migrate analyses. Thus, Moroccan populations would have originated after overseas dispersal from the Iberian Peninsula in the Pleistocene, with subsequent gene flow in more recent times, implying at least two trans-marine dispersal events. We modeled the distribution of the species and of each lineage, and projected these models back in time to infer climatically favourable areas during the mid-Holocene, the last glacial maximum (LGM) and the last interglacial (LIG), to reconstruct more recent population dynamics. We found minor differences in climatic favourability across lineages, suggesting intraspecific niche conservatism. Genetic diversity was significantly correlated with the intersection of environmental favourability in the LIG and LGM, indicating that populations of *P. waltl* are genetically more diverse in regions that

have remained environmentally favourable through the last glacial cycle, particularly southern Iberia and northern Morocco. This study provides novel insights into the relative roles of geology and climate on the biogeography of a biodiversity hotspot.

Introduction

Phylogeography aims to reconstruct the evolutionary history of taxa, both as a chronicle -- a narrative detailing major events and their relative chronological order -- and as a history, inferring the causal connections linking those events (Buckley, 2009). Different aspects of past population histories, such as population fragmentation, expansion/contraction, dispersal, or selection, can be inferred from the molecular signatures of past evolutionary processes in the genome (Nielsen, 2005; Arenas et al., 2011). However, in species with deep evolutionary histories, unraveling the relative order of events occurring at different temporal depths from their molecular signatures is often challenging, because different processes can produce a similar trace on the genome, and demographic processes often erase the signals of past events (Ellegren and Galtier, 2016).

Reconstructing an objective, robust narrative and inferring the underlying history requires an integrative, iterative research program in which the study of organismal biology has a central role (Buckley, 2009). In this respect, phylogeography has proven very successful in integrating information from a variety of sources, most recently including environmental data. This, coupled with recent methodological advances in species distribution models (SDMs) such as the application of fuzzy logic (Barbosa and Real, 2012; Barbosa, 2015), has greatly improved our understanding of ecological and climate correlates of species occurrence (Soberón and Peterson, 2005; Svenning et al., 2011; Takahashi et al., 2014). Projected backwards in time via models of past climate, SDMs have been used to produce estimates of potential species ranges during the Last Interglacial (LIG, ~130 ka BP) and the Last Glacial Maximum (LGM, ~21 ka BP) (Acevedo et al., 2012; Smith et al., 2013; Varela et al., 2010). Until recently, paleodistributions of animal species could only be inferred from scarce and often incomplete fossil records, but paleodistributions can now be used to validate projections

of SDMs to the past, opening new and exciting avenues in historical biogeography (Malaney et al., 2012; Gavin et al., 2014). Furthermore, in combination with genetic data, SDMs can help address problematic issues in phylogeographic inference, such as the detection of lineage extinction or the characterization of the routes and geographic extent of species range shifts (Weisrock and Janzen, 2000; Schmitt et al., 2014). The integration of SDMs in phylogeographic studies allows the identification of climatically stable areas through time, which may have served as reservoirs of genetic diversity (glacial refugia), thus producing more robust hypotheses of the evolutionary history of entire biotas (Demos et al., 2014).

The Mediterranean Basin is one of the major global biodiversity hotspots (Myers et al., 2000), and a significant part of its extraordinarily high level of endemism has been attributed to the action of geological and climatic changes through time, especially since the Miocene, when many of the extant taxa originated. For instance, the migration of the Corsica-Sardinia microplate and the fragmentation of the Betic-Rifean Massif had profound impacts on the origin of new taxa (Bidegaray-Batista and Arnedo, 2011; Maia-Carvalho et al., 2014). Later, during the Quaternary, successive climatic oscillations produced cycles of range contraction and expansion in the populations of many species in temperate areas, triggering speciation and extinction (Hewitt, 2000; Bennett and Provan, 2008). In Europe, major glacial refugia have been identified within each of the main southern peninsulas. These areas tend to have not only higher species richness and endemicity, but also higher intraspecific genetic diversity, reflected by parameters such as allelic richness and nucleotide diversity (Ursenbacher et al., 2015; Wielstra et al., 2015).

Reconstructing the evolutionary history of species originating in the Miocene implies unraveling the relative order of events from their molecular signatures over a deep temporal framework, and thus requires comprehensive geographic sampling and the use of multiple

unlinked genetic markers, with different modes of inheritance and mutation rates, to detect genetic signatures of demographic events and processes occurring at multiple temporal scales. For instance, mtDNA can reflect relatively deep phylogeographic breaks, whereas microsatellites are useful to describe more recent population structure. In addition, a growing number of studies are combining SDMs and phylogeographic inference to identify climatically stable areas that may have acted as refugia over the last glacial cycles. However, relatively few such integrative studies have focused on the Iberian Peninsula, which is otherwise well known from a comparative phylogeographic perspective (Gómez and Lunt, 2007).

In this study, we combine SDMs and molecular data to identify glacial refugia and reconstruct the population history of the Ribbed Newt *Pleurodeles waltl* Michahelles, 1830 (Amphibia, Caudata, Salamandridae), an Ibero-Maghrebian endemic with populations north and south of the Strait of Gibraltar. Previous phylogeographic studies have identified two divergent mitochondrial lineages within this species: one distributed across western and central Iberia, and another in eastern Iberia and northern Morocco (Batista et al., 2004; Carranza and Arnold, 2004; Veith et al., 2004). Different hypotheses differing in the estimated temporal depth of the split have been proposed to explain this phylogeographic break. Furthermore, the origin of north African populations, which are genetically very similar to southern Iberian populations, is unclear, which has led to the proposal of competing hypotheses, including recent human-mediated dispersal vs a natural, overseas dispersal colonization event (Batista et al., 2004; Carranza and Arnold, 2004; Veith et al., 2004). Here we analyze data from the nuclear and mitochondrial genomes in a geographically comprehensive sample, covering the complete distribution range of this species, and readdress biogeographic hypotheses through time-calibrated genealogies, continuous diffusion

phylogeographic models, and estimates of historical rates of gene flow between southern Iberia and north Africa. In addition, we characterize the environmental niche and population structure of the species, explicitly testing for ecological differentiation across mtDNA lineages. Finally, we identify potential refugial areas as the intersection between climatically favorable areas through last glacial and interglacial periods, and test for their correlation with different estimates of genetic diversity (nucleotide diversity, allelic richness, average number of private alleles, and heterozygosity), in order to reconstruct the historical demography and postglacial recolonization history of this species throughout its deep evolutionary history.

Methods

Sampling and DNA extraction, amplification and sequencing

We sampled 478 individuals of *P. waltl* from 55 localities across the species range (Fig. 1A, Table 1). Samples included tail tips of adults or larvae, which were released back in the place of capture. Tissues were preserved in absolute ethanol and kept in a freezer at -20 °C. Genomic DNA was later extracted with NucleoSpin Tissue-Kits (Macherey-Nagel).

We amplified and sequenced a mitochondrial fragment (856 bp) of the NADH dehydrogenase subunit 4 gene and adjacent tRNAs in 240 specimens from the 55 localities sampled (Fig. 1A, Table 1). We refer to this combined fragment (ND4+tRNAs) as ND4 throughout the text. We used primers ND4 and Leu (Arévalo et al., 1994), with PCR cycling conditions including an initial denaturation (94°C, 5 min), 40 cycles of denaturation (94°C, 30 s), annealing (53°C, 30 s), and extension (72°C, 1:30 min), and a final extension step (72ºC, 10 min). PCR reactions were performed in a 25 µl volume containing 25 ng of template DNA, 0.2 µM of each primer, 0.4 mM dNTP, 1 mM $MgCl_2$ and 2.5 µl of 5x GoTaq Flexi buffer

(PROMEGA) and 0.5 U GoTaq Flexi DNA polymerase (PROMEGA). PCR products were visualized in 1% agarose gels to verify amplification and possible contaminations through the use of negative controls. Positive bands were precipitated with sodium acetate and ethanol, and then re-suspended in 20 µl of ultrapure water prior to sequencing in an automatic sequencer ABI PRISM 3730. Since some relevant nodes remained unresolved in the ND4 genealogy (see Results), we amplified and sequenced four additional mitochondrial regions: 16S (561 bp, using primers 16Sar and 16Sbr, Palumbi, 1996), D-loop (778 bp, with primers PRO and PHE, Pereira et al., 2016), cytochrome oxydase subunit I (604 bp, with primers AnF1 and AnR1, Jungfer et al., 2013) and cytochrome b (1,069 bp, with primers MNCN-Glu F and Amp-P10 R, San Mauro et al., 2004), in a subsample of 11 *P. waltl* representing all major lineages in the ND4 genealogy plus one representative of the related *P. poireti* from Algeria. Sequences from the only other extant species in the genus, *P. nebulosus*, were downloaded from GenBank to complete this five-gene dataset (Accession: EU880329, this sample was incorrectly attributed to *P. poireti* in Zhang et al., 2008, but can be assigned to *P. nebulosus* based on collection locality data, see Escoriza et al., 2016).

Four nuclear DNA regions were amplified and sequenced for *P. poireti*, *P. nebulosus* and several representatives of the two major mitochondrial lineages in *P. waltl* (see Table S1 for details): chemokine receptor 4 gene (CXCR4, 507 bp, with the primers Cxcr4 F and Cxcr4 R, see sequences and PCR conditions in Nadachowska and Babik, 2009), recombination-activating protein 1 (RAG1, 900 bp, with the primers Amp-RAG1 F and Amp-RAG1 R1, San Mauro et al., 2004), solute carrier family 8 (sodium/calcium exchanger) member 3 (SLC8A3, 536 bp, with the newly designed primers SLC8A3F: 5'-GCA GAC AGG CGG TTA CTT TT-3' and SLC8A3R 5'-AGG ATG GCT CTG CCA ATG3') and sodium-calcium exchanger 1 (NCX1, 672 bp, with the primers Naca-L and Naca-P, Roelants and Bossuyt, 2005).

Microsatellites (SSR)

We genotyped 459 samples from 42 populations (Fig. 1C; Table 1) with ten previously published highly polymorphic microsatellites (Van de Vliet et al., 2009). These were grouped into three multiplex reactions: multiplex 1 (loci Ppl2, Ppl3, Ppl5), multiplex 2 (Ppl1, Ppl12, Ppl13, Ppl14) and multiplex 3 (Ppl6, Ppl7, Ppl10). Multiplex reactions were performed with Type-it Microsatellite PCR kits (Qiagen) following protocols described in Gutiérrez-Rodríguez et al. (2014). Genotyping was performed on an ABI PRISM 3730 sequencer with the GeneScan 500 LIZ size standard (Applied Biosystems), and peaks were scored manually using GENEMAPPER v4.0 (Applied Biosystems).

Analysis of mtDNA data

Sequences were edited with SEQUENCHER v4.10.1 and aligned manually in MEGA v6.06 (Tamura et al., 2013). The number of haplotypes was estimated using DNASP v5 (Librado and Rozas, 2009). Haplotype networks were generated with HAPLOVIEWER (available from: http:// www.cibiv.at/~greg/haploviewer), using a neighbor-joining tree reconstructed with PAUP. Nucleotide diversity (π) was calculated by grouping populations with a buffer of 30 km in the Iberian Peninsula and 20 km in Morocco, to increase sample size while maintaining a regular spatial sampling scheme. This analysis was carried out using SPADS v1.0 (Dellicour and Mardulyn, 2014), and the results were interpolated in ARCGIS 10 (ESRI Inc., Redlands, CA, USA) with the universal kriging function and a spherical semivariogram model. Average genetic distances (p-uncorrected) between and within lineages were calculated with MEGA. We tested for signs of demographic expansion in *P. waltl* and all

major mtDNA clades recovered in the five-gene dataset (see Results): Atlantic, Mediterranean, southern, and Morocco, by calculating mismatch distributions, Fu' Fs (Fu, 1997) and Tajima's D (Tajima, 1989) in DNASP. Significance was assessed through coalescent simulations with 10,000 replicates.

BEAST analyses

We used *BEAST to infer speciation times in *Pleurodeles*, including the deep intraspecific split in *P. waltl* (see Results), and estimate nucleotide substitution rates in the different genomic regions analyzed (five mitochondrial and four nuclear DNA genes), in particular in the ND4 partition for subsequent use in continuous diffusion analyses (see below). *BEAST simultaneously infers the species tree and the underlying gene trees based on the multispecies coalescent model (Heled et al., 2010). In order to calibrate the species tree, and considering that the closest extant relative of *Pleurodeles* is the clade formed by *Tylototriton* and *Echinotriton* (e.g., Marjanović & Laurin, 2013) and the limited availability of genetic data for the latter taxa, we used *Tylototriton* as the outgroup. GenBank accession numbers of newly generated sequences as well as those downloaded from GenBank for the analyses are listed in Table S1. The cladogenesis between *Pleurodeles* and *Tylototriton + Echinotriton* can be assigned a minimum age by assuming that the fossil taxon *Chelotriton* is a member of the group *Tylototriton + Echinotriton* (Marjanović and Laurin, 2013). Specifically, the minimum age would be the middle Eocene age of *Chelotriton robustus* at around 41 million years ago (Mya) (Marjanović & Laurin, 2013; Marjanović & Witzmann, 2015). While there is considerable uncertainty about the maximum age of this split, because the fossil record of the Cretaceous and Paleocene is not dense enough, we tentatively considered the split between Salamandrinae and Pleurodelinae at 66 Mya based on Marjanović and Laurin (2013, Fig. 2).

To apply this calibration, we specified a lognormal prior for the age of the root of the species tree, encompassing values between 41 and 66 Mya (mean: 53.0 Mya; st. dev.: 0.12; offset: 0; mean in real space=true). We selected the optimal nucleotide substitution model for each of the nine gene partitions based on JMODELTEST v2.5 (Darriba et al., 2012), choosing the best-ranked model that was also available in BEAST. We did not explore codon-partitioning strategies in this analysis to avoid overparametrization. Analyses were run under a strict molecular clock and the Yule speciation model. We also ran analyses assuming a relaxed (uncorrelated-lognormal) molecular clock for all genetic regions, but for all partitions the corresponding 95% highest posterior density intervals for the parameter "coefficient of Variation" included zero, suggesting that a strict clock model provides a better fit to the data. Convergence was assessed through inspection of the logfile in TRACER v1.6 (Rambaut et al., 2014), and a Maximum Clade Credibility Tree was reconstructed with TreeAnnotator (distributed as part of the BEAST package).

We used phylogeographic continuous diffusion models as implemented in BEAST v1.8 (Drummond et al., 2012) to reconstruct range dynamics through time in *P. waltl* based on mtDNA sequences. This analysis simultaneously reconstructs gene trees, ancestral population sizes, and the geographic ranges of inferred nodes, which can be visualized through time provided appropriate calibrations or substitution rates are supplied. For this analysis, we used all 240 ND4 *P. waltl* sequences and selected the optimal partitioning scheme and optimal nucleotide substitution models with PARTITIONFINDER v1.1.1 (Lanfear et al., 2012). Geographic coordinates for each sequence were included, with a jitter module generating random coordinates (window size: 0.1) for individuals from the exact same location. We chose this window size to be small enough that it was robust to the low overall dispersal and high genetic structure in the species and at the same time didn't impose

computational constraints associated with extremely low values. Analyses were run under a strict molecular clock, with the clock rate specified based on the results of the previous analysis (lognormal distribution; mean=0.006 substitutions/site/million years; st. dev.: 0.3; offset: 0; mean in real space=true), and using the Bayesian Skyline plot (linear, number of groups=5) as the coalescent prior. The Brownian model was used to describe the geographic diffusion process through time across branches of the inferred gene tree (Lemey et al., 2010). Convergence was assessed through inspection of the log file in TRACER v1.6, and a MCCT was reconstructed with TreeAnnotator. This tree was used to generate time-calibrated reconstructions of the diffusion process using the modules "Continuous Tree" and "Time Slicer" in SPREAD3 (Bielejec et al., 2016). BEAST analyses were run in the Cipres Science Gateway (Miller et al., 2010).

Analysis of microsatellite data

The presence of null alleles, stuttering and large allele dropout in each population was tested using MICROCHECKER v2.2.3 (Van Oosterhout et al., 2004), with a 99% confidence interval and 1,000 randomizations. We also carried out a genetic parentage analysis to eliminate possible full-siblings in each population using COLONY v2.0.5.1 (Jones and Wang, 2010). Colony analyses were performed assuming a mating system with monogamous females and polygamous males, implementing the full-likelihood method of Wang (2004) with 10 independent, medium-length runs. Deviations from Hardy-Weinberg equilibrium (HWE) and evidence of linkage disequilibrium (LD) were tested between all pairs of loci using GENEPOP on the web (http://genepop.curtin.edu.au; Raymond and Rousset, 1995). Significance values for all multiple tests were evaluated by applying a sequential Bonferroni correction (Rice, 1989).

We estimated the number of alleles (Na), and both observed (Ho) and expected heterozygosity (He) for each locus and population with GENALEX v6.5b5 (Peakall and Smouse, 2012). Allelic richness (Ar) and number of private alleles (P) for each population were calculated with ADZE (Szpiech et al., 2008). This program uses a rarefaction approach to allow the comparison of populations with different sample sizes. Ar, P, Ho and π were spatially interpolated using ARCINFO (ESRI, Redlands, CA, USA) with the universal kriging function and a spherical semivariogram model. Subsequently, we performed a principal components analysis on the four genetic diversity variables with PAST v3 (Hammer et al., 2001). Values for the first axis (PC1) were checked to correlate positively with all individual measures of genetic diversity, and then interpolated across the species' range following the steps above.

We used the individual-based Bayesian clustering method implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000) to describe range-wide patterns of genetic structure based on microsatellite data. STRUCTURE analyses consisted of ten independent runs for K=1 to 20 clusters (each run consisting of a burn-in of 500,000 generations followed by 1,000,000 MCMC iterations), assuming the admixture model with correlated allele frequencies between populations and the LOCPRIOR option off (Falush et al., 2003). The optimal number of genetic clusters was estimated through the ΔK method (Evanno et al., 2005) using STRUCTURE HARVESTER 0.6.94 (Earl and vonHoldt, 2012), although other K values showing consistent, biologically significant results were also considered for further discussion. Average admixture coefficients of optimal clusters were calculated with CLUMPP v1.1 (Jakobsson and Rosenberg, 2007). Graphs of assignment probabilities were produced with DISTRUCT v1.1 (Rosenberg, 2004).

We tested for evidence of recent demographic changes in each of the five major microsatellite clusters identified in previous analyses using the k and g tests (Reich et al., 1999). The k-test is based on the observed distribution of allele size within locus, while the g-test is based on allele length variance across loci (Reich et al. 1999). Analyses were carried using the kgtests Excel macro program (Bilgin, 2007).

Historical migration between North Africa and the Iberian Peninsula

In order to elucidate the evolutionary history of populations currently separated by the Strait of Gibraltar, historical migration analyses were carried out to confront the "retention of ancestral polymorphism" vs "gene flow" hypotheses between populations on opposite sides of the Strait. Historical migration rates ($M = m/\mu$) and effective population sizes ($\Theta = 4N_e\mu$) were estimated using MIGRATE-N v3.6.4 (Beerli, 2009). Three different migration models were tested: 1) unidirectional migration from southern Iberia to Morocco; 2) unidirectional migration from Morocco to southern Iberia; and 3) panmixia. These analyses were performed by grouping microsatellite genotypes from all Moroccan samples into one population (Morocco), whereas the Iberian population included all individuals in the Iberian South microsatellite cluster (the most closely related group, see BEAST Results in Fig. 2A). Analyses were run under the Bayesian approach (Beerli, 2006; Beerli and Felsenstein, 2001), assuming the Brownian motion mutation model with constant mutation across all loci. Uniform priors were chosen for Θ (Min. = 0, Max. = 1000, Delta = 100) and M (Min. = 0, Max. = 1000, Delta = 100). Ten replicates were run for each of the three models, starting with a random UPGMA tree and calculating the initial values of M and Θ from F_{ST} estimates. The runs consisted of one long chain with 2,500,000 steps, discarding 100,000 as a burn-in, with a static heating scheme using four chains. The best model was chosen based on Bezier marginal

likelihoods and applying the Bayes factor approach (Beerli and Palczewski, 2010; Kass and Raftery, 1995).

An alternative approach to estimating gene flow and effective population sizes among groups was also performed based on the isolation with migration model (IM: Hey and Nielsen, 2004) using IMA2 (Hey, 2010; Hey and Nielsen, 2007). We used the same groups as in the Migrate analyses, but with a combined microsatellite + mitochondrial dataset. The stepwise-mutation model (SMM: Kimura and Ohta, 1978) was used for microsatellites (mean mutation rate: $5x10^{-4}$), and the HKY model of DNA substitution (Hasegawa et al., 1985) was used for ND4 sequences. Two independent runs with different seed numbers and 20 heated metropolis-coupled Markov chains (a= 0.96 and b= 0.9) were perfomed. These runs consisted of a varying burn-in period of at least 5 million steps, followed by 15 million steps, with sampling every 100 steps. The effective population size of the ancestral (q_0), Morocco (q_1) and southern Iberian Peninsula (q_2) groups, the population migration rate (m_{1-2} and m_{2-1}), and the time of divergence (t) between groups were estimated. The maximum prior values for the parameters were q = 200, m = 1.0 and t = 200. Finally, a likelihood-ratio test was carried out to determine the statistical significance of inferred migration rates between groups.

Climatic favourability models

To infer climatically favourable areas (potential refugia and diversification/dispersal centres) under current and past climate, we built SDMs based on the current occurrence patterns of *P.waltl* and of each of its major mitochondrial lineages separately. Predictor variables were obtained from the WorldClim data set (Hijmans et al., 2005) for current climate, for the LIG, and for the three climatic simulations currently available for both the LGM and the Mid Holocene based on the same variables used for model building: CCSM4, MIROC-ESM, and

MPI-ESM-P . The dataset included 19 bioclimatic variables with potential direct or indirect relationships with species occurrence (Table S2).

The modelling area included the entire distribution range of the species, spanning three countries (Portugal, Spain and Morocco; Fig. 1). The modelling units were 10 km x 10 km UTM cells, matching most national atlases of amphibian distributions (e.g., Pleguezuelos et al., 2002; Loureiro et al., 2008). While this spatial resolution is relatively coarse given the limited mobility of the studied species, it is appropriate for the spatial extent and the climatic data under analysis, and the best compromise for minimizing positional error in occurrence records while encompassing the species' distribution range. Vector maps of countries and of the UTM grids were downloaded from the EDIT geoplatform (Sastre et al., 2009) and clipped to the study area. Distribution data from Portugal (Loureiro et al., 2008) and Spain (MAGRAMA, 2015) were imported directly from the source databases. Occurrence points in the Moroccan herpetological atlas (Bons and Geniez, 1996) were digitized, geo-referenced and intersected with the UTM grid. Additional occurrence data from Morocco were obtained from Beukema et al. (2013). We further completed the dataset by adding the locations of our own samples (Table 1).

For modelling the distribution of each lineage while taking advantage of the previously published, though not genetically identified distribution data (Bons and Geniez, 1996; Loureiro et al., 2008; Beukema et al., 2013; MAGRAMA, 2015), we needed to assign available presence records to a lineage wherever possible. We used the *phylin* R package (Tarroso et al., 2015) to compute a genetic semi-variogram based on our sampled populations (31 Eastern and 26 Western lineage occurrences), fit a spherical model, and calculate the potential distribution of each lineage, taking into account spatial autocorrelation and both the geographic and genetic distance between samples. *P. waltl* records with at least 0.8

probability (i.e., odds of at least 4:1) of belonging to one of the lineages were then assigned to that lineage.

Models were built in R v3.0.2 (R Core Team, 2013) with the *fuzzySim* package (Barbosa, 2015), using the favourability function (Real et al., 2006), which allows obtaining prevalence-independent values directly comparable across taxa (Acevedo and Real, 2012). Variables with a significant bivariate relationship with the distribution of the species or lineage were first selected based on the false discovery rate (Benjamini and Hochberg, 1995), and then included in multivariate models with a forward-stepwise procedure based on Akaike's Information Criterion. Non-significant variables left in the model were further removed (Crawley, 2007). Model performance was then evaluated with the *modEvA* R package (Barbosa et al., 2013). We measured both discrimination (using 0.5 as the threshold value, which for favourability models equates to using prevalence; Real et al., 2006; Acevedo and Real, 2012) and calibration, i.e. the fit of predicted probabilities to oberved occurrence frequencies (see Jiménez-Valverde et al., 2013).

The models were extrapolated to hindcast climatically favourable areas based on the Worldclim paleoclimatic simulations. Given that favourability values can be handled directly with fuzzy logic (Real et al., 2006; Acevedo and Real, 2012), we used the *fuzzySim* package to calculate the fuzzy intersection of LIG and LGM favourability (Zadeh, 1965), to infer how favourable each cell remained along the glacial cycle. We then calculated the correlations between these sustained favourability values and current genetic diversity.

We mapped and quantified favourability changes between past and current climates, for *P. waltl* and for each mitochondrial lineage, using the fuzzy range change measures (including overall proportional gain, loss and stability) available in *fuzzySim* package. We also compared the favourability patterns obtained for the main lineages, using niche similarity

measures such as Schoener's D and Warren's I (Warren et al., 2008), as well as fuzzy similarity indices adapted from classical binary similarity, including Jaccard's and Baroni-Urbani and Busers's index (Barbosa, 2015). The latter two have the advantage of having associated significance values.

Results

Mitochondrial Analyses

Partial sequences of the ND4 gene plus adjacent tRNAs (856 bp) were obtained for 240 individuals from 55 populations (GenBank accession numbers in Table S1). Twenty-nine different haplotypes were found (Table 1), defined by 72 variable sites, of which 64 were parsimony-informative. Nucleotide diversity ranged from 0 to 0.0031 (Table 1). The reconstructed haplotype network showed two main haplogroups, diagnosing two major lineages in *P. waltl*, henceforth referred to as Western and Eastern (Fig. 1B). The Western lineage is distributed across the Iberian west and center, whereas the Eastern lineage extends across eastern Iberia and northern Morocco (Fig. 1A). Each haplogroup can be subdivided into several geographically consistent sub-haplogroups (Fig. 1A, B). The highest values of nucleotide diversity were found in northern Morocco (Eastern haplogroup) and south-western Iberia (Western haplogroup, populations Membrio and Jerez), and the lowest values occurred across northern, central and eastern Iberia (Fig. S1). Average genetic distance (p-uncorrected) between haplogroups was 5.4%, while average distances within haplogroups were 0.26% (Eastern, range: 0.01-0.8%) and 0.18% (Western, range: 0.1-1.2%). Scenarios of historical demographic expansion were not consistently supported by mismatch distributions or

neutrality tests (Fu' Fs, Tajima's D and R^2) for the species or for any of the main four haplogroups considered separately (Table 2 and Fig. S2).

*BEAST analyses produced a fully resolved tree with high support in all nodes (Bayesian Posterior Probabilities, BPPs = 1; Fig. S3). According to these results, the split between *P. waltl* and *P. nebulosus* + *P. poireti* took place in Miocene (median: 18.00 million years ago -Mya-, 95% highest posterior density interval -HPDi-: 12.66-23.57 Mya). The vicariant event between *P. poireti* and *P. nebulosus* occurred during the Miocene (median: 9.21 Mya, 95% highest posterior density interval -HPDi-: 5.11-13.67 Mya). Finally, the inferred split between the Western and Eastern lineages of *P. waltl* dates back to Pliocene (median: 3.51 Mya, 95% highest posterior density interval -HPDi-: 0.75-5.39 Mya).

The topology of the 5-gene mtDNA dataset recovered in the *BEAST analysis includes two major clades and several well-supported intraspecific subclades in *P. waltl* (Fig. 2A). The Western clade includes two well-supported subclades, corresponding to the Atlantic and Algarve haplogroups. The TMRCA of these subclades dates back to the Pleistocene (median: 0.78 Mya, 95% HPDi: 0.44-1.18 Mya). The Eastern clade is divided into three well-supported subclades (BPP:1): Mediterranean, Southern and Morocco (Fig. 2A). The TMRCA of these clades dates back to the Pleistocene (median: 0.5 Mya, 95% HPDi: 0.29-0.77 Mya). The Southern and Morocco subclades are recovered as sister groups (BPP:0.98). The TMRCA of these clades dates back to the Pleistocene (median: 0.35 Mya, 95% HPDi: 0.19-0.54 Mya). Finally, the four independent nDNA genealogies recovered in the *BEAST analysis recovered the monophyly of *P. waltl* and of *P. poireti* + *P. nebulosus* (BPP = 1.0, Figs. S4-S7). The Eastern and Western clades in *P. waltl* were also recovered as monophyletic, though with low support (BPP<0.90).

Continuous diffusion analyses produced robust results, with ESSs values >200 for all parameters. The ND4 tree is largely consistent with the 5-gene mtDNA topology, except with lower support values at some nodes (Fig. 2B and Fig. S8). The ancestral area of the Western lineage was inferred to be located in the southwestern Iberian Peninsula, with subsequent expansions to the south and the north, reaching the North Plateau in the LIG (88 ka.; Fig. S9). In the Eastern lineage, the inferred ancestral area included southern Iberia and North Africa. The Eastern lineage would have colonized Morocco by 270 ka, and subsequently expanded in Iberia to the north and east (Fig. S9).

Microsatellite Analyses

Potential null alleles were detected in the loci Ppl2 at Peñarroya-Pueblonuevo and Ppl10 at Moulay Bousselhaim. Forty-six individuals were removed after the sibship analysis in COLONY, leaving only one representative per sibship group (Table 3). The observed number of alleles ranged from 1 to 10 per locus and population. Summary statistics of genetic diversity for each population are shown in Table 3. The lowest observed heterozygosities were found north of the Central System mountains, in the North Plateau populations, and also in northeastern Iberia, whereas the highest values were recorded in southern Iberia and Morocco (Fig. S1D). Allelic richness ranged from 1.45 to 6.09, with the highest values found in south-western Iberia and in Morocco (Fig. S1B). The numbers of private alleles were also highest in the vicinity of the Strait of Gibraltar (Fig. S1C).

The first principal component of genetic diversity (PC1) correlated positively with all summarized variables (π, Ho, Ar and P) and accounted for 98.9 % of their variance. Mapping the interpolated values of PC1 across the species' range revealed a decrease of genetic

diversity from south to north, with genetically impoverished populations in the North Plateau and the eastern litoral of the Iberian Peninsula (Fig. 3 and Fig. S1).

STRUCTURE analyses supported an optimal clustering level at K=2, according to the Evanno method (Fig. 1D), with replicate runs producing consistent results (standard deviation of likelihood scores across replicates near zero). At K=2, one cluster is distributed north of the Central System mountains in the North Plateau, whereas the other occupies the rest of the species' range (Fig. 1C). An additional hierarchical level of genetic structuring was detected at K=4, with a substantial increase in ΔK and all 10 replicates producing congruent, biologically meaningful results. At K=4, the previously recovered second cluster was subdivided into three groups: one in the southwest, in the Tagus and Guadiana river basins; another extending throughout the Guadalquivir river basin and Morocco; and a third one in eastern Iberia (Fig. 1C). Most populations had high average assignment probabilities to a single cluster, with the exception of population 8 (Valdemanco), which is admixed, suggesting gene flow between two different clusters. At higher values of K, we recovered finer structuring within groups, including an Algarve genetic cluster appearing in most replicates at K values ≥11.

Based on k-test results, we detected significant signatures of demographic expansions in the North-Plateau (P-value=0.001) and Southwestern (P-value=0.045) clusters and in the Morocco group (P-value=0.008). In contrast, results of inter-locus g-tests were not significant for any genetic cluster.

Historical migration between North Africa and the Iberian Peninsula

The Bayes factor approach implemented in MIGRATE provided best support for the second model, i.e., unidirectional migration from Morocco to the southern Iberian Peninsula (Table

4). Similar results were obtained with IMA2, where the likelihood ratio test (LRT) supports non-zero gene flow from Morocco to the Iberian Peninsula (2Nm = HiPt: 1.893, 95% HPD: 0.8213 – 3.338). Estimates of effective population sizes largely overlapped between Morocco (q_1 = HiPt: 58.10, 95% HPD: 42.90 - 77.90) and southern Iberia (q_2 = HiPt: 57.30, 95% HPD: 42.70 - 76.10). No precise divergence time estimates could be obtained in the analyses due to lack of convergence.

Climatic favourability

The obtained distribution models (Table S3) had a good overall fit to the current distribution data, especially for *P. waltl* and for the Western lineage (Fig. 4). Performance measures reached relatively high scores (Fig. S10), with e.g. areas under the receiver operating characteristic curve (AUC) considered "excellent" for *P. waltl* and the Western lineage and "good" for the Eastern lineage (Swets, 1988), and with McFadden's pseudo-R^2 values reflecting "excellent fit" in all three cases (McFadden, 1978). The Eastern lineage revealed less clear environmental patterns, with several climatically favourable areas outside and detached from its current distribution, but often inside the current distribution of the Western lineage (Fig. 4). Fuzzy intersection between climatic favourability values revealed a range of areas that are highly favourable for both lineages simultaneously (Fig. 4).

Niche similarity measures revealed some affinity between the environmental patterns of the two lineages, with Schoener's D = 0.42 and Warren's I = 0.69 (Warren et al., 2008). Fuzzy comparison of favourability values using adapted binary similarity indices (Barbosa 2015) revealed similarity values compatible with chance (i.e., neither larger nor smaller than expected at random) for Jaccard's similarity index (J = 0.28), and higher than expected by

chance for Baroni-Urbani and Busers's index (B = 0.61), which accounts for shared absences as well as shared presences.

According to the models applied to paleoclimatic simulations (Fig. S111), despite some quantitative differences, past favourability values were significantly correlated with current ones – i.e., the most and least favourable areas, within the available climatic conditions, were largely the same as today, especially during the LGM and the Mid Holocene (Fig. S12). The intersections between favourability in the LIG and under each LGM simulation, indicating the maintenance of favourable conditions through the glacial cycle (Fig. 3), were significantly although not highly correlated with genetic diversity within the species' range; the strongest correlation was obtained with the MIROC-ESM scenario (Pearson's r = 0.56 under MIROC-ESM, r = 0.11 under LGM scenario CCSM4, and r = 0.05 under MPI-ESM-P; $p < 0.001$ with 5540 degrees of freedom).

Favourability change between time periods included both increases and decreases, with eastern Spain generally showing the sharpest changes in climatic favourability across time (Fig. S13). Overall proportional gain, loss, stability and change in favourability between past and current climate are shown in Figure 5.

Discussion

Inference of population history based on molecular markers is challenging, especially in species with deep evolutionary histories, because different processes can leave similar genetic signatures, and recent processes may erase the signal of older events. Integrative approaches combining molecular and environmental data have great potential for solving these problems, at least in part, and for unravelling the relative importance of different events through the evolutionary history of a species. Our approach, based on a combination of time-calibrated

gene tree and ancestral area reconstruction, cluster analysis on genotypic data, estimation of ancestral population sizes and historical patterns of gene flow, and reconstruction of climatic favourability through time, allowed the identification of both paleogeographic and climatic factors involved in the population differentiation of *P. waltl*, with an overall larger importance of the former in shaping current patterns of genetic diversity and structure. Our study provides evidence for key events affecting this species at three major temporal windows. First, a deep vicariant event, dating back to the Pliocene, left a strong genetic signal of population differentiation. Later, in the Pleistocene, trans-marine dispersal events, probably associated with eustatic changes during glacial periods, allowed the colonization of North Africa from the Iberian Peninsula, with at least one additional, more recent event of trans-marine dispersal. Finally, climatic fluctuations in the past 130 kya appear to have had only limited effects on population structure and genetic diversity.

Deep history: Neogene tectonics and relationships between Moroccan and Iberian biotas

The Maghreb has been identified as an important center for species diversification during the Pliocene and Pleistocene, with the area around the Strait of Gibraltar acting as an important biogeographical link between North Africa and Europe at different times (Husemann et al., 2013). Previous studies have revealed complex relationships between species on both sides of the Strait of Gibraltar, but North African populations have often been neglected. Species with a deep evolutionary history that are distributed on both sides of this major barrier, such as *P. waltl*, can thus provide valuable information about the biogeography of this region.

Paleogeographic events in the Pliocene are probably associated with the split between the two main mitochondrial lineages (Western and Eastern) recovered in *BEAST analyses (Figs. 1 and 2). These lineages were identified in previous studies (Carranza and Arnold,

2004; Veith et al., 2004) and are largely separated by the Guadalquivir river basin. Based on our calibration, the intraspecific split in *P. waltl* can be tentatively associated with the paleogeographic events associated with the formation of this basin (Doadrio, 1988). This event has also been linked with phylogeographic splits in other taxa with limited dispersal, including invertebrates (*Carabus*: Andújar et al., 2012), amphibians (*Salamandra*: García-París et al., 1998; *Discoglossus*: Martínez-Solano, 2004), and reptiles (*Podarcis*: Kaliontzopoulou et al., 2011; *Vipera*: Velo-Antón et al., 2012). It is worth noting that nucleotide differences among the two main mitochondrial clades of *P. waltl* involved up to ten fixed non-synonymous substitutions.

Within the two major mtDNA lineages, we also identified several sublineages with a marked geographic pattern. In the case of the Western lineage, two sublineages were recovered: one formed by the population of the Algarve (S Portugal) and the other grouping the remaining populations (Atlantic). The split between them occurred in the Pleistocene (TMRCA median: 0785 Ka, 95%HPDi: 0.44-1.79 Mya). This phylogeographic break is present in many other taxa, including plants (Trindade et al., 2012; Caldas, 2013), invertebrates (Mendes, 1992; Serrano, 1995; Barat, 2013), fish (Mesquita et al., 2005, 2007), amphibians (Martínez-Solano et al., 2006; Gonçalves et al., 2009; Reis et al., 2011), reptiles (Godinho et al., 2008), and mammals (Centeno-Cuadros et al., 2009). This shared break has been associated to the formation of the Serra do Caldeirão mountain range during the lower Pliocene (Gonçalves et al., 2009). In the case of the Eastern lineage, with a TMRCA estimated at a median age of 504 Ka (95%HPDi: 286-773 Ka), three sub-clades were recovered (Mediterranean, Southern and Morocco), with their estimated TMRCAS concentrated within the past 240 Ka.

Using sequences of the ND4 mitochondrial gene, we found that all haplotypes on either side of the Strait of Gibraltar were exclusive to each area (Fig. 1B), in contrast with previous findings of shared haplotypes based on more slowly evolving markers. This led to the hypothesis of a recent anthropogenic introduction from the Iberian Peninsula into Morocco (Carranza and Arnold, 2004) that can be rejected based on the new data, which argue in favor of a relatively old independent history for Moroccan populations. In fact, our time estimates suggest a TMRCA for Moroccan haplotypes of 240 Ka (95%HPDi: 122-394 Ka), in the Pleistocene. Overall similarity between Moroccan and southern Iberian populations based on microsatellite data can be explained by shared history and post-divergence gene flow, as inferred from MIGRATE and IMA2 analyses. Whereas the timing of this event (or events) of post-divergence gene flow could not be estimated in the present study, some opportunities may have existed during the cooler phases of the Pleistocene, when the 14.4 km distance between both sides of the Strait of Gibraltar were reduced to 3.5 km (Pleguezuelos et al., 2008 and references therein) as a consequence of sea level fluctuations. This, together with the reduction of salinity in the Western Mediterranean Sea during the glaciations (Bozec et al., 2007), may have facilitated trans-marine dispersal events through rafts of vegetation, as hypothesized for other amphibians (Bell et al., 2015; Measey et al., 2007; Vences et al., 2003).

Concordance between mtDNA haplogroups and microsatellite clusters

Microsatellite data can reveal finer, more recent population structure than slower-evolving markers and, when combined with mtDNA sequences, detect admixture between divergent lineages, illustrating important aspects of the process of population differentiation and speciation (Dufresnes et al., 2015; Thomé et al., 2016). In *P. waltl*, cluster analyses of

microsatellite genotype data are generally consistent with the major mtDNA divisions, with few exceptions. Structure results for K=4 recovered a North-Plateau and a South-Western cluster (both of which have mtDNA haplotypes in the Western mtDNA lineage), and also a Mediterranean and a Southern cluster, the latter including the Moroccan populations (all individuals in these two clusters carry Eastern mtDNA haplotypes). The North Plateau group emerged as a distinct cluster from K=2. This group is characterized by extremely low levels of genetic diversity and was inferred to have colonized its current range relatively recently, perhaps undergoing bottlenecks, to which Structure seems to be particularly sensitive (Manel et al., 2003). This pattern of low genetic diversity in this area has also been reported in the co-occurring anuran *Pelobates cultripes* (Gutiérrez-Rodríguez et al., 2017). The Algarve mtDNA lineage was recovered as a genetically distinct group, although at higher values of K, in line with the results of van de Vliet et al. (2014), who found significantly reduced cross-amplification success when comparing populations from Algarve (south of the Mira river) with populations further north. The different Iberian groups were largely delimited by major river basins: the North-Plateau group in the Duero Basin, the West-South group in the Tagus and Guadiana basins, the South group in the Guadalquivir Basin, and the Mediterranean group in the eastern basins. Strikingly, based on Structure analyses, populations from Morocco were not recovered as a distinct group. This contrasts with Busack's (1985, 1986) finding of Nei distances of 0.1 between Moroccan and Iberian populations in his comparative allozyme study, which is also consistent with reported morphological differences between Iberian and Moroccan populations (Pasteur, 1958; but see also Beukema et al., 2013). Although sample sizes were low in Busack's study, there were fixed allelic differences at two loci between populations in southern Iberia and Morocco, and calibration of an allozyme-based molecular clock suggested differentiation within the past 2 Mya.

Recent history: the role of glacial events in population structure

The cycling of glacial and interglacial periods during the Pleistocene produced large-scale demographic changes in temperate biotas (Hewitt, 2004). These demographic changes, often involving bottlenecks and founder effects, could potentially erase the signal of deeper evolutionary history, for instance through the random extinction of geographically restricted lineages. This is a limit of traditional phylogeographic inference that can be alleviated with the integration of species distribution models, especially in morphologically and ecologically conservative species. In our study species, the recent evolutionary history, as inferred from the integration of climate-based distribution models and current patterns of genetic diversity, seems to have been characterized by overall stability since the LIG, in contrast with the deeper evolutionary events uncovered by the analyses of molecular data (especially mtDNA). Climatic favourability during the LGM was significantly correlated with current favourability, which, combined with the results of continuous diffusion analyses and population expansion tests, does not provide clear and consistent signs of recent demographic expansions (Fig. S2, Table 2), but rather suggests overall demographic stability associated with climatically favourable conditions throughout the species' range. However, closer inspection of net changes in climatic favourability across the analysed time periods (Fig. 5) shows that moderate decreases in the LIG-LGM transition were followed by larger increases in the LGM-MH, especially for the Eastern mtDNA haplogroup (Fig. S11). Considering the results of tests for signatures of demographic changes based on mtDNA and microsatellites, these increases in favourability through time do not seem to have been tracked by the species or its major genetic clusters with recent demographic expansions (Table 2 and Fig. S2), with the exception of the North-Plateau cluster. This may be the result of more recent declines in

climatic favourability from the MH to the present (Fig. 5), with successions of favourable and unfavourable periods potentially erasing the genetic signatures of previous demographic events.

The intersection of climatic favourability in the LIG and LGM, which highlights steadily favourable areas for the species along the last glacial cycle, was significantly correlated with current genetic diversity in *P. waltl*. This is in spite of wide differences across available paleoclimatic simulations for the LGM. Although the correlation coefficients were not particularly high, they argue in favor of the existence of southern refugia in the Iberian Peninsula that acted as genetic reservoirs through time, with moderate coefficients probably reflecting a longer timeframe for the role of these areas as refugia (Fig. S9).

The current climatic favourability models were well adjusted to the occurrence patterns of the species and major lineages (Fig. S10), with little evidence for substantial differences in favourability among lineages, as shown by niche overlap and fuzzy similarity measures. This results in large areas of overlapping favourability, mainly in the central area of the Iberian Peninsula (Fig. 4), suggesting a high environmental potential for co-occurrence and thus admixture across lineages. Genetic evidence is consistent with this interpretation, as potentially admixed populations (21 (mtDNA-nDNA discordance) and 22 (admixture in Structure analyses); Fig. 1) are mostly located in areas identified as favourable for both lineages (Fig. 4).

The application of species distribution modelling to intraspecific genetic lineages holds great promise for clarifying the role of niche evolution in shaping patterns of population divergence and for predicting future responses to climate change (Gotelli and Stanton-Geddes, 2015; Brown et al., 2016). However, defining lineage boundaries is often complicated, even when genetic sampling is comprehensive. In this respect, the procedure

implemented in the *phylin* package (Tarroso et al., 2014) provided a significant improvement over previously available methods for modelling the distributions of intra-specific lineages or cryptic species (e.g., Real et al., 2005), as it accounts for the autocorrelation structure of genetic information, interpolates genetic data based on a statistical model, estimates the uncertainty of the predictions, and identifies the most likely occurrence areas of each lineage based on a subset of genetically identified individuals. In turn, the procedures implemented in package *fuzzySim* (Barbosa, 2015) make it possible to 1) use multiple criteria simultaneously for selecting the appropriate variables for each model; 2) obtain environmental favourability values directly comparable across lineages and time periods, independently of sampled prevalence (Real et al., 2006); and 3) assess the significance of similarity among favourability patterns without using arbitrary thresholds to separate predicted presences from predicted absences. Thus, it allows direct comparison between continuous model predictions, thus avoiding loss of quantitative information (Barbosa, 2015). Package *modEvA* (Barbosa et al., 2013) provided a unified framework for evaluating models using multiple metrics, not only of discrimination and classification capacity, but also of calibration or reliability of the continuous predictions, which is just as important in assessing model performance (Pearce and Ferrier, 2000; Jiménez-Valverde et al., 2013).

Acknowledgments

We thank I. Acevedo, E. Ayllón, J. Barbadillo, A. Bermejo, D. Buckley, P. Cabezas, M.J. Fernández, I. Gomez-Mestre, E. González, P. Hernández, M. Lapeña, P. Mas, P. Pavón, E. Recuero, R. Ribeiro, A. Sánchez, V. Sancho, and G. Velo-Antón for help during field work or providing samples. D. Buckley, E. Jockusch, R. Pereira and two anonymous reviewers provided valuable feedback on a preliminary draft. R. Pereira kindly shared primer information for SLC8A3, developed while funded by the European Science Foundation

(Frontiers of Speciation Research, Exchange grant 3318), and by the European Commission (Synthesys grant ES-TAF-1486). D. Marjanović provided valuable feedback on molecular clock calibration strategies and the Salamandridae fossil record. JGR was supported by the Consejo Superior de Investigaciones Científicas of Spain (CSIC) and the European Social Fund (ESF) (JAE-pre PhD fellowship). AMB is supported by FCT and FEDER / COMPETE 2020 through contract IF/00266/2013, exploratory project CP1168/CT0001, and funds POCI-01-0145-FEDER-006821 to unit UID/BIA/50027. This research was funded by grants CGL2008-04271-C02-01/BOS, and CGL2011-28300 (Ministerio de Ciencia e Innovación -MICINN-), Ministerio de Economía y Competitividad -MEC-, Spain, and FEDER) to IMS, who was funded by the Severo Ochoa Program (SEV-2012-0262).

References

Acevedo, P., Melo-Ferreira, J., Real, R., Alves, P.C., 2012. Past, Present and Future Distributions of an Iberian Endemic, *Lepus granatensis*: Ecological and Evolutionary Clues from Species Distribution Models. PLoS ONE 7, e51529.

Acevedo, P., Real, R., 2012. Favourability: concept, distinctive characteristics and potential usefulness. Naturwissenschaften 99, 515–522.

Andújar, C., Gómez-Zurita, J., Rasplus, J-Y., Serrano, J., 2012. Molecular systematics and evolution of the subgenus *Mesocarabus* Thomson, 1875 (Coleoptera: Carabidae: *Carabus*), based on mitochondrial and nuclear DNA. Zool. J. Linn. Soc. 166, 787–804.

Arenas, M., Ray, N., Currat, M., Excoffier, L., 2011. Consequences of Range Contractions and Range Shifts on Molecular Diversity. Mol. Biol. Evol. 29, 207–218.

Barat. J., 2013. Revisión de la identidad de *Neocallicrania serrata* (Bolívar, 1885) y descripción de dos táxones afines: *Neocallicrania serrata pfaui* ssp. n. y *Neocallicrania barrosi* sp. n.(Orthoptera, Tettigoniidae, Bradyporinae, Ephippigerini). Boletín de la SEA 52, 1–16.

Barbosa, A.M., Real, R., 2012. Applying Fuzzy Logic to Comparative Distribution Modelling: A Case Study with Two Sympatric Amphibians. The Scientific World Journal 2012, 1–10. doi:10.1100/2012/428206.

Barbosa, A.M., Real, R., Muñoz, A.R., Brown, J.A., 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. Divers. Distrib. 19, 1333–1338.

Barbosa, A.M., 2015. fuzzySim: applying fuzzy logic to binary similarity indices in ecology. Methods Ecol. Evol. 6, 853–858.

Batista, V., Harris, D.J., Carretero, M.A., 2004. Genetic variation in *Pleurodeles waltl* Michaelles, 1830 (Amphibia: Salamandridae) across the Strait of Gibraltar derived from mitochondrial DNA sequences. Herpetozoa 16, 166–168.

Beerli, P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22, 341–345.

Beerli, P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. In: *Population Genetics for Animal Conservation* (eds Bertorell G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C), pp. 42–79. Cambridge University Press, Cambridge.

Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. U. S. A. 98, 4563–4568.

Beerli, P., Palczewski, M., 2010. Unified Framework to Evaluate Panmixia and Migration Direction Among Multiple Sampling Locations. Genetics 185, 313–326.

Bell, R.C., Drewes, R.C., Zamudio, K.R., 2015. Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. Evolution 69, 904–915.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B-Stat. Methodol. 57, 289–300.

Bennett, K.D., Provan, J., 2008. What do we mean by 'refugia'? Quat. Sci. Rev. 27, 2449–2455.

Beukema, W., De Pous, P., Donaire-Barroso, D., Boaerts, S., Garcia-Porta, J., Escoriza, D., Arribas, O.J., El Mouden E.H., Carranza, S., 2013. Review of the systematics, distribution, biogeography and natural history of Moroccan amphibians. Zootaxa 3661.

Bidegaray-Batista, L., Arnedo, M.A., 2011. Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. BMC Evol. Biol. 11, 317.

Bielejec, F., Baele, G., Vrancken, B., Suchard, M.A., Rambaut, A., and Lemey, P., 2016. SpreaD3: interactive visualisation of spatiotemporal history and trait evolutionary processes. Mol. Biol. Evol. 33, 2167-2169.

Bilgin, R., 2007. Kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites. Mol. Ecol. 7, 416–417.

Bons, J., Geniez, P., 1996. *Anfibios y Reptiles de Marruecos*. Asociación Herpetológica Española, Barcelona. 319 pp

Bozec, A., Kageyama, M., Ramstein, G., Crépon, M., 2007. Impact of a Last Glacial Maximum sea‐level drop on the circulation of the Mediterranean Sea. Geophys. Res. Abstr. 9, 03935.

Brown, J.L., Weber, J.J., Alvarado-Serrano, D.F., Hickerson, M.J., Franks, S.J., Carnaval, A.C., 2016. Predicting the genetic consequences of future climate change: The power of coupling spatial demography, the coalescent, and historical landscape changes. Am. J. Bot. 103, 153–163.

Busack, S.D., 1985. *A biogeographical analysis of a vicariant event: the herpetofauna of the Gibraltar Strait*. Ph. D. diss., University of California, Berkeley.

Busack, S.D., 1986, Biogeographic analysis of the herpetofauna separated by the formation of the Strait of Gibraltar. Natl. Geogr. Res. 2, 17–36.

Buckley, D., 2009. Towards an organismal, interative, and iterative Phylogeography. BioEssays, 31, 784–793.

Caldas, F.B., 2013. *Thymus lotocephalus*. The IUCN Red List of Threatened Species 2013: e.T161974A5522381. http://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T161974A5522381.en. Downloaded on 07 March 2016.

Carranza, S., Arnold, E.N., 2004. History of West Mediterranean newts, *Pleurodeles* (Amphibia: Salamandridae), inferred from old and recent DNA sequences. Syst. Biodivers. 1, 327–337.

Centeno-Cuadros, A., Delibes, M., Godoy, J.A., 2009. Phylogeography of Southern Water Vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? Mol. Ecol. 18, 3652–3667.

Crawley, M.J., 2007. *The R Book*. John Wiley & Sons, Chichester.

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

Dellicour, S., Mardulyn, P., 2014. SPADS 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. Mol. Ecol. Resour. 14, 647–651.

Delplancke, M., Yazbek, M., Arrigo, N., Espíndola, A., Joly, H., Alvarez, N., 2016. Combining conservative and variable markers to infer the evolutionary history of *Prunus* subgen. *Amygdalus* s.l. under domestication. Genet. Resour. Crop Evol. 63, 221–234.

Dieringer, D., Schlötterer, C., 2003. microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes 3, 167–169.

Doadrio, I., 1988. Delimitation of areas in the Iberian peninsula on the basis of freshwaterfishes. Bonn. Zool. Beitr. 39, 113-128.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Dufresnes, C., Brelsford, A., Crnobrnja-Isailović, J., Tzankov, N., Lymberakis, P., Perrin, N., 2015. Timeframe of speciation inferred from secondary contact zones in the European tree frog radiation (*Hyla arborea* group). BMC Evol. Biol., 15, 155.

Earl, D., vonHoldt, B., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour 4, 359–361.

Ellegren, H., Galtier, N., 2016. Determinants of genetic diversity. Nat Rev Genet 17, 422–433. doi:10.1038/nrg.2016.58

Escoriza, D., Gutiérrez-Rodríguez, J., Ben Hassine, J., Martínez-Solano, I., 2016. Genetic assessment of the threatened microendemic *Pleurodeles poireti* (Caudata, Salamandridae), with molecular evidence for hybridization with *Pleurodeles nebulosus*. Conserv. Genet. 17, 1445-1458.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14, 2611–2620.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics 164, 1567–1587.

Felsenstein, J., 2008. Phylogenies Inference Package (PHYLIP) version 3.69. Department of Genome Sciences and Department of Biology, University of Washington, Seattle.

Fu, Y-X., 1997. Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. Genetics 147, 915–925.

García-París, M., Alcobendas, M., Alberch, P., 1998. Influence of the Guadalquivir River basin on mitochondrial DNA evolution of *Salamandra salamandra* (Caudata: Salamandridae) from southern Spain. Copeia 173–176.

Godinho, R. Crespo, E.G., Ferrand, N., 2008. The limits of mtDNA phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. Mol. Ecol. 17, 4670–4683.

Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L., Feldman, M.W., 1995a. An evaluation of genetic distances for use with microsatellite loci. Genetics 139**,** 463–471.

Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L., Feldman, M.W., 1995b. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. U. S. A**. 92**, 6723–6727.

Gómez, A., Lunt, D., 2007. Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula. In: *Phylogeography of Southern European Refugia* (eds Weiss S, Ferrand N), pp. 155–188. Springer Netherlands.

Gonçalves, H., Martínez-Solano, I., Pereira, R.J., Carvalho, B., 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. Mol. Ecol. 18, 5143–5160.

Gutiérrez-Rodríguez, J., Gonzalez, E.G., Martínez-Solano, I., 2014. Development and characterization of twelve new polymorphic microsatellite loci in the Iberian ribbed newt, (Caudata: Salamandridae), with data on cross-amplification in *P. nebulosus*. Amphib. Reptil. 35, 129–134.

Gutiérrez-Rodríguez, J., Barbosa, A.M., Martínez-Solano, I., 2017. Present and past climatic effects on the current distribution and genetic diversity of the Iberian spadefoot toad (*Pelobates cultripes*): an integrative approach. J. Biogeogr. 44, 245–258.

Hasegawa, M., Kishino, H., Yano, T-a., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontol. electron. 4, 9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580. doi:10.1093/molbev/msp274

Hewitt, G., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 359, 183–195.

Hey, J., 2010. Isolation with Migration Models for More Than Two Populations. Mol. Biol. Evol. 27, 905–920.

Hey, J., Nielsen, R., 2004. Multilocus Methods for Estimating Population Sizes, Migration Rates and Divergence Time, With Applications to the Divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167, 747–760.

Hey, J., Nielsen, R., 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. U. S. A. 104, 2785–2790.

Hijmans, R., Cameron, S., Parra, J., Jones, P., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801–1806.

Jiménez-Valverde, A., Acevedo, P., Barbosa, A.M., Lobo, J.M., Real, R., 2013. Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. Glob. Ecol. Biogeogr. 22, 508–516.

Jones, O.R., Wang, J., 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol. Ecol. Resour. 10, 551–555.

Jungfer, K.H., Faivovich, J., Padial, J., Castroviejo-Fisher, S., Lyra, M.M., Berneck, B., Iglesias, P.P., Kok, P.J.R., MacCulloch, R.D., Rodrigues, M.T., Verdade, V.K., Torres Gastello, C.P., Chaparro, J.C., Valdujo, P.H., Reichle, S., Moravec, J., Gvozdik, V., Gagliardi-Urrutia, G., Ernst, R., De la Riva, I., Means, D.B., Lima, A.P., Señaris, J.C., Wheeler, W.C., Haddad, C.F.B., 2013. Systematics of spiny-backed treefrogs (Hylidae: *Osteocephalus*): an Amazonian puzzle. Zool. Scr. 42, 351–380.

Kaliontzopoulou, A., Pinho, C., Harris. D.J., Carretero, M.A., 2011. When cryptic diversity blurs the picture: a cautionary tale from Iberian and North African *Podarcis* wall lizards. Biol. J. Linnean Soc. 103, 779–800.

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. J. Am. Stat. Assoc. 90, 773–795.

Kimura, M., Ohta, T., 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc. Natl. Acad. Sci. U. S. A. 75, 2868–2872.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Mol. Biol. Evol. 29, 1695–1701.

Lemey, P., Rambaut, A., Welch, J.J., Suchard, M.A., 2010. Phylogeography Takes a Relaxed Random Walk in Continuous Space and Time. Mol. Biol. Evol. 27, 1877–1885.

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

Loureiro, A., Carretero, M.A., Ferrand, N., Paulo, O., 2008. *Atlas dos Anfíbios e Répteis de Portugal Continental*. Instituto da Conservação da Natureza, Lisboa (Portugal). 257 pp.

MAGRAMA., 2015. Inventario Español de Especies Terrestres. Fauna de vertebrados: Anfibios y reptiles. Ministerio de agricultura, alimentación y Medio Ambiente, w*ebsite* Http://www.magrama.gob.es/es/biodiversidad/temas/conservacion-de-especies-amenazadas/vertebrados [*accessed 2 January 2015]*.

Maia-Carvalho, B., Gonçalves, H., Ferrand, N., Martínez-Solano, I., 2014. Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). Mol. Phylogenet. Evol. 79, 270–278.

Manel, S., Schwartz, M.K., Luikart, G., Taberlet, P., 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. 189–197.

Marjanović, D., Laurin, M., 2013. An updated paleontological timetree of lissamphibians,with comments on the anatomy of Jurassic crown-group salamanders (Urodela). Historical Biology 1–16. doi:10.1080/08912963.2013.797972.

Marjanović, D., Witzmann, F., 2015. An Extremely Peramorphic Newt (Urodela: Salamandridae: Pleurodelini) from the Latest Oligocene of Germany, and a New Phylogenetic Analysis of Extant and Extinct Salamandrids. PLoS ONE 10, e0137068. doi:10.1371/journal.pone.0137068.s009.

Martínez-Solano, I., 2004. Phylogeography of Iberian *Discoglossus* (Anura: Discoglossidae). J. Zool. Syst. Evol. Res. 42, 298–305.

Martínez-Solano, I., Teixeira, J., Buckley, D., García-París, M., 2006. Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. Mol. Ecol. 15, 3375–3388.

Measey, G.J., Vences, M., Drewes, R.C., Chiari, Y., Melo, M., Bourles, B., 2007. Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. J. Biogeogr. 34, 7–20.

Mendes, L., 1992. New data on the Thysanuran (*Microcoryphia* and *Zygentoma*: Insecta) from the Guadiana River Valley in Algarve (Portugal). Arq. Mus. Bocage (Nova Ser.) 2, 275–286.

Mesquita, N., Hänfling, B., Carvalho, G., Coelho, M., 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. Mol. Ecol. 14, 1939–1954.

Mesquita, N., Cunha, C., Carvalho, G.R., Coelho, M.M., 2007. Comparative phylogeography of endemic cyprinids in the south-west Iberian Peninsula: evidence for a new ichthyogeographic area. J. Fish Biol. 71, 45–75.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE)*, 1–8.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.

Nadachowska, K., Babik, W., 2009. Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). Mol. Biol. Evol. 26, 829–841.

Nielsen, R., 2005. Molecular signatures of natural selection. Annu. Rev. Genet. 39, 197–218.

Palumbi, S.R., 1996. Nucleic acids II: the polymerase chain reaction. In : *Molecular Systematics* (eds. Hillis DM, Moritz C, Mable BK), pp . 205–247. Sinauer & Associates Inc., Sunderland, Massachusetts.

Pasteur, G., 1958. Sur la systématique des espèces du genre *Pleurodeles* (Salamandridés). Bull. Soc. Sci. Nat. Maroc. 38, 157–165.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28, 2537–2539.

Pearce, J., Ferrier, S., 2000. Evaluating the Predictive Performance of Habitat Models Developed using Logistic Regression. Ecol. Model. 133, 225–245.

Pereira, R.J., Martínez-Solano, I., Buckley, D., 2016. Hybridization during altitudinal range shifts: nuclear introgression leads to extensive cyto-nuclear discordance in the fire salamander. Mol. Ecol. 25, 1551–1565.

Pleguezuelos, J.M., Márquez, R., Lizana, M., (eds.) 2002. *Atlas y Libro Rojo de los Anfíbios y Reptiles de España*, 2nd edn. Dirección General de Conservación de la Naturaleza — Asociación Herpetologica Española, Madrid.

Pleguezuelos, J.M., Fahd, S., Carranza, S., 2008. El papel del Estrecho de Gibraltar en la conformación de la actual fauna de anfibios y reptiles en el Mediterráneo Occidental. Boletín de la Asociación Herpetológica Española 19**,** 2–17.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155, 945–959.

Rambaut, A., Suchard, M., Xie, D., Drummond, A., 2014. Tracer v1. 6. *Computer program and documentation distributed by the author, website http://beast. bio. ed. ac. uk/Tracer [accessed 27 July 2014]*.

Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J. Hered. 86, 248–249.

Real, R., Barbosa, A.M., Martínez-Solano, I., García-París, M., 2005. Distinguishing the distributions of two cryptic frogs (Anura: Discoglossidae) using molecular data and environmental modeling. Can. J. Zool. 83, 536–545.

Real, R., Barbosa, A.M., Vargas, J.M., 2006. Obtaining Environmental Favourability Functions from Logistic Regression. Environ. Ecol. Stat. 13, 237–245.

Reich, D.E., Feldman, M.W., Goldstein, D.B., 1999. Statistical properties of two tests that use multilocus data sets to detect population expansions. Mol. Biol. Evol. 16, 453–466.

Reis, D.M., Cunha, R.L., Patrão, C., Rebelo, R., Castilho, R., 2011. *Salamandra salamandra* (Amphibia: Caudata: Salamandridae) in Portugal: not all black and yellow. Genetica 139, 1095–1105.

Rice, W.R., 1989. Analyzing Tables of Statistical Tests. Evolution 43, 223–225.

Roelants, K., Bossuyt, F., 2005. Archaeobatrachian paraphyly and Pangaean diversification of crown-group frogs. Syst. Biol. 54, 111–126.

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138.

San Mauro, D., Gower, D.J., Oommen, O.V., Wilkinson, M., Zardoya, R., 2004. Phylogeny of caecilian amphibians (Gymnophiona) based on complete mitochondrial genomes and nuclear RAG1. Mol. Phylogenet. Evol. 33, 413–427.

Sastre, P., Roca, P., Lobo, J.M., co-workers E., 2009. A Geoplatform for improving accessibility to environmental cartography. J. Biogeogr. 36, 568–568.

Serrano, A.R.M., 1995. Description and natural history of tiger beetles larvae (Coleoptera Cicindelidae) from Castro-Marim-Vila Real de Santo Antonio region (Algarve, Portugal). Arq. Mus. Bocage (Nova Ser.) 2, 555–606.

Smith, S.E., Gregory, R.D., Anderson, B.J., Thomas, C.D., 2013. The past, present and potential future distributions of cold-adapted bird species. Divers. Distrib. 19, 352–362.

Szpiech, Z.A., Jakobsson, M., Rosenberg, N.A., 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics 24, 2498–2504.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tarroso, P., Velo-Antón, G., Carvalho, S.B. 2015. phylin: an r package for phylogeographic interpolation. Mol. Ecol. Resour. 15, 349–357.

Thomé, M.T.C., Sequeira, F., Brusquetti, F., Carstens, B., Haddad, C.F., Rodrigues, M.T., Alexandrino, J., 2016. Recurrent connections between Amazon and Atlantic forests shaped diversity in Caatinga four-eyed frogs. J. Biogeogr. 43, 1045–1056.

Trindade, H., Sena, I., Gonçalves, S., Romano, A., 2012. Genetic diversity of wild populations of *Tuberaria major* (Cistaceae), an endangered species endemic to the Algarve region (Portugal), using ISSR markers. Biochem. Syst. Ecol. 45, 49–56.

Ursenbacher, S., Guillon, M., Cubizolle, H., Dupoué, A., Blouin‐Demers, G., Lourdais, O., 2015. Postglacial recolonization in a cold climate specialist in western Europe: patterns of genetic diversity in the adder (*Vipera berus*) support the central‐marginal hypothesis. Mol. Ecol. 24, 3639–3651.

Van de Vliet, M.S., Diekmann, O.E., Machado, M., Beebee, T.J., Beja, P., Serrão, E.A., 2014. Genetic Divergence for the Amphibian *Pleurodeles waltl* in Southwest Portugal: Dispersal Barriers Shaping Geographic Patterns. J. Herpetol. 48, 38–44.

Van de Vliet, M.S., Diekmann, O.E., Serrão, E.A., Beja, P., 2009. Isolation of highly polymorphic microsatellite loci for a species with a large genome size: sharp-ribbed salamander (*Pleurodeles waltl*). Mol. Ecol. Resour. 9, 425–428.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4, 535–538.

Varela, S., Lobo, J.M., Rodríguez, J., Batra, P., 2010. Were the Late Pleistocene climatic changes responsible for the disappearance of the European spotted hyena populations? Hindcasting a species geographic distribution across time. Quat. Sci. Rev. 29, 2027–2035.

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

Velo-Antón, G., Godinho, R., Harris, D.J., Santos, X., Martínez-Freiria, F., Fahd, S., Larbes S., Pleguezuelos J.M., Brito, J.C., 2012 Deep evolutionary lineages in a Western Mediterranean snake (*Vipera latastei/monticola* group) and high genetic structuring in Southern Iberian populations. Mol. Phylogenet. Evol. 65, 965–973.

Vences, M., Vieites, D.R., Glaw, F., Brinkmann, H., Kosuch, J., Veith, M., Meyer, A., 2003. Multiple overseas dispersal in amphibians. Proc. R. Soc. Lond. Ser. B-Biol. Sci. 270, 2435–2442.

Wang, J., 2004. Sibship Reconstruction From Genetic Data With Typing Errors. Genetics 166, 1963–1979.

Warren, D.L., Glor, R.E., Turelli, M., 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62, 2868-2883.

Wielstra, B., Babik, W., Arntzen, J.W., 2015. The crested newt *Triturus cristatus* recolonized temperate Eurasia from an extra‐Mediterranean glacial refugium. Biol. J. Linnean Soc. 114, 574–587.

Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353.

Zhang, P., Papenfuss, T.J., Wake, M.H., Qu, L., Wake, D.B., 2008. Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. 49, 586–597.

Table 1. Locality information for the samples used in this study, including latitude and longitude, sample size for microsatellite (SSR) and mtDNA analyses, number of haplotypes, sample codes, mtDNA lineages, and nucleotide diversity. Locality ID corresponds to Figure 1.

ID	Locality	Longitude	Latitude	n SSR	mtDNA			
					n	Sample Codes	Haplotypes	Lineage	(π)
1	Albires, León, Spain	-5.280	42.276	10	4	IMS3032-IMS3035	H1	Western	0.00000
2	Revilla de Campos, Palencia, Spain	-4.710	42.006	10	4	IMS3097-IMS3100	H1	Western	0.00000
3	Valdefinjas, Zamora, Spain	-5.467	41.435	10	3	IMS3182, IMS3184-IMS3185	H1	Western	0.00000
4	Navales-Valdecarros, Salamanca, Spain	-5.450	40.785	10	4	PW565-PW568	H1, H23	Western	0.00117
5	Ciudad Rodrigo, Salamanca, Spain	-6.576	40.569	10	4	PW663- PW666	H1	Western	0.00000
6	Fermoselle, Zamora, Spain	-6.367	41.328	20	4	IMS2841-IMS2844	H1	Western	0.00000
7	Santo Tomé del Puerto, Segovia, Spain	-3.589	41.200	19	4	IMS1827-IMS1830	H1	Western	0.00000
8	Valdemanco-La Cabrera, Madrid, Spain	-3.645	40.853	10	5	PW09001-PW09004, PW09006	H1	Western	0.00000
9	Faro, Faro, Portugal	-7.979	37.051	9	4	PW738-PW741	H26	Western	0.00000
10	Sedas, Beja, Portugal	-7.599	37.537	10	4	PW442, PW444-PW446	H23	Western	0.00000
11	Abela, Setúbal, Portugal	-8.558	38.042	10	4	PW470-PW473	H23, H24	Western	0.00117
12	Évora, Évora, Portugal	-7.928	38.555	10	4	PW493-PW496	H23	Western	0.00000
13	Coruche, Santarém, Portugal	-8.519	38.999	10	4	PW685- PW688	H1, H3	Western	0.00058
14	Tocha, Coímbra, Portugal	-8.807	40.375	10	4	PW705- PW708	H4	Western	0.00000
15	Jerez de los Caballeros, Badajoz, Spain	-6.696	38.340	10	3	PW513- PW514, PW516	H23, H25	Western	0.00234
16	Campanario, Badajoz, Spain	-5.613	38.840	10	4	PW541-PW544	H1	Western	0.00000
17	Peñarroya-Pueblonuevo, Córdoba, Spain	-5.092	38.285	10	4	PW380-PW383	H1	Western	0.00000
18	Membrio, Cáceres, Spain	-7.048	39.532	9	4	IMS2833-IMS2836	H15, H16, H17	Western	0.00312
19	Navalmoral de la Mata, Cáceres, Spain	-5.547	39.908	9	4	IMS2789-IMS2791, IMS2793	H1, H14	Western	0.00058
20	Menasalbas, Toledo, Spain	-4.357	39.662	10	3	PW595-PW597	H1	Western	0.00000
21	El Toboso, Toledo, Spain	-2.982	39.485	10	7	PW620- PW623, Toboso1-Toboso3	H2	Eastern	0.00000
22	Malagón, Ciudad Real, Spain	-3.907	39.178	10	6	PW651- PW654, La_Nava1-La_Nava2	H1	Western	0.00000
23	Cabezarrubias del Puerto, Ciudad Real, Spain	-4.257	38.569	10	5	IMS3410-IMS3414	H1	Western	0.00000

24	Tanger Asilah, Tanger-Tetuan, Morocco	-5.979	35.624	15	9	IMS1153-IMS1157, PW769-PW772	H8, H18	Eastern	0.00163
25	Larache-Ksar el Kebir, Tanger-Tetuan, Morocco	-6.039	35.040	8	4	PW849-PW852	H28, H18	Eastern	0.00190
26	Moulay Bousselhaim, Garb-Chrarda-Beni Hsen, Morocco	-6.086	34.770	10	4	PW786-PW789	H22, H18	Eastern	0.00190
27	Kenitra, Garb-Chrarda-Beni Hsen, Morocco	-6.576	34.208	10	3	PW808-PW810	H22, H27	Eastern	0.00187
28	Had Soualem, Chauía-Uardiga	-7.829	33.406	10	4	PW837- PW840	H22	Eastern	0.00000
29	Doñana, Huelva, Spain	-6.455	36.989	10	4	PW718- PW721	H5, H6, H7	Eastern	0.00109
30	Chiclana, Cádiz, Spain	-6.032	36.434	20	20	IMS1194 - IMS1213	H5, H9, H10, H11	Eastern	0.00131
31	Cabra, Córdoba, Spain	-4.368	37.507	10	4	PW362- PW365	H5	Eastern	0.00117
32	El Pedroso, Sevilla, Spain	-5.785	37.828	10	6	IMS1220-IMS1221, PW400-PW403	H12	Eastern	0.00109
33	Villanueva de San Juan, Sevilla, Spain	-5.289	37.060	10	3	PW424, PW427-PW428	H5	Eastern	0.00000
34	Bienservida, Albacete, Spain	-2.712	38.550	10	5	IMS3308-IMS3312	H2	Eastern	0.00123
35	Navalcudia, El Bonillo, Albacete, Spain	-2.489	38.939	10	5	IMS3399-IMS3403	H2	Eastern	0.00123
36	El Perelló, Tarragona, Spain	0.670	40.858	10	10	IMS2357-IMS2366	H2	Eastern	0.00000
37	Morella, Castellón, Spain	-0.071	40.502	10	4	IMS3247-IMS3250	H2	Eastern	0.00000
38	Toll Nou, Beneixama, Alicante, Spain	-0.737	38.750	10	5	IMS3369-IMS3373	H2	Eastern	0.00000
39	Cabañas de El Salobral, Albacete, Spain	-1.965	38.791	10	5	IMS3268-IMS3272	H2	Eastern	0.00123
40	Las Nogueras, Valencia, Spain	-1.082	39.588	11	11	IMS1989-IMS1999	H2	Eastern	0.00000
41	Sinarcas, Valencia, Spain	-1.239	39.763	10	5	IMS3288-IMS3292	H2	Eastern	0.00000
42	Los Palancares, Cuenca, Spain	-1.959	40.010	10	5	IMS3379-IMS3383	H2	Eastern	0.00000
43	Gárgoles, Guadalajara, Spain	-2.631	40.757	-	4	IMS2623-IMS2625, IMS2627	H13	Eastern	0.00000
44	Los Escoriales, Jaén, Spain	-3.924	38.180	-	7	PW345- PW351	H5	Eastern	0.00000
45	Sierra de Loja, Granada, Spain	-4.173	37.114	-	2	PW875-PW876	H5, H29	Eastern	0.00117
46	Moulay Abdesalam, Tanger-Tetuan, Morocco	-5.488	35.271	-	1	GVA214	H21	Eastern	0.00190
47	Mazarambroz, Toledo, Spain	-4.096	39.653	-	3	Castanar2, Castanar4-Castanar5	H1	Western	0.00000
48	La Mierla, Guadalajara, Spain	-3.261	40.923	-	3	La_Mierla1-La_Mierla3	H1	Western	0.00000
49	Pelahustán, Toledo, Spain	-4.585	40.180	-	1	MNCN10075	H1	Western	0.00000
50	Ksar el Kebir, Tanger-Tetuan, Morocco	-5.916	35.004	-	1	MNCN7184	H18	Eastern	0.00190
51	14 km S Asilah, Tanger-Tetuan, Morocco	-6.017	35.323	-	2	MVZ231893, MVZ186103	H21, H18	Eastern	0.00163
52	5.5km SE Rabat, Rabat-Salé-Zemmour-Zaer, Morocco	-6.824	33.960	-	1	MVZ162384	H19	Eastern	0.00187

53	10 km E Rabat-Sale Bridge, Rabat-Salé-Zemmour-Zaer, Morocco	-6.702	34.012	-	1	MVZ186090	H22	Eastern	0.00187
54	2.5 km E Puerto Real, Cádiz, Spain	-6.115	36.561	-	1	MVZ231894	H20	Eastern	0.00131
55	Tanger, Tanger-Tetuan, Morocco	-5.833	35.792	-	1	MVZ230275	H8	Eastern	0.00163

Table 2. Results of tests for demographic expansion in *P.waltl* (Pw), including the Atlantic (A), Mediterranean (Md), Southern (S), and Moroccan (M) haplogroups. Values of Fu's Fs, Tajima's D, R^2 and the raggedness index (R) from mismatch distributions are provided, including their associated *p* values.

	Fu's Fs	*p*	Tajima's D	*p*	R^2	*p*	R	*p*
Pw	15.803	0.970	2.8453	0.998	0.1604	0.999	0.1131	1.0
A	-4.1894	0.029	-1.4277	0.042	0.0451	0.089	0.2432	0.633
Md	-0.3342	0.433	-0.5643	0.189	0.0578	0.207	0.6044	0.619
S	-4.359	0.005	-1.2637	0.090	0.0616	0.070	0.1594	0.328
M	-0.4011	0.451	-0.3521	0.409	0.1091	0.372	0.0182	0.005

Table 3. Genetic diversity in 42 populations of *P. waltl* based on 10 microsatellite loci. N=sample size; Nc=sample size after exclusion of potential siblings from the sample; Na=mean number of alleles per locus; *Ho* and *He*: observed and expected heterozygosity; F_{IS}: inbreeding coefficient; Ar: allelic richness standardized to the lowest sample size (4). Locality ID corresponds to Figure 1 and to Table 1.

ID	N	Nc	Na	*Ho*	*He*	F_{IS}	Ar (4)
1	10	9	1.297	0.178	0.184	0.033557047	1.51554
2	10	9	1.382	0.238	0.218	-0.088834844	1.58325
3	10	8	1.203	0.175	0.142	-0.230769231	1.44564
4	10	10	1.884	0.430	0.408	-0.053921569	2.30782
5	10	10	1.705	0.450	0.373	-0.206434316	2.04987
6	20	9	1.466	0.278	0.258	-0.076555024	1.73336
7	19	16	2.461	0.550	0.555	0.009148487	2.84315
8	19	19	3.780	0.725	0.703	-0.03077398	3.86226
9	9	9	3.148	0.700	0.659	-0.061797753	4.02687
10	10	10	5.769	0.864	0.812	-0.064020545	5.21722
11	10	10	4.761	0.882	0.771	-0.143899921	4.68988
12	10	5	4.853	0.915	0.764	-0.197448062	5.16222
13	10	10	4.207	0.830	0.749	-0.108144192	4.42524
14	10	10	3.412	0.590	0.627	0.059011164	3.7525
15	10	10	4.277	0.809	0.755	-0.07108769	4.3841
16	10	10	4.259	0.790	0.745	-0.061114842	4.46695
17	10	10	3.723	0.668	0.701	0.046095579	4.0978
18	9	7	4.759	0.900	0.777	-0.158525573	4.93636
19	9	9	4.354	0.856	0.761	-0.124315555	4.47977
20	10	10	3.420	0.790	0.673	-0.17384844	3.68966
21	10	9	1.987	0.465	0.405	-0.147505887	2.3225
22	10	10	4.140	0.726	0.723	-0.003748901	4.31657
23	10	7	3.165	0.886	0.661	-0.339506173	3.61662
24	15	15	5.504	0.853	0.800	-0.067259589	4.92443
25	8	8	5.365	0.850	0.803	-0.058365759	5.29472
26	10	10	7.100	0.850	0.857	0.007589025	5.90853
27	10	10	3.462	0.800	0.699	-0.145311382	3.90016
28	10	10	5.104	0.790	0.797	0.008160703	4.96101
29	10	10	7.373	0.926	0.855	-0.083156093	5.8904
30	20	12	8.956	0.917	0.876	-0.046373365	6.08765
31	10	10	3.385	0.720	0.681	-0.057268722	3.67601
32	10	10	4.069	0.828	0.709	-0.167040886	4.26372
33	10	10	6.200	0.919	0.832	-0.104925626	5.45794
34	10	10	3.238	0.690	0.675	-0.022979985	3.80034
35	10	10	3.909	0.750	0.727	-0.032346869	4.18354
36	10	8	1.315	0.200	0.163	-0.224880383	1.52487
37	10	9	1.426	0.278	0.259	-0.07398568	1.87571
38	10	9	1.499	0.267	0.275	0.029213483	1.76742
39	10	10	2.777	0.684	0.605	-0.131936788	3.31429
40	11	6	1.532	0.450	0.276	-0.628140704	1.73899
41	10	10	2.208	0.450	0.440	-0.02371494	2.61162
42	10	10	1.642	0.390	0.340	-0.148748159	1.83139

Table 4. Comparison of three models of migration between Morocco (M) and the southern Iberian Peninsula (S).

Model	Bezier IML	Ln Bayes factor	Choice (Bezier)	Model Probability
S to M	-21463.13	-6012.53	2	0
M to S	-15450.60	0	1 (best)	1
Panmictic	-104801.28	-89350.68	3	0

Figure 1. A) Geographic distribution of the sampled populations, showing the distribution of the two major mtDNA haplogroups (Eastern and Western). The Eastern haplogroup comprises three groups: Mediterranean (green), Southern (dark blue), and Morocco (light blue). The Western haplogroup comprises two groups: Algarve (orange), and Atlantic (red). B) Haplotype network of mtDNA (ND4) sequences in *P. waltl*. Two main mtDNA haplogroups were identified, separated by 43 mutations. C) Results of Structure, showing the spatial distribution of genetic clusters based on values of K=2 and K=4. D) Values of ΔK, showing peaks at K=2 and K=4.

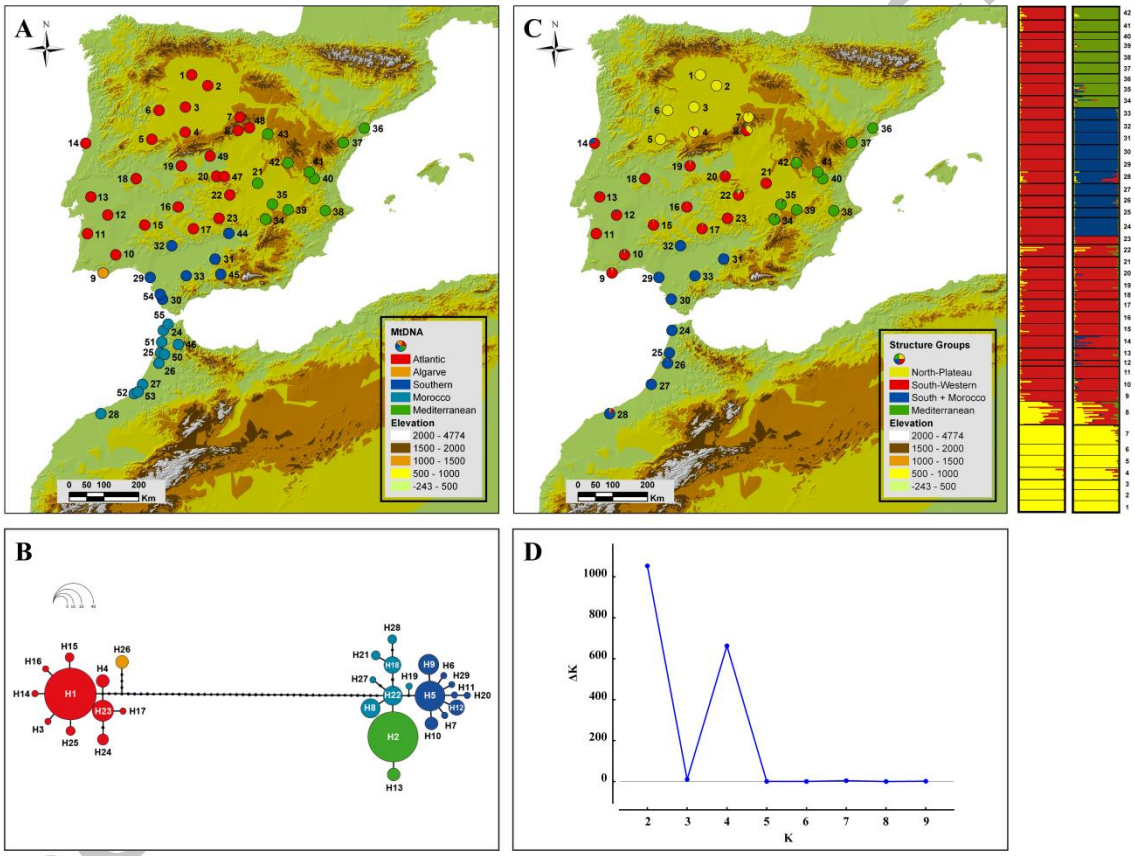


Figure 2. A. Time-calibrated mtDNA tree resulting from the *BEAST analysis. Values above branches represent Bayesian Posterior Probabilities (BPPs) higher than 0.9. Bars show 95% Highest Posterior Density Intervals (HPDIs) of time estimates. Scale bar in millions of years. B. Time-calibrated tree from ND4 sequences of *P. waltl* recovered in continuous diffusion analyses. Asterisks indicate samples included in the *BEAST analysis.

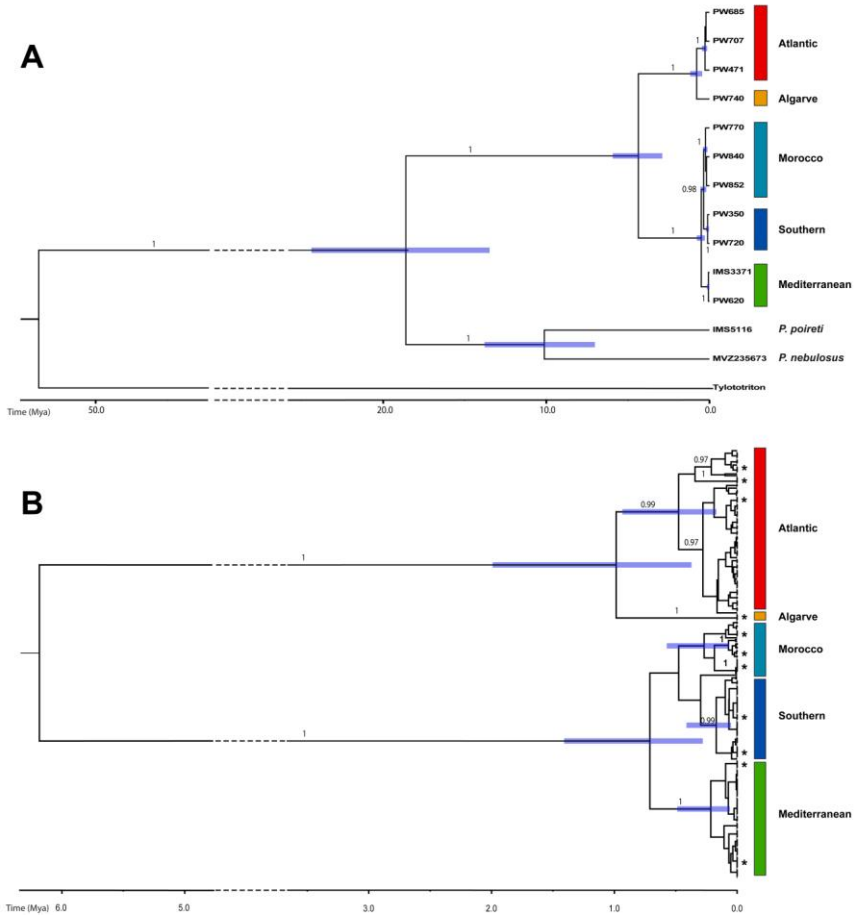


Figure 3. Genetic diversity of *P. waltl*, based on the first axis of a principal components analysis (PC1) summarizing nucleotide diversity, observed heterozygosity, allelic richness, and number of private alleles (positive PC1 scores indicate higher diversity); and the intersection between climatic favourability for this species in the Last Inter-Glacial period (LIG) and in each of the three climatic simulations (CCSM4, MIROC-ESM and MPI-ESM-P) available for the Last Glacial Maximum (LGM).

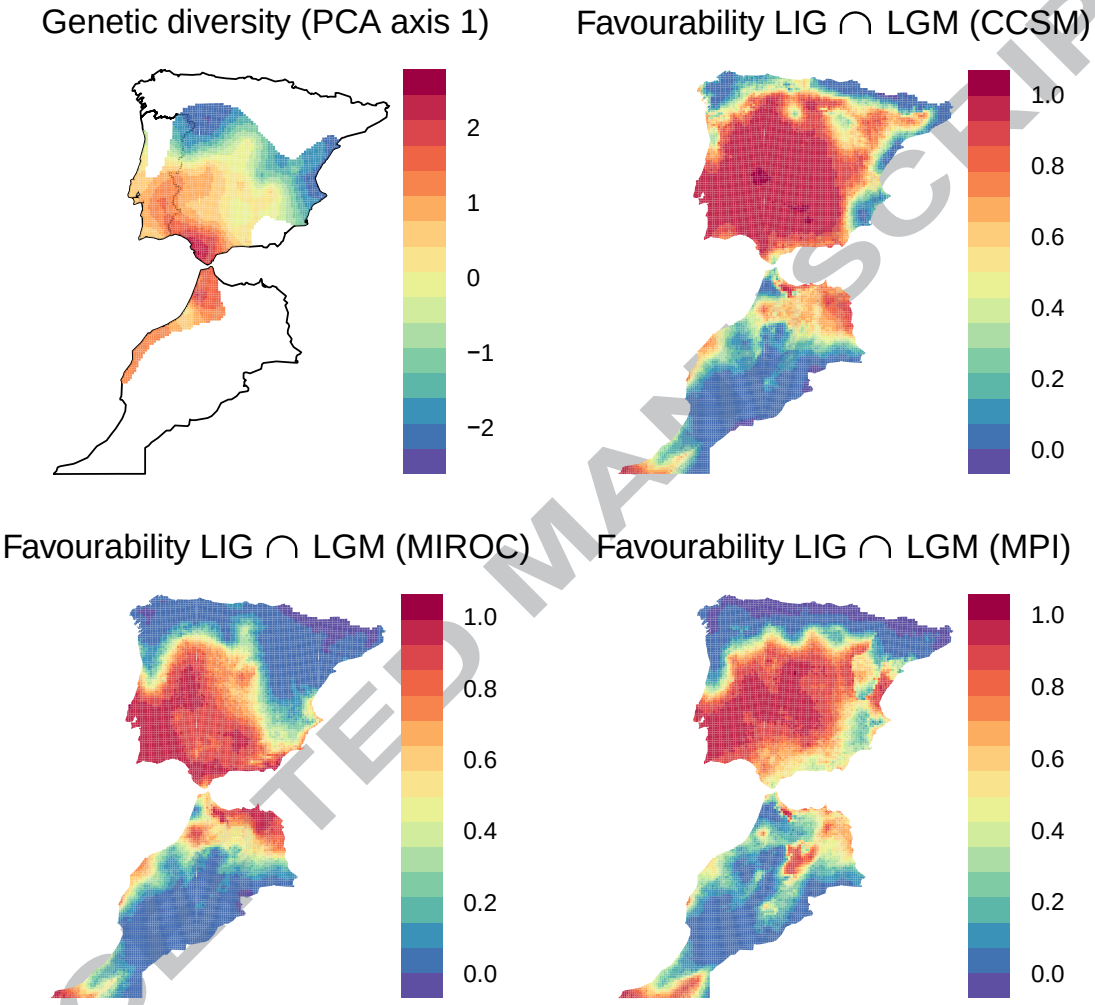


Figure 4. Species distribution models showing current climatic favourability for *P. waltl* and for each of the two major mtDNA lineages, including areas that are climatically favourable for both lineages (intersection). Black dots represent presence records.

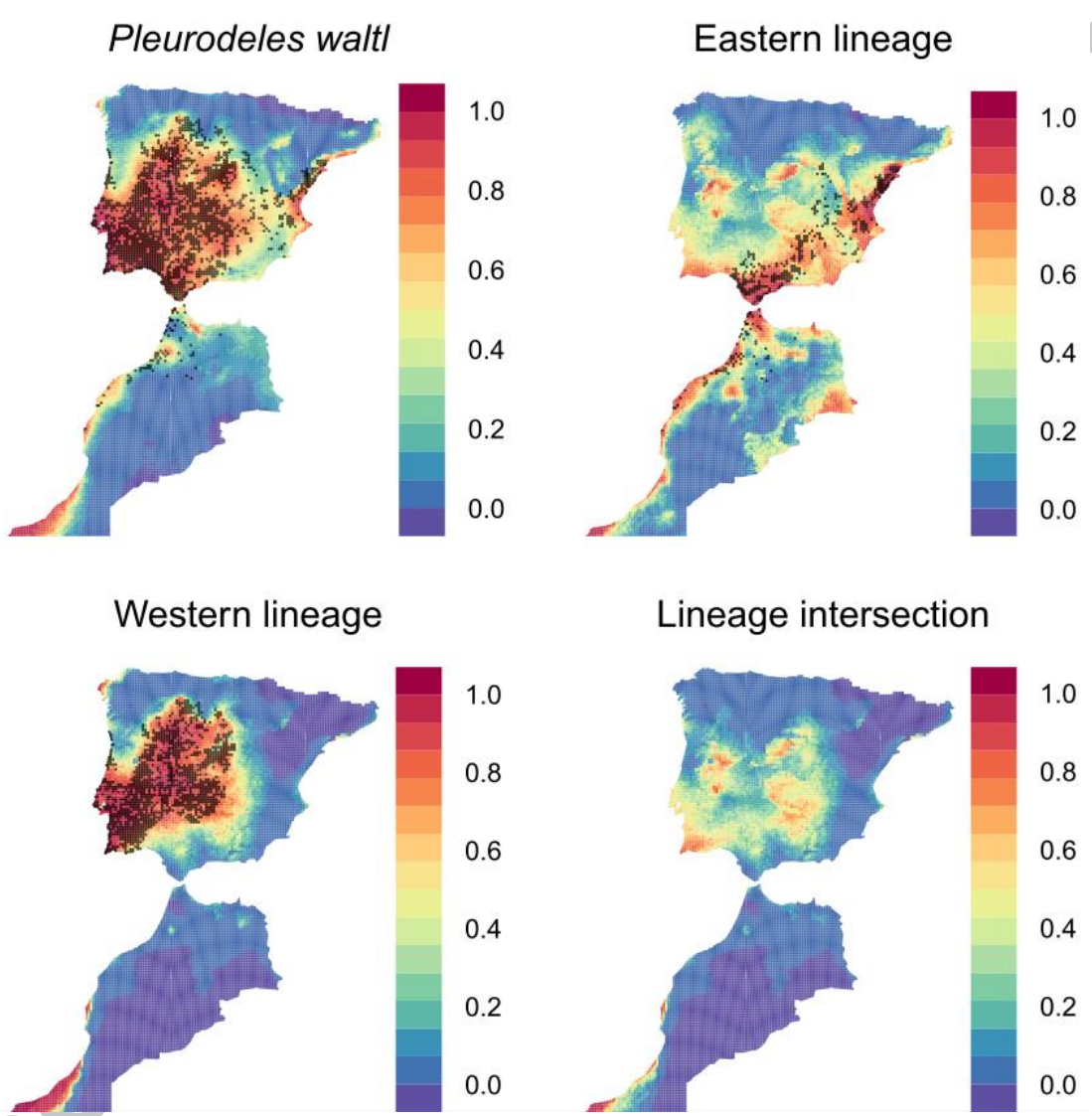
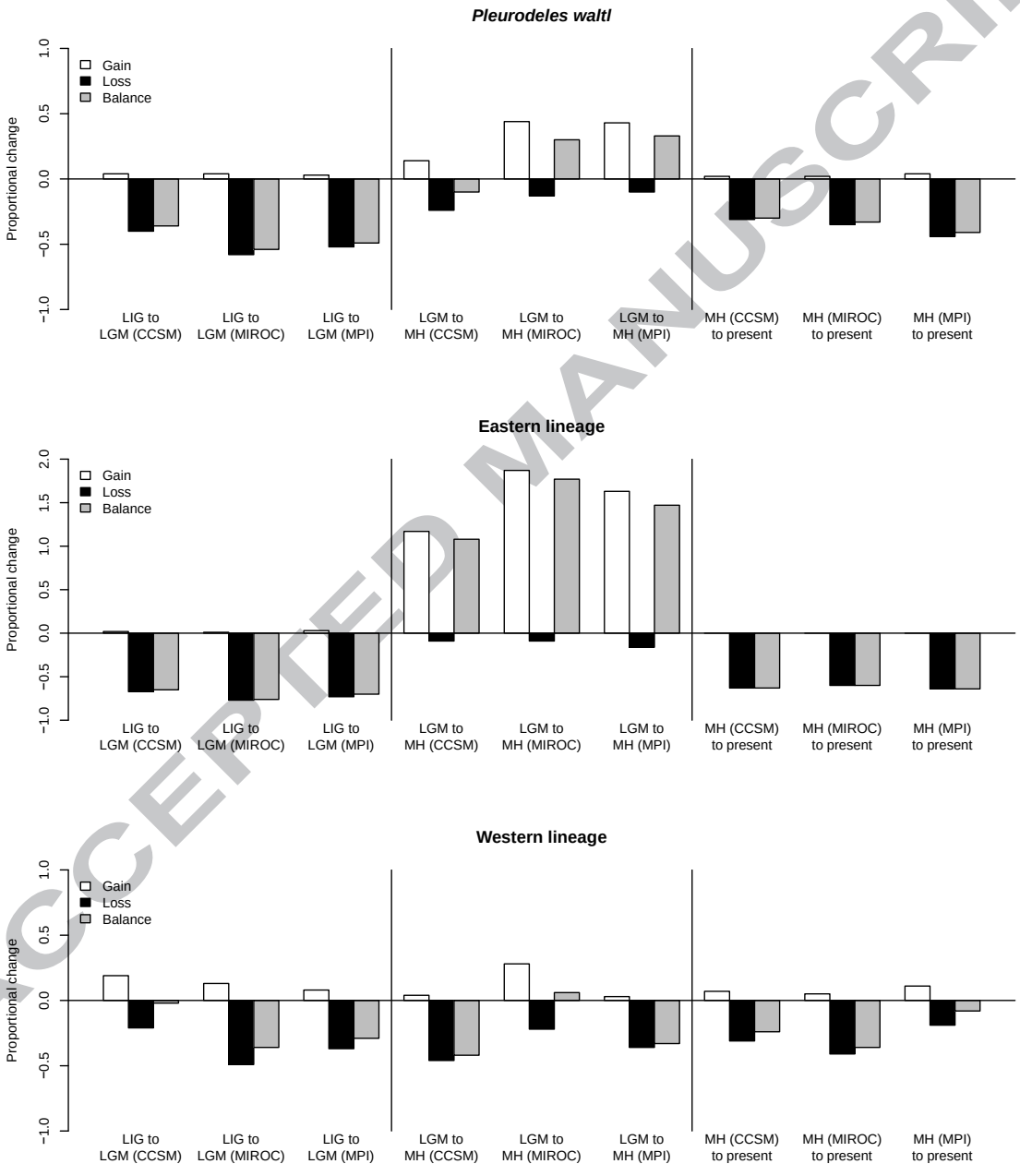
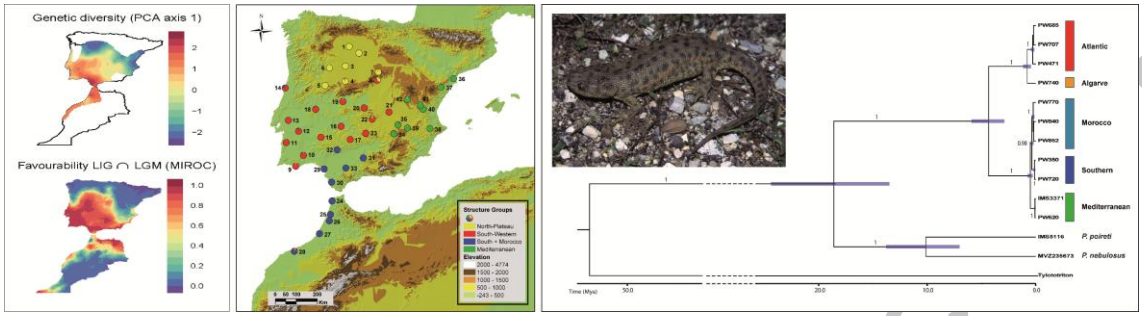


Figure 5. Fuzzy range change measures (fuzzy equivalents of the proportional gain, loss and overall change in areas favourable for presence) among time periods, from the Last Inter-Glacial (LIG) to the Last Glacial Maximum (LGM), the Mid Holocene (MH) and the present, including the three paleoclimatic simulations currently available across periods on WorldClim (CCSM4, MIROC-ESM and MPI-ESM-P).



Graphical abstract



Highlights

Plio-Pleistocene east-west mtDNA break in Pleurodeles waltl

Exclusive mtDNA haplotypes and monophyly of Moroccan populations

Post-divergence gene flow across the Gibraltar Strait

Temporal stability in climatic favourability correlates with genetic diversity