**Combining phylogeography and landscape genetics to infer the
evolutionary history of a short-range Mediterranean relict,
*Salamandra salamandra longirostris***

B. Antunes^{1,2}, A. Lourenço^{1,2}, G. Caeiro-Dias^{1,2,3}, M. Dinis^{1,2}, H. Gonçalves^{1,4}, I.
Martínez-Solano⁵, P. Tarroso¹, & G. Velo-Antón^{1*}

¹ CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do
Porto. Instituto de Ciências Agrárias de Vairão. Rua Padre Armando Quintas 7. 4485-661 Vairão
Portugal

² Departamento de Biologia da Faculdade de Ciências da Universidade do Porto. Rua Campo Alegre,
4169-007 Porto, Portugal.

³ CEFE UMR 5175, CNRS – Université Montpellier - Université Paul-Valéry Montpellier - EPHE, 1919
route de Mende, 34293 Montpellier, Cedex 5, France

⁴ Museu de História Natural e da Ciência da Universidade do Porto (MHNC-UP), Praça Gomes Teixeira,
4099-002 Porto, Portugal

⁵ Departamento de Biodiversidad y Biología Evolutiva. Museo Nacional de Ciencias Naturales MNCN-
CSIC, c/José Gutiérrez Abascal 2, 28006 Madrid, Spain.

*Corresponding author: Guillermo Velo-Antón (guillermo.velo@cibio.up.pt)

Fax: (+351) 252661780

Abstract

Examining historical and contemporary processes underlying current patterns of genetic variation is key to reconstruct the evolutionary history of species and implement conservation measures promoting their long-term persistence. Combining phylogeographic and landscape genetic approaches can provide valuable insights, especially in regions harboring high levels of biodiversity that are currently threatened by climate and land cover changes, like southern Iberia. We used genetic (mtDNA and microsatellites) and spatial data (climate and land cover) to infer the evolutionary history and contemporary genetic connectivity in a short-range endemic salamander subspecies, *Salamandra salamandra longirostris*, using a combination of ecological niche modelling, phylogeographic, and landscape genetic analyses. Ecological-based analyses support a role of the Guadalquivir River Basin as a major vicariant agent in this taxon. The lower genetic diversity and greater differentiation of peripheral populations, together with analyses of climatically stable areas throughout time, suggest the persistence of a population in the central part of the current range since the Last Inter Glacial [LIG; ~120,000 - 140,000 years BP], and a micro refugium in the eastern end of the range. Habitat heterogeneity plays a major role in shaping patterns of genetic differentiation in *S. s. longirostris*, with forests representing key areas for its long-term persistence under scenarios of environmental change. Our study stresses the importance of maintaining population genetic connectivity in low-dispersal organisms under rapidly changing environments, and will inform management plans for the long-term survival of this evolutionarily distinct Mediterranean endemic.

Keywords: amphibians, climate change, connectivity, integrative studies, land cover

Introduction

Identifying causative factors underlying patterns of genetic variation is a major goal in the fields of ecology, evolution, and conservation (Allendorf and Luikart 2013). Geographic and environmental heterogeneity are regarded as major drivers of gene flow and, by extension, of spatial patterns of genetic diversity and population divergence (Lee and Mitchell-Olds 2011; Wang et al. 2013; Sexton et al. 2014). Reduced gene flow (and high genetic divergence) among populations is expected when they are separated by increasing geographic distances (isolation by distance [IBD]; Wright 1943) and/or landscape and climatic barriers (i.e. landscape complexity is explicitly accounted for: isolation-by-resistance [IBR]; McRae 2006). Additionally, gene flow may also decrease due to ecological processes (e.g. natural and/or sexual selection against immigrants; reduced hybrid fitness) driven by environmental dissimilarity among sites, regardless of the geographic distance (or landscape complexity) separating them (isolation-by-environment [IBE]; Wang and Bradburd 2014).

Combining historical (i.e. phylogeography) and contemporary (i.e. landscape genetics) approaches provides a powerful framework for understanding the sequence of events and the relative role of geographic and ecological factors underlying genetic diversity (e.g. Zellmer et al. 2009; He et al. 2013; Velo-Antón et al. 2013; Rissler, 2016; Nogueras et al. 2016; Zhang et al. 2016; Gutiérrez-Rodríguez et al. 2017a). These studies represent a window to the past and an opportunity to predict the fate of species, which is a conservation priority in regions with high levels of biodiversity that are also threatened by current climatic and landscape changes, like the southern Iberian Peninsula.

This region, located in one of the world's biodiversity hotspots (Myers et al. 2000), is a geologically, climatically and topographically heterogeneous region. The

interplay between these factors through geological time has promoted the persistence of genetically differentiated lineages in multiple isolated climatic refugia in response to Quaternary climatic oscillations. This resulted in high levels of endemism and strong levels of population subdivision at the intra-specific level, documented for example across different species of amphibians and reptiles (Gonçalves et al. 2009; Velo-Antón et al. 2012a; Dias et al. 2015; Díaz-Rodríguez et al. 2015). However, rapid contemporary changes in habitat and climate in this region due to anthropogenic activities (mainly agricultural and deforestation activities) are threatening the long-term persistence of this biodiversity (Carvalho et al. 2011; Ferreira and Beja 2013). Amphibians are vulnerable to environmental changes because their ectothermic physiology is constrained by climate-related factors. Additionally, their relatively low dispersal ability (e.g. Smith and Green 2005) and dependence on humid environments and water bodies, which are essential to complete their life cycle, further exacerbate their vulnerability to changes in the surrounding environment (Stuart et al. 2004; Cushman 2006). Assessing the future viability of amphibian populations in this region requires understanding how extrinsic factors affect current patterns of genetic diversity and connectivity while accounting for their long and complex evolutionary history.

The fire salamander, *Salamandra salamandra* (Linnaeus 1758), is a terrestrial amphibian distributed throughout most of Europe. It occurs in humid and shaded environments with a wide range of vegetation communities, generally near ponds and streams where females deliver aquatic larvae (Velo-Antón et al. 2015). Much of the genetic and phenotypic intra-specific variation is found in the Iberian Peninsula (see Velo-Antón and Buckley 2015), with the southern tip of this region harboring a highly divergent lineage, *S. s. longirostris*. This lineage exhibits peculiar phenotypic traits, like a pointed snout, a characteristic pattern of dorsal coloration, and linear bone growth

(Joger and Steinfartz 1998; Alcobendas and Castanet 2000; Donaire-Barroso et al. 2009). There is an ongoing debate about its taxonomic status (see Frost, 2018) due to its geographic isolation, separated from other subspecies by the Guadalquivir river Basin (GRB), and deep divergence at the mtDNA level (~6% *cyt-b* divergence with *S. s. crespoid* and *S. s. morenica*; García-París et al. 1998). However, its morphological, ecological and nuclear genetic differentiation have not been thoroughly studied to date. Divergence between *S. s. longirostris* and *S. s. crespoid* / *morenica* likely began with the formation of the Guadalquivir river valley during the Pliocene (García-París et al. 1998). Warm climate periods in the Quaternary, possibly associated with the existence of climatic refugia along the Betic range, could have also contributed to the diversification of southern Iberian *S. salamandra* lineages. The current allopatric and restricted distribution of *S. s. longirostris* in an area facing pronounced habitat loss has raised concern about its conservation, being catalogued as vulnerable to extinction (VU) according to the International Union for the Conservation of Nature (IUCN; Pleguezuelos et al. 2004).

We used a combination of ecological niche modelling [ENM], phylogeographic, population and landscape genetic analyses to infer how historical and contemporary geographic and environmental heterogeneity have shaped current patterns of genetic variation in *S. s. longirostris*. Specifically, we used genetic (mtDNA and microsatellites) and spatial data (climate and land cover), covering the entire range of *S. s. longirostris*, to: (1) characterize patterns of genetic diversity and structure at historical (mtDNA) and contemporary (microsatellites) scales; (2) identify climatically stable areas that might have acted as climatic refugia during the Late Pleistocene; and (3) investigate how environmental and geographic factors (climate, habitat and topography)

have driven genetic divergence at deeper (phylogeographic) and shallower (landscape genetic) evolutionary scales.

Material and methods

Sampling and laboratory procedures

Sampling was conducted in areas with potentially suitable habitat for *S. s. longirostris* (i.e. humid areas in woodlands and mountains, and rocky areas with abundant vegetation and water bodies) across its entire range (south of the GRB; Figure 1). A total of 156 tissue samples (tail or toe clips), obtained mostly from larvae, were collected and georeferenced at 26 localities (Online Resource 2.1).

DNA was extracted from tissue using the EasySpin® Genomic DNA Tissue Kit (Citomed, Lisbon, Portugal), following the manufacturer's protocol. Quality of extracted DNA was checked by electrophoresis on 0.8% agarose gels. Extracted DNA was used as template in polymerase chain reactions (PCR) to amplify: I) one mitochondrial fragment, the mitochondrially encoded *cytochrome b* and adjacent tRNAs (hereafter *cyt-b*; ca. 1400 bp); and II) nine microsatellite loci (Sal29, Sal23, SalE7, SalE5, SalE2, SalE06, Sal3, SalE8, SalE12; Steinfartz et al. 2004).

The *cyt-b* fragment was amplified for 42 samples covering the entire range of *S. s. longirostris*, using the primers Glu14100L (forward, 5' GAA AAA CCA AYG TTG TAT TCA ACT ATA A 3') and Pro15500H (reverse, 5' AGA ATT YTG GCT TTG GGT GCCA 3') (Zhang et al. 2008) following the protocol described in Beukema et al. (2016). Sequences were edited and aligned by eye in GENEIOUS version 11.1.4 (<http://www.geneious.com/>). One individual of *S. s. crespai* and another one representing *S. s. morenica* were also sequenced for *cyt-b* to calculate genetic distances

between subspecies (Online Resource 2.1). The nine microsatellites were amplified for most samples (n=146). They were distributed in four optimized multiplex reactions (panels S2, S3, S4 and S5; see also Online Resource 2.2) and amplified following the conditions described in Álvarez et al. (2015). Quality and quantity of PCR products for *cyt-b* and microsatellites were verified on 2 % agarose gels. Cycle sequencing reactions for *cyt-b* were carried out using the ABI Prism[®] BigDye[®] Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) standard protocol. An ABI Prism[®] 3130xl Genetic Analyser Sequencer (Applied Biosystems/HITASHI) was then used to sequence the *cyt-b* fragments for both strands and to genotype microsatellites. Allele scoring was performed using GeneMapper version 4.0 (Applied Biosystems/HITASHI).

Phylogenetic analyses

We used the software BEAST version 1.8.4 (Drummond et al. 2012) to perform coalescent-based Bayesian phylogenetic inference on *S. s. longirostris cyt-b* sequences. We selected the optimal nucleotide substitution model (TrN) with JMODELTEST version 2.1.4 (Darriba et al. 2012), under the Bayesian information criterion (BIC). For BEAST analyses, we ran analyses using both a lognormal relaxed clock and a strict clock, with a constant population size model as the coalescent tree prior. Three independent runs for each clock model were performed and combined with a total of 50 million generations (burn-in: 10%). Tree files of all runs were combined using the software LOGCOMBINER version 1.7.5, and parameter convergence was verified by examining the effective sample sizes (ESSs) in TRACER version 1.6. A maximum clade credibility summary tree with Bayesian posterior probabilities (BPP) for each node was obtained using TREEANNOTATOR version 1.8.4. The resulting tree was

visualized and edited with FigTree version 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree>). BEAST analyses were run in the Cipres Science Gateway (Miller et al. 2010). Genetic distances (uncorrected p-distances) were calculated between subspecies and between clades within *S. s. longirostris* using GENEIOUS.

Microsatellite pre-treatment procedures

The microsatellite data set was subjected to several procedures to reduce potential statistical biases. First, genotyped samples presenting more than 20% of missing data were discarded. Second, locations with a sample size ≥ 10 were considered local random mating units (i.e. demes or populations). Following these criteria, a total of seven populations / demes were considered (MSI, PLAG, LMO, AVI, VLR, TOR and VVR; Table 1; Online Resource 2.1) for subsequent population and landscape genetic analyses. Because of low sample sizes in some locations ($n < 10$; Online Resource 2.1), we drew 2-km radius buffers (roughly the maximum dispersal distance recorded in *Salamandra salamandra*; Hendrix et al. 2017) around each sample in ArcGIS version 10.1 (ESRI) and pooled samples if their buffers intersected. Third, sampling individuals in early life stages (e.g. amphibian larvae) may introduce significant biases in routine microsatellite quality control tests and genetic diversity estimates (e.g. Hardy-Weinberg equilibrium [HWE] and Linkage Equilibrium [LE]; Sánchez-Montes et al. 2017), as well as in landscape genetics analyses (Peterman et al. 2016; Wang 2018). This is because larvae are often spatially clustered, increasing the likelihood of sampling relatives and consequently, inflating inter-population genetic divergence. To address this shortcoming, we screened familial relationships independently for each of the seven populations using COLONY 2.0.6.1 (Jones and Wang 2010). Analyses were performed

using the full-likelihood method, with high likelihood precision and medium run length, assuming polygamy for both sexes. No *a priori* information regarding known parents was provided. The remaining parameters were kept as default. If a pair of related individuals (parent-offspring, half- or full-siblings) had a posterior probability higher than 0.50, then one of them was removed from the dataset. After excluding related individuals, we re-estimated population genetic diversity and differentiation parameters to assess the effect of including relatives in results. Related individuals were discarded from downstream genetic analyses when estimates including (full dataset) and excluding relatives correlated below $r = 0.9$. Finally, the presence of null alleles was assessed using a maximum likelihood estimator implemented in INEst version 2.0 (Chybicki and Burczyk 2009). Analyses for each of the seven demes were run under the individual inbreeding model, with a total of 200,000 cycles thinned every 50 cycles and a burn-in of 10%. GENEPOP version 4.5.1 (Rousset 2008) was used to test for deviations from HWE and LE (10,000 dememorization steps, 1,000 batches and a batch length of 10,000 iterations). The p-value ($p < 0.05$) of the HWE and LE multiple exact tests was corrected using the false discovery rate (Benjamini and Hochberg 1995).

Contemporary genetic diversity and structure

To characterize genetic diversity in the seven demes, observed heterozygosity (H_O), expected heterozygosity (H_E), mean number of alleles (N_A) and mean population relatedness (R ; Queller and Goodnight 1989) were estimated in GenAlEx version 6.502 (Peakall and Smouse 2012). Additionally, unbiased allelic richness (AR) and private allelic richness ($P-AR$) were calculated in HP-RARE version 1.0 (Kalinowski 2005). Two measures of population genetic differentiation were estimated and compared: (1) pairwise F_{ST} (Weir and Cockerham 1984), calculated in R (R Development Core Team

2015) package *diveRsity* version 1.9.89 (Keenan et al. 2013); and (2) the conditional genetic distance (cGD) (Dyer and Nason 2004), estimated in R package *gstudio* version 1.3 (Dyer 2014). Respective 95% confidence intervals (CIs) were calculated based on 1,000 bootstrap replicates, and pairwise estimates were considered significant when 95% CIs did not overlap zero.

Population genetic structure was assessed using the Bayesian algorithm implemented in STRUCTURE version 2.3.4 (Pritchard et al. 2000). Ten independent runs were performed for a number of clusters (K) ranging between 1-8. A burn-in period of 100,000 iterations, followed by 1,000,000 iterations, with a correlated allele frequencies admixture model and no prior information regarding population of origin was set for each run. The best K describing the observed genetic data was identified with STRUCTURE HARVESTER version 0.6.94 (Earl et al. 2012), following two independent criteria: (1) the K exhibiting the highest mean value of likelihood (Pritchard et al. 2000); and (2) the K showing the lowest ΔK statistic value (Evanno et al. 2005). For the optimal K, runs were summarized and graphically displayed using Pophelper version 1.0.10 (Francis 2016).

Additionally, a discriminant analysis of principal components (DAPC; Jombart et al. 2010), a multivariate method implemented in R package *adegenet* version 2.0.1 (Jombart et al. 2008), was used for an independent description of population structure, because this method does not assume HWE or LE. DAPC summarizes the data to minimize genetic differentiation within groups while maximizing it between groups, potentially unravelling complex structures, like hierarchical clustering or clinal differentiation (e.g. isolation by distance; Jombart et al. 2010). A prior number of genetic clusters of K=7 (the lowest value obtained in the analysis comparing clustering solutions using the Bayesian Information Criterion, BIC) was chosen, while the number

of retained principal components and discriminant functions were selected following the package's manual guidelines to avoid overfitting.

Ecological niche-based modelling

Ecological niche-based models (ENMs) were used to reconstruct present and past (Last Interglacial [LIG; ~120,000 - 140,000 years BP], Last Glacial Maximum [LGM; ~21,000 years BP] and Mid-Holocene [~6,000 years BP]) climatic niches of *S. s. longirostris* with the maximum entropy algorithm implemented in MAXENT version 3.3.3k (Phillips et al. 2006). MAXENT has shown to be highly reliable when using presence-only data and small datasets (Elith et al. 2010).

ENMs for the present were built using 28 occurrence records obtained from our own fieldwork, collaborators, and from the Spanish Herpetological Atlas (Online Resource 2.3). To avoid statistical biases arising from spatial autocorrelation (Fourcade et al. 2014), a maximum of one presence record per 10 x 10 km grid cell and a minimum distance of 2 km (maximum distance recorded for the species; Hendrix et al. 2017) was allowed. The Nearest Neighbour index function of ArcGIS (ESRI 2012) was used to assess the degree of data clustering (z -score = -0.435). From the 19 climatic variables with a 1km² resolution downloaded from WorldClim (Hijmans et al. 2005, www.worldclim.org), we retained six climate variables (with pairwise Pearson's $r \leq 0.75$) to represent contemporary climatic conditions and avoid the possible statistical effects of collinearity among predictor variables on niche modeling: Isothermality, Mean Temperature of Wettest Quarter, Mean Temperature of Driest Quarter, Mean Temperature of Coldest Quarter, Annual Precipitation and Precipitation of Warmest Quarter. Models built with current conditions were projected to past climates (LIG, LGM, Mid-Holocene) using the same six variables. LGM variables were downloaded at

2.5 arc minutes resolution (~5x5km), while the remaining variables were downloaded at 30 arc-seconds resolution (~1x1km).

The distribution area for model construction, validation and projections was created using a 200-km radius buffer around occurrence records. Bootstrap subsampling was used to build 50 model replicates, retaining 30% of the data records in order to test each replicate. The “Fade by clamping” option was used for past projections to avoid spurious model projections (Elith et al. 2011). Default values were used for all other settings. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC) in MAXENT, which ranges from 0.5 (complete randomness) to 1 (perfect discrimination) (Phillips et al. 2006).

Consensus models and projections to past climatic conditions were converted to binary projections (i.e. presence/absence) using the minimum training presence threshold (suitability values higher or equal to the lowest value given to a distribution record are considered suitable areas). The four obtained suitability maps (Current, Mid-Holocene, LGM (CCSM and MIROC) and LIG) were then overlapped in ArcGIS to identify stable climatic areas over time (i.e. common potential areas of occurrence which could serve as refugia in all time periods).

Landscape genetic analyses

We assessed the effect of contemporary geographic heterogeneity on gene flow by calculating matrices of pairwise resistance distances in CIRCUITSCAPE version 4.0 (McRae 2006) to test for patterns of IBD and IBR. CIRCUITSCAPE calculates averaged pairwise resistances to gene flow among populations based on all possible paths (unlike the least-cost path), thus better explaining the movement of genes among

widely separated regions over many generations (McRae et al. 2008). For the IBD model, a matrix of geographic distances was estimated from a raster layer depicting a “flat” landscape (i.e. a value of one in all cells). Patterns of IBR were examined through the calculation of three independent matrices of pairwise resistance distances ($IBR_{CLIMATIC}$, IBR_{SLOPE} and IBR_{NDVI}). We first used MAXENT to generate independent ENMs based on: (1) the six bioclimatic variables; (2) slope; and (3) the Normalized Difference Vegetation Index (NDVI; 250 x 250 m). Then, each ENM map was used as input for the program CIRCUITSCAPE to generate pairwise resistance matrices for each model (Online Resources 1.1 and 2.4).

Prior to matrix calculation of the $IBR_{CLIMATIC}$, IBR_{SLOPE} and IBR_{NDVI} models, we converted suitability scores (S) of each ENM layer’s cell (S ranges from 0, unsuitable, to 1, suitable) to resistance scores (R) using a simple arithmetic operation ($R=(1-S)/100$). In all CIRCUITSCAPE analyses, the pairwise mode and a cell connection scheme of eight neighbours were set, while remaining parameters were kept as default.

The role of environmental dissimilarity (IBE) on genetic differentiation was assessed by estimating climate ($IBE_{CLIMATIC}$) and NDVI (productivity index; IBE_{NDVI}) dissimilarity matrices between deme locations. Local environmental variation of the six bioclimatic variables was captured by a principal component analysis, with “varimax” rotation applied to the values of the six climatic variables extracted from the location of populations, occurrence records and 1000 randomly generated points in the study area. Then, the three principal components of each population location were selected and climatic dissimilarities were obtained by calculating pairwise Euclidean distances among populations. Principal component analyses were performed using IBM SPSS version 24.0 (IBM Corp., Armonk, NY, USA). Habitat dissimilarities were obtained by

calculating Euclidean distances between population location values extracted from the previously built NDVI-based ENM model.

Matrices of geographic (IBD), resistance (IBR_{CLIMATIC}, IBR_{NDVI} and IBR_{SLOPE}) and environmental (IBE_{CLIMATIC} and IBE_{NDVI}) heterogeneity were regressed against matrices of pairwise genetic differentiation (pairwise F_{ST} and cGD). Univariate regressions with randomizations were used, which is a suitable approach when the hypothesis to be tested strictly concerns distances (Legendre and Fortin 2010). Furthermore, to quantify the independent effects of each factor on genetic differentiation, we implemented a multivariate matrix regression with randomization approach (MMRR, Wang 2013). We first examined multicollinearity among our predictor variables using Pearson's correlation tests and variance inflation factors (VIF) (Zuur et al. 2010; Prunier et al. 2015).

Results

Phylogenetic analyses

The *cyt-b* dataset comprised 979 base pairs (17 variable and 11 parsimony informative), and we recovered 19 haplotypes among the 42 samples analyzed. An identical phylogenetic tree topology was obtained in analyses assuming a lognormal or a strict molecular clock. In both cases, we recovered two well-supported clades (BPP > 0.95) and a third moderately supported clade (BPP = 0.9) (Fig. 2). One main clade groups most of the samples (n = 31) and covers the central range of *S. s. longirostris*, with a sister clade (n = 5) including samples at the southern end of the range. The third clade (n = 6) is sister to the other two clades and includes the easternmost populations. *Salamandra s. longirostris* sequences differed by 6% and 7% from *S. s. crespoides* and *S.*

s. morenica, respectively, while *S. s. crespoi* and *S. s. morenica* differed 2%. Intra-clade distances within *S. s. longirostris* were low (1%).

Microsatellite pre-treatment procedures

A total of 146 samples were successfully genotyped. Nine individuals were identified as full-siblings, but including them did not significantly affect genetic diversity and population differentiation estimates ($r > 0.99$ in the full vs reduced dataset). Therefore, all samples were retained for downstream genetic analyses (Online Resource 2.1).

No microsatellite loci showed evidence of null alleles. Three pairs of loci showed signs of linkage disequilibrium, though deviations of LE were not consistent among populations. Additionally, significant deviations from HWE were only found in the easternmost population (VVR) due to the presence of four monomorphic loci. Since deviation from LE and HWE were not consistent across populations, we used all loci for subsequent analyses.

Contemporary genetic diversity and genetic structure

The easternmost population (VVR) showed extremely low genetic diversity ($N_A = 2.22$, $H_O = 0.31$, $H_E = 0.27$, and $AR = 2.01$) and high relatedness values ($R = 0.82$) in comparison with the central group including the other six populations (Table 1). Within this group, the easternmost population (TOR) had the lowest genetic diversity values, followed by the westernmost population (MSI) (Table 1).

Population genetic differentiation was high overall (mean $F_{ST} = 0.24$; mean cGD = 2.69), with VVR exhibiting the highest genetic differentiation (range $F_{ST} = 0.38$ -0.56; range cGD = 4.10-5.19), followed by TOR (range $F_{ST} = 0.22$ -0.30; range cGD = 1.93-

3.51) (Table 2). STRUCTURE identified K=2 as the number of clusters best describing genetic structure (based on both the Ln Pr(X/K) and the ΔK methods), separating the easternmost population VVR from the remaining populations (Fig. 3). The second most supported K (K=5) showed high levels of admixture among central populations in the main group (MSI, PLAG, LMO, AVI), with low or no admixture at the range margins (Fig. 3). DAPC results were congruent with STRUCTURE's output, with the highest differentiation in the easternmost population (VVR), followed by the two easternmost populations from the central group (TOR and VLR) (Online Resource 3.1).

Ecological niche-based modelling

The ENM had high AUC values (AUC = 0.977), suggesting good predictive power. Among variables used in the model, Annual precipitation and Mean temperature of the driest quarter presented the highest contribution values (Online Resource 3.2). Climatically suitable areas identified for the present (Online Resource 3.3) largely depict the known current distribution of *S. s. longirostris*, with the exception of some areas north of the GRB. Climatically suitable areas identified for past scenarios (LIG, LGM and mid-Holocene) varied considerably, with LIG showing a suitable area similar to that for the present interglacial (Online Resource 3.3), while projections to the mid-Holocene and especially the LGM predicted larger suitable areas (Online Resource 3.3). Stable climatic areas derived from the intersection of these projections identified a continuous area concomitant with the current distribution range of *S. s. longirostris* as the main potential climatic refuge. In addition, small and partially disconnected areas were identified at the eastern end of the range and north of the GRB (Fig. 2).

Habitat and climatic-based ENMs had moderate to high AUC values, respectively (Online Resource 2.4). The connectivity analysis performed with CIRCUITSCAPE revealed overall moderate or high connectivity among populations of *S. s. longirostris* for all studied variables (climate, habitat, topography and distance), with the exception of population VVR, which is isolated by an area of ca. 50 Km from its closest population (Fig. 4).

Univariate regressions between ecological (IBD, $IBR_{CLIMATIC}$, IBR_{NDVI} , IBR_{SLOPE} , $IBE_{CLIMATIC}$ and IBE_{NDVI}) and genetic matrices (F_{ST} and cGD) revealed significant ($p < 0.05$) positive relationships, with the exception of both IBE matrices (Table 3). Univariate regressions using IBR_{NDVI} provided the highest coefficients of determination ($r^2 = 0.96$ and $r^2 = 0.88$ for F_{ST} and cGD, respectively; Table 3). Moreover, MMRR analyses retrieved an overall significant ($p < 0.05$) coefficient of determination of 0.99 and 0.92 for F_{ST} and cGD, respectively, though IBR_{NDVI} was the only significant ($p < 0.05$) predictor when regressed against pairwise F_{ST} (Table 4). Multicollinearity, however, might have influenced these models, as the input variables were not independent, showing pairwise correlation values and VIF values above the traditionally considered thresholds (i.e. $r > 0.7$ and $VIF > 10$), especially in $IBR_{CLIMATIC}$ and IBR_{NDVI} (Online Resource 2.5).

Discussion

Contemporary patterns of genetic diversity and structure are the result of complex evolutionary processes occurring at different temporal depths (e.g. Velo-Antón et al. 2013; Zhang et al. 2016; Gutiérrez-Rodríguez et al. 2017a). Our results highlight the

importance of both historical and current factors in shaping patterns of genetic variation in a Mediterranean relict, *S. s. longirostris*, with implications for the conservation of this taxon.

The Guadalquivir River Basin as a vicariant agent in S. s. longirostris

Previous phylogenetic studies have shown deep mtDNA divergence in *S. s. longirostris* with respect to other population lineages in *S. salamandra* (e.g. García-París et al. 1998, 2003; Steinfartz et al. 2000; Vences et al. 2014; Beukema et al. 2016). This, coupled with their disjunct distribution, led some authors to postulate GRB as the major vicariant agent promoting isolation and subsequent genetic differentiation of *S. s. longirostris* south of the GRB, a pattern in common with other co-distributed species (e.g. Santos et al. 2012; Velo-Antón et al. 2012a). Our palaeoclimatic niche modeling analyses have revealed the potential occurrence of areas suitable for *S. s. longirostris* north and south of the GRB since the LIG. Thus, it is possible that areas north of the GRB might have been occupied by *S. s. longirostris* during the LGM, which later became extinct because of environmental changes and/or competition with other species, including other lineages of *S. salamandra*. Alternatively, areas north of the GRB may have never been colonized by *S. s. longirostris* due to the physical barrier imposed by the GRB, which likely favored an allopatric differentiation process from neighbouring lineages of *Salamandra* (*S. s. crespai* / *S. s. morenica*). These scenarios can be further tested through the assessment of historical patterns of nuclear gene flow between *S. s. longirostris* and these lineages, which may have also been affected by the physical barrier imposed by the GRB.

Distribution dynamics and historical and contemporary patterns of genetic variation

Patterns of genetic structure at both historical (mtDNA) and contemporary scales (microsatellites) were mostly concordant, and support the predictions based on ENMs, with peripheral populations (VVR, TOR, VLR and MSI) exhibiting lower genetic diversity and greater genetic differentiation when compared to populations near the core of the distribution (Fig. 3; Tables 1 and 2). This pattern has also been reported for eastern European populations of *S. salamandra* at similar geographic scales (Najbar et al. 2015; Konowalik et al. 2016).

Climatic oscillations during the Pleistocene promoted cycles of population expansion and contraction, which likely caused demographic and genetic effects on peripheral vs central populations of *S. s. longirostris*. During range expansions, peripheral populations at the expansion front are subjected to serial founder events, with small effective population sizes resulting in spatial patterns of decreasing genetic diversity and increasing genetic divergence through processes like drift or allele surfing (e.g. Excoffier et al. 2009; Pereira et al. 2018). The high genetic diversity observed in central populations is consistent with demographic stability associated with the existence of a climatic refugium in this area. In contrast, peripheral populations of *S. s. longirostris* are located in areas where climatic suitability varied through time, which probably promoted recurrent population extinctions followed by recolonization (founder) events by individuals from central populations, resulting in the low levels of genetic diversity and high genetic differentiation observed (Excoffier et al. 2009). Among peripheral populations, the easternmost population (VVR) shows divergent mtDNA haplotypes (1%; Fig. 2), although one haplotype is shared with one individual in the geographically closest population (TOR; ~100 km distant). The level of genetic divergence between *S. s. longirostris* clades (1%) suggests their separation occurred

during the Pleistocene. The population VVR also showed extremely low levels of genetic diversity ($H_O = 0.3$) and high differentiation ($F_{ST} > 0.5$). This pattern reflects the genetic effects of long-term isolation and demographic instability, which have also been assessed in other populations of *S. salamandra*, including insular populations isolated during the Holocene (Velo-Antón et al. 2012b), or small-sized urban populations (Lourenço et al. 2017).

In addition to the homogeneous, climatically stable area inferred by ENMs in the core of the range, from the southernmost coast (strait of Gibraltar) to the Grazalema mountains at the Cádiz and Málaga border, an isolated climatically stable area was inferred to occur at the easternmost end of the range of *S. s. longirostris* (northwest of the Tejeda and Almjara mountains in Málaga), which is presently occupied by VVR. Thus, climatic fluctuations have probably had a major effect on historical patterns of population connectivity in *S. s. longirostris* by alternatively promoting isolation and recolonization dynamics through changes in habitat suitability.

Environmental factors shaping current genetic structure in S. s. longirostris

Models characterizing geographical heterogeneity (i.e. IBD and IBR models) were better predictors of patterns of genetic differentiation in *S. s. longirostris* than IBE models. This suggests that gene flow between populations in different ecological settings is not limited by selection against dispersers or by individual preferences to remain in a particular environment (Wang 2013; Wang and Bradburd 2014). Moreover, the IBR_{NDVI} showed the strongest statistical association with genetic structure. This result highlights the greater importance of landscape composition and structure (in this case the presence of forested or agricultural areas, Online Resource 3.4) on genetic connectivity relative to climate, stressing the crucial role of forests in maintaining gene

flow among salamander populations (e.g. Cushman 2006; Todd et al. 2009). The importance of preserving forest habitats is particularly pressing for southern Iberian populations of *S. salamandra*. Forests are extremely important for salamander survival, providing food resources, shelters, and aquatic habitats for reproduction (Egea-Serrano et al. 2006; Romero et al. 2013). The importance of forests and vegetation acting as microclimate buffers in this area and in neighboring regions has also been emphasized for other amphibian species (e.g. Dias et al. 2014; Escoriza and Ben Hassine 2014).

Overall, our landscape genetics approach suggests that models incorporating both geographic distances and ecologically relevant climate or landscape features provide better predictions of observed genetic patterns at fine spatial scales (e.g. Muñoz-Pajares et al. 2016; Gutiérrez-Rodríguez et al. 2017b). Unfortunately, it was not possible to disentangle the relative role of each variable in our statistical framework. This is a recurring problem in landscape genetic studies arising from the non-independence of predictor variables (i.e. collinearity) in regression analyses, which inflates the variance of regression parameters and ultimately leads to the spurious identification of relevant predictors in the model (Dormann et al. 2013).

Implications for conservation under rapid changing environments

Salamandra s. longirostris is a good example of the ongoing biodiversity crisis faced by Iberian amphibians due to the rapid climatic (Carvalho et al. 2011) and landscape changes (Ferreira and Beja 2013). It also represents the high levels of endemism found in southern Iberia, and the evolutionary distinctiveness of allopatric lineages with restricted ranges occurring in this area (e.g. Santos et al. 2012; Velo-Antón et al. 2012a, Dias et al. 2014), many of which have low vagility, which makes them particularly vulnerable to rapid environmental changes. Maintaining population genetic connectivity

is key to preserve evolutionary potential and, consequently, to ensure their long-term persistence (Baguette et al. 2013). Reduced gene flow can result in lower effective population sizes and genetic diversity, causing genetic erosion and eventually local extinction events (Frankham 2005; Cushman 2006). Our results reveal limited genetic connectivity (and low genetic diversity) in peripheral populations (mainly the easternmost population). It is thus crucial to preserve forested patches between central and peripheral populations, but especially among the latter, to maintain population connectivity and help to ensure the long-term conservation of *S. s. longirostris*. Such forested habitats also play a crucial role as microrefugia, which may reduce the risk of local extinction of fire salamanders and other species under predicted scenarios of climate change in the Iberian Peninsula (Carvalho et al. 2011). These results have general implications for biodiversity conservation under rapidly changing environments (e.g. Velo-Antón et al. 2013; Nogueras et al. 2016; Carvalho et al. 2017; Razgour et al. 2017) but can also inform management plans for the long term survival of this short-range endemic salamander. Two management units corresponding to the central and eastern population groups can be established, with the latter requiring urgent action to compensate the negative effects of reduced genetic variation on long-term population persistence.

Acknowledgments

We thank David Buckley, David Donaire, Francisco Jiménez Cazalla, Jesús Díaz-Rodríguez, Luis García-Cardenete and Saúl Yubero for providing samples and help with field work. S. Lopes helped with genotyping. Fieldwork for obtaining tissue samples was done with the corresponding permits from the regional administrations. This work was funded by FEDER funds through the Operational Programme for

Competitiveness Factors—COMPETE—and by National Funds through FCT—
 Foundation for Science and Technology—under the PTDC/BIA-EVF/3036/2012,
 PTDC/BIA-BEC/099915/2008, POCI-01-0145-FEDER-006821 and FCOMP- 01-0124-
 FEDER-028325. GVA, HG, GCD, AL and PT are supported by FCT (IF/01425/2014,
 SFRH/BPD/102966/2014, SFRH/BD/89750/2012, PD/BD/106060/2015,
 SFRH/BPD/93473/2013, respectively.

References

- Alcobendas M, Castanet J (2000) Bone growth plasticity among populations of
Salamandra salamandra: interactions between internal and external factors.
Herpetologica 56:14-26.
- Allendorf FW, Luikart GH, Aitken SN (2013) Conservation and the Genetics of
 Populations, 2nd edn. Wiley-Blackwell, West Sussex.
- Álvarez D, Lourenço A, Oro D, Velo-Antón G (2015) Assessment of census (N) and
 effective population size (Ne) reveals consistency of Ne single-sample estimators
 and a high Ne/N ratio in an urban and isolated population of fire salamanders.
Conserv Genet Resour 7:705-712.
- Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C (2013) Individual
 dispersal, landscape connectivity and ecological networks. *Biol Rev* 88:310-326.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and
 powerful approach to multiple testing. *J Roy Stat Soc B Met* 57:289-300.
- Beukema W, Niecieza AG, Lourenço A, Velo-Antón G (2016) Colour polymorphism in
Salamandra salamandra (Amphibia: Urodela), revealed by a lack of genetic and
 environmental differentiation between distinct phenotypes. *J Zool Syst Evol Res*
 54:127-136.
- Carvalho SB, Brito JC, Crespo EG, Watts ME, Possingham HP (2011) Conservation
 planning under climate change: toward accounting for uncertainty in predicted
 species distributions to increase confidence in conservation investments in space and
 time. *Biol Conserv* 144: 2020-2030.
- Carvalho SB, Velo-Antón G, Tarroso P, Portela AP, Barata M, Carranza S, Mortiz C,
 Possingham HP (2017) Spatial conservation prioritization of biodiversity spanning
 the evolutionary continuum. *Nat Ecol Evol* 1: 0151.
- Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding
 coefficients. *J Hered* 100:106-113.
- Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: a review
 and prospectus. *Biol Conserv* 128:231-240.

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

574 Dias G, Beltrán JF, Tejedo M, Benítez M, González-Miras E, Ferrand N, Gonçalves H
575 (2015) Limited gene flow and high genetic diversity in the threatened Betic midwife
576 toad (*Alytes dickhilleni*): evolutionary and conservation implications. Conserv Genet
577 16: 459-476

578 Díaz-Rodríguez J, Gonçalves H, Sequeira F, Sousa-Neves T, Tejedo M, Ferrand N,
579 Martínez-Solano I (2015) Molecular evidence for cryptic candidate species in Iberian
580 *Pelodytes* (Anura, Pelodytidae). Mol Phylogenetics Evol 83:224-241.

581 Donaire-Barroso D, González de la Vega JP, Barnestein JAM (2009) Aportación sobre
582 los patrones de diseño pigmentario en *Salamandra longirostris* Joger and Steinfartz,
583 1994, y nueva nomenclatura taxonómica. Butll Soc Cat Herp 18:10-17.

584 Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber
585 B, Lafourcade B, Leitão PJ, Münkemüller T (2013) Collinearity: a review of
586 methods to deal with it and a simulation study evaluating their performance.
587 Ecography 36:27-46.

588 Drummond A J, Suchard MA, Xie D, Rambaut A (2012). Bayesian phylogenetics with
589 BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969-1973.

590 Dyer RJ, Nason JD (2004) Population graphs: the graph- theoretic shape of genetic
591 structure. Mol Ecol 13:1713-1728.

592 Dyer RJ (2014) Gstudio: a Package for the Spatial Analysis of Population Genetic
593 Marker Data. Virginia Commonwealth University: Richmond, VA, USA.

594 Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program
595 for visualizing STRUCTURE output and implementing the Evanno method. Conserv
596 Genet Resour 4:359-361.

597 Egea-Serrano A, Oliva-Paterna FJ, Torralva M (2006) Breeding habitat selection of
598 *Salamandra salamandra* (Linnaeus, 1758) in the most arid zone of its European
599 distribution range: application to conservation management. Hydrobiologia 560:363-
600 371.

601 Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species.
602 Methods Ecol Evol 1:330-342.

603 Escoriza D, Hassine JB (2014) Microclimatic variation in multiple *Salamandra algira*
604 populations along an altitudinal gradient: phenology and reproductive strategies.
605 Acta Herpetol 9:33-41.

606 ESRI (2012) ArcGIS 10.1. Redlands, California.

607 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals
608 using the software STRUCTURE: a simulation study. Mol Ecol 14:2611-2620.

609 Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. Annu
610 Rev Ecol Evol Syst 40:481-501.

611 Ferreira M, Beja P (2013) Mediterranean amphibians and the loss of temporary ponds:
612 Are there alternative breeding habitats?. Biol Conserv 165:179-186.

613 Fourcade Y, Engler JO, Rödder D, Secondi J (2014) Mapping species distributions with
 614 MAXENT using a geographically biased sample of presence data: a performance
 615 assessment of methods for correcting sampling bias. *PloS ONE* 9: e97122.

616 Francis RM (2017) Pophelper: an R package and web app to analyse and visualize
 617 population structure. *Mol Ecol Resour* 17:27-32.

618 Frankham R (2005) Genetics and extinction. *Biol Conserv* 126:131-140.

619 Frost DR (2018) Amphibian Species of the World. Online resource. Available at:
 620 <http://research.amnh.org/herpetology/amphibia/index.html>.

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

624 García-París M, Alcobendas M, Buckley D, Wake DB (2003). Dispersal of viviparity
 625 across contact zones in Iberian populations of fire salamanders (*Salamandra*)
 626 inferred from discordance of genetic and morphological traits. *Evolution* 57:129-143.

627 Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho B, García-París M, Ferrand N
 628 (2009) High levels of population subdivision in a morphologically conserved
 629 Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence
 630 from mtDNA, microsatellites and nuclear genealogies. *Mol Ecol* 18:5143-5160.

631 Gutiérrez-Rodríguez J, Barbosa AM, Martínez-Solano I (2017a) Present and past
 632 climatic effects on the current distribution and genetic diversity of the Iberian
 633 Spadefoot toad (*Pelobates cultripes*): an integrative approach. *J Biogeog* 44:245-258.

634 Gutiérrez-Rodríguez J, Goncalves J, Civantos E, Martínez-Solano I (2017b)
 635 Comparative landscape genetics of pond-breeding amphibians in Mediterranean
 636 temporal wetlands: the positive role of structural heterogeneity in promoting gene
 637 flow. *Mol Ecol* 26:5407-5420.

638 He Q, Edwards DL, Knowles LL (2013) Integrative testing of how environments from
 639 the past to the present shape genetic structure across landscapes. *Evolution* 67:3386-
 640 3402.

641 Hendrix R, Schmidt BR, Schaub M, Krause ET, Steinfartz S (2017) Differentiation of
 642 movement behavior in an adaptively diverging salamander population. *Mol Ecol*
 643 26:6400-6413.

644 Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution
 645 interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965-1978.

646 Joger, U, Steinfartz S (1994) Zur subspezifischen Gliederung der siidiberischen
 647 Feuersalamander (*Salamandra salamandra*-Komplex). *Abh. Ber. Nat.* 17:83-98.

648 Jombart T, Devillard S, Dufour A-B, Pontier D (2008) Revealing cryptic spatial patterns
 649 in genetic variability by a new multivariate method. *Heredity* 101:92–103.

650 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal
 651 components: a new method for the analysis of genetically structured populations.
 652 *BMC Genet* 11: 94.

653 Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference
 654 from multilocus genotype data. *Mol Ecol Resour* 10:551-555.

655 Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction
656 on measures of allelic richness. *Mol Ecol Resour* 5:187-189.

657 Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: An R
658 package for the estimation and exploration of population genetics parameters and
659 their associated errors. *Methods Ecol Evol* 4:782-788.

660 Konowalik A, Najbar A, Babik W, Steinfartz S, Ogielska M (2016) Genetic structure of
661 the fire salamander *Salamandra salamandra* in the Polish Sudetes. *Amphib Reptil*
662 37:405-415.

663 Lee C-R, Mitchell-Olds T (2011) Quantifying effects of environmental and
664 geographical factors on patterns of genetic differentiation. *Mol Ecol* 20:4631-4642.

665 Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches
666 for detecting complex multivariate relationships in the spatial analysis of genetic
667 data. *Mol Ecol Resour* 10:831-844.

668 Lourenço A, Álvarez D, Wang IJ, Velo-Antón G (2017) Trapped within the city:
669 Integrating demography, time since isolation and population-specific traits to assess
670 the genetic effects of urbanization. *Mol Ecol* 26:1498-1514.

671 Martínez-Freiría F, Velo-Antón G, Brito JC (2015) Trapped by climate: interglacial
672 refuge and recent population expansion in the endemic Iberian adder *Vipera seoanei*.
673 *Divers Distributions* 21:331-344.

674 McRae BH (2006) Isolation by resistance. *Evolution*, 60:1551-1561.

675 McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model
676 connectivity in ecology, evolution, and conservation. *Ecology*, 89:2712-2724.

677 Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for
678 inference of large phylogenetic trees. In *Gateway Computing Environments*
679 *Workshop (GCE)*, 2010:1-8.

680 Muñoz-Pajares AJ, García C, Abdelaziz M, Bosch J, Perfectti F, Gómez JM (2017)
681 Drivers of genetic differentiation in a generalist insect-pollinated herb across spatial
682 scales. *Mol Ecol* 26:1576-1585.

683 Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca, GA, Kent J (2000)
684 Biodiversity hotspots for conservation priorities. *Nature* 403:853-858.

685 Najbar A, Babik W, Najbar B, Ogielska M (2015) Genetic structure and differentiation
686 of the fire salamander *Salamandra salamandra* at the northern margin of its range in
687 the Carpathians. *Amphib Reptil* 36:301-311.

688 Nogueras V, García-Navas V, Cordero PJ, Ortego J (2016) The role of environment
689 and core-margin effects on range-wide phenotypic variation in a montane
690 grasshopper. *J Evol Biol* 29:2129-2142.

691 Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population
692 genetic software for teaching and research – an update. *Bioinformatics* 28:2537-
693 2539.

694 Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species
695 geographic distributions. *Ecol Model* 190:231-259.

696 Pleguezuelos JM (2004) Atlas y libro rojo de los anfibios y reptiles de España (p. 584).
697 Márquez R, Lizana M (Eds.). Organismo Autónomo de Parques Nacionales. Spain.

698 Pereira P, Teixeira J, Velo-Antón G (2018) Allele surfing shaped the genetic structure
699 of the European pond turtle via the colonization and population expansion across the
700 Iberian Peninsula from Africa. *J Biogeogr.* doi.org/10.1111/jbi.13412

701 Peterman W, Brocato ER, Semlitsch RD, Eggert LS (2016) Reducing bias in population
702 and landscape genetic inferences: the effects of sampling related individuals and
703 multiple life stages. *PeerJ* 4:e1813; DOI 10.7717/peerj.1813.

704 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
705 multilocus genotype data. *Genetics* 155:945-959.

706 Prunier JG, Colyn M, Legendre X, Nimon KF, Flamand MC (2015) Multicollinearity in
707 spatial genetics: separating the wheat from the chaff using commonality analyses.
708 *Mol Ecol* 24:263-283.

709 Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers.
710 *Evolution* 43:258-275.

711 Razgour O, Taggart JB, Manel S, Juste J, Ibáñez C, Rebelo H, Alberdi A, Gareth J, Park
712 K (2017) An integrated framework to identify wildlife populations under threat from
713 climate change. *Mol Ecol Resour.* doi.org/10.1111/1755-0998.12694

714 R Development Core Team (2015) R: A Language and Environment for Statistical
715 Computing. R Foundation for Statistical Computing, Vienna, Austria.

716 Rissler LJ (2016) Union of phylogeography and landscape genetics. *Proc Natl Acad Sci*
717 *USA* 113:8079-8086.

718 Romero D, Olivero J, Real R (2013) Comparative assessment of different methods for
719 using land-cover variables for distribution modelling of *Salamandra salamandra*
720 *longirotris*. *Environmental Conservation* 40:48-59.

721 Rousset F (2008) genepop'007: a complete re-implementation of the genepop software
722 for Windows and Linux. *Mol Ecol Resour* 8:103-106.

723 Sánchez-Montes G, Ariño AH, Vizmanos JL, Wang J, Martínez-Solano I (2017) Effects
724 of sample size and full sibs on genetic diversity characterization: a case study of
725 three syntopic Iberian pond-breeding amphibians. *J Hered* 108:535–543.

726 Santos X, Rato C, Carranza S, Carretero MA, Pleguezuelos JM (2012) Complex
727 phylogeography in the Southern Smooth Snake (*Coronella girondica*) supported by
728 mtDNA sequences. *J Zool Syst Evol Res* 50:210-219.

729 Sexton JP, Hangartner SB, Hoffmann AA (2014) Genetic isolation by environment or
730 distance: which pattern of gene flow is most common?. *Evolution* 68:1-15.

731 Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian
732 ecology and conservation: are all amphibian populations metapopulations?
733 *Ecography* 28: 110–128.

734 Steinfartz S, Veith M, Tautz D (2000) Mitochondrial sequence analysis of *Salamandra*
735 taxa suggests old splits of major lineages and postglacial recolonizations of Central
736 Europe from distinct source populations of *Salamandra salamandra*. *Mol Ecol*
737 9:397-410.

738 Steinfartz S, Kuesters D, Tautz D (2004) Isolation and characterization of polymorphic
739 tetranucleotide microsatellite loci in the Fire salamander *Salamandra salamandra*
740 (Amphibia: Caudata). *Mol Ecol Resour* 4:626-628.

741 Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues AS, Fischman DL, Waller R. W
742 (2004) Status and trends of amphibian declines and extinctions worldwide. *Science*
743 306:1783-1786.

744 Todd BD, Luhring TM, Rothermel BB, Gibbons JW (2009) Effects of forest removal on
745 amphibian migrations: implications for habitat and landscape connectivity. *J Appl*
746 *Ecol* 46:554-561.

747 Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER:
748 software for identifying and correcting genotyping errors in microsatellite data. *Mol*
749 *Ecol Resour* 4:535-538.

750 Velo-Antón G, Parra JL, Parra-Olea G, Zamudio KR (2013) Tracking climate change in
751 a dispersal-limited species: reduced spatial and genetic connectivity in a montane
752 salamander. *Mol Ecol* 22:3261-3278.

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

757 Velo-Antón G, Zamudio KR, Cordero-Rivera A (2012b) Genetic drift and rapid
758 evolution of viviparity in insular fire salamanders (*Salamandra salamandra*).
759 *Heredity* 108:410-418.

760 Velo-Antón G, Santos X, Sanmartín-Villar I, Cordero-Rivera A, Buckley D (2015)
761 Intraspecific variation in clutch size and maternal investment in pueriparous and
762 larviparous *Salamandra salamandra* females. *Evol Ecol* 29:185-204.

763 Velo-Antón G, Buckley D (2015) *Salamandra* común–*Salamandra salamandra*
764 (Linnaeus, 1758). Enciclopedia virtual de los vertebrados españoles (<http://www.vertebradosibericos.org>). Madrid: Museo Nacional de Ciencias Naturales (MNCN),
765 CSIC.

767 Vences M, Sanchez E, Hauswaldt JS, Eikelmann D, Rodríguez A, Carranza S, Donaire
768 D, Gehara M, Helfer V, Lötters S, Werner P, Schulz S, Steinfartz S (2014) Nuclear
769 and mitochondrial multilocus phylogeny and survey of alkaloid content in true
770 salamanders of the genus *Salamandra* (Salamandridae). *Mol Phylogenetics Evol*
771 73:208-216.

772 Wang IJ, Bradburd GS (2014) Isolation by environment. *Mol Ecol* 23:5649-5662.

773 Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic
774 variation: a multiple matrix regression approach for quantifying geographic and
775 ecological isolation. *Evolution* 67:3403-3411.

776 Wang IJ, Glor RE, Losos JB (2013) Quantifying the roles of ecology and geography in
777 spatial genetic divergence. *Ecol Lett* 16:175-182.

778 Wang J (2018) Effects of sampling close relatives on some elementary population
779 genetics analyses. *Mol Ecol Resour* 18:41–54.

780 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population
781 structure. *Evolution* 38:1358-1370.

782 Wright S (1943) Isolation by distance. *Genetics* 28:114-138.

783 Zellmer AJ, Knowles LL (2009). Disentangling the effects of historic vs. contemporary
784 landscape structure on population genetic divergence. *Mol Ecol* 18:3593-3602.

785 Zhang P, Papenfuss TJ, Wake MH, Qu L, Wake DB (2008) Phylogeny and
786 biogeography of the family Salamandridae (Amphibia: Caudata) inferred from
787 complete mitochondrial genomes. *Mol Phylogenetics Evol* 49:586-597.

788 Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX (2016) Contributions of historical
789 and contemporary geographic and environmental factors to phylogeographic
790 structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). *Sci Rep*
791 6:24041.

792 Zuur AF, Ieno EN, Elphick, CS (2010) A protocol for data exploration to avoid
793 common statistical problems. *Methods Ecol Evol* 1:3-14.

794

Figures

Figure 1. Distribution range of *S. s. longirostris* (in green) and neighbouring *S. salamandra* populations (in red) separated by the Gualdalquivir River Basin. Cells denote 10x10 km distribution records of *S. salamandra* based on herpetological atlases from Spain. Inset represents the distribution of *S. salamandra* with the allopatric distribution of *S. s. longirostris* in the southern Iberian Peninsula. Black circles indicate sampling locations for molecular analyses.

Figure 2. Map depicting climatically stable areas, suitable for *S. s. longirostris*, obtained through overlapping past (LIG, LGM and mid-Holocene) and present ENMs. A Bayesian phylogenetic tree based on *cyt-b* is also shown, with colors representing the three mtDNA clades discussed in the text. Black and grey dots represent nodes with posterior probability higher than 0.95 or equal to 0.9, respectively. The scale indicates the proportion of substitutions/site along the branch. Salamander image credit: Francisco Jiménez Cazalla.

Figure 3. Map representing the NDVI raster layer of the study area. The STRUCTURE barplots below the map represent the two most supported numbers of clusters (K=2 and K=5) for the seven analysed populations. Population pie charts represent individual cluster memberships for K=5.

Figure 4. Current connectivity among populations of *S. s. longirostris* according to CIRCUITSCAPE analyses based on climatic ($IBR_{CLIMATIC}$), slope (IBR_{SLOPE}), productivity (IBR_{NDVI}), and distance (IBD) variables.

816 **Figure 1**

817

818

819

Figure 2

Figure 3

841 **Figure 4**

842

843

Tables

Table 1. Sampling information and genetic diversity statistics calculated from microsatellite markers for each population. The following parameters are displayed: sample size (N), mean number of alleles (N_A), observed heterozygosity (H_O), expected unbiased heterozygosity (H_E), allelic richness (AR), private allelic richness (P-AR) and relatedness (R).

Population	N	N _A	H _O	H _E	AR	P-AR	R
MSI	12	4.67	0.67	0.67	4.20	0.24	0.29
PLAG	19	6.44	0.66	0.71	5.33	0.40	0.15
LMO	10	6.33	0.68	0.77	5.99	0.62	0.05
AVI	23	8.00	0.68	0.73	5.77	1.08	0.13
VLR	15	6.11	0.54	0.65	5.12	0.94	0.28
TOR	12	3.89	0.50	0.56	3.59	0.55	0.45
VVR	19	2.22	0.31	0.27	2.01	0.46	0.82

Table 2. Matrix of pairwise genetic differentiation between populations. Below and above the diagonal are pairwise F_{ST} and cGD, respectively. Statistically non-significant pairwise comparisons are highlighted in italics.

F _{ST} /cGD	MSI	PLAG	LMO	AVI	VLR	TOR	VVR
MSI		1.575	1.085	1.688	2.876	1.935	5.014
PLAG	0.111		<i>0.490</i>	1.094	2.281	3.510	5.194
LMO	0.082	<i>0.022</i>		0.603	1.791	3.020	4.704
AVI	0.135	0.090	0.040		1.188	3.145	4.101
VLR	0.208	0.119	0.109	0.134		1.957	4.201
TOR	0.267	0.245	0.230	0.225	0.299		5.067
VVR	0.510	0.405	0.439	0.378	0.454	0.561	

Table 3. Results from univariate regressions with randomizations among matrices of genetic differentiation (pairwise F_{ST} and cGD values) and matrices of isolation among populations (IBD, IB_{RCLIMATIC}, IB_{RNDVI}, IB_{RSLOPE}, IB_{ECLIMATIC} and IB_{ENDVI}). The

coefficient of determination (r^2), regression coefficient (β), (t) and p-value (p) are displayed. Statistically significant predictors ($p < 0.05$) are highlighted in bold.

	F _{ST}				cGD			
	r^2	β	t	p	r^2	β	t	p
IBD	0.798	1.89	8.65	0.005	0.795	1.77	8.57	0.003
IBR CLIMATIC	0.881	1.14	11.88	0.014	0.857	1.05	10.68	0.002
IBR NDVI	0.956	1.18	20.22	0.002	0.877	1.06	11.66	0.002
IBR SLOPE	0.644	1.74	5.86	0.005	0.683	1.68	6.40	0.001
IBE CLIMATIC	0.077	0.33	1.26	0.557	0.067	0.29	1.17	0.469
IBE NDVI	0.157	0.39	1.88	0.380	0.104	0.30	1.49	0.384

Table 4. Results from multivariate regression with randomizations among matrices of genetic differentiation (pairwise F_{ST} and cGD values) and matrices of isolation among populations (IBD, IBR_{CLIMATIC}, IBR_{NDVI}, IBR_{SLOPE}, IBE_{CLIMATIC} and IBE_{NDVI}). The coefficient of determination (r^2), regression coefficient (β), (t) and p-value (p) are also displayed. Statistically significant predictors ($p < 0.05$) are highlighted in bold.

	F _{ST}				cGD			
	r^2	β	t	p	r^2	β	t	p
IBD	-	-0.21	-0.91	0.740	-	0.22	0.40	0.811
IBR CLIMATIC	-	-0.33	-1.35	0.645	-	-0.04	-0.07	0.979
IBR NDVI	-	1.85	7.60	0.001	-	1.15	1.99	0.257
IBR SLOPE	-	-0.18	-0.95	0.727	-	-0.03	-0.07	0.974
IBE CLIMATIC	-	-0.15	-1.76	0.176	-	0.10	0.47	0.674
IBE NDVI	-	-0.15	-2.33	0.322	-	-0.30	-1.96	0.270
Global	0.988	-	-	0.012	0.922	-	-	0.008

Figure 1

Figure 4

[Click here to view linked References](#)

[Click here to access/download](#)

Supplementary Material

[Antunes_etal_longirostris_OnlineResource1.doc](#)

[Click here to view linked References](#)

[Click here to access/download](#)

Supplementary Material

[Antunes_etal_longirostris_OnlineResource2.doc](#)

[Click here to view linked References](#)

[Click here to access/download](#)

Supplementary Material

[Antunes_etal_longirostris_OnlineResource3.doc](#)