

Received Date : 14-Oct-2015

Revised Date : 11-Jan-2016

Accepted Date : 01-Feb-2016

Article type : Original Article

**Title:** Hybridization during altitudinal range shifts: nuclear introgression leads to extensive cyto-nuclear discordance in the fire salamander

**Authors:** Ricardo J. Pereira<sup>1</sup>, Iñigo Martínez-Solano<sup>2,3</sup>, David Buckley<sup>4,5</sup>

1- Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350, Copenhagen, Denmark;

2- Instituto de Investigación en Recursos Cinegéticos (IREC-UCLM-CSIC-JCCM), Ronda de Toledo, s/n, 13005 Ciudad Real, Spain;

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

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

5- Current address: 1 Bailey Mews, Auckland Road, Cambridge CB5 8DR, UK

**Corresponding author:** David Buckley

**email:** davidbuckley15@gmail.com

**running title:** Hybridization during altitudinal range shifts

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.13575

This article is protected by copyright. All rights reserved.

Accepted Article

**keywords:** Climate change, global warming, gene flow, adaptation, postglacial expansion, ecomorphs.

**Abstract:**

Ecological models predict that, in the face of climate change, taxa occupying steep altitudinal gradients will shift their distributions, leading to the contraction or extinction of the high-elevation (cold-adapted) taxa. However, hybridization between eco-morphologically divergent taxa commonly occurs in nature and may lead to alternative evolutionary outcomes, such as genetic merger or gene flow at specific genes. We evaluate this hypothesis by studying patterns of divergence and gene flow across three replicate contact zones between high- and low-elevation ecomorphs of the fire salamander (*Salamanca salamandra*) that have experienced altitudinal range shifts over the current postglacial period. Strong population structure with high genetic divergence in mitochondrial DNA suggests that vicariant evolution has occurred over several glacial-interglacial cycles, and that it has led to cryptic differentiation within ecomorphs. In current parapatric boundaries we do not find evidence for local extinction and replacement upon postglacial expansion. Instead parapatric taxa recurrently show discordance between mitochondrial and nuclear markers, suggesting nuclear mediated gene flow across contact zones. Isolation with migration models support this hypothesis by showing significant gene flow across all five parapatric boundaries. Together, our results suggest that, while some genomic regions, such as the mitochondria, may follow morphologic species traits and retreat to isolated mountain tops, other genomic regions, such as nuclear markers, may flow across parapatric boundaries, sometimes leading to a complete genetic merger. We show that despite high ecologic and morphologic divergence over prolonged periods of time, hybridization allows for evolutionary outcomes alternative to extinction and replacement of taxa in response to climate change.

**Introduction**

In the face of climate change, species either adapt to new environments, move to adjust their range to suitable habitats, or face local or global extinction (Parmesan 2006). These dynamics are particularly strong along steep elevation gradients, where during warming scenarios low-altitude species are expected to expand while high-altitude taxa will retreat to isolated mountain tops (Moritz *et al.* 2008; Rovito *et al.* 2009; Morueta-Holme *et al.* 2015). Bioclimatic models predict that, at the ongoing pace of current climate change, many of these high-altitude species will suffer extirpation or extinction within the current century, particularly when dispersal is limited (Araújo *et al.* 2006; Carvalho *et al.*

2010). An underlying assumption of these models is that the species physiological limits are constant over the time period concerned, an assumption that holds when adaptation depends exclusively on slow evolutionary processes such as de novo mutation or fixation of ancestral alleles. Yet, species do not always evolve independently from each other. Hybridization can lead to rapid adaptation by generating novel phenotypes that might be favorable in extreme environments (e.g. transgressive segregation in sunflowers; Rieseberg & Willis 2007), can allow adaptive traits to move into a foreign genetic background (e.g. adaptive introgression in gray wolves; Anderson *et al.* 2009), or simply lead to a complete genetic and ecological merger (e.g. reversal of speciation, as reviewed in Seehausen *et al.* 2008). Importantly, the outcomes of hybridization are not uniform across the genome; while some genomic regions may remain distinct in the face of hybridization, other regions may flow between species, leading to differential introgression across the genome (Wu 2001; Harrison & Larson 2014). In order to establish predictions on how species and populations respond to warming climates, it is essential to consider the effect of hybridization during altitudinal range shifts, and assess whether hybridization leads to a predictable evolutionary outcome.

A direct approach to assess the effects of hybridization during range shifts is to analyze past population dynamics in species that survived past glacial-interglacial cycles. During postglacial warming, latitudinal shifts on species distributions have been associated with introgression of particular genes. For example, expansion of hare species from the Iberian refugium has led to the capture of mitochondrial DNA from a cold-adapted species by warm-adapted species in parts of the range characterized by colder climate (Melo-Ferreira *et al.* 2006). Range shifts in bear species during the last interglacial period have led to the introgression and fixation of the mitochondria from the cold-adapted polar bear into the warm-adapted brown bear (Cahill *et al.* 2013). Although in the former example the biogeographic history of the taxa suggests that the mitochondria remained in its ancestral range while the nuclear genome was substituted by that of the invading species (Melo-Ferreira *et al.* 2006; Excoffier *et al.* 2009), in the latter example it is more likely that the mitochondria has introgressed into a foreign genomic background along with few nuclear genes (Liu *et al.* 2014). Regardless of which genomic regions flow between species, the presence of cyto-nuclear discordances is an emerging pattern in studies of latitudinal range shifts between hybridizing species. It is possible, therefore, that altitudinal range shifts between hybridizing taxa occupying steep altitudinal gradients may lead to similar patterns of differential introgression and cytonuclear discordance.

The ecologically distinct subspecies of the fire salamander distributed along the Iberian Central System provide an excellent biological system to test this hypothesis. *Salamandra salamandra* (Linnaeus 1758) has diversified extensively since the Late Miocene/Pliocene in the Iberian Peninsula (Steinfartz *et al.* 2000; García-París *et al.* 2003; Vences *et al.* 2014), where up to 9 taxa with distinct morphologies (Raffaelli 2007), reproductive modes (Buckley *et al.* 2009), and

Accepted Article

habitat requirements (Kuzmin *et al.* 2009) are currently recognized (Velo-Antón & Buckley 2015). Despite a remarkably high genetic divergence (up to 6.3% in the *cyt b* mitochondrial gene), these taxa (Fig. 1) are currently recognized as subspecies since they often interbreed in zones of secondary contact (García-París *et al.* 2003). Yet, it remains unclear whether different subspecies correspond to distinct evolutionary units. Here, we refer to taxonomically recognized ‘subspecies’ as ‘ecomorphs’, to acknowledge that these units are solely based on eco-morphologic divergence. Along the Iberian Central System (hereafter ICS), two of these ecomorphs interact along a steep altitudinal gradient. The low- and mid-elevation *S. s. bejarae* (Wolterstorff 1934) inhabits forested areas in Central Spain; stream larvae develop into large adults with yellow-blotched coloration and red dots at the parotid glands and head. The high-elevation *S. s. almanzoris* (Müller & Hellmich 1935) inhabits sub-alpine meadows and glacial lakes; larvae develop into smaller adults with fewer yellow blotches and no red spots (Bosch & López-Bueis 1994; Montori & Herrero 2004). Although it is unclear how developmental plasticity contributes to eco-morphologic divergence, these subspecies are known to belong to different evolutionary lineages (1.4-2 % divergent in *cyt b*; Martínez-Solano *et al.* 2005a) and to have shifted their ranges multiple times during glacial-interglacial cycles. Until the end of the last glacial period, major mountains along the ICS were covered by thick glacial ice (Hughes & Woodward 2006; 2008; Domínguez-Villar *et al.* 2013; Pedraza *et al.* 2013). During the current postglacial warming, these subspecies have experienced an upward shift of their distribution, occupying areas that were previously glaciated. The high-elevation *almanzoris* is currently isolated in three mountain tops, surrounded by the low-elevation *bejarae* (Martínez-Solano *et al.* 2005a). The contact zone between low- and high-elevation taxa has likely moved tracking climate change as known in other ecologically divergent taxa (Taylor *et al.* 2015), but historical ranges of these ecomorphs are unknown. Concerning ongoing and future warming scenarios, bioclimatic models for several co-distributed species of amphibians and reptiles predict extirpation of local populations and even extinction, particularly along the ICS (Carvalho *et al.* 2010).

Here, we take advantage of three replicated contact zones between high- and low-elevation ecomorphs of *Salamandra salamandra* along the ICS to understand how hybridization can affect the extinction/colonization dynamics during altitudinal shifts. First, we use phylogeographic methods on mitochondrial data from several subspecies to test for the presence of cryptic lineages within the two focal ecomorphs. Second, we use population genetics methods on combined nuclear and mitochondrial data to assess the demographic history of these lineages and test for possible cyto-nuclear discordances. Third, we use coalescent methods to assess the relative roles of incomplete lineage sorting and gene flow in generating cyto-nuclear discordance. By comparing results along three independent altitudinal transects in different parts of the ICS, we expect to derive general conclusions about the role of gene flow (or lack thereof) at different genomic regions in the persistence of locally adapted ecomorphs during altitudinal shifts.

## Methods

### *Specimen sampling*

To assess evolutionary consequences of altitudinal range shifts during the current postglacial period, we sampled three naturally replicated contacts between the low- and high-elevation ecomorphs *bejarae* and *almanzoris*. These ecomorphs contact in three isolated mountaintops: Sierras of Gredos, San Vicente, and Guadarrama; respectively contact zones 1, 2 and 3 (CZ1-CZ3, Fig. 1, Suppl. Fig. 1). Geologically, Sierra of San Vicente is considered a southern isolate of the main Gredos massif. We collected 6 to 7 individuals per population, as far as possible from contact zones delineated in previous studies (Martínez-Solano *et al.* 2005a) to avoid biases caused by recent generations of hybrids that do not necessarily contribute to gene flow. High-elevation *almanzoris* from CZ1 and CZ3 currently live in areas that were glaciated during the last glacial maximum (Fig. 1 in Domínguez-Villar *et al.* 2013). Although *bejarae* is continuously distributed at low-elevations and thus are expected to present genetic connectivity along its range, pairs of high- and low-elevation ecomorphs from each contact zone are closest to each other, relative to populations from different contact zones.

To understand the phylogeographic history of these lineages we also sampled their respective type localities as taxonomic reference, five high-elevation isolates, as well as allopatric populations that are unlikely to have experienced recent gene flow (e.g. Toledo; Fig. 1). In addition, we sampled three populations of the parapatric ecomorph *S. s. gallaica*, and the more distantly related and geographically isolated *S. s. longirostris* (two samples). In total, our sampling included 79 individuals, all sequenced for one or two mitochondrial markers and 57 also sequenced for five nuclear markers (Suppl. Table 1).

### *Genetic sampling*

To estimate patterns of divergence and gene flow between ecomorphs we used two mitochondrial markers and developed five new independent nuclear markers. For the mitochondrial markers, we amplified and sequenced a 763 bp fragment of the D-loop region using the primers PRO and PHE (Suppl. Table 2), as well as a 1116 bp fragment of the *cyt b* gene using the primers Glu14100L and Pro15500H (Zhang *et al.* 2008) and following the published protocol. For the nuclear markers, we designed new specific primers targeting the genes beta fibrinogen intron 7 ( $\beta$ fibint7, 697 bp), chemokine receptor 4 (CXCR4, 645 bp), sodium/calcium exchanger 1 (NCX1, 652 bp), recombination activating protein 1 (RAG1, 386 bp), and solute carrier family 8 (sodium/calcium exchanger) member 3 (SLC8A3, 576 bp) (Suppl. Table 2). The primers were designed to target

conserved areas of the gene, based on alignments of published sequences of *Salamandra salamandra* and phylogenetically related taxa (Sequeira *et al.* 2006; Roelants *et al.* 2007). Amplifications were carried out as follows, adjusting the annealing temperature (Ta) for each locus: 94°C for 5 min (initial denaturation); 35 cycles at 94°C for 30 sec, Ta°C for 30 sec, and 72°C for 45 sec; 72°C for 5 min (final extension). Detailed information on primer sequences, fragment length and adjusted Ta is reported in Suppl. Table 2. PCR products were purified with ExoSAP-IT (USB Corporation, Cleveland, OH), and sequenced in both directions on an ABI 3730 capillary sequencer. DNA sequences were edited and aligned using Geneious Pro v5.0 (Drummond *et al.* 2009). For each locus, sequence ends were trimmed to include the largest fragment sequenced consistently. To resolve the haplotype phases of nuclear DNA sequences we employed the Bayesian algorithm implemented in PHASE software (Stephens *et al.* 2001) and used the haplotypes resolved with highest posterior probability (pp) in subsequent analyses.

### ***Phylogeographic history of high- and low-elevation ecomorphs***

To test if different ecomorphs (subspecies) evolved from a single or multiple historical lineages we performed a phylogeographic analysis using the two mitochondrial genes (cyt *b* and D-loop). For this analysis, we included all populations mentioned above (Fig. 1; Suppl. Table 1). This study includes a total of 17 populations of the ecomorphs distributed along the ICS (*gallaica*, *bejarae* and *almanzoris*) with a total of 64 individuals, plus 2 *longirostris* individuals as outgroup.

A mitochondrial gene tree was inferred with software BEAST v1.8.2 (Drummond *et al.* 2012). Optimal partitioning strategies for each marker and associated models of nucleotide substitution were simultaneously selected with the software PartitionFinder v1.1.1 (Lanfear *et al.* 2012). Accordingly, three partitions were specified, corresponding to first codon positions in cyt *b* plus D-loop (HKY+G), second codon positions in cyt *b* (HKY+I), and third codon positions in cyt *b* (TrN). Analyses in BEAST were run specifying a Bayesian Skyline (linear, 10 groups) as the coalescent prior and assuming a strict molecular clock. We assessed convergence by inspection of parameter estimates in Tracer v1.6 (Rambaut *et al.* 2014). Subsequently, after removing 10% of the resulting trees as burn-in, the resulting trees were summarized in a Maximum Clade Credibility (MCC) consensus tree with TreeAnnotator v1.8.2 (Drummond & Rambaut 2007).

### ***Demographic history and cyto-nuclear discordances***

To test how the postglacial range shifts shaped the genetic structure of the evolutionary lineages of the ICS, we used a combination of one mitochondrial (D-loop) and the five independent nuclear markers. We pooled individuals into the five phylogeographic groups detected above (see Results): *almanzoris*,

*bejarae*-S (South of ICS), *bejarae*-W (West on the ICS), *bejarae*-E (East on the ICS), and *gallaica*. This study included a total of 13 populations of the focal ecomorphs with a total of 55 individuals, plus 2 other individuals from the outgroup.

We calculated the following diversity statistics: number of haplotypes (H), number of segregating sites (S), haplotype diversity (Hd), nucleotide diversity  $\pi$ , and population mutation parameter  $\theta$ . To test if these groups were affected by recent demographic expansions, we tested for departures from mutation-drift equilibrium by computing Tajima's D (Tajima 1989); either positive selection or demographic expansions would lead to negative values of this statistic. Furthermore, we tested each population for signal of recent population expansion by calculating Fu's  $F_S$  (Fu 1997) and  $R_2$  (Ramos-Onsins & Rozas 2002). Population growth is expected to generate an excess of rare alleles, which would lead to negative values of  $F_S$  and low values of  $R_2$ . We tested significance of the three statistics using 100,000 coalescent simulations. All these analyses were conducted in DNASP v5.10 (Librado & Rozas 2009).

We built Neighbor Joining (NJ) trees for each nuclear marker, using the Geneious Tree Builder, and with the Jukes-Cantor genetic distance model. Because incomplete lineage sorting (ILS) can result in shared alleles between the ingroup and outgroup, NJ trees were rooted to their midpoint.

Discordance among gene trees can arise due to ILS, particularly in closely related taxa with high effective population sizes such as these salamanders. We applied a coalescent framework that uses multiple loci to simultaneously estimate independent "gene-trees" at each locus and the "species-tree" they are embedded into. Because such methods assume no intragenic recombination, we ran the Phi-Pack software (Bruen *et al.* 2006) on phased data to test this assumption (window size 100 bp,  $\alpha=0.000001$ ). To assess possible cyto-nuclear discordances, we used the Bayesian multispecies coalescent approach (\*BEAST; Heled & Drummond 2010), implemented in BEAST v1.8.2, to estimate Maximum Clade Credibility species trees based on the analysis of 1) the five nuclear loci; and 2) the mitochondrial markers. We ran the two analyses for 200,000,000 generations, specifying for each nuclear gene an HKY+I substitution model and the partition scheme and substitution models previously defined in PartitionFinder for the mtDNA. We used a strict molecular clock with a gamma distribution (shape= 0.01; scale= 100; offset= 0) as prior for all the clock.rate parameters. We also used a Birth & Death model as a prior for the population-trees. The MCC population-trees were then annotated and generated in TreeAnnotator v1.8.2. To get a temporal estimate of divergence times among phylogeographic groups, we added a temporal prior to the root of the tree in the mtDNA analysis, the split between *S. s. longirostris* and the other salamander lineages, which was dated in previous studies between 2 and 4 mya (prior used: Lognormal distribution, initial value= 3; mean= 3, stdev= 0.2) (García-París *et al.* 1998; Steinfartz *et al.* 2000). The population-tree analyses were run in the Cipres Science Gateway (Miller *et al.* 2010).

### *Isolation with migration models*

To test if the postglacial range shifts along the ICS resulted in genetic exchange between evolutionary lineages we employed the isolation with migration (IM) model implemented in IMA2 (Hey 2010). In a two-populations model, at certain time ( $t$ ) an ancestor population with population size  $\theta_A$  splits into two extant populations that may differ in population size ( $\theta_1$  and  $\theta_2$ ), and where migration in both directions ( $m_1$  and  $m_2$ ) may occur after divergence. Comparing this 6-parameter model to a simpler 4-parameter model without migration can be used to test for ‘migration’ versus ‘isolation’ between descendent taxa (Nielsen & Wakeley 2001). To achieve this, the model involves several simplifying assumptions (Pinho & Hey 2010): no recombination within each locus, free recombination among loci, mutations follow a simple model, no population structure within each taxa, no genetic contribution from unsampled taxa, and selective neutrality.

Although in principle this method is applicable to multiple taxa comparison, the high number of co-estimated parameters requires 1) large amounts of data beyond those presented in this study and 2) well-resolved phylogenetic relationships among all taxa. Therefore, in the absence of these requirements, to test for significant gene flow we employed a two-populations model to all five pairwise comparisons of the evolutionary lineages that are currently parapatric. The ecomorphs *almanzoris* and *bejarae* currently establish secondary contact in three contact zones addressed in three independent analyses: 1) *almanzoris* vs *bejarae*-W in contact zone 1; 2) *almanzoris* vs *bejarae*-E in contact zone 2; and 3) *almanzoris* vs *bejarae*-E in contact zone 3. We tested for gene flow between the ecomorphs *gallaica* and *bejarae* in a fourth analysis (*gallaica* vs *bejarae*-W) by pooling individuals from the three populations of *gallaica* (localities Peredo, Buçaco and Monteclaro) and using the *bejarae*-W population of Monforte (contact zone 4, CZ4). Finally, to address potential gene flow between the two cryptic lineages of *bejarae* discovered here (see Results) we performed a fifth analysis (*bejarae*-W vs *bejarae*-E, CZ5). All the analyses above were performed with 5 to 7 individuals per population (see Suppl. Table 1 for details), sequenced for the five nuclear loci plus the mitochondrial marker (D-loop), excluding alignment gaps and trimming sequences (664, 601, 607, 336, 513, and 761bp for  $\beta$ fibint7, CXCR4, NCX1, RAG1, SLC8A3, and D-loop, respectively).

After several experimental runs to assess appropriate parameter settings and ensure proper mixing, we ran 10 million steps along the Markov Chain, following 1 million steps of burn-in. This analysis was repeated four times using different priors and random seeds to assure convergence of parameter estimates. We estimated joint posterior density curves, highest value estimated (HiPt), and 95% confidence intervals (HPD95Lo and HPD95Hi) for all demographic parameters based on 300,000 saved genealogies. To investigate whether estimates of population migration rates ( $2Nm$ ) were significantly different from zero, i.e. indicating significant gene flow among populations, we performed a likelihood ratio test (Nielsen & Wakeley 2001).



## Results

### *Phylogeographic history of high- and low-elevation ecomorphs*

BEAST analyses recovered a well-resolved tree (Fig. 2), with two major haplotype clades with the highest statistical support (Bayesian Posterior Probabilities, BPPs: 1.0) corresponding to *S. s. almanzoris* (encompassing high elevation populations across the three replicate contact zones), and a *S. s. bejarae* + *S. s. gallaica* clade. The *S. s. almanzoris* clade was further subdivided into two well-supported haplotype clades, corresponding to the two major mountain systems where this subspecies occurs: Guadarrama (CZ3; BPP: 0.99), and Gredos (including the type locality, CZ1 in the main Gredos massif, and CZ2 in Sierra de San Vicente; BPP: 0.98). In the *S. s. bejarae* + *S. s. gallaica* clade, three *bejarae* samples from the westernmost contact zone (CZ1) grouped with the five *gallaica* samples from central and northern Portugal (BPP: 1.0), whereas the remaining *bejarae* samples formed another haplotype clade with moderate support (BPP: 0.93). This *bejarae* haplotype clade included three well-supported subclades (BPPs: 1.0 in all cases). The first included samples from the allopatric population of Toledo (hereafter *bejarae*-South). This haplotype clade was sister to the *bejarae* samples from the Iberian Central System (BPP: 1.0). These *bejarae* samples are further substructured in two well-supported clades (BPP: 1.0): a western clade including samples from the westernmost population in Monforte and CZ1 (hereafter *bejarae*-West) and a clade including all *bejarae* individuals from CZ2 and CZ3 (hereafter *bejarae*-East).

For downstream analyses, instead of the three ecomorphs recognized taxonomically, we considered these five well-supported (BPP > 0.99) phylogenetic groups described here; i.e. *gallaica*, *bejarae*-East, *bejarae*-West, *bejarae*-South and *almanzoris*.

### *Demographic history and cyto-nuclear discordances*

Genetic diversity varied both spatially along the Iberian Central System and between nuclear and mitochondrial markers (Table 1). Nuclear markers show a consistent decrease of polymorphisms eastwards along the ICS, across all polymorphism indexes measured (H, S, Hd,  $\pi$ , and  $\theta$ ). The *gallaica* group has the highest genetic diversity across all polymorphism indexes, followed by *bejarae*-West. The easternmost phylogeographic groups *bejarae*-East and *almanzoris* have the lowest genetic diversity (Table 1), along with the geographic isolate *bejarae*-South. The mitochondrial marker D-loop shows a similar eastward decrease of genetic diversity, with *gallaica* being the most diverse population. Yet, *bejarae*-West shows remarkably low levels of polymorphism, similarly to *bejarae*-East. The high-elevation *almanzoris* and the isolate *bejarae*-South show moderate levels of diversity in mtDNA, showing a large number of haplotypes (relatively high H and S) that are very

similar to each other (relatively low  $\pi$  and  $\theta$ ). Likely due to insufficient levels of genetic diversity in each gene, neutrality and population growth indexes were not significant for any gene, except for the SLC gene in *bejarae*-West, where a significant signal of recent population expansion was detected in a single test (Fu's  $F_s$ :  $p < 0.001$ ).

The low genetic divergence observed in nuclear genes resulted in short branches across all individual gene trees (Suppl. Fig. 2). Yet, haplotypes from some phylogeographic groups tended to cluster together, such as the most divergent taxa *longirostris* and the allopatric *bejarae*-South, while groups along the ICS were often intermixed, with for instance *almanzoris* being virtually indistinguishable from *bejarae*-East (Suppl. Fig. 2).

The population-trees based on \*BEAST analyses of mitochondrial and nuclear markers are shown in Fig. 3 and Suppl. Fig. 3. Both analyses mixed and converged properly, with ESS well over 200 for all the parameters estimated, as visualized in Tracer 1.6. Genealogical relationships among groups in the two analyses are, however, discordant. The *gallaica* and all the *bejarae* lineages grouped together in the mitochondrial population-tree, with the two parapatric *bejarae* groups along the ICS (*bejarae*-East and *bejarae*-West) forming a robust and well-supported sub-clade. However, this pattern is not recovered in the nuclear population-tree, which consistently grouped the *almanzoris* phylogeographic group with the easternmost *bejarae*-East, and *gallaica* with *bejarae*-West. Cyto-nuclear discordances, therefore, are especially conspicuous in the parapatric lineages along the ICS but do not seem to involve the allopatric groups.

We calibrated the root for the mitochondrial-population tree with a temporal estimate (2-4 my) to estimate divergence times between the phylogeographic groups. The Most Recent Common Ancestor (MRCA) of the *almanzoris* and *bejarae*+ *gallaica* groups occurred around 1 mya (0.6-1.75 mya). The latter shared a MRCA around 0.5 mya (0.25-1 mya), while the MRCA of the parapatric lineages *bejarae*-West and *bejarae*-East was dated around 0.25 mya (0.05-0.5 mya).

### ***Isolation with migration models***

Our estimation of demographic parameters using coalescent models reached very high values of ESS and consistent estimates across all runs, indicating a good sampling of the parameter space and convergence in the same result. For all the five pairwise comparisons tested, migration as reflected by  $2Nm$  is significantly different from zero in at least one direction, suggesting gene flow across all parapatric boundaries (Table 2). Bidirectional gene flow is significant in almost all contact zones. The exception was the CZ3, where despite a moderate amount of gene flow from *bejarae*-East into

*almanzoris* (highest probability density, HiPt, of  $2Nm=0.1$ ), this was not significantly better than a simpler model with no migration (likelihood ratio test = 0.964).

To test for potential asymmetry of gene flow we estimated posterior density curves for  $2Nm$  in each direction. The HiPt values of  $2Nm$  converged to the same value across replicated runs, suggesting reliable point estimates of gene flow in all contact zones. Yet, estimations of  $2Nm$  across 300,000 saved genealogies had a broad distribution, resulting in wide confidence intervals around the point estimates of gene flow in each direction (Table 2). As expected, the power to detect asymmetry of gene flow is higher (i.e. with narrower 95% highest posterior density intervals) between populations with largest genetic divergence such as in CZ1 and CZ5, relative to genetically similar populations from CZ2 and CZ3 (Suppl. Fig. 4). Thus, the prevalent direction of gene flow is undistinguishable in CZ2 and CZ3 (both between *almanzoris* and *bejarae*-E), seems to be asymmetric but not significantly in CZ1 and CZ4 (respectively between *almanzoris* and *bejarae*-W and between *gallaica* and *bejarae*-W), and it is significantly asymmetric towards East in CZ5 (between *bejarae*-W and *bejarae*-E).

## Discussion

### *Cryptic historical lineages within ecomorphs*

Previous genetic studies revealed remarkably high levels of genetic divergence among different ecomorphs of *Salamandra salamandra* (Vences *et al.* 2014), suggesting that they have been evolving independently through the Pleistocene climate cycles in the Iberian Peninsula (Steinfartz *et al.* 2000; García-París *et al.* 2003). The same holds true for the ecomorphs currently distributed along the Iberian Central System; the high- and low-elevation ecomorphs *bejarae* and *almanzoris* are 1.4-2 % divergent in *cyt b* (Martínez-Solano *et al.* 2005a). Yet, it remains unclear whether ecomorphs directly correspond to evolutionarily independent lineages, or if there is cryptic differentiation between lineages with similar ecology and morphology.

Our phylogeographic study based on a comprehensive sampling of the taxa interacting along the ICS showed that the oldest split in this region is between the high-elevation *almanzoris* and the remaining low-elevation taxa *gallaica* and *bejarae* (Fig. 2). *Salamandra s. almanzoris* could thus represent a relatively older branch in the diversification of *Salamandra* in Iberia, whereas the ecomorphs “*bejarae*” and “*gallaica*” (*sensu lato*) comprise a more recent radiation. Using the temporal estimate for the split between *S. s. longirostris* and the rest of Iberian subspecies calculated in two independent studies (García-París *et al.* 1998; Steinfartz *et al.* 2000) to calibrate a molecular clock, the MRCA of *almanzoris* and *bejarae*+ *gallaica* groups occurred around 1 mya (0.6-1.75 mya).

During the glacial period, it is likely that a widespread *almanzoris* ancestor occupied most of the ICS, probably even more extensively than its present range. The *bejarae* ancestor was likely in warmer habitats, west or south of the ICS, where the more ancestral clades *gallaica* and *bejarae*-South are found today. Most of the range currently occupied by *almanzoris* has been covered by thick glacial ice until some 30,000-25,000 years ago (Hughes & Woodward 2006; 2008; Domínguez-Villar *et al.* 2013). Climate warming, therefore, drove *almanzoris* populations upwards to high-elevation tracking colder habitats, leading to population fragmentation and to genetic substructure in the two major sublineages observed today (Fig. 2). Conversely, *bejarae* would have expanded eastwards along the Central System, pushing the contact zone to the isolated mountain tops sampled here. The age and initial location of the contact zone between *almanzoris* and *bejarae* are uncertain but, given the high genetic divergence between these clades in mitochondrial DNA (Fig. 2; ~1 mya), it is likely that it has moved multiple times during glacial and interglacial periods.

Most importantly, our results show definitive evidence for cryptic differentiation within the taxonomically recognized ecomorphs (Fig. 2). The high-elevation ecomorph *almanzoris* is substructured into geographically isolated groups, with Gredos (CZ1) and San Vicente (CZ2) isolates forming a clade and Guadarrama (CZ3) forming another. This split is recent and thus might be related to population fragmentation during the current postglacial warming. In contrast, we find multiple lines of evidence supporting that the phylogenetic split within the low-elevation ecomorph *bejarae* is necessarily older than the last glacial maximum. First, we find that our samples of *bejarae* are substructured into three geographically distinct clades with high support (BPPs: 1.0 in all cases): i) a southern clade that has not been in recent contact with the populations from the ICS (*bejarae*-South), ii) a western clade containing the type locality (Lagunilla, see Suppl. Table 1) and contacting with the ecomorphs *gallaica* in the West, and *almanzoris* at CZ1 (*bejarae*-West), and iii) an eastern clade contacting with *almanzoris* at CZ2 and 3 (*bejarae*-East). Second, we find that *bejarae*-South has diverged from the remaining clades around 0.5 mya, suggesting that these lineages have been evolving in allopatry during the most recent glacial cycles. Although the independent nuclear gene trees need to be interpreted with caution due to the lower mutation rate, they generally supported this finding by showing that *bejarae*-South haplotypes always grouped together (Suppl. Fig. 2). Lastly, we do not find strong support for the monophyly of the *bejarae* group as a whole. The estimated gene trees are particularly discordant at this node of our population-tree (Suppl. Fig. 3), suggesting that the taxonomically recognized subspecies *bejarae* might be paraphyletic. Further data from the entire distribution of the *bejarae* and *gallaica* ecomorphs are needed to test this hypothesis. Nevertheless, the current data is sufficient to demonstrate that the ecomorph *bejarae* is composed by at least three old evolutionary lineages

Accepted Article

Together, our results show that the ecomorphs interacting along the ICS do not directly correspond to evolutionarily independent units. Periods of geographic isolation, often lasting for hundreds of thousands of years have led to vicariant divergence of evolutionary lineages that preserve similar ecology and color pattern. Similar patterns of genetic divergence in the absence of ecomorphological divergence have been also reported in radiations in other cryptic, terrestrial salamander species, for instance in the Appalachians mountains (Adams *et al.* 2009), suggesting that species with low vagility might be prone to cryptic differentiation. When examining the effect of postglacial range shifts in the ICS, therefore, one needs to consider at least five well-supported phylogeographic groups: the high-elevation *almanzoris*, and the low-elevation *gallaica*, *bejarae-S*, *bejarae-W*, and *bejarae-E*.

#### ***Cyto-nuclear discordance between parapatric taxa***

Genetic introgression between different species of plants and animals has been increasingly reported in the literature. For instance, organelle genes such as those from the mitochondria are often introgressed across species boundaries, leading to the replacement of genetic variation in part (Melo-Ferreira *et al.* 2006) or in the entire species range (Cahill *et al.* 2013). Such cyto-nuclear discordances are often associated with species colonization of new habitat, in particular during latitudinal expansions that followed glacial periods. Species occupying different elevation bands in mountain ranges have been experiencing similar range shifts during the current postglacial period, but in altitude rather than in latitude. We evaluated whether postglacial altitudinal range shifts experienced by phylogeographic groups of *Salamandra salamandra* along the Iberian Central System have also resulted in cyto-nuclear discordances.

The observed patterns of genetic variability are congruent with the predicted postglacial expansion. Genetic variability is higher in the western phylogeographic groups *gallaica* and *bejarae-West*, relative to the eastern groups *bejarae-East* and *almanzoris* (Table 1). The western part of the ICS is a well-known sub-refugium within the major Iberian refugium where lineages of several plants and animals have survived during the Pleistocene glacial periods (Gómez & Lunt 2007). On the other hand, the eastern part of the ICS was covered by permanent glaciers until 30,000-25,000 years ago (Hughes & Woodward 2006; 2008; Domínguez-Villar *et al.* 2013) and has been recently colonized by several vertebrate species that show a similar eastwards decrease of genetic variability (Martínez-Solano *et al.* 2005b; Godinho *et al.* 2008). As discussed above, it is likely that our focal phylogeographic groups of *Salamandra salamandra* have also survived glacial periods in the western part of the ICS, and expanded eastwards following habitat amelioration.

Our species trees analyses reveal two major discordances and one concordance between the mitochondrial and the nuclear data (Fig. 3). While in the mitochondrial tree, the high-elevation *almanzoris* is the most divergent phylogeographic group, well separated from all low-elevation groups, in the nuclear tree *almanzoris* and *bejarae*-East are virtually undistinguishable. In addition, while in the mitochondrial tree *gallaica* is the most divergent low-elevation lineage, it becomes sister group to *bejarae*-West in the nuclear species tree. Notably, the geographic isolate *bejarae*-South remains genetically distinct in the nuclear species tree. Because these methods explicitly integrate the coalescent in species tree estimation (Heled & Drummond 2010), it is unlikely that the observed discordances can be explained by incomplete lineage sorting. Yet, the geographic distribution of the groups involved in cyto-nuclear discordances and concordance (Fig. 1) might shed some light on the processes underlying such discordance. The two well-supported clades causing cyto-nuclear discordance (*almanzoris* + *bejarae*-E and *bejarae*-W + *gallaica*) involve pairs of taxa that are currently parapatric. Notably, the *bejarae*-South lineage that has not been in recent contact with the ICS remains highly divergent. These observations suggest that genetic introgression across parapatric boundaries might strongly contribute to cyto-nuclear discordances. Similar spatial discordance between mitochondrial and nuclear markers has been described in other parapatric subspecies of *Salamandra salamandra* (*bernardezi* and *fastuosa*; García-París *et al.* 2003). Our results support the previous hypothesis that hybridization following postglacial expansion might underlie this common pattern.

Although the phylogeographic groups of *Salamandra salamandra* have been diverging for thousands or millions of years, they have not evolved complete reproductive isolation. During multiple range shifts driven by glacial cycles, these lineages have interacted genetically, exchanging a significant percentage of their genetic material to the point of shifting the estimated genetic relationships based on different parts of the genome. Similar to what was previously observed in large continental scales, our results show that postglacial range shifts eastwards and upwards along the ICS can also result in cyto-nuclear discordances.

#### ***Nuclear-mediated introgression during postglacial expansions***

Theoretical studies based on simulations of range expansions have shown that if interbreeding between taxa is not severely prevented, i.e. if evolution of invading and local taxa are not independent, two outcomes are expected (Currat *et al.* 2008). First, massive introgression of neutral genes will occur, even if species hybridize at a low rate and selection is absent. Second, introgression will be asymmetric and conditioned by the demographically expanding species, resulting in a massive introgression of local genes into the genome of the invading taxon (Currat *et al.* 2008; Excoffier *et al.*

2009). Here, we use the replicated nature of the contact zones in *Salamandra salamandra* along the ICS to test these two theoretical expectations.

As discussed above, our results show conflicting patterns between the multiple data sets used to study the evolution of *Salamandra salamandra* in the ICS. While morphologic analysis based on color pattern distinguishes three ecomorphs categorized as subspecies, mitochondrial data show clear evidence for five evolutionary units. Although most of these lineages are well supported by nuclear data, parapatric taxa are phylogenetically more similar to each other (*gallaica* and *bejarae*-West) to the point of becoming genetically indistinguishable (*almanzoris* and *bejarae*-East). It is important to note that the lack of genetic differentiation can be explained either by complete fusion of the two ancestral genomes, extinction of the *bejarae*-East nuclear genome, or extinction of the *almanzoris* nuclear genome; further genomic and spatial sampling is needed to differentiate between these alternative scenarios. Isolation with migration models helped to integrate across multiple data sets and explicitly test if these discordant patterns are indeed associated with gene flow across parapatric boundaries or if they can be explained by different coalescence times between mitochondrial and nuclear markers (Fig. 4). Despite the near fixation of mtDNA haplotypes (Table 1) and reduced variability in individual nuclear markers (Suppl. Fig. 2), our estimates of  $2Nm$  allowed rejection of the null hypothesis of no gene flow in all five contact zones (Table 2), and across replicated analyses. Moreover, significant gene flow occurs in both directions in 4 contact zones. Only in CZ3, despite a modest point estimate of  $2Nm$ , gene flow is significant in one direction, but not in the reciprocal, possibly due to reduced power when estimating  $2Nm$  between populations with low genetic divergence (Suppl. Fig. 4). These results confirm the theoretical expectation that “massive introgression follows range expansions” (Currat *et al.* 2008), and strongly support our hypothesis that introgression has resulted in the cyto-nuclear discordances observed between parapatric taxa.

Regarding the direction of introgression, simulation studies predict that invasion scenarios with interbreeding result in a net asymmetry in effective introgression from the local taxon to the invading one (Currat *et al.* 2008; Excoffier *et al.* 2009). Thus, in the context of the postglacial expansion of *S. salamandra* ecomorphs along the ICS, we would predict asymmetric gene flow in the direction opposite of demographic expansion, i.e. westwards between low-elevation lineages (CZ4 and CZ5) and downwards between high- and low-elevation lineages (CZ1, CZ2 and CZ3). From the five contact zones analyzed, the posterior density curves for  $2Nm$  were largely overlapping (Suppl. Fig. 4), generally encompassing the highest estimated value of gene flow in the reverse direction (Table 2). In comparisons between most divergent lineages (CZ1 and CZ5) where confidence intervals were narrower, only CZ5 shows significant asymmetric gene flow, but in the direction of demographic expansion. The nuclear markers developed here were based on genes that are conserved across different genera of salamanders (Sequeira *et al.* 2006; Roelants *et al.* 2007). Although these

markers offered enough power to conclusively test for the presence of gene flow in the five contacts (LLR test, Table 2), their power is not always enough to test asymmetry of introgression.

Statistical limitations of the isolation with migration model may also affect the results presented here. This model involves several simplifying assumptions (see Methods), some of which are likely violated in virtually every natural system. A simulation study showed that parameter estimates are generally quite robust to small to moderate violations of those simplifying assumptions, comparable with what is often encountered in real-world scenarios (Strasburg & Rieseberg 2010). Most relevant to our use of the model in *Salamandra salamandra*, that study has shown that the presence of a third unsampled taxon might cause biases in estimates of  $2Nm$  when using a simpler 2-population model. In our analyses, each pair of populations from a contact zone is geographically closer to each other than to populations from different pairwise comparisons. Yet, given that *bejarae* is more or less continuously distributed at low-elevation throughout the ICS, we cannot reject the possibility that our estimation of gene flow at one contact zone might be affected by introgression at different contacts. In addition, recent simulation studies have reported a high rate of false-positives in tests of gene flow using the isolation with migration model (Cruickshank & Hahn 2014), but these biases are most common when using a small number of loci and recent splitting times (Hey *et al.* 2015). Given that most taxa in our focal contact zones have likely diverged around 1 mya (Fig. 2), it is unlikely that our estimates of  $2Nm$  are false positives. With the advent of high-throughput sequencing, limitations on number of loci and degree of genetic variability will soon disappear for non-model organisms, allowing the use of more complex statistical models involving multiple populations, and testing for asymmetry of introgression across contact zones, taking further advantage of replicated contact zones in biological systems such as *Salamandra salamandra*.

Despite current methodological limitations, our integration of phylogeographic and population genetic analyses indicate that gene flow is common across contact zones of *S. salamandra*, that it is extensive, and that it is most prominent in nuclear markers relative to the mitochondria.

#### ***Altitudinal range shifts lead to extensive cyto-nuclear discordance***

In altitudinal gradients, warming conditions will favor expansion of low-elevation taxa into habitats occupied by high-elevation species. Models assuming independent evolution of species predict that, during these altitudinal range shifts, high-elevation species will experience population extirpation or extinction, particularly when dispersal is limited (Araújo *et al.* 2006; Carvalho *et al.* 2010). Yet, empirical work in post-glacial range expansions along latitudinal gradients have show that, when taxa can interbreed, organelle genes typically experience reduced gene flow compared to nuclear genes



and thus tend to be captured by an invading nuclear genome leading to cyto-nuclear discordances (reviewed in Petit & Excoffier 2008).

We have shown that low- and high-elevation ecomorphs of *Salamandra salamandra* that have been diverging for around 1 mya (Fig. 2) still remain capable of interacting genetically at contact zones. Although local rates of hybridization are unknown, we found that this rate is high enough to result in extensive nuclear-mediated gene flow during postglacial altitudinal range shifts (Table 2), associated with cyto-nuclear discordances (Fig. 3). Differential introgression between mitochondrial and nuclear markers can potentially lead to the complete fusion of nuclear genomes while historical mitochondrial lineages are still preserved (as observed in CZ2 and CZ3 with *almanzoris* and *bejarae*-East; Fig. 4). Further work is needed to assess where ecologically relevant traits, and underlying genetic loci, fall within this spectrum of differential introgression.

Examples of cyto-nuclear discordances caused by differential gene flow between mitochondrial and nuclear markers have been described between other ecomorphs of *Salamandra salamandra* (García-París *et al.* 2003) and are abundant in animal species (reviewed in Toews & Brelsford 2012). Yet, little is known about their underlying mechanism. Although the central role of the mitochondria in cell metabolism invites an adaptive interpretation of this repeated pattern, such as the adaptation of mitochondrial lineages to different temperature regimes (Welch *et al.* 2014), this scenario is difficult to distinguish from a purely neutral one. For example, a native mitochondrial lineage may be captured by an invading nuclear genome due to demographic effects during a range expansion, favored by the smaller effective population size of the mitochondrial genomes (Currat *et al.* 2008; Excoffier *et al.* 2009). Other potential causes for nuclear introgression would be sex-biased dispersal (Petit & Excoffier 2009), asymmetric incompatibilities in hybrids (Turelli & Moyle 2007) and sexual selection (While *et al.* 2015). Despite evidence of male-biased dispersal in the closely related *Salamandra atra* (Helfer *et al.* 2012), this has not been observed in *S. salamandra terrestris* (Schulte *et al.* 2007). Thus, it is unclear how these different mechanisms may contribute to cyto-nuclear discordances in ecomorphs of *Salamandra salamandra* along the ICS.

Regarding the consequences of altitudinal range shifts, our study of replicated contact zones between the low- and high-elevation ecomorphs of *Salamandra salamandra* demonstrates that hybridization can shift the extinction/recolonization dynamics expected for species occupying elevation gradients. Instead, differential introgression across genomic regions is the most likely evolutionary outcome (Harrison & Larson 2014). Although genomic integrity (*sensu* Mayr 1963) of high- and low-elevation taxa is quickly lost during elevation range shifts due to nuclear-mediated introgression, historical mitochondrial lineages might persist over longer periods of time.

## Acknowledgments:

This research was supported by the European Science Foundation (Frontiers of Speciation Research, Exchange grant 3318), and by the European Commission (Synthesys grant ES-TAF-1486□), granted to RJP. Partial funds were also provided by a Junta de Comunidades de Castilla la Mancha and FEDER (PPII10-0097-4200) grant to IMS, who is currently supported by funding from the Spanish Severo Ochoa program (SEV-2012-0262). During the preparation of this manuscript RJP was funded by the European Union's Horizon 2020 research and innovation program, under the Marie Skłodowska-Curie grant agreement No 658706. We thank MJ Fernández Benítez for technical support and A Álvarez, C Grande, J Gutiérrez-Rodríguez, P Hernández Sastre, EL Jockusch, N Martínez-Solano, JR Mayor, M Modrell, E Recuero, A Sánchez Vialas, F Smith, S Yubero, and DB Wake for help in sample collection. We thank G Velo-Antón for help with computational resources and P Tarroso for providing the baseline map in Suppl. Fig. 1. We also thank the editor S Spear and three anonymous reviewers for their thoughtful insights in an earlier version of this manuscript.

## References:

- Adams DC, Berns CM, Kozak KH, Wiens JJ (2009) Are rates of species diversification correlated with rates of morphological evolution? *Proceedings Biological Sciences / The Royal Society*, **276**, 2729–2738.
- Anderson TM, vonHoldt BM, Candille SI *et al.* (2009) Molecular and evolutionary history of melanism in North American gray wolves. *Science*, **323**, 1339–1343.
- Araújo MB, Thuiller W, Pearson RG (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, **33**, 1712–1728.
- Bosch J, López-Bueis I (1994) Comparative study of the dorsal pattern in *Salamandra salamandra bejarae* (Wolterstorff, 1934) and *S. s. almanzoris* (Müller & Hellmich, 1935). *Herpetological Journal*, **4**, 46–48.
- Bouckaert RR (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, **26**, 1372–1373.
- Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics*, **172**, 2665–2681.
- Buckley D, Alcobendas M, García-París M (2009) The evolution of viviparity in salamanders (Amphibia, Caudata): Organization, variation, and the hierarchical nature of the evolutionary process. In: *Evolución y Adaptación. 150 años después del Origen de las Especies* (eds Dopazo H, Arcadi A), pp. 145–154. SESBE, Obrapropia, S. L., Valencia.
- Cahill JA, Green RE, Fulton TL, *et al.* (2013). Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. *PLoS Genet.* **9**, e1003345.
- Carvalho SB, Brito JC, Crespo EJ, Possingham HP (2010) From climate change predictions to actions - conserving vulnerable animal groups in hotspots at a regional scale. *Global Change Biology*, **16**, 3257–3270.
- Curat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

- Domínguez-Villar D, Carrasco RM, Pedraza J *et al.* (2013) Early maximum extent of paleoglaciers from Mediterranean mountains during the last glaciation. *Scientific Reports*, **3**, 2034.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214–218.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Drummond AJ, Ashton B, Cheung M *et al.* (2009) Geneious v4.7. *geneious.com*.
- Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 481–501.
- Fu YX (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*, **1998**, 173–176.
- García-París M, Alcobendas M, Buckley D, Wake DB (2003) Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (*Salamandra*) inferred from discordance of genetic and morphological traits. *Evolution*, **57**, 129–143.
- Godinho R, Crespo E, 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. *Molecular Ecology*, **17**, 4670–4683.
- Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. *Phylogeography of southern European refugia*, 155–188.
- Harrison RG, Larson EL (2014) Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, **105**, 795–809.
- Helfer V, Broquet T, Fumagalli L (2012) Sex-specific estimates of dispersal show female philopatry and male dispersal in a promiscuous amphibian, the alpine salamander (*Salamandra atra*). *Molecular Ecology*, **21**, 4706–4720.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.
- Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.
- Hughes PD, Woodward JC (2006) Quaternary glacial history of the Mediterranean mountains. *Progress in Physical Geography*, **30**, 334–364.
- Hughes PD, Woodward JC (2008) Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, **23**, 575–588.
- Kuzmin S, Papenfuss T, Sparreboom M, Ugurtas IH, Anderson S, Beebee T, Denoël M, Andreone F, Anthony B, Schmidt B, Ogirowski A, Ogielska M, Bosch J, Tarkhnishvili D, Ishchenko V (2009) *Salamandra salamandra*. The IUCN Red List of Threatened Species 2009: e.T59467A11928351. Downloaded on 12 October 2015.
- Lanfear R, Calcott B, Ho SYW, Guindon S (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Linnaeus CA (1758) *Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. Editio Decima, reformata. Tomus 1.
- Liu S, Lorenzen ED, Fumagalli M *et al.* (2014) Population genomics reveal recent speciation and rapid evolutionary adaptation on polar bears. *Cell*, **157**, 785–794.

- Martínez-Solano I, Alcobendas M, Buckley D, García-París M (2005a) Molecular characterisation of the endangered *Salamandra salamandra almanzoris* (Caudata, Salamandridae). *Annales Zoologici Fennici*, **42**, 57–68.
- Martínez-Solano I, Rey I, García-París M (2005b) The impact of historical and recent factors on genetic variability in a mountain frog: the case of *Rana iberica* (Anura: Ranidae). *Animal Conservation*, **8**, 431–441.
- Mayr E (1963) *Animal species and evolution*. Belknap Press of Harvard University Press, Cambridge, MA.
- Melo-Ferreira J, Boursot P, Randi E *et al.* (2006) The rise and fall of the mountain hare (*Lepus timidus*) during Pleistocene glaciations: expansion and retreat with hybridization in the Iberian Peninsula. *Molecular Ecology*, **16**, 605–618.
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE)*, 1–8.
- Montori A, Herrero P (2004) Caudata. In: *Amphibia, Lissamphibia*. García-París M, Montori A, Herrero P. In: *Fauna Ibérica*, vol 24. Ramos MA *et al.* (Eds). Museo Nacional de Ciencias Naturales, CSIC, Spain.
- Moritz C, Patton JL, Conroy CJ *et al.* (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science*, **322**, 261–264.
- Morueta-Holme N, Engemann K, Sandoval-Acuña P *et al.* (2015) Strong upslope shifts in Chimborazo's vegetation over two centuries since Humboldt. *Proceedings of the National Academy of Sciences of the United States of America*, 201509938–5.
- Müller L, Hellmich W (1935) Mitteilungen über die Herpetofauna der Iberischen Halbinsel. I. Über *Salamandra salamandra almanzoris* n. ssp. und *Bufo bufo gredosicola* n. ssp., zwei neue Amphibienrassen aus der Sierra de Gredos. *Zoologischer Anzeiger*, **112** (3/4), 49–57.
- Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. *Genetics*, **158**, 885–896.
- Parmesan C (2006) Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.
- Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Ann. Rev. Ecol. Evol. Syst.*, **41**, 215–230.
- Pool J, Nielsen R (2007) Population size changes reshape genomic patterns of diversity. *Evolution*, **61**, 3001–3006.
- Raffaelli J (2007) *Les Urodèles du Monde*. Penclen Editions, 480pp.
- Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6. Available from <http://beast.bio.ed.ac.uk/Tracer>
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.
- Rieseberg LH, Willis JH (2007) Plant speciation. *Science*, **317**, 910–914.
- Roelants K, Gower DJ, Wilkinson M *et al.* (2007) Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 887–892.
- Rovito S, Parra-Olea G, Vásquez-Almazán C, Papenfuss T, Wake D (2009) Dramatic declines in neotropical salamander populations are an important part of the global amphibian crisis. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 3231.
- Schulte U, Küsters D, Steinfartz S (2007) A PIT tag based analysis of annual movement of adult fire salamanders (*Salamandra salamandra*) in a Middle European habitat. *Amphibia-Reptilia*, **28**, 531–536.
- Seehausen O, Takimoto G, Roy D, Jokela J (2008) Speciation reversal and biodiversity dynamics

- with hybridization in changing environments. *Molecular Ecology*, **17**, 30-44.
- Sequeira F, Ferrand N, Harris DJ (2006) Assessing the phylogenetic signal of the nuclear  $\beta$ -Fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia*, **27**, 409–418.
- Sessions SK (2008) Evolutionary cytogenetics in salamanders. *Chromosome Research*, **16**, 183–201.
- Steinfartz S, Veith M, Tautz D (2000) Mitochondrial sequence analysis of *Salamandra* taxa suggests old splits of major lineages and postglacial recolonizations of Central Europe from distinct source populations of *Salamandra salamandra*. *Molecular Ecology*, **9**, 397–410.
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.
- Strasburg JL, Rieseberg LH (2010) How robust are "Isolation with Migration" analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, **27**, 297-310.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.
- Turelli M, Moyle L (2007) Asymmetric postmatingisolation: Darwin's corollary to Haldane's rule. *Genetics*, **176**, 1059-1088.
- Velo-Antón G, Buckley D (2015) Salamandra común – *Salamandra salamandra*. In: *Enciclopedia Virtual de los Vertebrados Ibéricos*. Salvador A, Martínez-Solano I (Eds.). Museo Nacional de Ciencias Naturales, Madrid, Spain. <http://www.vertebradosibericos.org/>
- Vences M, Sánchez E, Hauswaldt JS *et al.* (2014) Nuclear and mitochondrial multilocus phylogeny and survey of alkaloid content in true salamanders of the genus *Salamandra* (Salamandridae). *Molecular Phylogenetics and Evolution*, **73**, 208–216.
- Welch AJ, Bedoya-Reina OC, Carretero-Paulet L *et al.* (2014) Polar bears exhibit genome-wide signatures of bioenergetic adaptation to life in the arctic environment. *Genome Biology and Evolution*, **6**, 433-450.
- While GM, Michaelides S, Haethcote RJP *et al.* (2015) Sexual selection drives asymmetric introgression in wall lizards. *Ecology Letters*, **18**, 1366-1375.
- Wolterstorff W (1934) Über *Salamandra maculosa* Laur. (= salamandra L.) *molleri* und ein neue Form, subsp *bejarae*. Blät. Aquar. Terrar. Kunde **43**, 41-43.
- Wu C (2001) The genic view of the process of speciation. *J. Ecol. Biol.*, **14**, 851-865.
- Zhang P, Papenfuss TJ, Wake MH, Qu L, Wake DB (2008) Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **49**, 586–597.

#### **Data Accessibility:**

Cleaned unphased sequences were deposited in the NCBI GenBank (Acc. KU249657-KU250056 and KU295291-KU295332, see Suppl. Table 1 for details). The dryad archive (doi:10.5061/dryad.j10r6) contains: unphased and phased sequences, gene alignments, tree files, input and output files for IMA2 and BEAST xml files. The precise GPS coordinates for collection sites will be provided for research purposes upon request.

## Tables

**Table 1.** Polymorphism indexes in phylogeographic groups of *Salamandra salamandra* along the Iberian Central System.

taxa	locus	N	H	S	Hd	$\pi$	$\theta$	
<i>almanzoris</i>	D-loop	39	8	9	0.719	0.00098	0.00114	
	$\beta$ -fibrinogen	40	6	6	0.673	0.00266	0.00213	
	CXCR4	40	3	3	0.347	0.00098	0.00117	
	NCX1	40	2	1	0.224	0.00037	0.00039	
	RAG1	40	2	1	0.05	0.00015	0.0007	
	SLC	40	1	0	0	0	0	
	nDNA mean			2.8	2.2	0.2588	0.000832	0.000878
	<i>bejarae E</i>	D-loop	5	1	0	0	0	0
$\beta$ -fibrinogen		26	3	4	0.218	0.00057	0.00158	
CXCR4		26	3	3	0.335	0.00071	0.00131	
NCX1		26	1	0	0	0	0	
RAG1		26	2	1	0.077	0.00023	0.00078	
SLC		26	1	0	0	0	0	
nDNA mean				2	1.6	0.126	0.000302	0.000734
<i>bejarae W</i>		D-loop	8	2	0.536	0.536	0.00029	0.00021
	$\beta$ -fibrinogen	16	2	1	0.525	0.00079	0.00045	
	CXCR4	16	2	1	0.458	0.00076	0.0005	
	NCX1	16	4	4	0.733	0.00253	0.00199	
	RAG1	16	3	6	0.242	0.00352	0.00538	
	SLC	14	5	4	0.505	0.00111	0.00245	
	nDNA mean			3.2	3.2	0.4926	0.001742	0.002154
	<i>bejarae S</i>	D-loop	4	4	5	1	0.00143	0.00146
$\beta$ -fibrinogen		12	1	0	0	0	0	
CXCR4		12	4	5	0.712	0.00287	0.00275	
NCX1		12	1	0	0	0	0	
RAG1		12	1	0	0	0	0	
SLC		12	1	0	0	0	0	
nDNA mean				1.6	1	0.1424	0.000574	0.00055
<i>gallaica</i>		D-loop	8	6	10	0.929	0.00214	0.00206
	$\beta$ -fibrinogen	8	1	0	0	0	0	
	CXCR4	16	6	6	0.783	0.00269	0.00301	
	NCX1	16	4	4	0.65	0.00206	0.00199	
	RAG1	16	4	7	0.675	0.00868	0.00628	
	SLC	16	4	4	0.742	0.00387	0.00235	
	nDNA mean			3.8	4.2	0.57	0.00346	0.002726

N- sample size, H- number of haplotypes, S- number of segregating sites, Hd- haplotype diversity,  $\pi$ - nucleotide diversity,  $\theta$ - population mutation rate

**Table 2.** Estimation of gene flow (population migration rates,  $2Nm$ ) across all parapatric boundaries between historical lineages of *Salamandra salamandra* in the Iberian Central System.

Contact Zone	parental taxa (1 vs 2)	from taxa 1 into taxa 2		from taxa 2 into taxa 1	
		HiPt	LLRtest	HiPt	LLRtest
1	<i>almanzoris</i> vs <i>bejarae</i> -W	0.078 (0.007-0.296)	7.324**	0.210 (0.044-0.642)	8.439**
2	<i>almanzoris</i> vs <i>bejarae</i> -E	0.345 (0.000-3.036)	4.571*	0.319 (0.000-3.786)	2.901*
3	<i>almanzoris</i> vs <i>bejarae</i> -E	0.286 (0.000-3.636)	3.060*	0.102 (0.000-1.237)	0.964ns
4	<i>gallaica</i> vs <i>bejarae</i> -W	0.853 (0.000-2.936)	3.690*	0.208 (0.000-0.902)	2.864*
5	<i>bejarae</i> -W vs <i>bejarae</i> -E	0.211 (0.043-0.594)	8.091**	0.031 (0.002-0.168)	8.153**

HiPt- Highest value in the histogram of probability densities, followed by 95% HPD confidence intervals; LLRtest- Likelihood ratio test (Nielsen & Wakeley 2001); Statistical significance is indicated by asterisks (\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ), or lack thereof by ‘ns’.

### Figures:

**Figure 1.** Sampling of the populations of *Salamandra salamandra* interacting along the Iberian Central System, plus an outgroup (ssp. *longirostris*). The colored areas correspond to the approximate range of the nine ecomorphs (subspecies) within the Iberian Peninsula (based on García-París *et al.* 2003). The colored circles locate the populations sequenced for mitochondrial and nuclear markers (see Suppl. Table 1 for precise localities and additional samples). Samples at the Contact Zones 1, 2 and 3 correspond to taxon pairs *bejarae*-*almanzoris* collected at low- and high-elevation, respectively.

**Figure 2.** Phylogeographic structure of the populations of *Salamandra salamandra* interacting along the Iberian Central System based on two mitochondrial genes. The colors of the clades correspond to the different ecomorphs (subspecies). The labels and colored circles define six well-supported (BPP > 0.99) phylogeographic groups. The black and grey asterisks indicate clade support (respectively BPP > 0.99 and BPP > 0.90).

**Figure 3.** Population-trees of the six focal phylogeographic groups of *Salamandra salamandra* inferred from \*BEAST analyses of mitochondrial (A) and nuclear markers (B). Asterisks indicate support (BPP > 0.99).

**Figure 4.** Integration of patterns of divergence and gene flow during the postglacial range shifts in the Iberian Central System. There are 3 recognized ecomorphs (subspecies) interacting along the ICS: *gallaica* (G), *bejarae* (B), *almanzoris* (A), plus geographically isolated population of *bejarae*; transitions between ecomorphs are diffuse along the continuous range and the precise boundaries remain unknown. Mitochondrial DNA shows that there are three cryptic lineages of *bejarae*: West (B<sub>W</sub>), East (B<sub>E</sub>), and South (B<sub>S</sub>); parapatric boundaries are sharp and geographically well delimited. Nuclear loci show that parapatric taxa are no longer distinct (A and B<sub>E</sub>) or show decreased differentiation (G and B<sub>W</sub>), while allopatric taxa remain highly divergent; this results in cyto-nuclear discordances. Isolation with migration models confirm significant gene flow across all contact zones.

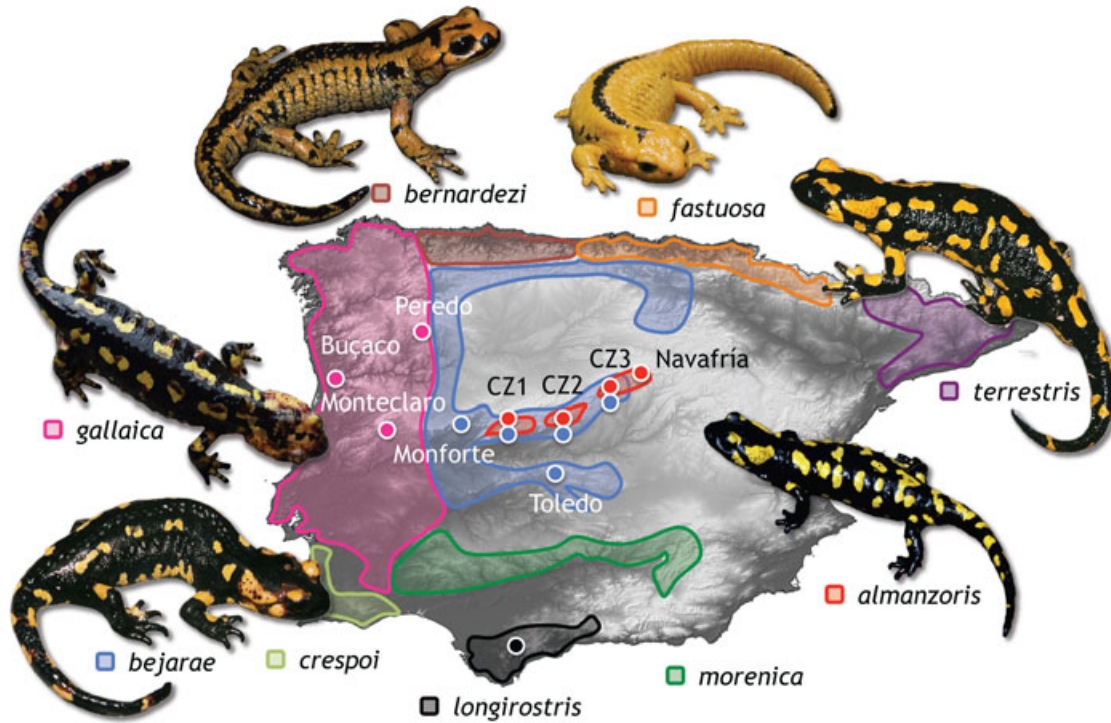
**Supplementary Figure 1.** Detailed distribution of sampling sites at the three replicated contact zones in the Iberian Central System. High- and low-elevation ecomorphs of *Salamandra salamandra* are colored as in Fig. 1. Polygons reflect the presumed distribution based on the mitochondrial study from Martinez-Solano *et al.* (2005). Dots represent the precise location of sampling localities included in pairwise comparisons to assess gene flow (with white counter line) and in the mitochondrial phylogeographic analysis (no counter line). Populations of ecomorph *gallaica* analyzed in CZ4 are located outside the zoomed area, whereas in CZ5, the three populations of ecomorph *bejarae* at the western end of the figure are tested against the five populations in the east.

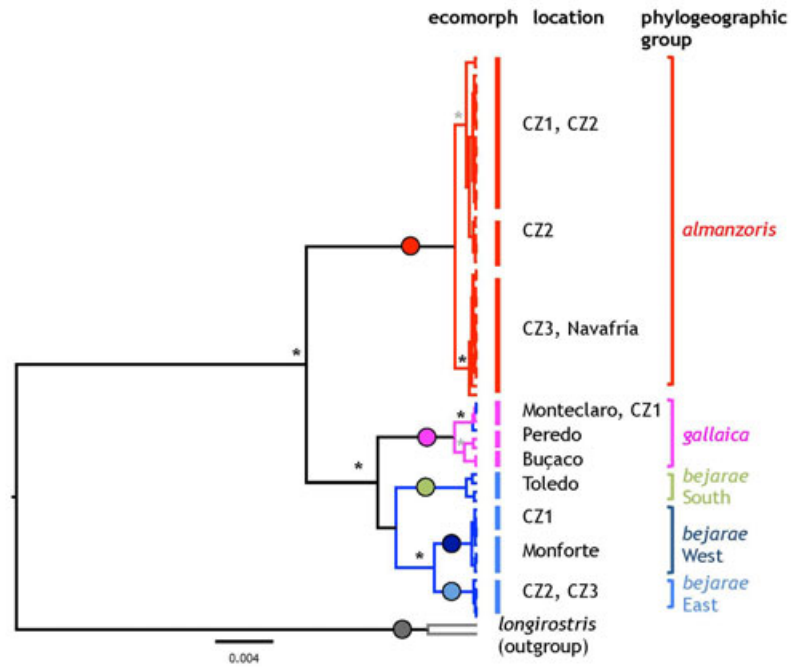
**Supplementary Figure 2.** Sorting of genetic variability across five independent nuclear markers. The circles represent all sequenced haplotypes and are colored accordingly to the respective phylogeographic groups observed in mtDNA (Figure 2).

**Supplementary Figure 3.** Post-burnin distribution of the ‘population-trees’ for the mtDNA (A) and nuclear (B) markers, as visualized in DensiTree 2.0 (Bouckaert 2010). DensiTree plots the 9000 trees for each distribution. All these trees have a finite number of topologies, which are represented with different colors. The most common topology appears in blue, the second most common in red, then in green and dark green.

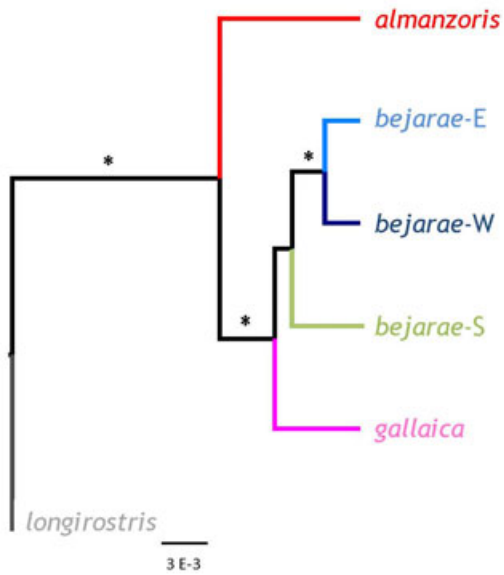


**Supplementary Figure 4.** Joint posterior density curves for  $2Nm$  for the five contact zones. Curves are colored according to the population that is the source of gene flow, as depicted in the legend.





A. mtDNA



B. nDNA

