ORIGINAL ARTICLE

No signs of inbreeding despite long-term isolation and habitat fragmentation in the critically endangered Montseny brook newt (*Calotriton arnoldi*)

E Valbuena-Ureña[1,2], A Soler-Membrives[1], S Steinfartz[3], P Orozco-terWengel[4] and S Carranza[5]

Endemic species with restricted geographic ranges potentially suffer the highest risk of extinction. If these species are further fragmented into genetically isolated subpopulations, the risk of extinction is elevated. Habitat fragmentation is generally considered to have negative effects on species survival, despite some evidence for neutral or even positive effects. Typically, non-negative effects are ignored by conservation biology. The Montseny brook newt (*Calotriton arnoldi*) has one of the smallest distribution ranges of any European amphibian (8 km^2) and is considered critically endangered by the International Union for Conservation of Nature. Here we apply molecular markers to analyze its population structure and find that habitat fragmentation owing to a natural barrier has resulted in strong genetic division of populations into two sectors, with no detectable migration between sites. Although effective population size estimates suggest low values for all populations, we found low levels of inbreeding and relatedness between individuals within populations. Moreover, *C. arnoldi* displays similar levels of genetic diversity to its sister species *Calotriton asper*, from which it separated around 1.5 million years ago and which has a much larger distribution range. Our extensive study shows that natural habitat fragmentation does not result in negative genetic effects, such as the loss of genetic diversity and inbreeding on an evolutionary timescale. We hypothesize that species in such conditions may evolve strategies (for example, special mating preferences) to mitigate the effects of small population sizes. However, it should be stressed that the influence of natural habitat fragmentation on an evolutionary timescale should not be conflated with anthropogenic habitat loss or degradation when considering conservation strategies.

INTRODUCTION

Among the threatened species, those that are endemic to a restricted spatial area should *per se* experience a higher risk of extinction. Such risk derives from either stochastic environmental processes (for example, extreme climatic conditions, fires, and so on) or effects of genetic drift and inbreeding (Allendorf and Luikart, 2007). Consequently, the preservation of genetic diversity is important for maintaining the evolutionary (adaptive) potential to overcome environmental changes and enable the population growth and survival that is crucial for the fitness of a species (Allentoft and O'Brien, 2010). In general, fragmentation of a species range into smaller subunits by external factors such as anthropogenic activities (Blank *et al.*, 2013; Storfer *et al.*, 2014) or climatic events (Veith *et al.*, 2003) reduces gene flow and compromises the population's long-term survival (Sunny *et al.*, 2014). A central goal of conservation biology is to identify the genetic structure and diversity of species at the population level (Apodaca *et al.*, 2012 and references therein) and characterize the gene flow between populations in relation to the species' dispersal propensity (that is, the probability of dispersal between habitat patches) and rates (Slatkin, 1994). Organisms with lower dispersal

rates are more susceptible to isolation than those with higher dispersal rates. Thus dispersal may counteract the loss of gene flow among populations and, therefore, has been shown to be an important factor for the long-term survival of species (Allentoft and O'Brien, 2010).

Strong genetic differentiation among populations is a sign of interrupted gene flow, with non-natural external factors such as human-induced disturbance causing habitat fragmentation and hindering dispersal (Templeton *et al.*, 1990). However, strong differentiation can also be the outcome of non-human mediated processes, such as naturally occurring habitat fragmentation, local adaptation (for example, Steinfartz *et al.*, 2007; Nosil *et al.*, 2009) or incipient speciation on a small spatial scale (for example, MacLeod *et al.*, 2015). Although it has been generally argued that fragmentation can lead to isolation and thus increase extinction risks, it has also been suggested that in some instances habitat fragmentation can have neutral or even positive effects (Fahrig, 2003). A fragmented species may develop populations that individually harbor low levels of genetic variation, but when all populations are considered together, the species does not present low levels of diversity. Consequently, fragmented species may

[1]Unitat de Zoologia, Facultat de Biociències, Universitat Autònoma de Barcelona, Cerdanyola del Vallès (Barcelona), Spain; [2]Centre de Fauna Salvatge de Torreferrussa (Catalan Wildlife Service – Forestal Catalana), Finca de Torreferrusa, Barcelona, Spain; [3]Zoological Institute, Department of Evolutionary Biology, Technische Universität Braunschweig, Braunschweig, Germany; [4]School of Biosciences, Cardiff University, Cardiff, UK and [5]Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Barcelona, Spain
Correspondence: Dr E Valbuena-Ureña or Dr A Soler-Membrives, Unitat de Zoologia, Facultat de Biociències, Universitat Autònoma de Barcelona, Cerdanyola del Vallès (Barcelona) 08193, Spain.
E-mail: Emiliojavier.Valbuena@uab.cat or Anna.Soler@uab.cat

preserve high levels of total genetic variation, similar to equally sized species with panmitic population (Templeton *et al.*, 1990).

Amphibians are generally considered to have limited dispersal abilities, causing genetic differentiation across small geographic scales (Monsen and Blouin, 2004 and references therein), although more recent studies indicate that, in some cases, dispersal propensities have been vastly underestimated (for example, Smith and Green, 2005). The notable sensitivity of amphibians to environmental change and habitat fragmentation are other factors that may reinforce patterns of sharp genetic discontinuation over short distances (Savage *et al.*, 2010; Storfer *et al.*, 2014; Velo-Antón *et al.*, 2013). Therefore, data on gene flow between populations of endangered amphibians should have a direct influence on management programs and decisions regarding conservation strategies, such as determining the number of breeding lines and translocation actions (Sunny *et al.*, 2014).

The genus *Calotriton* (Gray, 1858 and recently resurrected by Carranza and Amat, 2005) includes only two species, both inhabiting the Iberian Peninsula and adapted to live in cold and permanent-flowing streams: the Pyrenean brook newt (*Calotriton asper*) and the Montseny brook newt (*Calotriton arnoldi*). *Calotriton asper* is widely distributed across the Pyrenean mountain chain, with some populations extending northwards and southwards, reaching the Pre-Pyrenees, and occupying an area $>20\,000\,\text{km}^2$. In contrast, *C. arnoldi* is only known from the Montseny Natural Park in the NE Iberian Peninsula, and its disconnected populations are found within a restricted altitudinal range in seven geographically proximate brooks. Although the historic range of this species is unknown, it currently occupies a total area of only $8\,\text{km}^2$. Moreover, its habitat is naturally fragmented into two watersheds, on the eastern and western sectors of the Tordera River valley, separated by unsuitable terrestrial habitat between them (see Figure 1a). The current census population size of this species is estimated to be 1500 adult individuals (Carranza and Martínez-Solano, 2009). Additionally, recent human activities (for example, extraction of large amounts of water for commercial purposes, deforestation and the building of forest tracks and roads) have had a significant negative effect on *C. arnoldi*'s habitat (Amat *et al.*, 2014). Hence, *C. arnoldi* is one of the most spatially restricted and endangered vertebrates in Europe, and it is classified as critically endangered by the International Union for Conservation of Nature (Carranza and Martínez-Solano, 2009).

A previous study based on mitochondrial (Cyt *b*) and nuclear (recombination-activating gene 1 (RAG-1)) sequences, as well as morphological characters suggested a high degree of differentiation between populations in the eastern and western sectors of *C. arnoldi*'s distribution range (Valbuena-Ureña *et al.*, 2013). Therefore, the observed fragmentation of this species into highly genetically isolated populations is probably the result of an ancient, naturally driven, intrinsic fragmentation process rather than the result of recent human disturbances. However, a detailed exploration of population structure, gene flow among populations and estimates of ancient and current effective population sizes is lacking. Such studies are crucial for the understanding of past and ongoing evolutionary processes and their implications for the conservation of such a spatially restricted endemic species.

Although there exists several studies of species with very limited distribution ranges (for example, Sunny *et al.*, 2014), as well as of amphibians with highly structured populations (Monsen and Blouin, 2004; Blouin *et al.*, 2010; Savage *et al.*, 2010; Blank *et al.*, 2013), *C. arnoldi* represents an exceptional example of a critically endangered amphibian species with limited dispersal capabilities inhabiting a very small fragmented habitat. Here we present an analysis of the genetic diversity and evolutionary history of *C. arnoldi*, which provides general insights into management priorities of species with a very limited distribution range. In order to estimate the effects of natural habitat fragmentation for this species in terms of fitness-related genetic parameters (for example, genetic diversity, inbreeding coefficients and so on), we compared these parameters directly in populations from the non-fragmented range of its sister species *C. asper* in the central Pyrenees. We discuss the absence of anticipated negative consequences for these parameters in *C. arnoldi* in the light of species conservation in naturally fragmented species.

MATERIALS AND METHODS

Sampling and DNA extraction

A total of 160 adult *C. arnoldi* were analyzed, including samples from all 7 known locations of this species (Figure 1a). In recognition of the low dispersal capacity (Carranza and Martínez-Solano, 2009) and the absence of migrants between sites (see Results below), individuals from the seven locations are considered herein as demographic populations. Genetic populations will be referred to henceforth as clusters. For conservation reasons, the three eastern populations are herein referred to as A1, A2 and A3 and the four western populations as B1, B2, B3 and B4. Samples included 77 individuals from the eastern sector (23 from A1 and 27 from each A2 and A3 populations) and 83 individuals from the western sector (25 from B1, 28 from B2, 26 from B3 and 4 from B4). The small number of individuals from B4 is due to the low abundance of individuals at this site. Therefore, results from this population should be treated with caution. Tissue samples consisted of small tail or toe clips preserved in absolute ethanol. Genomic DNA was extracted using the Qiagen (Valencia, CA, USA) DNeasy Blood and Tissue Kit, following the manufacturer's protocol.

Phylogenetic analyses and estimation of divergence times

A data set of mitochondrial and nuclear genes was assembled to estimate the divergence times between *C. asper* and *C. arnoldi*, as well as between *C. arnoldi*'s populations. This data set consisted of three samples of *C. asper* from Irati, northwestern Pyrenees, Spain (see Milá *et al.*, 2010) and a randomly selected set of three samples from each of the seven known wild populations of *C. arnoldi* (Figure 1a, Supplementary Table S1). The following regions of four mitochondrial and three nuclear genes were amplified and sequenced for both strands, totaling 3553 base pairs (bp; 84 variable positions): 374 bp (16 variable) of the mtDNA gene cytochrome *b* (Cyt *b*) using primers Cytb1EuprF and Cytb2EuprR from Carranza and Amat (2005) and conditions as in Carranza *et al.* (2000); 556 bp (42 variable) of the mtDNA gene NADH dehydrogenase subunit 4 (ND4) using primers from Arèvalo *et al.* (1994) and conditions as in Martínez-Solano *et al.* (2006); 370 bp (5 variable) of the mtDNA gene 12S rRNA (12S) and 553 bp (5 variable) of the mtDNA gene 16S rRNA (16S) with the same primers and conditions as in Carranza and Amat (2005); 695 bp (7 variable) of the nucDNA gene proopiomelanocortin and 475 bp (3 variable) of the nucDNA gene brain-derived neurotrophic factor using primers and conditions as in Recuero *et al.* (2012); and 530 bp (6 variable) of the nucDNA gene RAG-1 with primers and conditions as in Šmíd *et al.* (2013). GENEIOUS v. R6.1.6 (Biomatters Ltd., Auckland, New Zealand) was used for assembling and editing the chromatographs. Heterozygous positions for the nuclear-coding gene fragments were identified based on the presence of two peaks of approximately equal height at a single nucleotide site in both strands and were coded using IUPAC ambiguity codes. The nuclear-coding fragments were translated into amino acids and no stop codons were observed. DNA sequences were aligned for each gene independently using the online application of MAFFT v.7 (Katoh and Standley, 2013) with default parameters (auto strategy, gap opening penalty: 1.53, offset value: 0.0). In order to optimize the alignment of the ribosomal genes, we did not include any outgroups and used Bayesian methods for inferring the root of the phylogenetic tree (Huelsenbeck *et al.*, 2002).

Best-fitting models of nucleotide evolution were inferred using jModeltest v.0.1.1 (Darriba *et al.*, 2012) under the Akaike information criterion (Akaike, 1973). The HKY model was selected for the 12S and 16S genes and the TrN for all the remaining genes. Phylogenetic analyses were performed using BEAST v.1.8.0 (Drummond and Rambaut, 2007). For the time calibration, we used the
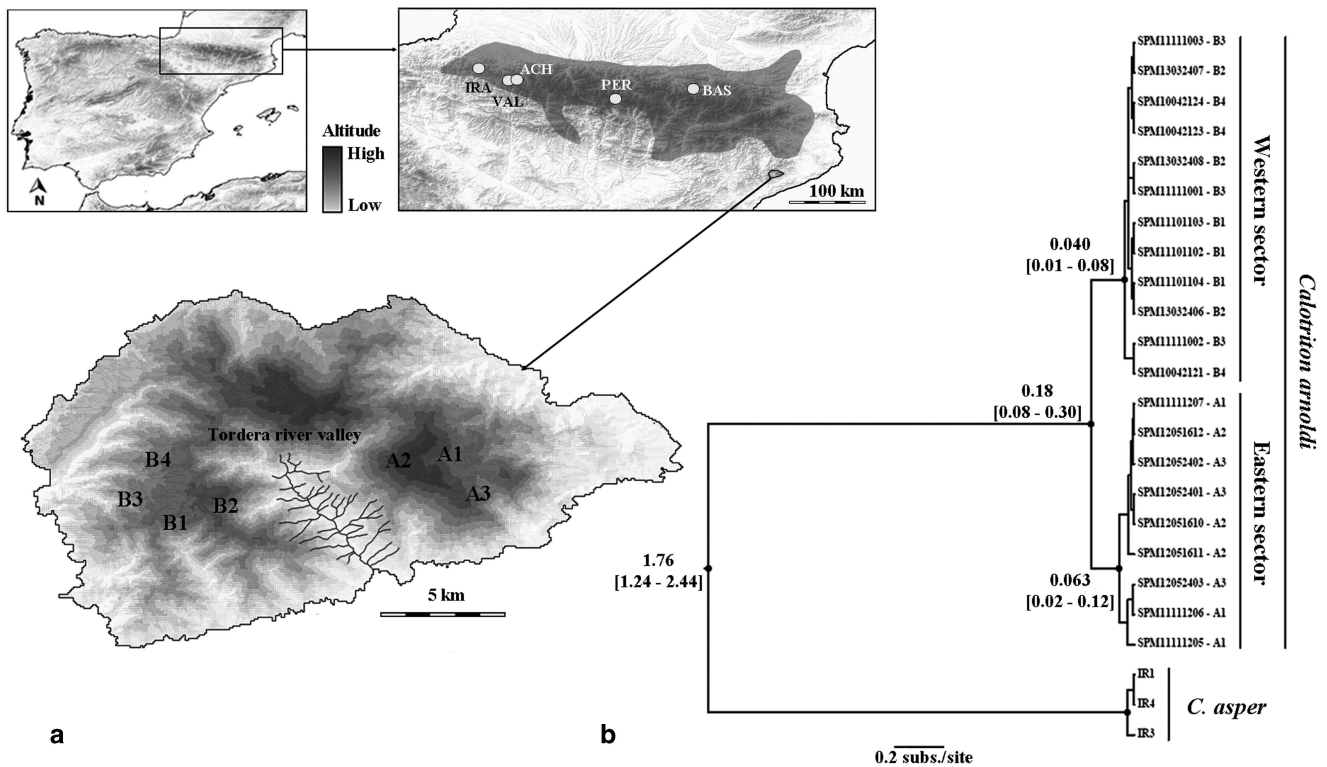
Figure 1 (**a**) The distribution range of the Montseny brook newt, *Calotriton arnoldi*. Populations located in the eastern sector and in the western sector are separated by the Tordera river valley (Valbuena-Ureña *et al.*, 2013). All localities have been sampled for this study. Shade indicates the actual distribution range of *Calotriton asper*. Locations of the *C. asper* populations used in this study are also shown (ACH: Ibón de Acherito; BAS: Bassies; IRA: Irati; PER: Ibón de Perramó; VAL: Barranco de Valdragás). (**b**) Bayesian inference tree of *Calotriton* inferred using BEAST with the concatenated data sets. A list of details of all the specimens is presented in Supplementary Table S1. Black-filled circle indicates pp > 0.95 in the BEAST analysis. Ages of some relevant nodes are shown by the nodes with the 95% HPD underneath between square brackets.

rate of molecular evolution of the Cyt *b* gene by Hauswaldt *et al.* (2014), inferred for the Urodelan genus *Salamandrina*, based on four fossil/geological calibration points. Three individual runs of 5×10^7 generations were performed, sampling every 10 000 generations. Models and prior specifications applied were either program defaults or as follows: model of sequence evolution for each gene, as indicated above; substitution models and clock models unlinked; trees linked; coalescent constant size tree prior; random starting tree; and strict clock rate for all the partitions. The molecular evolution rate of Hauswaldt *et al.* (2014) was implemented in our analyses in the clock rate prior of Cyt *b* using a normal distribution centered at 0.0102 subst/site/Myr, and with an s.d. that captured 95% of the high probability density of the posterior reported by Hauswaldt *et al.* (2014) (0.0085–0.018 subst/site/Myr). Posterior trace plots and effective sample sizes of the runs were monitored in TRACER v1.5 (Rambaut and Drummond, 2007) to ensure convergence. The results of the individual runs were combined in LogCombiner, discarding the initial 10% of the samples, and the maximum clade credibility ultrametric tree was produced with TreeAnnotator (both provided with the BEAST package). Nodes were considered strongly supported if they received posterior probability (pp) support values ⩾ 0.95.

Microsatellite loci genotyping and basic population genetic parameters

Individuals were genotyped for a total set of 24 microsatellite loci: 15 specifically developed for *C. arnoldi* (Valbuena-Ureña *et al.*, 2014) and 9 additional loci originally developed for the closely related sister species *C. asper* and which cross-amplify successfully in *C. arnoldi* (Drechsler *et al.*, 2013). Microsatellite loci were multiplexed in five mixes using the Type-it multiplex PCR (Qiagen). Primer combinations of the five mixes are provided in Supplementary Table S2. PCR conditions and genotyping of loci followed the descriptions provided in Drechsler *et al.* (2013).

The MICRO-CHECKER software (Van Oosterhout *et al.*, 2004) was used to check for potential scoring errors, large allele dropout and the presence of null alleles. Pairwise linkage disequilibrium between loci was checked using the software GENEPOP version 4.2.1 (Rousset, 2008). The same program was used to calculate deviations from Hardy–Weinberg equilibrium in each population and for each locus, which provides an exact probability value (Guo and Thompson, 1992). Genetic diversity was measured for each sampling site as the mean number of alleles (A), observed (H_O) and expected heterozygosity (H_E) and allelic richness (Ar) using FSTAT version 2.9.3.2 (Goudet, 1995). The observed number of private alleles for each locus and each population was calculated with GDA (Lewis and Zaykin, 2000), and a rarified measure of private allele richness was obtained with HP-RARE (Kalinowski, 2005). FSTAT was used to estimate the populations' inbreeding coefficients (F_{IS}) following Weir and Cockerham (1984).

In order to compare the genetic diversity measures of *C. arnoldi* to its more widely distributed sister species in the central Pyrenees, we used four *C. asper* populations from the study of Drechsler *et al.* (2013) (Figure 1a) as a reference, adding some samples and re-sequencing others for some markers. The genetic diversity estimates in terms of differences in heterozygosity estimates (H_E and H_O) and number of alleles per locus (A) between *C. arnoldi* populations (and clusters) and these four populations of *C. asper* were tested pairwise using a nonparametric Wilcoxon signed-rank test, with Bonferroni's correction for multiple comparisons. The inbreeding coefficients F_{IS} were also estimated for the *C. asper* populations.

Microsatellite loci-derived population structure analysis

The pairwise population divergence between *C. arnoldi*'s seven sampling localities was estimated with the F_{ST} as calculated in FSTAT and with Jost's *D* (Jost, 2008) using the R package DEMEtics (Gerlach *et al.*, 2010). We also used a Bayesian approach to examine population structure of *C. arnoldi* across

its distribution range, as implemented in STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000). STRUCTURE's Bayesian clustering algorithm assigns individuals to clusters without using prior information on their localities of origin. Settings used included an admixture model with correlated allele frequencies, and the number of inferred clusters (K) ranged from one (complete panmixia) to eight (that is, the number of sample locations plus one). STRUCTURE was run for each value of K 10 times, with one million Markov Chain Monte Carlo (MCMC) iterations, discarding the first 100 000 MCMC steps as burn-in phase. We also ran STRUCTURE with the same parameters for each sector (eastern and western) separately to check for possible genetic substructure within sectors. The optimal number of clusters was inferred using ΔK method by Evanno *et al.* (2005), as implemented in STRUCTURE HARVESTER (Earl and vonHoldt, 2012). The average from all the outputs of each K was obtained with CLUMPP version 1.1.2 (Jakobsson and Rosenberg, 2007) and plotted with DISTRUCT version 1.1 (Rosenberg, 2004). Additionally, we employed a model-independent clustering approach using GENETIX, version 4.05.2 (Belkhir *et al.*, 2004), by performing a factorial correspondence analysis on the allelic frequencies obtained for the seven Montseny brook newt populations. This analysis was performed across the distribution range of *C. arnoldi*, as well as in each sector separately, to examine the existence of substructure within them. Analysis of molecular variance (AMOVA) was performed in ARLEQUIN 3.5.1.2 (Excoffier and Lischer, 2010) by grouping the sampling localities as indicated by STRUCTURE. Isolation by distance was evaluated by examining the relationship between geographical and genetic distances between populations with a Mantel test (Mantel, 1967). As the lifestyle of *C. arnoldi* is strictly aquatic (Carranza and Amat, 2005), geographic distances were calculated following the watercourse and log-transformed to linearize the relationship between geographic distances and F_{ST} values (see Rousset, 1997). Genetic distances were calculated as $F_{ST}/(1 - F_{ST})$, and the significance of matrix correlation coefficients was estimated with 2000 permutations in ARLEQUIN. Analyses were performed between all sampled populations and by grouping populations by sector using ARLEQUIN.

Analysis of recent gene flow

Recent gene flow between sectors and populations within sectors were assessed using three programs: GENECLASS 2.0 (Piry *et al.*, 2004), STRUCTURE, and BIMr (Faubet and Gaggiotti, 2008). The Bayesian assignment approach implemented in GENECLASS was used following Paetkau *et al.* (2004). STRUCTURE was rerun to detect migrants by calculating a Q value, which is the proportion of that individual's ancestry from a population. An individual is a putative migrant when the Q value for its origin site (Q_o) is lower than the Q value for its site of assignment (Q_a). BIMr was used to estimate migration rates within the last two generations ($N_{gen} \leqslant 2$) between populations within sectors. BIMr uses a Bayesian assignment test algorithm to estimate the proportion of genes derived from migrants within the last generation, assuming linkage equilibrium and allowing for deviation from Hardy–Weinberg equilibrium. We estimated migration rates among populations within sectors separately. For each analysis, we ran a Markov chain with a burn-in period of 50 000 iterations, followed by 50 000 samples that were collected using a thinning interval of 50. Convergence of the Markov Chain was assessed by repeating the analyses independently five times. Pairwise migration rates between and within populations across runs were averaged.

Inference of demographic history

The effective population size (N_e) for each *C. arnoldi* population and cluster resulting from STRUCTURE and for the four *C. asper* populations were calculated using three single-sample N_e estimators: ONeSAMP (Tallmon *et al.*, 2008), COLONY version 2.0.4.4 (Jones and Wang, 2010), and LDNe version 1.31 (Waples and Do, 2008). ONeSAMP employs approximate Bayesian computation and calculates eight summary statistics to estimate N_e from a sample of microsatellite loci genotypes. The analyses were submitted online to the ONeSAMP 1.2 server (http://genomics.jun.alaska.edu/asp/Default.aspx). A variety of input priors were tested, with minimum N_e as low as 2 and maximum N_e as high as 1000. After convergence of test runs was achieved, the prior distributions were set between a minimum N_e of 2 and a maximum value of 100 for populations or 500 for clusters. COLONY implements a maximum

likelihood method to conduct sibship assignment analyses, which are used to estimate N_e under the assumption of random mating. COLONY was run using the maximum likelihood approach for a dioceous/diploid species, with medium length runs and random mating, assuming polygamy for both males and females (as is the case for most salamanders) with no sibship prior. We did not use the option 'update allelic frequencies' and other parameters used as default. Finally, LDNe employs a linkage disequilibrium method (Hill, 1981) using a jackknife approach to estimate confidence intervals (CIs) and assuming a minimum allele frequency of 2% in order to reduce the bias caused by rare alleles.

In order to characterize changes in the demographic history, additional analysis was performed in MSVAR version 1.3 (Storz and Beaumont, 2002). This analysis was undertaken for all populations with the exception of B4, owing to the low sample size. MSVAR uses a Bayesian approach with coalescent simulations to estimate three population demographic parameters: (i) the ancestral population size (N_t) of a population, (ii) its current effective population size (N_0), and (iii) the time (t) at which the change from N_t to N_0 occurred. Three scenarios, a bottleneck, an expansion and a stable demography, were tested for each population in order to assess whether the posterior distributions of the three parameters of interest were independent of the prior distributions used to run the analyses. As no microsatellite mutation rate for this species has been described, an average vertebrate rate of 10^{-4} was used (Bulut *et al.*, 2009), allowing the rate to vary by up to two orders of magnitude above (10^{-2}) and below (10^{-6}). Prior distributions are shown in Supplementary Table S3. Each MSVAR run consisted of 4×10^8 iterations of the MCMC algorithm, discarding the first 25% of the coalescent simulations. Gelman and Rubin's diagnostic (Brooks and Gelman, 1998) was used to asses convergence between the independent MSVAR runs using the library CODA (Plummer *et al.*, 2006) in R. Finally, the demographic analysis with MSVAR was complimented with bottleneck analyses in BOTTLENECK v1.2.02 (Piry *et al.*, 1999), under the stepwise mutation and the two-phased mutation models with default parameters.

Genetic relatedness of individuals

In order to measure levels of inbreeding, the software MLRELATE (Kalinowski *et al.*, 2006) was used, which estimates the relatedness among individuals within each population. This program is appropriate as it is designed for microsatellite loci, is based on maximum likelihood tests and considers null alleles. Furthermore, GenAlEx v. 6 (Peakall and Smouse, 2006) was used to obtain pairwise relatedness among individuals in each population separately using the r_{qg} estimator (Queller and Goodnight, 1989). Mean pairwise relatedness values and their 95% CI estimates were calculated for the east and west sectors separately, and the statistical differences in mean population relatedness between populations were assessed with a permutation test following Peakall and Smouse (2006). These CI intervals of r_{qg} from the simulations represent the range of r_{qg} that would be expected under random mating across all populations within sectors. Population r_{qg} values that fell above the expected 95% CI values indicate a higher relatedness than anticipated and are possibly due to reproductive skew, inbreeding or genetic drift among populations within the same sector. These estimates of genetic relatedness were also computed for the four *C. asper* populations and are used as a reference of non-fragmented populations.

RESULTS

Estimation of divergence times

Convergence was confirmed by examining the likelihood and posterior trace plots of the three runs with TRACER v.1.5. Effective sample sizes of the parameters were >200, indicating a good representation of independent samples in the posterior. The phylogenetic relationships are shown in Figure 1b. *Calotriton asper* and *C. arnoldi* form two independent clades, and within *C. arnoldi* there are two well-supported reciprocally monophyletic groups that include the populations from the eastern and western sectors. Nevertheless, none of the sampling localities within either of the two sectors were monophyletic, likely indicating a lack of resolution of the gene fragments used and/or gene flow between the localities in each sector or the retention of ancestral polymorphisms between sectors. According to the present dating estimates, *C. asper* and *C. arnoldi* diverged approximately 1.76

Mya (95% HPD 1.24–2.44 Ma) and the eastern and western sectors of *C. arnoldi* 0.18 Mya (95% HPD 0.08–0.30 Ma).

Genetic diversity

Genetic diversity for each sampled population and cluster obtained from the genetic structure analyses are given in Table 1 (Supplementary Table S4, for locus-specific results). Loci Us3 and Us7 were monomorphic for the populations within the western sector. We further found that some alleles were fixed for some populations: Calarn15906 was found to be monomorphic in population B2, seven loci (Calarn 29994, Calarn06881, Calarn36791, Calarn52354, Calarn31321, Calarn15136 and Us2) were fixed in population B3, and loci Calarn15906 and Ca32 showed no polymorphisms for individuals of population B4. The observed number of alleles per locus ranged from 4 to 12, with a mean of 7.08, and the mean number of alleles in the eastern and western populations were 5.50 and 3.96, respectively. There was no sign of linkage disequilibrium between any pair of loci, with the only exception of Calarn02248 and Calarn50748 in population B3 after Bonferroni correction ($P < 0.00018$). Only two loci in two different populations showed signs of null alleles (Us7 in A1 and Ca22 in B1). Private alleles (PA)—defined here as alleles exclusively found in a single population throughout the study site, that is, the species range—are also listed in Table 1. Populations of the eastern sector had 75 PAs, while the western populations had 38 PAs. Allelic richness (AR) per population ranged from 1.77 to 4.22, and expected heterozygosity ranged from 0.197 to 0.559 (weighted average: 0.441), with the lowest value found in B3 and the highest in A3. No significant departures from Hardy–Weinberg equilibrium ($P > 0.0003$) were found after applying Bonferroni's correction. Overall, F_{IS} was estimated to be 0.380 ($P = 0.0021$), but this parameter did not show values significantly different from zero for each population after applying Bonferroni's correction (see Table 1).

Similar levels of genetic diversity were observed between *C. arnoldi* populations or clusters and the four *C. asper* populations (Table 1,

Supplementary Table S5). None of the F_{IS} values were significantly different from zero for these four populations after applying Bonferroni's correction. In general, the total number of alleles per locus and expected and observed heterozygosity values were similar between *C. asper* and *C. arnoldi* populations (and clusters). The differences detected in the western sector are mostly due to population B3, which has low levels of genetic diversity. This population showed some significant differences, mainly when compared with *C. asper* populations at Barranco de Valdragás and Ibón de Acherito (Supplementary Table S6). We only found significant differences in one of the diversity indices explored (*number of alleles per locus*) between two populations (B2 and B4) and two out of the four *C. asper* populations. We did not detect differences in terms of expected nor observed heterozygosities between any *C. arnoldi* and *C. asper* populations but population B3.

Determining population structure

Population differentiation was significant for each pair of population combinations ($P < 0.001$) for both the F_{ST} and Jost's D (Table 2). F_{ST} values between the eastern versus western sector populations ranged from 0.443 to 0.617, and Jost's D values from 0.801 to 0.877. Pairwise comparisons between populations within sectors were much lower, with F_{ST} and D values within sectors ranging from 0.086 to 0.372 and from 0.100 to 0.299, respectively. Population B4 was not included in the F_{ST} and D estimations owing to its low sample size. Populations A3 of the eastern and B3 of the western sector were the most differentiated populations when compared with their respective sector populations.

Consistent with the results of phylogenetic analyses, STRUCTURE revealed two highly distinct genetic clusters corresponding to populations constituting the eastern and the western sectors (Figure 2). The existence of two clusters was highly supported by the analysis of ΔK values corresponding to $K = 2$ (Figure 2b). Some evidence for additional substructure is also indicated by a second weak peak at $K = 4$. When each sector was analyzed independently, two clusters

Table 1 Estimates of genetic parameters for each *Calotriton arnoldi* population and cluster defined by STRUCTURE analysis and for the four *Calotriton asper* populations

	Grouping	N	A	Ar	PA	PAAr	H_O	H_E	F_{IS}
C. arnoldi	Population								
	A1	23	4.167	4.167	7	0.311	0.545	0.538	0.017
	A2	27	4.042	3.954	5	0.214	0.526	0.516	−0.015
	A3	27	4.292	4.222	5	0.215	0.560	0.559	−0.007
	B1	25	3.542	3.500	3	0.137	0.467	0.469	−0.005
	B2	28	2.917	2.860	2	0.087	0.371	0.380	0.028
	B3	26	1.792	1.768	2	0.079	0.230	0.197	−0.121
	B4	4	2.333	—	0	—	0.438	0.433	−0.023
	Clusters								
	Eastern	77	4.167	4.112	75	3.157	0.544	0.538	**0.090**
	A1–A2	50	4.099	4.052	19	0.746	0.535	0.526	0.029
	Western	83	2.724	2.703	38	1.646	0.359	0.352	**0.184**
	B1–B2–B4	57	3.150	3.162	19	0.750	0.418	0.423	**0.073**
C. asper	Population								
	Ibón de Perramó	48	3.947	4.210	19	0.710	0.438	0.444	0.025
	Barranco de Valdragás	39	6.000	6.010	18	0.770	0.641	0.619	−0.022
	Ibón de Acherito	40	5.895	5.940	18	0.520	0.593	0.588	0.005
	Bassies	162	4.071	5.690	33	1.130	0.500	0.558	0.107

Abbreviations: A, number of alleles per locus; Ar, allelic richness; F_{IS}, inbreeding coefficient; H_E, expected heterozygosity; H_O, observed heterozygosity; N, sample size; PA, number of private alleles; PAAr, allelic richness of private alleles. Values represent averages across 24 loci. Values in bold indicate statistical significance after Bonferroni's correction.

were further identified in each sector, grouping A3 separately from A1 and A2 and B3 separately from B1, B2 and B4 (Figure 2a). The same general results were also found with the factorial correspondence analysis (Supplementary Figure S1), which demonstrated the clear separation between the two sectors. In these results, A3 and B3 were also the most distinct populations in their respective sectors. The results of the *a posteriori* AMOVA revealed that the clusters resulting from STRUCTURE ($K = 2$) explained 40.81% of the molecular variance, 11.61% was explained by among populations within groups

and 48.31% by within-population variation. These results agree with the population differentiation analysis (F_{ST} values; Table 2).

A relationship between genetic differentiation and geographical distance (Supplementary Table S7) was found among all sampled populations (Supplementary Figure S2, $r = 0.735$, $P = 0.020$), suggesting a strong isolation by distance effect at the level of all populations. However, at a finer scale, when both sectors were analyzed independently, no isolation by distance was observed.

Recent gene flow and migration rates

No migration between the eastern and western sectors could be detected by any of the methods used. All individuals of the eastern sector were assigned by GENECLASS with a probability of $\geqslant 90\%$ to their population of origin. In the western sector, 96% of individuals originating from B1 and 88% of the individuals originating from B3 were correctly assigned to their population of origin. Among the samples from B2, a total of 10.7% of individuals were assigned to B4 and 17.9% to B1, while the remaining 71% of individuals were assigned to their population of origin. About 8.7% of the individuals could not be assigned to any of the sampled populations. No first-generation migrants were detected among populations from the eastern sector, and in the western sector only one individual from B2 was detected to be a migrant from B1 ($P = 0.001$). Although this individual had a low probability of being a migrant from B1

Table 2 Genetic differentiation among populations

F_{ST}/D	A1	A2	A3	B1	B2	B3	B4
A1	—	0.131	0.243	0.814	0.816	0.852	0.782
A2	0.086	—	0.299	0.855	0.868	0.877	0.822
A3	0.151	0.178	—	0.806	0.801	0.858	0.762
B1	0.457	0.473	0.443	—	0.100	0.249	0.122
B2	0.509	0.524	0.491	0.096	—	0.248	0.146
B3	0.614	0.617	0.599	0.336	0.372	—	0.305
B4	0.443	0.460	0.419	0.109	0.145	0.488	—

Pairwise F_{ST}, below the diagonal; D estimator values above the diagonal. All P-values were significant ($P < 0.001$).
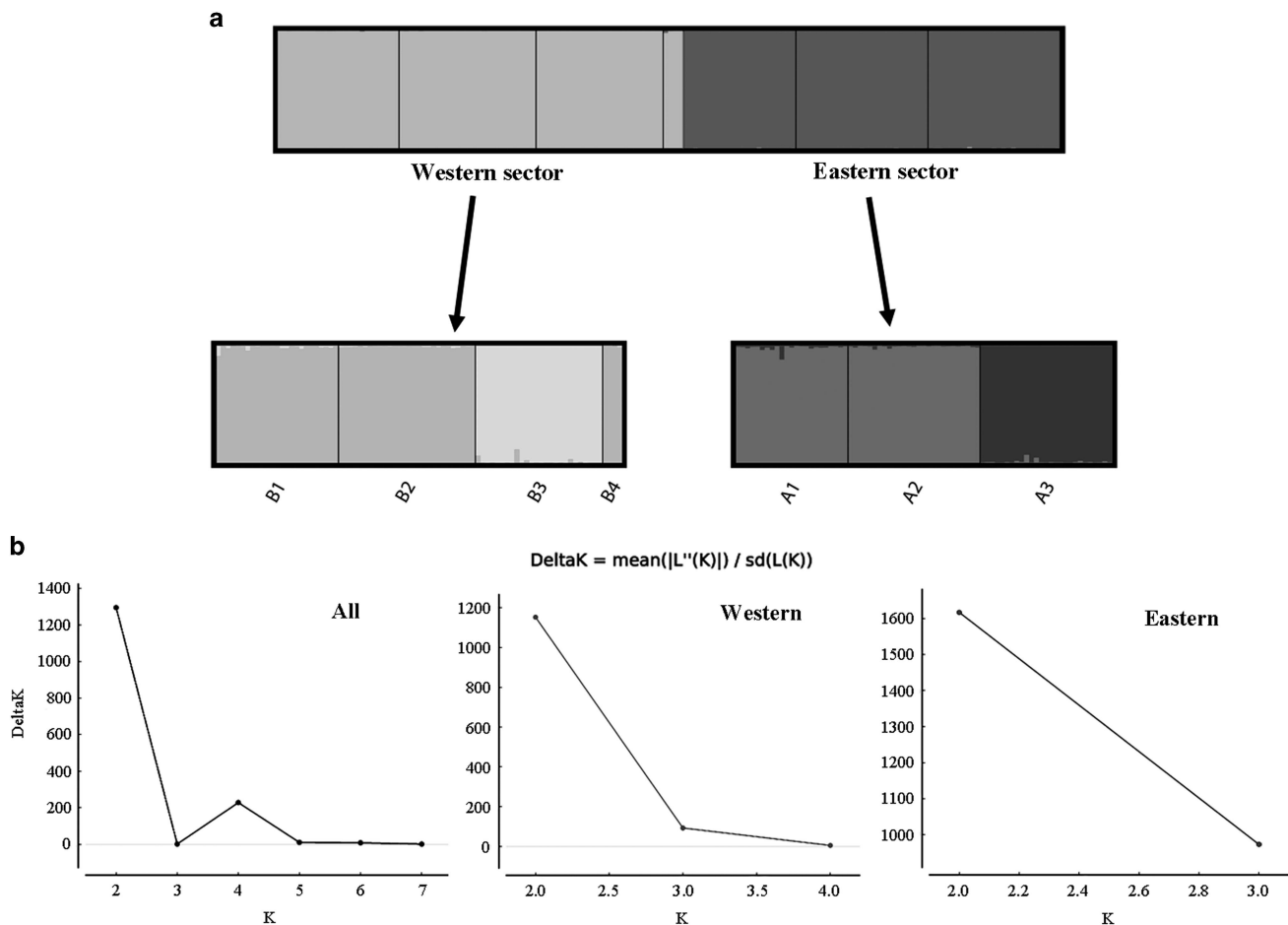

Figure 2 (**a**) Results of Bayesian clustering and individual assignment analysis obtained with STRUCTURE after running the program with all populations (above) and by sector (below); vertical bars delimit sampling locations. (**b**) Inference for the best value of K based on the ΔK method among runs for all populations and by sector.

(probability of migration of 0.038 according to analysis in STRUC-TURE), it had an estimated pp of 0.231 to have a single parent from population B1. STRUCTURE results were similar to GENECLASS assignments, with >98% of the sampled individuals being assigned to their population of origin. Estimates of recent gene flow using BIMr were consistent among the five independent runs, suggesting that convergence of the Markov chain had been reached. Recent migration rates showed no detectable recent gene flow between populations within either the eastern or western sectors (Supplementary Table S8). The overall outcome of the recent gene flow and migration rate analyses are in keeping with the population structure detected above.

Demographic history—effective population size (N_e)

All three methods used to estimate the effective population sizes (N_e) of the seven populations and of the two sectors (that is, the two clusters identified by STRUCTURE) resulted, in general, in low values, ranging from 7 to 342 (see Table 3). Estimation of the 95% CI upper limit for populations A1 and A3 was problematic (estimated at infinity or incongruence values) in the LDNe. Despite slight differences between methods, in general all estimators showed narrow CIs, thereby supporting the accuracy of these estimates. Effective population sizes were particularly low in population B3, with a 95% CI estimate of $N_e = 2$–30 regarding the three methods used. Effective population sizes for the *C. asper* populations were similar or higher than those of the *C. arnoldi* populations (Table 3).

Consistent with the previous results, MSVAR analysis also indicated relatively small current N_e values for each of the six populations tested (B4 was not tested). For all populations, the current effective population size seems to have been the outcome of a reduction in N_e some time between 1000 and 10 000 years ago, with the populations' ancestral N_e being at maximum, one to two orders of magnitude larger than the current one, for example, N_0 A2 ~ 100, N_t A2 ~ 1000 (Figure 3 and Supplementary Table S9). The ratio of the

ancestral N_e divided by the current N_e consistently result in values >1 for all populations, providing further evidence for a decrease in N_e in the past of these populations (Supplementary Figure S3). All MSVAR analyses showed convergent results, as indicated by a Gelman and Rubin statistic being under 1.2, with the exception of the clustered populations identified by STRUCTURE, where the MCMC did not converge. Finally, analyses performed in BOTTLENECK did not identify a significant excess of heterozygosity in any of the sampled populations (nor in the sectors) (Wilcoxon one-tailed test for excess of heterozygosity $P > 0.05$ for all tests), suggesting that the bottleneck indicated by MSVAR probably did not cause a dramatic loss of genetic diversity.

Relatedness of individuals

All populations presented a similar proportion of relatedness (Table 3), with most individuals being highly unrelated to each other (80%). The only exception was population B3, which had a lower percentage of unrelated specimens (68%). Full sibling and parent–offspring relations in B3 were 12% and 11%, respectively, while the other populations showed much lower percentages, none exceeding 4%. The estimated Queller and Goodnight (1989) index of relatedness, calculated between individuals in each population separately, indicated random mating among individuals within each population, that is, panmixia within populations (A1, $r_{qg} = -0.045$; A2, $r_{qg} = -0.038$; A3, $r_{qg} = -0.038$; B1, $r_{qg} = -0.042$; B2, $r_{qg} = -0.037$; B3, $r_{qg} = -0.067$; B4, $r_{qg} = -0.333$). Conversely, at the sector level (that is, when testing whether there is random mating between populations within a sector), values ranged from 0.150 to 0.206 and from 0.019 to 0.231 within the eastern western sectors, respectively, with the only exception being population B3, which showed an average pairwise relatedness (r_{qg}) of 0.745 (upper and lower CI estimates at 95% of 0.759 and 0.730, respectively). Most populations demonstrated significantly higher relatedness than could be expected if each sector represented a

Table 3 Estimates of effective population size (N_e) for each population and cluster of *Calotriton arnoldi* and for the four *Calotriton asper* populations, calculated with three programs: LDNe, ONeSAMP, and COLONY; estimations of the upper and lower 95% CI estimates for each method are indicated

		LDNe			OneSamp			COLONY			Relationship			
		N_e	95% CIs		N_e	95% CIs		N_e	95% CIs		Unrelated	Half siblings	Full siblings	Parent offspring
C. arnoldi	Population													
	A1	342.30	77.00	Infinite	27.65	24.51	34.81	46.00	26.00	90.00	91.70	7.51	0.40	0.40
	A2	49.40	32.20	93.20	33.94	29.94	41.89	40.00	25.00	71.00	90.31	7.12	0.85	1.71
	A3	142.10	61.80	Infinite	36.85	33.33	43.20	44.00	28.00	80.00	91.17	7.98	0.28	0.57
	B1	55.80	34.10	126.40	31.59	27.77	40.69	35.00	20.00	68.00	86.33	12.00	0.67	1.00
	B2	62.20	27.50	15091.10	36.39	30.46	53.44	31.00	18.00	57.00	81.75	13.23	1.85	3.17
	B3	7.30	2.40	21.70	14.97	12.61	19.61	13.00	7.00	30.00	68.31	8.00	12.31	11.38
	B4	Infinite	Infinite	Infinite	5.54	4.87	6.64	—	—	—	100	—	—	—
	Clusters													
	A1–A2	44.50	36.00	56.50	85.95	66.52	127.10	60.00	41.00	92.00	84.16	14.37	0.57	0.90
	B1–B2–B4	30.00	23.60	39.00	80.14	55.44	157.01	42.00	27.00	66.00	80.89	15.91	1.50	1.69
C. asper	Population													
	Ibón de Perramó	349.80	100.30	Infinite	42.80	33.96	62.07	40.00	26.00	65.00	84.13	12.68	1.60	1.60
	B. de Valdragás	1293.00	201.40	Infinite	41.41	35.29	58.22	58.00	37.00	98.00	90.69	8.50	0.40	0.40
	Ibón de Acherito	172.80	89.50	1078.90	60.80	46.42	99.78	68.00	44.00	111.00	88.85	9.62	0.51	1.03
	Bassies	92.00	61.90	149.50	28.73	21.65	41.06	72.00	53.00	100.00	78.02	16.10	3.57	2.31

Abbreviation: CI, confidence interval. Relationship indicates the percentage of individual relatedness within each population and cluster.
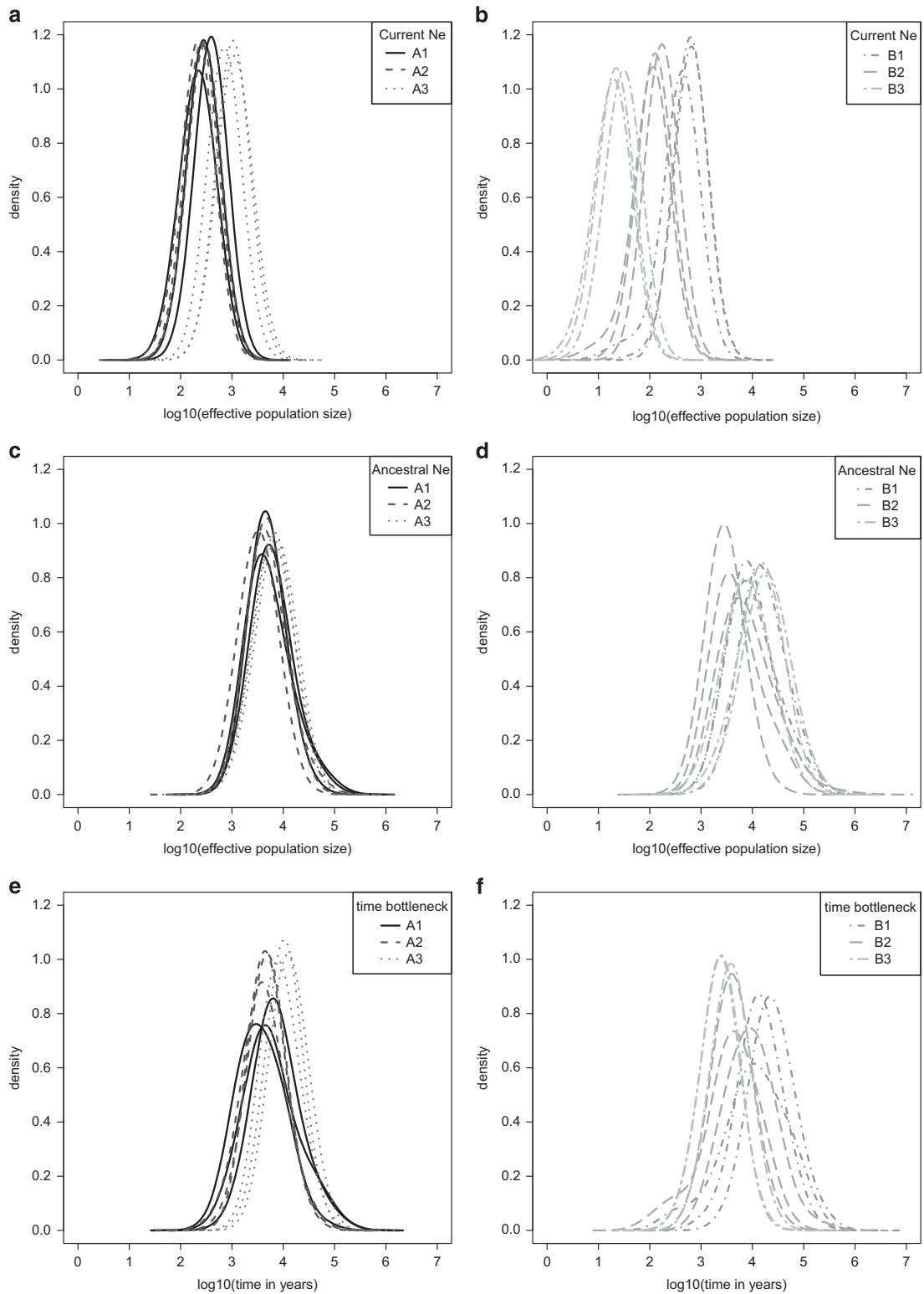
Figure 3 Demographic analysis using MSVAR. For populations A1–A3 and B1–B3, their current effective population size is shown (**a**, **b**), the ancestral effective population size before the bottleneck (**c**, **d**) and the time of the bottleneck (**e** and **f**). The posterior distributions of each parameter for the A populations are shown in shades of dark gray, and for the B populations in shades of light gray. For each population, three distributions are shown for each parameter, as each population was analyzed using three alternative priors. The x axis is in log(10) scale, for example, 2 represents 100, and 4 represents 10 000.

panmictic population. Thus, consistent with our results of the gene flow and migration rate analysis, this suggests that random mating among individuals from distinct populations within sectors does not occur (Supplementary Figure S4). These results are concordant with the lack of migration among populations indicated by other analyses. Similar values of relatedness were detected for the reference *C. asper* populations (Table 3) and r_{qg} (Ibón de Perramó, $r_{qg} = -0.031$; Barranco de Valdragás, $r_{qg} = -0.026$; Ibón de Acherito, $r_{qg} = -0.046$; Bassies, $r_{qg} = -0.005$).

DISCUSSION

Habitat fragmentation is typically expected to lead to a decrease in genetic diversity owing to stochastic processes (for example, genetic drift), which have a stronger effect in smaller populations (Leimu et al., 2006). Therefore, species restricted to small geographic areas may experience a high risk of extinction if populations become fragmented and isolated from each other. However, there is also some evidence to suggest that habitat fragmentation can give rise to neutral or even positive effects (Templeton et al., 1990; Fahrig, 2003). Here we show that extreme subdivision in an amphibian species (*C. arnoldi*) has not negatively affected certain genetic parameters that are supposed to be important indicators for fitness, such as genetic diversity and inbreeding coefficients, when compared with non-fragmented populations of its sister species (*C. asper*).

Phylogenetic divergence

The estimated divergence between the two species of *Calotriton* confirms previous dating analyses (Carranza and Amat, 2005) and indicates that these species split approximately 1.5 Mya, during the Pleistocene epoch. Speciation within *Calotriton* may have been initiated by a geographical barrier or could have resulted from climatic fluctuations during the Pleistocene. Following the challenging climatic conditions of the last glacial maximum, the high dispersal capabilities of the Pyrenean brook newt help to explain its rapid dispersion through the Pyrenean axial chain and the Prepyrenees; this allowed the connection of populations and subsequent genetic homogenization as a consequence of gene flow (Valbuena-Ureña et al., 2013). In contrast to *C. asper*, a juvenile dispersal phase is absent in *C. arnoldi*, therefore hindering its capacity for colonization. This species probably found refuge in the Montseny massif, being unable to colonize areas beyond the Montseny mountain. The differentiation into two sectors seems to be relatively ancient (~180 000 years ago), coinciding with the Riss glaciation (300 000–130 000 ya). This glaciation is characterized by a significant temperature drop and dry climate, which may have decreased the water flow of the Tordera River, causing the extinction of intermediate populations between the current populated sites. Such a scenario is further corroborated by differences in morphology, as well as mitochondrial and nuclear coding genes between the sectors (Valbuena-Ureña et al., 2013), suggesting that the fragmentation into subpopulations is not a recent event driven by anthropogenic activities but rather by natural processes. As our results indicate (Figure 3), the effective population sizes of *C. arnoldi* have remained low as the species split.

Neutral genetic diversity is shaped by the balance of evolutionary forces (mutation, genetic drift and migration) over contemporary and historical timescales (Dalongeville et al., 2016). Although genetic diversity greatly depends on the age of the population concerned, the *Calotriton* species have diverged relatively recently and both have experienced similar historical climatic events (Valbuena-Ureña et al., 2013); therefore, the comparison between them is appropriate.

Patterns of genetic diversity

Owing to lower vagility, loss of genetic diversity in amphibians is likely to be greater than in many other taxa and is highly correlated with declines in population fitness and the diminishment of their adaptive potential (Allentoft and O'Brien, 2010). Overall, it appears that across its small and restricted distribution range, moderate levels of genetic diversity and high genetic differentiation among sites characterize the Montseny brook newt. The comparison of genetic variation between this species and its closely related and ecologically similar sister species *C. asper* indicates that, in general, *C. arnoldi* harbors similar levels of genetic diversity despite its far smaller distribution range. The differences detected in the western sector are mostly due to population B3. We can state that population B3 differs from all other populations of the same species, not only from *C. asper* populations. Therefore, we believe that these data do not support a general pattern in which the western sector significantly differs from the four randomly selected *C. asper* populations. It seems that population B3 is an example of a fragile population in terms of low genetic diversity and low effectives rather than a situation in which the entire sector suffers the effects of habitat fragmentation. Moreover, neither species show signs of inbreeding. *Calotriton asper* shows similar or slightly higher values of N_e than *C. arnoldi*. Although *C. asper* has a juvenile dispersal phase that may reduce risk of inbreeding, *C. arnoldi* is exclusively aquatic with no dispersal phase and may therefore have developed other mechanisms to counteract the genetic consequences of small populations sizes. It is surprising that populations of *C. arnoldi* display similar levels of genetic variation to *C. asper* despite their differences in range size, despite *C. arnoldi* effective population size, and the evidence of a past bottleneck. However, when comparing the expected heterozygosity of *C. arnoldi* to that of other salamanders and temperate amphibians, *C. arnoldi* is within the typical range (0.4–0.6; Chan and Zamudio, 2009 and references therein).

Our results clearly show that *C. arnoldi* populations are highly structured over short geographic distances, and that the species is differentiated into an eastern and a western sector (Figure 2). Interestingly, the eastern sector presents higher levels of genetic variability than the western sector both in terms of microsatellite loci and nuclear and mitochondrial DNA sequences (Valbuena-Ureña et al., 2013). The most likely explanation for this pattern is the larger effective population sizes of the eastern populations in comparison to the western ones.

That the two sectors are highly genetically differentiated, with no gene flow between them, is indicated by multiple lines of evidence, including: a large number of private alleles in each sector (75 and 38 in the eastern and western sectors, respectively); significantly different patterns of genetic variation between sectors (for example, AMOVA, allelic richness and the number of fixed alleles in each sector); outcome of the principal coordinate analysis; high F_{ST} values; unambiguous genetic assignment of individuals to their population of origin; and observed isolation by distance effect. As *C. arnoldi* is exclusively aquatic, dispersal can only occur along watercourses, therefore reducing dispersal capabilities with respect to similar species capable of terrestrial dispersal. This is reflected in the levels of genetic differentiation observed, which are notably higher than values typically found for amphibians that use both aquatic and terrestrial habitats (Spear et al., 2005). The sectors of *C. arnoldi* are effectively isolated by a 37 km long watercourse, whereas distance by land is only 6 km. The watercourse between the two sectors passes through long stretches of river that includes a low altitude (<600 m) section with high water temperatures and potential predators; it therefore constitutes an adverse environment for these aquatic newts and thus presents a strong

migration barrier. Accordingly, we can assume that, in this system, natural fragmentation has had a strong impact on observed and associated microevolutionary processes (see Templeton et al., 1990).

Although the strong population subdivision is clearly detected between sectors, the low dispersal capability of this species is also detected among populations within sectors. The significant F_{ST} values indicate that dispersal between populations is low, as confirmed by the differentiation of populations/clusters A3 and B3 from the other populations within their respective sectors. These results were consistent with the outcome of principal coordinate analysis and migration tests. Moreover, at a sector level, most populations showed significantly higher degrees of relatedness (r_{qg}) than expected if sectors were in panmixia. This pattern is expected when migration among populations is not sufficiently high to counteract the relatedness resulting from nonrandom mating among populations. This notable sector structuring could suggest high levels of relatedness and inbreeding of individuals within populations. However, this is not found, as non-relatedness values within populations remain high, and random mating within populations seems to occur (see discussion below).

Our results indicate that the overall genetic diversity of C. arnoldi has been maintained at relatively high levels across its small and fragmented distribution range. This species comprises of highly genetically differentiated populations that display moderate levels of genetic diversity. Therefore, both intrapopulation genetic diversity levels and the strong differentiation among them allow this species to retain enough genetic variation to persist despite the vulnerability inherent in its small distribution range.

The impact of natural fragmentation on C. arnoldi

It is broadly accepted that habitat fragmentation (either naturally occurring or human driven) will result in the subdivision of populations, and if migration of individuals is not possible, subpopulations will start to diverge genetically (Templeton et al., 1990; Frankham et al., 2010). However, Templeton et al. (1990) suggested that, despite the negative effects deriving from fragmentation, genetic variation is not completely lost but often presents as fixed differences between local populations. Although this aspect is relevant for species conservation, it is little considered at present, and pertinent case studies are lacking. In our view, the surprising results obtained herein, involving an endangered species affected by natural habitat fragmentation, provide an excellent study system to promote discussion on this overlooked aspect.

Both the census and effective population sizes in C. arnoldi rank it as a critically endangered species; current N_e values for all C. arnoldi populations are critically low (<50) and are consistent with the small census size (Carranza and Martínez-Solano, 2009). The divergence time estimated between C. asper and C. arnoldi, and between the two C. arnoldi sectors, indicates that these splits were not recent events (over 1 Mya for the former and over 100 Kya for the latter). Moreover, the N_e values estimated for C. arnoldi indicate a small population size throughout its comparable short evolutionary history of roughly 1.76 Mya. These facts support the hypothesis that, after the divergence of the two species, C. asper went through a rapid expansion phase, while C. arnoldi remained geographically restricted. The current distribution range of C. asper populations cover an area of roughly 20 000 km², whereas populations of C. arnoldi are restricted to an area of only 8 km²; such differences are expected to be reflected in genetic parameters, and yet they are not.

In general, different behavioral strategies can be assumed for animals to avoid inbreeding. The most easy and obvious strategy would be postnatal dispersal of individuals to reduce the probability of inbreeding, the second would be mating preferences for non-related individuals (Blouin and Blouin, 1988). Based on the high degree of genetic differentiation and the lack of migration between C. arnoldi subpopulations between the two sectors, we can basically exclude postnatal dispersal as a mechanism to avoid inbreeding. It is therefore likely that special mating preferences exist in C. arnoldi to minimize the effects of potential inbreeding. As we have not observed an excess of heterozygosity for analyzed microsatellite loci across sectors, we can further conclude that females—assuming that they are the choosing sex—might not only prefer to mate with unrelated males but also with related ones. Indeed, more recent empirical studies—in contrast to early ones—indicate that animals sometimes show no avoidance or even prefer to mate with relatives (see Szulkin et al., 2013). In crickets, for example, Tregenza and Wedell (2002) could show that females mating multiply with different males avoid low egg viability, which occurs when solely mating with non-related or only with related males, if they mate with both unrelated and related males. Sperm storage in special cloacal glands of the female (called spermathecae) in combination with multiple paternity is widespread and well documented for salamander and newt species of the suborder Salamandroidea, to which also Calotriton newts belong (Kühnel et al., 2010; Caspers et al., 2014). Although we are lacking direct evidence, it is very likely that females of C. arnoldi mate multiply with different males, resulting in multiple paternity. Assuming similar mating patterns as described above for crickets, C. arnoldi newts could avoid the negative consequences of inbreeding without displaying an excess of heterozygosity. Of course, at the moment it is completely unclear and needs further investigation by which behavioral mechanisms these newts can cope with small sizes of fragmented populations.

Overall, our results suggest that, in terms of maintaining genetic diversity, small effective population sizes do not necessarily pose a problem, as there may be other reproductive or behavioral mechanisms that can counteract the effects of genetic drift (Allentoft and O'Brien, 2010). In C. arnoldi, such mechanisms are likely to have prevented a substantial loss of alleles through the bottleneck experienced during the Holocene. Evidence suggests that life-history strategies can explain a considerable proportion of the variation in genetic diversity, as polymorphism levels are influenced by species biology (Romiguier et al., 2014; Fouquet et al., 2015; Paz et al., 2015; Dalongeville et al., 2016). Ecological factors affecting genetic diversity may include migration capability, morphological or physiological adaptations and reproductive strategy, among others.

Our results indicate that the overall genetic diversity of C. arnoldi has been maintained at a relatively high level despite its small and fragmented distribution range. Therefore, species fragmentation should not be regarded in this case as primarily detrimental. Populations of C. arnoldi do not show the low levels of intrapopulation genetic diversity or signs of inbreeding that are typical byproducts of habitat fragmentation. However, data regarding the potential effect of the fragmentation on a species potential to adapt to environmental changes, which again may be influenced by the life-history strategies, are currently lacking (Romiguier et al., 2014; Dalongeville et al., 2016). Further studies are needed to understand the relationship between genetic diversity, adaptive potential and life-history traits in this species.

Species characterized by independent and isolated populations may avoid species-level extinction, as local (population-level) extinctions, resulting from local demographic stochasticity or small-scale environmental catastrophes, are unlikely to be simultaneously experienced by all populations. Furthermore, in terms of infectious diseases (for example, parasite infections or bacterial pathogens such as those causing the 'Red-leg' syndrome; Daszak et al., 2003; Allentoft and

O'Brien, 2010), populations that are completely isolated might survive an outbreak as there is little or no exchange of individuals between single populations. Therefore, the persistence of some populations facilitates the survival of the species, and recolonization may occur over time, thus reversing extirpations.

Implications for conservation

Impacts of habitat fragmentation must be measured independently from effects of habitat loss or degradation. The effects of habitat loss may outweigh the effects of habitat fragmentation and can have important implications for conservation. Habitat loss is widely recognized to have strong and consistently negative effects on biodiversity, reducing species richness, population abundance and distribution and genetic diversity (Fahrig, 2003 and references therein).

In conservation biology, an N_e of 500 has been suggested as a minimum value for the long-term survival of a species, whereas N_e values <50 in isolated populations are of major concern (Frankham et al., 2014), as these populations have an increased probability of extinction resulting from genetic effects, such as inbreeding (Allendorf and Luikart, 2007) and stochastic environmental processes. Inbreeding is exacerbated by small N_e values. However, it is possible that populations with low N_e may survive over long periods of time as they can successfully and rapidly purge detrimental allelic variants, such a scenario has been proposed for other species (for example, Orozco-terWengel et al., 2015). However, the current low effective population sizes of C. arnoldi mean that habitat loss or degradation could rapidly drive these small populations to extinction. Stochastic factors can cause a disproportionately high mortality rate when species have very small distribution ranges. Moreover, the effects of habitat loss may be greater when the habitat is highly and rapidly fragmented. This implies that a key question concerning the conservation of a species is 'how much habitat is enough?'. The conservation of a vulnerable or endangered species requires estimating the minimum habitat required for persistence of the given species. In addition, many species require more than one kind of habitat within a life cycle. Therefore, landscape patterns that maintain the required habitat proportions should be conserved (Fahrig, 2003).

Studies that enhance understanding of genetic population structure and the gene flow between them contribute valuable information to management and conservation programs. The definition of appropriate conservation units are crucial for maintaining the distinct evolutionary lineages and the species' evolutionary potential (Frankham et al., 2010). In C. arnoldi, the evolutionary potential is not only manifested within the species as a whole but also within each sector. Conservation strategies should be adopted to ensure that the evolutionary potential and the genetic diversity within the distinct groups is not lost. Therefore, such strategies should focus on habitat preservation and restoration of each sector, with the aim of maintaining the strong population structure highlighted by this study.

DATA ARCHIVING

Data are deposited in the Dryad repository: http://dx.doi.org/10.5061/dryad.p4s53.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Akaike H (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F (eds). Second International Symposium on Information Theory. Akademiai Kiado: Budapest, Hungary, pp 267–281.

Allendorf FW, Luikart G (2007). Conservation and the Genetics of Populations. Blackwell: Oxford, UK.

Allentoft M, O'Brien J (2010). Global amphibian declines, loss of genetic diversity and fitness: a review. Diversity 2: 47–71.

Amat F, Carranza S, Valbuena-Ureña E, Carbonell F (2014). Saving the Montseny brook newt (Calotriton arnoldi) from extinction: an assessment of eight years of research and conservation. Froglog 22: 55–57.

Apodaca J, Rissler L, Godwin J (2012). Population structure and gene flow in a heavily disturbed habitat: implications for the management of the imperilled Red Hills salamander (Phaeognathus hubrichti). Conserv Genet 13: 913–923.

Arèvalo E, Davis SK, Sites JW (1994). Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the Sceloporus grammicus complex (Phrynosomatidae) in Central Mexico. Syst Biol 43: 387–418.

Belkhir K, Chikhi L, Raufaste N, Bonhomme F (2004). GENETIX 4.05, Logiciel Sous Windows TM Pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II: Montpellier, France.

Blank L, Sinai I, Bar-David S, Peleg N, Segev O, Sadeh A et al. (2013). Genetic population structure of the endangered fire salamander (Salamandra infraimmaculata) at the southernmost extreme of its distribution. Anim Conserv 16: 412–421.

Blouin SF, Blouin M (1988). Inbreeding avoidance behaviours. Trends Ecol Evol 3: 230–233.

Blouin M, Phillipsen I, Monsen K (2010). Population structure and conservation genetics of the Oregon spotted frog Rana pretiosa. Conserv Genet 11: 2179–2194.

Brooks SP, Gelman A (1998). General methods for monitoring convergence of iterative simulations. J Comp Graph Stat 7: 434–455.

Bulut Z, McCormick C, Gopurenko D, Williams R, Bos D, DeWoody JA (2009). Microsatellite mutation rates in the eastern tiger salamander (Ambystoma tigrinum tigrinum) differ 10-fold across loci. Genetica 136: 501–504.

Carranza S, Amat F (2005). Taxonomy, biogeography and evolution of Euproctus (Amphibia: Salamandridae), with the resurrection of the genus Calotriton and the description of a new endemic species from the Iberian Peninsula. Zool J Linn Soc 145: 555–582.

Carranza S, Arnold EN, Mateo JA, López-Jurado LF (2000). Long-distance colonization and radiation in gekkonid lizards, Tarentola (Reptilia: Gekkonidae), revealed by mitochondrial DNA sequences. Proc R Soc Lond Ser B Biol Sci 267: 637–649.

Carranza S, Martínez-Solano I (2009). Calotriton arnoldi. IUCN Red List of Threatened Species. e.T136131A4246722. Available at: http://dx.doi.org/10.2305/IUCN.UK.2009.RLTS.T136131A4246722.en.

Caspers BA, Krause ET, Hendrix R, Kopp M, Rupp O, Rosentreter K et al. (2014). The more the better – polyandry and genetic similarity are positively linked to reproductive success in a natural population of terrestrial salamanders (Salamandra salamandra). Mol Ecol 23: 239–250.

Chan LM, Zamudio KR (2009). Population differentiation of temperate amphibians in unpredictable environments. Mol Ecol 18: 3185–3200.

Dalongeville A, Andrello M, Mouillot D, Albouy C, Manel S (2016). Ecological traits shape genetic diversity patterns across the Mediterranean Sea: a quantitative review on fishes. J Biogeogr 43: 845–857.

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

Daszak P, Cunningham AA, Hyatt AD (2003). Infectious disease and amphibian population declines. Divers Distrib 9: 141–150.

Drechsler A, Geller D, Freund K, Schmeller DS, Künzel S, Rupp O et al. (2013). What remains from a 454 run: estimation of success rates of microsatellite loci development in selected newt species (Calotriton asper, Lissotriton helveticus, and Triturus cristatus) and comparison with Illumina-based approaches. Ecol Evol 3: 3947–3957.

Drummond A, Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

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

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

Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564–567.

Fahrig L (2003). Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34: 487–515.

Faubet P, Gaggiotti OE (2008). A new Bayesian method to identify the environmental factors that influence recent migration. Genetics 178: 1491–1504.

Fouquet A, Courtois EA, Baudain D, Lima JD, Souza SM, Noonan BP *et al.* (2015). The trans-riverine genetic structure of 28 Amazonian frog species is dependent on life history. *J Trop Ecol* **31**: 361–373.

Frankham R, Ballou JD, Briscoe DA (2010). *Introduction to Conservation Genetics*. 2nd edn.Cambridge University Press: Cambridge, UK.

Frankham R, Bradshaw CJ, Brook BW (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol Conserv* **170**: 56–63.

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010). Calculations of population differentiation based on GST and D: forget GST but not all of statistics!. *Mol Ecol* **19**: 3845–3852.

Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* **86**: 485–486.

Guo SW, Thompson EA (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361–372.

Hauswaldt JS, Angelini C, Gehara M, Benavides E, Polok A, Steinfartz S (2014). From species divergence to population structure: a multimarker approach on the most basal lineage of Salamandridae, the spectacled salamanders (genus *Salamandrina*) from Italy. *Mol Phylogenet Evol* **70**: 1–12.

Hill WG (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet Res* **38**: 209–216.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002). Potential applications and pitfalls of bayesian inference of phylogeny. *Syst Biol* **51**: 673–688.

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

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

Jost L (2008). GST and its relatives do not measure differentiation. *Mol Ecol* **17**: 4015–4026.

Kalinowski ST (2005). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* **5**: 187–189.

Kalinowski ST, Wagner AP, Taper ML (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* **6**: 576–579.

Katoh K, Standley DM (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Kühnel S, Reinhard S, Kupfer A (2010). Evolutionary reproductive morphology of amphibians: an overview. *Bonn Zool Bull* **57**: 119–126.

Leimu R, Mutikainen PIA, Koricheva J, Fischer M (2006). How general are positive relationships between plant population size, fitness and genetic variation? *J Ecol* **94**: 942–952.

Lewis PO, Zaykin D (2000). *Genetic Data Analysis: Computer Program for the Analysis of Allelic Data, Version 1.0. (d15)*. University of Connecticut: Storrs, CT, USA.

MacLeod A, Rodríguez A, Vences M, Orozco-terWengel P, García C, Trillmich F *et al.* (2015). Hybridization masks speciation in the evolutionary history of the Galápagos marine iguana. *Proc R Soc B* **282**: 20150425.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**: 209–220.

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

Milá B, Carranza S, Guillaume O, Clobert J (2010). Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Mol Ecol* **19**: 108–120.

Monsen KJ, Blouin MS (2004). Extreme isolation by distance in a montane frog *Rana cascadae*. *Conserv Genet* **5**: 827–835.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009). Divergent selection and heterogeneous genomic divergence. *Mol Ecol* **18**: 375–402.

Orozco-terWengel P, Barbato M, Nicolazzi E, Biscarini F, Milanesi M, Davies W *et al.* (2015). Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Front Genet* **6**: 191.

Paetkau D, Slade R, Burden M, Estoup A (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* **13**: 55–65.

Paz A, Ibáñez R, Lips KR, Crawford AJ (2015). Testing the role of ecology and life history in structuring genetic variation across a landscape: a trait-based phylogeographic approach. *Mol Ecol* **24**: 3723–3737.

Peakall R, Smouse PE (2006). GENELEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.

Piry S, Luikart G, Cornuet JM (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* **90**: 502–503.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J Hered* **95**: 536–539.

Plummer M, Best N, Cowles K, Vines K (2006). Coda: convergence diagnosis and output analysis for MCMC. *R News* **6**: 7–11.

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

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

Rambaut A, Drummond A (2007). *Tracer v1.5 < http://beast.bio.ed.ac.uk/Tracer >*.

Recuero E, Canestrelli D, Vörös J, Szabó K, Poyarkov NA, Arntzen JW *et al.* (2012). Multilocus species tree analyses resolve the radiation of the widespread *Bufo bufo* species group (Anura, Bufonidae). *Mol Phylogenet Evol* **62**: 71–86.

Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A *et al.* (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **515**: 261–263.

Rosenberg NA (2004). DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* **4**: 137–138.

Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219–1228.

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

Savage WK, Fremier AK, Bradley Shaffer H (2010). Landscape genetics of alpine Sierra Nevada salamanders reveal extreme population subdivision in space and time. *Mol Ecol* **19**: 3301–3314.

Slatkin M (1994)Gene flow and population structureIn:LA Real (ed). *Ecological Genetics*. Princeton University Press: Princeton, New Jersey, USA. pp 3–17.

Šmíd J, Carranza S, Kratochvíl L, Gvoždík V, Nasher AK, Moravec J (2013). Out of Arabia: a complex biogeographic history of multiple vicariance and dispersal events in the gecko genus *Hemidactylus* (Reptilia: Gekkonidae). *PLoS One* **8**: e64018.

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

Spear SF, Peterson CR, Matocq MD, Storfer A (2005). Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Mol Ecol* **14**: 2553–2564.

Steinfartz S, Weitere M, Tautz D (2007). Tracing the first step to speciation: ecological and genetic differentiation of a salamander population in a small forest. *Mol Ecol* **16**: 4550–4561.

Storfer A, Mech S, Reudink M, Lew K (2014). Inbreeding and strong population subdivision in an endangered salamander. *Conserv Genet* **15**: 137–151.

Storz JF, Beaumont MA (2002). Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution* **56**: 154–166.

Sunny A, Monroy-Vilchis O, Fajardo V, Aguilera-Reyes U (2014). Genetic diversity and structure of an endemic and critically endangered stream river salamander (Caudata: *Ambystoma leorae*) in Mexico. *Conserv Genet* **15**: 49–59.

Szulkin M, Stopher KV, Pemberton JM, Reid JM (2013). Inbreeding avoidance, tolerance, or preference in animals? *Trends Ecol Evol* **28**: 205–211.

Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008). COMPUTER PROGRAMS: onesamp: a program to estimate effective population size using approximate Bayesian computation. *Mol Ecol Resour* **8**: 299–301.

Templeton AR, Shaw K, Routman E, Davis SK (1990). The genetic consequences of habitat fragmentation. *Ann Missouri Bot Gard* **77**: 13–27.

Tregenza T, Wedell N (2002). Polyandrous females avoid costs of inbreeding. *Nature* **415**: 71–73.

Valbuena-Ureña E, Amat F, Carranza S (2013). Integrative phylogeography of *Calotriton* newts (Amphibia, Salamandridae), with special remarks on the conservation of the endangered Montseny brook newt (*Calotriton arnoldi*). *PLoS One* **8**: e62542.

Valbuena-Ureña E, Steinfartz S, Carranza S (2014). Characterization of microsatellite loci markers for the critically endangered Montseny brook newt (*Calotriton arnoldi*). *Conserv Genet Resour* **6**: 263–265.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.

Veith M, Kosuch J, Vences M (2003). Climatic oscillations triggered post-Messinian speciation of Western Palearctic brown frogs (Amphibia, Ranidae). *Mol Phylogenet Evol* **26**: 310–327.

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

Waples RS, Do C (2008). LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* **8**: 753–756.

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