

Received Date : 24-Feb-2014

Revised Date : 05-Nov-2014

Accepted Date : 14-Nov-2014

Article type : Original Article

Congruent signals of population history but radically different patterns of genetic diversity between mitochondrial and nuclear markers in a mountain lizard.

Running title: Demographic inference from genetic data in an endemic lizard

Anne-Laure Ferchaud^{1,2*}, Rémy Eudeline³, Véronique Arnal², Marc Cheylan², Gilles Pottier⁴, Raphaël Leblois^{5,6} and Pierre-André Crochet³

¹Department of Bioscience, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C

²EPHE-UMR5175 CEFE, Centre d'Écologie Fonctionnelle et Évolutive, 1919 route de Mende, 34293 Montpellier cedex 5, France

³ CNRS-UMR5175 CEFE, Centre d'Écologie Fonctionnelle et Évolutive, 1919 route de Mende, 34293 Montpellier cedex 5, France

⁴ Nature Midi-Pyrénées · Comité local 65, 21 rue des Thermes, 65 200 Bagnères-de-Bigorre, France

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.13011

This article is protected by copyright. All rights reserved.

⁵ INRA-UMR1062 CBGP, Centre de Biologie pour la Gestion des populations, F-34988
Montferrier-sur-Lez, France

⁶ IBC, Institut de Biologie Computationnelle, Montpellier, France

***Correspondence:** Anne-Laure Ferchaud, Department of Bioscience, Aarhus University, Ny
Munkegade 116, building 1535, room 221, 8000 Aarhus C, Denmark.

E-mail: annelaureferchaud@gmail.com

ABSTRACT

Historical factors, current population size, population connectivity and selective processes at linked loci contribute to shaping contemporary patterns of neutral genetic diversity. It is now widely acknowledged that nuclear and mitochondrial markers react differently to current demography as well as to past history, so the use of both types of markers is often advocated to gain insight on both historical and contemporary processes. We used 12 microsatellite loci genotyped in 13 populations of a mountain lizard (*Iberolacerta bonnali*) to test if the historical scenario favoured by a previous mitochondrial study was also supported by nuclear markers and thereby evaluated the consequences of post-glacial range movements on nuclear diversity. Congruent signals of recent history were revealed by nuclear and mitochondrial markers using an Approximate Bayesian Computation approach but contemporary patterns of mtDNA and nuclear DNA diversity were radically different. Although dispersal in this species is probably highly restricted at all spatial scales, colonisation abilities have been historically good, suggesting capability for reestablishment of locally extinct populations except in fully disconnected habitats.

INTRODUCTION

Genetic diversity is widely regarded as crucial for the persistence of small populations and for their potential to adapt to new environments (Frankham 2002, 2003, 2005), but see Jamieson & Allendorf (2012). Since the diversity of genes with phenotypic effects is still difficult to capture, neutral genetic diversity is the usual surrogate used in conservation genetics studies. Such contemporary patterns of neutral genetic diversity result from the interplay of historical factors (colonisation scenarios, demographic fluctuations), current population size and connectivity, and selective processes at linked loci (Frankham 2012).

Neutral molecular markers are also widely used to infer past processes of populations history. One of the most successful applications of the use of molecular markers to reconstruct past history has been the identification of glacial refugia and post-glacial recolonisation routes in northern hemisphere organisms. In Europe, these studies have identified a few general patterns (with the usual exceptions and variations): persistence in southern or central European refugia followed by rapid expansions for many temperate species; a complex history of long-term persistence in multiple “refugia within refugia” for species with Mediterranean contemporary distribution; range contraction in localised northern refugia for boreal species; and short-distance, down slope migration for many mountain species (reviewed in Hewitt (2004); Schmitt (2007, 2009); Stewart *et al.* (2010); Schmitt & Varga (2012)).

A previous study (Mouret *et al.* 2011) used mitochondrial sequences from the non-coding control region and coding cytochrome *b* to examine genetic population structure and phylogeography of the Pyrenean Mountains endemic lizard Pyrenean Rock Lizard *Iberolacerta bonnali*. Currently, the populations of the species only occur at elevations

between 1600 and 3300 m.a.s.l. (Arribas 2000) and their distribution is fragmented at two spatial scales. First, the species inhabits several distinct, and sometimes disconnected, massifs. Second, within each massif, populations are highly fragmented due to the patchy occurrence of their rocky habitats and to intervening areas of low or high elevation unsuitable for the species (Arribas 2000; Pottier *et al.* 2013).

The mitochondrial data analysed by Mouret *et al.* (2011) suggested a history of recent range fragmentation after the Last Glacial Maximum (LGM, which ended roughly in the Pyrenees 15,000 to 10,000 years ago, (Arribas 2004)) when lower elevations became unsuitable and the species retreated to high altitude habitats in the main Pyrenean massifs following climate warming. A large part of the current distribution, corresponding to areas that were unsuitable during the LGM, was inferred to have been recolonised from a small number of source populations. This complex history resulted in a counter-intuitive distribution of mitochondrial diversity, where small isolated populations retained a large amount of genetic diversity, probably because they were located close to glacial refugia, while in the largest, continuous part of the distribution along the main Pyrenean ridge most populations are fixed for the same mtDNA haplotype (Mouret *et al.* 2011). The result is that populations can now be divided into two groups based on mitochondrial diversity. A first group of populations (in white in Figure 1) inhabits mainly the northern peripheral massifs (Néouvielle massif and an isolated mountain further north) plus the Lustou massif in the east of the range; these are supposed to have resulted from the fragmentation of the large refugial populations and all harbour a significant amount of mtDNA diversity (Mouret *et al.* 2011). A second group of populations (in grey in Figure 1) inhabits the western-central part of the Pyrenean chain and the Sauvergarde massif (SAU) in the east; these are hypothesised to represent the part of the range recolonised after the LGM, and all individuals carry the same

control region haplotype (except one single individual in the ARR population). While it was hard to completely exclude that a selective sweep explains this loss of diversity, drift in a spatially expanding population seemed a better explanation in this case (Mouret *et al.* 2011). Whatever the mechanisms involved, current levels of intra-population mitochondrial diversity are better explained by postglacial population history than by current habitat fragmentation. The effects of habitat fragmentation at local scale were also evident as some close by populations exhibited highly different genetic composition, indicating virtually no female-mediated gene flow.

In most animal species nuclear and mitochondrial markers differ in effective size, presence or absence of recombination and biparental *versus* maternal inheritance. They should thus react differently to current demography as well as to past history affecting populations. For these reasons, and because of the large variance between genealogies of individual loci introduced by coalescence processes, the use of both types of markers is now widely advocated in phylogeographic studies (e.g. Edwards & Bensch (2009) and Nielsen & Beaumont (2009)). Even if the amount of discrepancy to be expected between mitochondrial-only studies and studies based on nuclear multilocus datasets remains somewhat debated (Barrowclough & Zink 2009; Zink & Barrowclough 2008), it cannot be taken for granted that mtDNA will provide a reliable assessment of neutral genetic diversity, and thus of demographic and historical inferences.

In this paper we use a dataset of 12 microsatellite loci genotyped in a large number of individuals (typically 28 to 31 per population, see Table 1) from a subsample of the populations used by Mouret *et al.* (2011) to examine (1) whether the historical scenario favoured by the mitochondrial dataset is also supported by nuclear markers, notably using an

Approximate Bayesian Computation approach; and (2) whether the geographical pattern of mitochondrial diversity is a good predictor for neutral nuclear diversity, assuming that microsatellites are better proxy of neutral genomic diversity than mtDNA. Particularly, we want to evaluate the consequences of post-glacial movements on nuclear diversity and compare it with its consequences on mtDNA diversity. This will be done firstly by correlating measures of mtDNA and nuclear DNA diversity. We will also examine the following hypothesis: assuming that refugia are home to high genetic diversity (Hewitt 2000; Petit *et al.* 2003) and that refugia in the Pyrenean Mountains were at low elevation, is there a loss of diversity as elevation increases? Finally, we intend to examine (3) whether the restricted gene flow among neighbouring populations suggested by the mitochondrial population structure was also evident when using nuclear markers with biparental inheritance.

MATERIALS AND METHODS

Tissue sampling and populations information

Tail tips from 372 *Iberolacerta bonnali* were collected between July 2000 and September 2000 and between July 2001 and September 2001 in 13 localities covering the entire French distribution of the species (Fig. 1). In fact, our sampling covers most of the range of the species (see Fig. 1) except for two isolated massifs in Spain, forming the south-eastern limit of the distribution (Maladeta and Aigüestortes, dark grey part of the distribution in Fig. 1). In these two massifs, populations harbour highly divergent mtDNA lineages when compared with the other populations, indicating that they constitute distinct ESUs and suggest they haven't shared any recent common ancestry with the other Pyrenean populations (Mouret *et al.* 2011). They thus do not share the same post-glacial history as most of the other populations and have been excluded from the present study. Tissue samples were placed in 95% ethanol while in the field and kept at room temperature until laboratory analysis. Sample

Accepted Article

size for each locality is around thirty individuals, except for the SAU population where $n=14$ (Table 1). Note that two localities (BIG and ARD) analysed by Mouret *et al.* (2011) were not included in this study because of low sample sizes (4 and 6 individuals, respectively).

A first set of populations, spread along the central parts of the main Pyrenean chain, is localised in the most continuous part of the species' range and all these populations are potentially connected by suitable habitats, as can be judged by occurrence of rocky habitats above the tree-line. These populations are, from west to east, Arriel (ARR), Géougue d'Arre (GEO), Fache (FAC), Vignemale (VIG) Munia (MUN), Lustou (LUS) and Sauvergarde (SAU) (Fig. 1). There is also potential habitat connectivity between these populations and the Néouvielle massif, including the populations Néouvielle (NEO), Long (LON), Estaragne (EST) and Piau Engaly (PIA). Lastly, the Montaigu population (MON) in the north and the Midi d'Ossau population (OSS) in the west are populations separated from the rest of the distribution by unsuitable habitats, and are thus currently isolated (see Mouret *et al.* 2011 for further description of populations and habitats and for geographic coordinates of each population).

DNA extraction and microsatellite analysis

DNA from 146 samples had been extracted in 2004 for mitochondrial DNA analysis (Mouret *et al.* 2011) and frozen at -20°C . The remaining 226 samples were DNA extracted by complete digestion of a piece of muscle using the DNeasy® Blood and Tissue kit following the manufacturer's recommended procedures (Spin-Column Protocol, QIAGEN).

Twelve microsatellite loci isolated and characterised in related species (*Lacerta agilis*, (Gullberg *et al.* 1997): La6; *Iberolacerta cyreni*, (Bloor 2006): Icy2, Icy4 and Icy5;

Iberolacerta monticola, (Remon *et al.* 2008): A8, B107, B114, B135, C5, C118, D101 and D115) were optimised in our species and were found to amplify reliably (see Table S1). Genotyping was carried out at the MBB Platform of the CeMEB LabEx "*Mediterranean Center for Environment and Biodiversity*" in Montpellier according to the instructions given in the papers mentioned above. PCR products were sized on an ABI Prism® 3130XL sixteen capillaries and analysed using GeneMapper v4.0 (Applied Biosystems®). Chromatograms were checked by eye by two independent viewers using GeneMapper v4.0. Finally, Micro-Checker v2.2.3 (Van Oosterhout *et al.* 2004) was used to test the presence of stuttering, large allele drop out and null allele's artefacts.

Genetic diversity

We tested for departure from the Hardy-Weinberg equilibrium and linkage disequilibrium among pairs of loci for each population using exact tests from Genepop v4.0.10 (Rousset 2008). Genetic diversity indices (observed heterozygosity H_o , expected heterozygosity H_e , and allelic richness A_r) as well as the inbreeding coefficient, F_{is} , were estimated using FSTAT v.2.9.3.2 (Goudet 1995). Correlation between genetic diversity and altitude was tested using a regression analysis implemented in a General Linear Model (R v.2.15.2, R development core team, 2010). Several indices of genetic diversity (H_e , F_{is} and A_r) were used for this analysis. Finally, we tested whether nuclear genetic diversity (H_e and A_r) was correlated to nucleotide mitochondrial diversity (π) with a Pearson correlation test in R v.2.15.2.

Population structure and gene flow

We first used the clustering method developed by Pritchard *et al.* (2000) and implemented in STRUCTURE v2.3 to examine population structure in our sample. We used an

admixture model, with correlated allele frequencies, without prior population information. This model assumes that the genome of each individual is a mixture of genes originating from K unknown ancestral populations. We ran 10 independent analyses for each K value ranging from 2 to 15, with 2×10^5 iterations and a burn-in period of 50%. We looked at the influence of the number of clusters, K , using different criteria: log-likelihood, the congruence between runs for the different K values and the change in the second order of likelihood, ΔK , proposed by Evanno *et al.* (2005).

We used F_{ST} as computed with Genepop V4.0.10 (Rousset 2008) to estimate pairwise population differentiation. Significance of all F_{ST} values was assessed by the exact tests for population differentiation in Genepop. The Neighbor and seqboot programs as implemented in Phylip v3.69 (Felsenstein 2004) were used respectively to infer a Neighbour-Joining tree based on the Reynold's distance ($D = -\ln(1-F_{ST})$, Reynolds *et al.* (1983) and to assess robustness of nodes using 1,000 bootstrap replications over loci. We tested for isolation by distance patterns by regressing $F_{ST}/1-F_{ST}$ between populations over the logarithm of geographical distances between populations (for straight-line distance, two dimensions) or geographical distance (for distance along favourable habitat, which is nearly one-dimensional in our case) as recommended by (Rousset 1997) and Rousset (2000). We used both straight-line distances between populations and a more realistic measure of distance corresponding to the minimum distance between populations through favourable habitats only (here areas above 1800m asl). Significance of the correlation between genetic and geographic distances was tested using Mantel tests with 30,000 permutations. We also computed standardised values of F_{ST} as (observed F_{ST})/(maximal possible F_{ST} value given the level of intra-population diversity) following the principles of Hedrick (2005) and the computation method of Meirmans (2006) using RecodeData v.0.1 (available at

<http://recodedata.sharewarejunction.com/>). Our results did not differ when using raw or standardised F_{ST} values and we thus based all analyses on raw F_{ST} values, although standardised values have been used for illustrative purpose.

We also computed pairwise R_{ST} , an index of differentiation based on differences in allele frequencies and allele size (Slatkin 1995) estimated between all pairs of sampled populations, as in Michalakis & Excoffier (1996). We compared observed R_{ST} values with pR_{ST} values obtained after 10,000 allele size permutations to test if the observed population structure is mainly due to recurrent migration or to population divergence. R_{ST} is expected to be significantly higher than the mean permuted pR_{ST} values under a phylogeographical pattern, *i.e.*, populations have diverged for a long time and exchanged migrants at a low rate compared to the mutation rate (Hardy *et al.* 2003). Conversely, R_{ST} and pR_{ST} should not be significantly different if mutation is not the main cause of differentiation, *i.e.* if gene flow was large relative to mutation rate. These analyses were performed with SPAGeDi version 1.2 (Hardy & Vekemans 2002).

Population history

In order to reconstruct the unknown history of divergence among these thirteen populations of *Iberolacerta bonnali*, we performed Approximated Bayesian Computations (ABC) on our microsatellite dataset with the software DIYABC v.2.0 (Cornuet *et al.* 2014). DIYABC allows for the comparison of different historical scenarios involving population divergence, admixture and population size changes, and then the inference of demographic and historical parameters under the best-supported scenario.

Since we wanted to test if the phylogeographic scenario proposed by Mouret *et al.* (2011) on the basis of mtDNA data was supported by nuclear markers, we compared two types of scenarios of historical divergence (Fig. 2) for our twelve samples. In the first type of scenario (“null hypothesis”, scenarios 1 & 2), all sampled populations diverged simultaneously. In the second type (scenarios 3 & 4), populations have diverged in two successive events, as suggested by mtDNA data (Mouret *et al.* 2011): “peripheral” populations with high mitochondrial diversity (MON, LUS, NEO, LON and EST, in white in Fig. 1) diverged first, while “central” populations with impoverished mitochondrial diversity (in grey in Fig. 1) diverged later when the central parts of the Pyrenean chain were recolonised during postglacial expansion.

We implemented each type of scenario twice, (i) once with constant population sizes and (ii) once allowing for past variation in population sizes, *i.e.* sudden expansions or contractions in all ancestral populations. In total four scenarios were thus implemented and are represented in Fig. 2. For scenarios 2 and 4, allowing for past variation in population sizes, the twelve populations had potentially different contemporary N_e value (**N1** to **N12**) and different past N_e value (**N14** to **N25**). **N13** represents the N_e of the ancestral population (see Fig. 2). In scenarios 1 and 2 all populations diverged simultaneously at time **t2**, whereas in scenario 3 & 4 the peripheral populations diverged first at **t2** and the central populations diverged later at **t1**. To better discriminate the different scenarios, which can be undistinguishable if **t1=t2**, we set a condition on the divergence parameters for the simulations under scenario 3 & 4: **t1** must to be smaller than **t2**. All these settings are summarised in Table 3.

In view of the limited dispersal capability, the geographic isolation of the eleven sampled populations and the strong genetic structure observed (see *Analyses of population structure and gene flow* in Results section), each population was treated as having evolved independently and no admixture rate was allowed between populations. In addition, DIYABC only allows for single events of admixture but no recurrent gene flow, precluding the analysis of scenarios implementing biologically realistic gene flow patterns. The potential effects of gene flow on our results are examined in the Discussion section.

Given their geographic proximity (3 km, probably close to the dispersal capacities of the species), the lack of ecological barriers between them and their low level of genetic differentiation ($F_{ST} = 0.07$, see Table 2), we decided to group together populations LON and EST in one large population named 'LES' because the assumptions of negligible gene flow (see below) seemed unrealistic. All other populations are either much further away or much more strongly differentiated.

The ABC analyses are based on the simulation of $5 \cdot 10^6$ genetic data sets under the four scenarios described above. Similarity between the simulated data sets and the real data set is based on the following summary statistics:

- To assess within-population genetic diversity, we used the mean over all loci of the following statistics for each sampled population: (1) number of alleles; (2) gene diversity (Nei, 1987); (3) allele size variance; (4) MGW-ratio index (Excoffier *et al.* 2005; Garza & Williamson 2001).

- To assess between-population genetic structure, we used the mean of the following statistics computed over all loci and for all pairs of sampled populations: (1) F_{ST} ; (2) number of alleles in two samples; (3) gene diversity (expected heterozygosity) in two samples; (4) allele size variance in two samples; (5) index of classification in two samples (Pascual *et al.* 2007; Rannala & Mountain 1997) and (6) shared allele distance between two samples (Chakraborty & Jin 1993). A total of 429 summary statistics was thus used to select a scenario and to infer the parameter values under the best-supported scenario. For details about the computation of each statistic, see the DIYABC manual (<http://www1.montpellier.inra.fr/CBGP/diyabc/download/file.php?id=197>).

After a few preliminary runs, analysed using the 'prior checking' option (see the DIYABC manual), the prior distributions for all N_e and divergence time parameters were adjusted step by step and finally set up as mentioned in Table 3. Log uniform distributions were used for all parameters because the size of the sample and the limited number of loci used allows only rough estimation of all parameters (i.e. precision is only about the order of magnitude).

For all preliminary and final runs, we evaluated each analysis using a Bayesian equivalent of goodness-of-fit of the selected scenario, using the "model checking" option of DIYABC (Cornuet *et al.* 2014; Cornuet *et al.* 2010), see DIYABC manual section 2.10. This option allows for the evaluation of to what extent the selected scenario and associated posterior distributions are corroborated by the observed data. Briefly, if a model-posterior combination fits correctly the observed data, then data simulated under this combination with parameters drawn from posterior distributions should be close to the observed data. In order for the model fit to be considered good, the observed statistics had to fall within the

distributions of simulated statistics. We simulated 10,000 data sets from the posterior distribution of parameters obtained under all scenarios to estimate such distributions. Principal Component Analysis (PCA) applied on summary statistics were also used as a mean to visualise the fit between simulated and observed datasets.

The four scenarios were compared using the logistic regression approach, and parameter estimation was performed for the scenario with the highest posterior probability only.

RESULTS

Patterns of nuclear genetic diversity

Table 1 summarises the genetic variability of microsatellite loci observed in our samples. No stuttering or large allele drop out and no null allele artefacts have been detected with the MICRO-CHECKER software (results not shown). Genepop did not show any significant linkage disequilibrium between pairs of loci and only one population (NEO) showed a significant heterozygote deficit with a small positive F_{IS} value ($p = 0.02$, $F_{IS} = 0.086$ Table 1) – note however that this is no longer significant after correction for multiple testing. Locus-by-locus tests did not show any significant Hardy-Weinberg disequilibrium (data not shown).

The geographic repartition of the nuclear genetic diversity is presented in Fig. 3a. No clear spatial pattern emerged from this picture, but the two most isolated populations (MON, an isolated peripheral population and OSS, the westernmost most isolated population of the central chain, see Figure 1 and Mouret *et al.* 2011) show the lowest values of allelic richness and expected heterozygosity (Table 1, Fig. 3a). The linear regression between two indices of

genetic diversity (H_e and A_r) and altitude revealed no effect of altitude (Linear Model: $t_{H_e} = -0.17$, $p_{H_e} = 0.86$; and $t_{A_r} = -1.32$, $p_{A_r} = 0.22$). No significant correlation has been found between mitochondrial diversity and two measures of nuclear diversity (Pearson's correlation: $cor_{H_e} = -0.51$, $p_{H_e} = 0.07$; and $cor_{A_r} = -0.44$, $p_{A_r} = 0.13$), although the low number of populations restrains the power of this analysis.

Analyses of population structure and gene flow

STRUCTURE results show a continuous increase of the log-likelihood from $K=2$ to $K=13$, but the log-likelihood seems to reach a plateau for $K=13$, 14 and 15 (see supplementary material Fig. S1a). A small variance between runs is observed for all K values except for $K=12$. Finally, ΔK values are somewhat chaotic but show two higher values for K going from 5 to 6 and from 12 to 13 (Fig. S1b). This behaviour is typical of strong isolation by distance patterns when analysed with STRUCTURE (RL, unpublished data). The clusters found for $K=13$, and represented in Fig. S1c and S1d, correspond exactly to the populations *a priori* defined from the sampling sites and used in Mouret *et al.* 2011 with very little occurrence of admixed individuals. Among all admixed individuals, two individuals from population ARR are assigned to the isolated MON population with very high probability (0.999 and 1): for these individuals a mistake during the processing of the samples seems the most likely explanation as i) no dispersal is possible from MON to ARR or probably to any other population, ii) true immigrants into ARR would probably be assigned to one of the geographically close populations (see discussion on gene flow in the next section) and iii) these two individuals are next to each other's in our dataset, a very unlikely situation if they were true immigrants. Other individuals with mixed assignment probabilities (see LON and EST in Fig. S1c) have probably a recent ancestry outside the sampled populations, suggesting occasional gene flow (see discussion).

In line with the reliable assignment of individuals to their population of origin, F_{ST} values (Table 2 and Fig. 3b) are relatively high for most pairs of populations (0.06 to 0.39 with a mean of 0.22), indicating a strong structure among populations. The Neighbour-Joining tree inferred from Reynolds' distances (Fig. 5) indicates that geographically close populations are generally grouped together, suggesting a pattern of isolation by distance. However, note here that the bootstrap support of node does not exceed 65, an unsurprising result given the low number of loci used to infer the tree. Indeed, a significant isolation by distance was also found between all pairs of populations with a slope of 0.0464 ($CI_{95}= 0.0042 - 0.1045$) and a Mantel test P-value of 0.023 when using straight-lines distances (Fig. 3c) and with a slope of 0.0016 ($CI_{95}= 0.0004 - 0.0035$) and a Mantel test P-value < 0.001 with more realistic geographical distances through favourable habitats (see materials and methods for details). In agreement with the occurrence of isolation by distance, the lowest F_{ST} values are mostly found between adjacent populations (Table 2). Populations which are known to be disconnected from other sampled areas because suitable habitats are lacking between them (OSS and MON here) have comparatively high pairwise F_{ST} values (Table 2 and Fig. 3b), translating into longer branches than average in the population tree (Fig. 4), but similarly high divergence can also be found between two adjacent peripheral populations (NEO and LON, situated only 2 km apart).

Finally, SPAGEDI's results showed that 46% of all pairwise R_{ST} and pR_{ST} estimates were significantly different (Table S2), meaning that mutations and drift that occurred after the divergence of the different populations contributed more than recurrent but limited migration to the observed differentiation among populations.

Approximate Bayesian computations on the population history

After a few preliminary runs, all ‘prior and posterior checking’ analyses confirmed that the scenarios and parameter prior distributions we chose fitted relatively well the data: only a low proportion (<5%) of observed statistics were in the tails of the simulated distributions, and visual checks by PCA also revealed that the statistics computed on the observed data were in line with those computed on the simulated scenarios (results not shown).

From our four competing scenarios (unique divergence *versus* two divergence events, stable *versus* varying population sizes), scenario 4, with two divergence events and past changes of N_e was strongly supported with a posterior probability of 0.8133 ($CI_{95} = [0.4640, 1]$), whereas scenarios 1, 2 and 3 (see Fig. 2) had much lower support (posterior probability of 0.1599 with $CI_{95} = [0, 0.4597]$, of 0.0014 with $CI_{95} = [0, 0.1875]$, and of 0.0254 with $CI_{95} = [0, 0.1792]$, respectively). Given this result, we then inferred all parameter posterior distributions under the fourth scenario only, thus considering past variation in population sizes and two successive divergence events (Table 3). Parameter inferences suggest that the current repartition of genetic diversity is the result of two divergence events: (1) populations MON, LUS, NEO, LES and PIA diverged 998 generations ago ($CI_{95} = [547; 2680]$) from the common ancestral population i.e. 9,980 years ago assuming a generation time of 10 years (Mouret *et al.* 2011), and (2) more recently (i.e. 398 generations, $CI_{95} = [298; 1780]$) or 3,980 years ago) populations OSS, ARR, GEO, FAC, VIG and MUN diverged. Table 3 shows that there is not much information in the data to precisely infer past changes in population size except for the timing of those past events, which seems relatively recent: all modes for time parameters t_3 to t_{13} are close to the lower bound, i.e. 10 generations, and upper bounds of CI_{95} 's are mostly below 900 generations, except for the MON population (t_{10}). Other than

that, most of the posterior distributions for population sizes are very flat (results not shown), resulting in CI₉₅'s bounds on the posterior distributions being very close to the bounds of the uniform prior distributions set for those parameters; little relevant information can thus be gained from inference of changes in population size.

DISCUSSION

Our main objective here was to use a set of microsatellite loci to check (1) whether the restricted gene flow among neighbouring populations suggested by the mitochondrial population structure in Mouret *et al.* (2011) was also evident when using nuclear markers with biparental inheritance, (2) whether mitochondrial diversity was a good predictor for nuclear neutral diversity, and most importantly (3) whether the recent population history scenario favoured by the mitochondrial-only dataset in Mouret *et al.* (2011) was also supported by a set of independent nuclear markers.

Our results indeed support a pattern of strongly reduced gene flow even at moderate to low geographic distances, which together with a moderate signal of isolation-by-distance indicate limited current dispersal that is too weak to erase the signal of population history. Indeed, mutations and drift after the populations' divergence contributed more than recurrent migration to the differentiation between populations. Furthermore, we did not find any correlation between nuclear diversity and mitochondrial diversity, yet another evidence that mtDNA is not a reliable indicator of global genomic diversity. Lastly, the microsatellite data support the scenario of fragmentation of peripheral populations and expansion of central populations suggested by the mitochondrial marker, suggesting that nuclear and mitochondrial genetic diversity can be affected in dramatically different ways by the same demographic events. We nevertheless acknowledge that a perfectly satisfactory modelling of

the combined effects of past demographic events and historical and contemporary gene flow is currently not possible with the type of data we have, as discussed below.

Dispersal, barriers to gene flow and conservation

Dispersal in this species is probably highly restricted at all spatial scales as evidenced by (1) the high F_{ST} values between all populations even when they are only 2 - 3 km apart (Fig. 3b and Table 2); (2) the highly structured pattern given by the STRUCTURE software; (3) the strong IBD signal found in the data (Fig 3c); and (4) the significant differences between pR_{ST} and R_{ST} values. Whether the low differentiation sometimes observed between distant populations (see Table 2 and Fig. 1) is due to ongoing gene flow or recent common origin of these populations cannot be decided based on our data. Indeed, it is not easy to decide whether our data conclusively demonstrates the occurrence of gene flow between populations after their divergence. The isolation by distance pattern and grouping of geographically close populations in the F_{ST} tree suggest post-divergence gene flow mediated by dispersal through favourable habitats over many generations (the distance between most of our neighbouring populations certainly exceeds the single generation dispersal abilities of the species, see for ex. Doughty *et al.* 1994). However, the same pattern could result from the sharing of more recent common origins between geographically close but totally isolated populations. More direct evidence of gene flow stems from the occurrence of individuals with low assignment probabilities to the populations where they were sampled, or with mixed inferred assignment. Such individuals should not be interpreted as originating directly from the other sampled populations (see above remarks on single-generation dispersal distance in the species); instead their genotypes suggest that they originated from outside the sampled populations, i.e. that they are recent immigrants or descendants from recent immigrants.

Accepted Article

Most severe limitations to gene flow are found either between distant populations or between populations separated by unfavourable habitats, as apparent from high F_{ST} values (Table 2) and consequently long branches in the population tree (Fig. 4). Populations OSS and MON are especially separated from the rest of the distribution by vast areas situated below the alpine zone where the species occurs. Note that the SAU population, sampled in an area where many other populations are known, is separated from the rest of the distribution by an area where targeted field trips have failed to find the species (hence the gap in the distribution map in Fig. 1) but which lies entirely within the ecological niche of the species. The species could be present in low densities in this area or could have recently disappeared, but the level of genetic differentiation between SAU and adjacent populations suggests that there hasn't been a gap in the distribution for a long time. Whether the low differentiation sometimes observed between distant populations is due to ongoing gene flow or recent common origin of these populations cannot be decided based on our data.

More surprising is the high level of genetic divergence between two populations only 2 km apart (NEO and LON, $F_{st} = 0.26$), when compared with LON and EST, also situated 2 km apart, which have much lower divergence ($F_{st} = 0.07$). Highly restricted gene flow between NEO and LON was already apparent as highly differentiated mtDNA composition in Mouret *et al.* (2011). While LON and EST are only separated by favourable habitats where the species probably occurs continuously (and thus are only isolated by distance), LON and NEO are separated by a forested valley and a strong mountain stream, mostly lying below the lowest altitude at which the species is normally found and whose higher parts were occupied by damp habitats before a dam was built in the 1950s. It thus seems that even narrow bands of unsuitable habitats can inhibit dispersal in this species.

In conclusion, most populations are probably demographically isolated and should be managed independently at short time scale but in the long term, recolonisation of areas affected by local extinction does not seem compromised by poor dispersal of the species, since most populations exhibit genetic evidence of gene flow.

Congruent signal of recent history between nuclear and mitochondrial markers

Based on analysis of mtDNA variation only, Mouret *et al.* (2011) proposed a scenario where populations in the peripheral massifs (LUS, NEO, MON & LES) originate from a more extensive past distribution north of the Pyrenees that became fragmented at the end of the last glacial episode when favourable habitats became restrained to mountain tops whereas most of the high altitude central Pyrenean chain (populations SAU, PIA, MUN, VIG, FAC, ARR, GEO & OSS) would have been recolonised more recently through demographic expansion from a small number of populations following the retreat of the Pyrenean glaciers. Our first aim was to validate this biogeographical hypothesis with a multilocus nuclear dataset. To do so we compared two scenarios of population history with an ABC method: a simultaneous fragmentation of all current populations versus a two-step history where the populations in the area of supposed postglacial expansion diverged later than the populations in the supposed fragmented refugial range.

The ABC results clearly favoured the two-step scenario, indicating that some populations indeed shared a more recent common ancestor than the others. Remarkably, the inferred date of divergence of the peripheral populations by the ABC method (10,000 years ago) is extremely close to the hypothesised date of fragmentation of the glacial refuge based on the information available on the paleoecology of the Pyrenees (15,000 to 10,000 years ago). More examples of ABC modelling of post-glacial expansions will be needed to decide

whether this strong concordance is a lucky coincidence or reflects the power of ABC methods. The biogeographical scenario retained for the Pyrenean Rock Lizard thus conforms to a model of downslope glacial refugia which has recently been identified in a series of other mountain organisms (see Mouret *et al.* (2011) for a more detailed discussion of the biogeographical scenario and Schmitt (2007, 2009) for other examples).

We would like to stress here that our ABC modelling is based on scenarios where no gene flow is allowed between populations after their initial divergence. This is an unrealistic assumption, however, as many of our analyses identified a clear signal of limited gene flow between many of the populations, undoubtedly mediated by many generations of short-distance dispersal through connecting habitats (see above). The main reason for this modelling option is that DIYABC does not allow for the consideration of recurrent gene flow due to dispersal, only the occurrence of single events of admixture. The only alternative option would have been to group all populations linked by gene flow but as we have shown above, gene flow is not strong enough to overcome the effects of isolation. The significant differences between pR_{ST} and R_{ST} values indicate that differentiation between populations is less affected by gene flow between populations than by drift and mutation within populations. We must therefore carefully examine how not incorporating gene flow could have affected our ABC results.

We are not aware of any theoretical works specifically addressing how gene flow affects the result of DIYABC analyses, but it seems natural that gene flow could affect the estimates of divergence time, resulting in younger estimated divergence times for populations exchanging migrants than in reality. If there was greater gene flow among “central” populations than among “peripheral” populations and no gene flow between the two groups,

Accepted Article

divergence times estimated by ABC modelling for populations in the “central” group could be younger than reality, suggesting a two-step divergence scenario (as in our results) even if all populations in fact diverged simultaneously. However, our analyses suggest the occurrence of limited gene flow between most adjacent populations in the two population groups, except for the isolated OSS (in the peripheral) and MON (central) populations. We also find evidence of gene flow between populations in different groups (PIA with LES and LUS, as well as SAU with LUS). Indeed, average F_{ST} values are similar between “central” (0.20) and “peripheral” (0.24) populations, demonstrating that genetic connectivity of the populations do not differ between the two groups of populations. We are therefore confident that the results of the DIYABC analysis are not driven by different amounts of gene flow between the two groups of populations and that microsatellite data indeed supports a more recent common origin of the “peripheral” populations as expected under a scenario of post-glacial recolonisation.

Different responses of mtDNA and nuclear DNA diversity to the same history: consequences for inference of past selective sweep.

Our ABC approach rejected a scenario where all sampled populations had the same recent history, which would make it difficult to explain the striking differences in mtDNA diversity among populations by purely neutral processes. Indeed, the concordance between the recent population history suggested by both mtDNA and nuclear datasets supports the idea that the pattern of mtDNA diversity results from the postglacial colonisation history of the Pyrenean Rock Lizard and not from selective processes such as selective sweeps (see Galtier *et al.* 2009 for a review). Mitochondrial DNA and nuclear DNA thus reacted in very different ways to the same population history. Nuclear DNA diversity hasn't been affected by the post-glacial expansion since diversity of populations in areas of inferred postglacial

Accepted Article

recolonisation and in areas close to the identified refugial populations do not differ, while a dramatic loss of mtDNA diversity followed postglacial expansion. Fixation of a single mtDNA variant during population expansions fits well with a model of allele surfing that has recently been evaluated by simulations (*e.g.* Klopstein *et al.* 2006) and identified in case studies (see Excoffier *et al.* 2009 for a review.).

Although colonising populations are generally expected to have reduced genetic diversity compared with source areas (see for ex. Roques *et al.* 2012), we detected no decrease in nuclear genetic diversity in the area of post-glacial colonisation. Many other studies of expanding populations found similar results (*e.g.* Banks *et al.* 2010; Bronnenhuber *et al.* 2011; Zenger *et al.* 2003) that are generally interpreted as the consequence of strong dispersal between populations in the colonised range preventing loss of genetic diversity. Indeed, theoretical results (summarised in Excoffier *et al.* 2009) show that loss of diversity in expanding population crucially depends on gene flow and population size in the front of the expansion (see also Roques *et al.* 2012). It would thus seem that the lower loss of diversity of mtDNA compared with nuclear DNA during the expansion process results from lower effective population size of mtDNA, lower dispersal of mtDNA loci, or a combination of both.

It is well-known that effective population size of mtDNA is four times lower than nuclear loci, but differences in dispersal rate between mtDNA (female mediated only) and nuclear DNA (subject to female and male mediated dispersal) are undocumented in the genus *Iberolacerta*. However, most studies investigating sex-biased dispersal in Squamates report that males disperse more than females (*Anolis*: Johansson *et al.* (2008); *Boa*: Rivera *et al.* (2008) ; *Chlamydosaurus*: Ujvari *et al.* (2008); *Egernia*: Chapple & Keogh (2005) and Stow

et al. (2001) ; *Eulamprus*: Dubey & Shine (2010); *Lacerta*: Olsson *et al.* (1996); *Podarcis*: Vignoli *et al.* (2012); *Sceloporus*: Massot *et al.* (2003); *Stegonotus*: Dubey *et al.* (2008); *Uta*: Doughty & Sinervo (1994); *Zootoca*: Clobert *et al.* (1994), with a single publication reporting female-based dispersal in a mountain Australian skink (*Niveoscincus*: Olsson & Shine (2003)) and another paper documenting location-dependent sex-bias in sea snakes of the genus *Laticauda* (Lane & Shine 2011). It is thus likely that male dispersal is stronger than female dispersal in the Pyrenean Rock Lizard, further increasing nuclear gene flow compared with mitochondrial gene flow.

If the difference in genetic diversity between mtDNA and nuclear DNA in the recently colonised area indeed results from different dispersal rate during postglacial range expansion and different effective population size of these markers instead of selective sweep of favourable mtDNA variants, it suggests neutral explanations might also need to be excluded in previous studies that have inferred selective sweeps in reptiles based on similar patterns of contrasted mtDNA *versus* nuclear diversity. In the gecko genera *Hemidactylus* (Rato *et al.* 2011) and *Tarentola* (Rato *et al.* 2010), similar differences in diversity between mtDNA and nuclear DNA have been interpreted as evidence of selective sweeps in the mtDNA. In both cases, the inferred selective sweeps have not affected the whole species range but only European areas which have been recently colonised by the species, either as natural post-glacial expansion, human-induced dispersal or a combination of both. This biogeographic pattern of reduced diversity in recently colonised areas would also fit well with the idea that the reduced mtDNA diversity in these species is a neutral consequence of reduced mitochondrial gene flow in expanding populations. Obviously, the two hypotheses (selective sweeps or unequal level of gene flow during colonisation) cannot be evaluated verbally and formal tests are needed to disentangle them, for example based on explicit

multilocus simulations of expanding populations to evaluate if the observed ratio of nuclear and mitochondrial diversity can be explained by purely neutral processes or not.

Conclusions

Multilocus microsatellite data support the same postglacial scenario as mtDNA in the Pyrenean Rock Lizard: a range fragmentation after the last glacial maximum isolating northern peripheral massifs and a postglacial colonisation of large parts of the main Pyrenean range. However, mitochondrial genetic diversity and nuclear genetic diversity were not affected similarly by this common history: mitochondrial diversity is highest in populations resulting from the fragmentation of the refugial area and lowest in areas of recent range expansion (historical processes), whereas nuclear diversity was not affected by postglacial expansion. The result is a contrasting pattern of strongly reduced mitochondrial diversity and unaffected nuclear diversity in those populations that have been founded during the postglacial expansion; this pattern is identical to what has been interpreted in other species of reptiles as evidence of selective sweep in mitochondrial DNA. Explicit modelling of population expansion incorporating mitochondrial and nuclear loci in a multilocus framework will be needed to confirm that differences in effective population size and/or dispersal rates can result in strikingly different patterns of genetic diversity in absence of selective processes.

Acknowledgements:

We are grateful to Jean-Pierre Vacher, Samuel Danflous, Claire Froidefond and Françoise Poitevin for help in collecting samples. This study was supported by the Parc National des Pyrénées as part of a project on the conservation of the Pyrenean rock lizards and by the PPF ('Populations fractionnées et insulaires') of the École Pratique des Hautes

Accepted Article

Études. Capture permits were delivered by the ‘Préfecture des Hautes-Pyrénées’ (permit numbers 2000-318-1 and 2001-109-7). Microsatellites genotyping were made at the MBB Platform of the CeMEB LabEx in Montpellier. Part of this work was carried out by using the resources from the INRA GENOTOUL and MIGALE bioinformatics platforms and the computing grid of CBGP lab. RL has been funded by the Agence Nationale de la Recherche (EMILE 09-blan-0145-01), and by the Institut National de Recherche en Agronomie (Project INRA Starting Group “IGGiPop”).

References

- Arribas OJ (2000) Taxonomic revision of the Iberian ‘Archaeolacertae’ III: diagnosis, morphology and geographic variation of *Iberolacerta bonnali* (Lantz, 1927). *Herpetozoa* **13**, 99-132.
- Arribas OJ (2004) *Fauna y paisaje de los Pirineos en la Era glaciaria* Lynx (eds), Barcelona.
- Banks SC, LING SD, JOHNSON CR, *et al.* (2010) Genetic structure of a recent climate change-driven range extension. *Molecular Ecology* **19**, 2011-2024.
- Barrowclough GF, Zink RM (2009) Funds enough, and time: mtDNA, nuDNA and the discovery of divergence. *Molecular Ecology* **18**, 2934-2936.
- Bloor P (2006) Polymorphic microsatellite markers for the Iberian rock lizard species, *Iberolacerta cyreni*, and cross-species priming in other *Iberolacerta* species. *Molecular Ecology Notes* **6**, 1252-1254.
- Bronnenhuber JE, Dufour BA, Higgs DM, Heath DD (2011) Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, *Neogobius melanostomus*, in Great Lakes tributaries. *Molecular Ecology* **20**, 1845-1859.
- Chapple DG, Keogh JS (2005) Complex mating system and dispersal patterns in a social lizard, *Egernia whitii*. *Molecular Ecology* **14**, 1215-1227.
- Chakraborty R, Jin L (1993) A Unified Approach to Study Hypervariable Polymorphisms - Statistical Considerations of Determining Relatedness and Population Distances. *DNA Fingerprinting : State of the Science* **67**, 153-175.
- Clobert J, Massot M, Lecomte J, *al. e* (1994) Determinants of dispersal behavior: the common lizard as a case study. In: *Historical and Experimental Perspective* (ed. Ecology L). Vitt, L. J. & Pianka, E. R., Princeton, NJ.
- Cornuet JM, Pudlo P, Veyssier J, *et al.* (2014) DIYABC v2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **in press**.
- Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11**.

- Doughty P, Sinervo B (1994) The Effects of Habitat, Time of Hatching, and Body-Size on the Dispersal of Hatchling *Uta Stansburiana*. *Journal of Herpetology* **28**, 485-490.
- Dubey S, Brown GP, Madsen T, Shine R (2008) Male-biased dispersal in a tropical Australian snake (*Stegonotus cucullatus*, Colubridae). *Molecular Ecology* **17**, 3506-3514.
- Dubey S, Shine R (2010) Restricted dispersal and genetic diversity in populations of an endangered montane lizard (*Eulamprus leuraensis*, Scincidae). *Molecular Ecology* **19**, 886-897.
- Edwards S, Bensch S (2009) Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology* **18**, 2930-2933.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.
- Excoffier L, Foll M, Petit RJ (2009) Genetic Consequences of Range Expansions. *Annual Review of Ecology Evolution and Systematics* **40**, 481-501.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**, 47-50.
- Felsenstein J (2004) *PHYLIP (Phylogeny Inference Package) version 3.6*. Department of Genome Sciences, University of Washington, Seattle.
- Frankham R (2002) Population viability analysis. *Nature* **419**, 18-19.
- Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies* **326**, S22-S29.
- Frankham R (2005) Genetics and extinction. *Biological Conservation* **126**, 131-140.
- Frankham R (2012) How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. *Heredity* **108**, 167-178.
- Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* **18**, 4541-4550.
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305-318.
- Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.
- Gullberg A, Tegelstrom H, Olsson M (1997) Microsatellites in the Sand lizard (*Lacerta agilis*): description, variation, inheritance and applicability. *Biochemical Genetics* **35**, 281-295.
- Hardy OJ, Charbonnel N, Freville H, Heuertz M (2003) Microsatellite allele sizes: A simple test to assess their significance on genetic differentiation. *Genetics* **163**, 1467-1482.
- Hardy OJ, Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**, 183-195.
- Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution* **27**, 578-584.
- Johansson H, Surget-Groba Y, Thorpe RS (2008) Microsatellite data show evidence for male-biased dispersal in the Caribbean lizard *Anolis roquet*. *Molecular Ecology* **17**, 4425-4432.
- Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* **23**, 482-490.
- Lane A, Shine R (2011) Phylogenetic relationships within laticaudine sea snakes (*Elapidae*). *Molecular Phylogenetics and Evolution* **59**, 567-577.

- Massot M, Huey RB, Tsuji J, van Berkum FH (2003) Genetic, prenatal, and postnatal correlates of dispersal in hatchling fence lizards (*Sceloporus occidentalis*). *Behavioral Ecology* **14**, 650-655.
- Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399-2402.
- Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 1061-1064.
- Mouret V, Guillaumet A, Cheylan M, *et al.* (2011) The legacy of ice ages in mountain species: post-glacial colonization of mountain tops rather than current range fragmentation determines mitochondrial genetic diversity in an endemic Pyrenean rock lizard. *Journal of Biogeography* **38**, 1717-1731.
- Nielsen R, Beaumont M (2009) Statistical inferences in phylogeography. *Molecular Ecology* **18**, 1034-1047.
- Olsson M, Gullberg A, Tegelstrom H (1996) Malformed offspring, sibling matings, and selection against inbreeding in the sand lizard (*Lacerta agilis*). *Journal of Evolutionary Biology* **9**, 229-242.
- Olsson M, Shine R (2003) Female-biased natal and breeding dispersal in an alpine lizard, *Niveoscincus microlepidotus*. *Biological Journal of the Linnean Society* **79**, 277-283.
- Pascual M, Chapuis MP, Mestres F, *et al.* (2007) Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**, 3069-3083.
- Petit RJ, Aguinalgalde I, de Beaulieu JL, *et al.* (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.
- Pottier G, Arthur C-P, Weber L, Cheylan M (2013) Répartition des lézards du genre *Iberolacerta* Arribas, 1997 (*Sauria* : *Lacertidae*) en France. 3/3 : le Lézard de Bonnal, *Iberolacerta bonnali* (Lantz, 1927). *Bulletin de la Société Herpétologique de France* **148**, 425-450.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9197-9201.
- Rato C, Carranza S, Harris DJ (2011) When selection deceives phylogeographic interpretation: The case of the Mediterranean house gecko, *Hemidactylus turcicus* (Linnaeus, 1758). *Molecular Phylogenetics and Evolution* **58**, 365-373.
- Rato C, Carranza S, Perera A, Carretero MA, Harris DJ (2010) Conflicting patterns of nucleotide diversity between mtDNA and nDNA in the Moorish gecko, *Tarentola mauritanica*. *Molecular Phylogenetics and Evolution* **56**, 962-971.
- Remon N, Vila M, Galan P, Naveira H (2008) Isolation and characterization of polymorphic microsatellite markers in *Iberolacerta monticola*, and cross-species amplification in *Iberolacerta galani* and *Zootoca vivipara*. *Molecular Ecology Resources* **8**, 1351-1353.
- Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the Co-Ancestry Coefficient - Basis for a Short-Term Genetic-Distance. *Genetics* **105**, 767-779.
- Rivera FN, Cisneros R, Hernandez-Rodriguez C, Zuniga G (2008) Genetic diversity and population structure of *Pichia guilliermondii* over 400 generations of experimental microevolution. *Biological Journal of the Linnean Society* **93**, 475-486.

- Accepted Article
- Roques L, Garnier J, Hamel F, Klein EK (2012) Allee effect promotes diversity in travelling waves of colonization. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8828-8833.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.
- Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.
- Rousset F (2008) Genepop'007: a complete reimplement of the Genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.
- Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* **4**.
- Schmitt T (2009) Biogeography and evolutionary importance of the Europe high mountain systems. *Frontiers in Zoology* **6**.
- Schmitt T, Varga Z (2012) Extra-Mediterranean refugia: The rule and not the exception? *Frontiers in Zoology* **9**.
- Slatkin M (1995) A Measure of Population Subdivision Based on Microsatellite Allele Frequencies (Vol 139, Pg 457, 1995). *Genetics* **139**, 1463-1463.
- Stewart JR, Lister AM, Barnes I, Dalen L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B-Biological Sciences* **277**, 661-671.
- Stow AJ, Sunnucks P, Briscoe DA, Gardner MG (2001) The impact of fragmentation on dispersal of Cunningham's sink (*Egernia unninghami*): evidence from allelic and genotypic analyses of microsatellites. *Molecular Ecology* **10**, 867-878.
- Ujvari B, Dowton M, Madsen T (2008) Population genetic structure, gene flow and sex-biased dispersal in frillneck lizards (*Chlamydosaurus kingii*). *Molecular Ecology* **17**, 3557-3564.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.
- Vignoli L, Vuerich V, Bologna MA (2012) Experimental study of dispersal behaviour in a wall lizard species (*Podarcis sicula*) (*Sauria: Lacertidae*). *Ethology Ecology & Evolution* **24**, 244-256.
- Zenger KR, Richardson BJ, Vachot-Griffin AM (2003) A rapid population expansion retains genetic diversity within European rabbits in Australia. *Molecular Ecology* **12**, 789-794.
- Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17**, 2107-2121.

Data accessibility

Genotype file and distances matrices are available on Dryad (doi:10.5061/dryad.mh6kq)

Author contributions

P-A. Crochet, M. Cheylan and G. Pottier developed the overall sampling plan. M. Cheylan and G. Pottier sampled most of the studied populations. R. Eudeline and V. Arnal isolated DNA and generated genotypes. R. Eudeline and A-L Ferchaud conducted preliminary analysis. A-L Ferchaud, P-A.Crochet and R. Leblois performed the final analyses and wrote the manuscript.

Figure 1:

Approximate distribution range of the Pyrenean rock lizard *Iberolacerta bonnali* (black and dark grey) and sampled localities (stars). For population names see Table 1. White and grey stars correspond to the two groups of populations identified in Mouret *et al.* (2011) on the basis of mtDNA diversity as potentially resulting from the fragmentation of a glacial refugium (white stars, “peripheral populations” in the text) or from the putative post-glacial range expansion (grey stars, “central populations” in the text). They correspond to the two groups with different divergence times in the ABC modelling. The thin white line is the French-Spanish border and thick white bars indicate inferred range fragmentation based on lack of suitable habitat. The dark grey part of the range includes the unsampled massifs of Maladeta and Aigüestortes that are inhabited by distinct ESUs (see “Tissue sampling and population information”).

Figure 2:

Graphical representation of the two scenarios used in the ABC analyses. N values are population sizes, and t values correspond to the timing of past divergence events ($t1$ and $t2$) or past change in population sizes ($t3$ to $t14$). Note that time is not scaled.

Figure 3:

- a. Distribution of genetic diversity within *I. bonnali* populations (12 microsatellites locus). Circles size refers to Allelic richness (Ar) whereas color gradient refer to expected heterozygosity (He).
- b. Pairwise standardised F_{ST} values between adjacent populations. Dark zone represents the approximate distribution range of the species as in Figure 1.
- c. Isolation by distance between the sampled populations, shown as the regression of pairwise genetic differentiation between populations against the distance computed as the shortest way through favourable habitats.

Figure 4:

Neighbor-Joining tree between populations based on Reynold's distance. Only bootstrap values above 10 are indicated on the tree. The scale bar indicates Reynold's distance between samples.

Table 1: Genetic diversity for each population

Population	<i>n</i>	<i>p</i>	<i>He</i>	<i>Hobs</i>	<i>Ar</i>	<i>Fis</i> (p-value)
OSS	30	1	0.31	0.324	2.45	-0.028 (0.68)
ARR	30	1	0.482	0.469	3.37	0.043 (0.16)
GEO	30	0	0.468	0.458	3.32	0.038 (0.19)
FAC	30	1	0.494	0.485	3.11	0.036 (0.18)
VIG	31	2	0.478	0.464	3.70	0.047 (0.1)
MUN	30	0	0.458	0.486	3.43	-0.045 (0.87)
SAU	14	0	0.389	0.381	2.912	0.058 (0.22)

MON	29	1	0.334	0.345	2.49	-0.014 (0.59)
LUS	30	3	0.439	0.434	3.30	0.029 (0.24)
NEO	31	0	0.400	0.372	3.02	0.086 (0.02)
LON	31	0	0.494	0.471	3.750	0.062 (0.06)
EST	28	1	0.482	0.500	3.15	-0.018 (0.67)
PIA	28	2	0.403	0.403	3.10	0.018 (0.36)

n = number of individuals analysed by population, p = number of private alleles by population H_e = expected heterozygosity, H_o = observed heterozygosity, A_r = Allelic richness per population (average of allelic richness across loci) based on minimum sample size of 14 diploid individuals, F_{IS} = inbreeding coefficient.

Table2: Genetic differentiation between populations.

	FAC	GEO	ARR	MON	LUS	MUN	EST	OSS	VIG	SAU	PIA	NEO	LON
FAC		0.377	0.37 6	0.502	0.425	0.387	0.451	0.597	0.22 3	0.469	0.409	0.574	0.392
GEO	0.193		0.12 3	0.394	0.288	0.201	0.370	0.443	0.423	0.408	0.371	0.431	0.267
ARR	0.18 7	0.06 3		0.294	0.317	0.216	0.323	0.30 5	0.373	0.343	0.314	0.482	0.276
MON	0.286	0.233	0.172		0.464	0.256	0.407	0.615	0.428	0.559	0.348	0.62 9	0.378
LUS	0.221	0.155	0.168	0.283		0.272	0.379	0.535	0.391	0.27 0	0.33 1	0.460	0.271
MUN	0.199	0.107	0.112	0.152	0.148		0.250	0.483	0.29 3	0.414	0.22 6	0.503	0.159
EST	0.225	0.191	0.164	0.238	0.201	0.130		0.553	0.262	0.420	0.19	0.500	0.14

OSS	0.350	0.267	0.18	0.415	0.332	0.295	0.331	0.534	0.476	0.530	0.582	0.541
			1									
VIG	0.11	0.220	0.191	0.249	0.208	0.15	0.134	0.320	0.435	0.161	0.488	0.274
	2					4						
SAU	0.251	0.225	0.186	0.356	0.15	0.232	0.229	0.310	0.237	0.447	0.457	0.383
					4							
PIA	0.219	0.206	0.172	0.217	0.18	0.12	0.10	0.339	0.089	0.264	0.477	0.165
					8	7	7					
NEO	0.315	0.241	0.266	0.39	0.263	0.285	0.276	0.372	0.270	0.272	0.282	0.47
				4								8
LON	0.194	0.136	0.139	0.219	0.141	0.082	0.07	0.320	0.138	0.206	0.089	0.26
							3					1

Pairwise F_{ST} estimates for 12 microsatellites locus between 13 French populations of *Iberolacerta bonnali* (below diagonal) and standardised values (above diagonal) based on Hedrick (2005) and Meirmans (2006). All F_{ST} values are highly significant (p values < 0.001, exact tests for population differentiation in Genepop). Bold values indicate neighbouring populations.

Table 3: Prior and posterior distributions of the ABC analysis.

Parameters	Prior distribution	Posterior distribution summary		
	[Lower bound; Upper bound]	mode	q025	q975
N1 (OSS)	LU[50; 5 000]	198	759	3960
N14	LU[50; 5 000]	85	520	2340
N2 (GEO)	LU[50; 10 000]	365	148	7050
N15	LU[50; 8 000]	831	139	8997
N3 (ARR)	LU[50; 8 000]	325	141	8645
N16	LU[50; 10 000]	1200	215	7490

N4 (FAC)	LU[50; 5 000]	77	58	2100
N17	LU[50; 10 000]	1500	321	9650
N5 (VIG)	LU[50; 10 000]	378	152	8950
N18	LU[50; 10 000]	720	182	9100
N6 (MUN)	LU[50; 8 000]	82	59	4020
N19	LU[50; 10 000]	4200	528	9890
N7 (SAU)	LU[50; 5 000]	156	76	4520
N20	LU[50; 10 000]	185	78	7850
N12 (PIA)	LU[50; 8 000]	298	148	4780
N25	LU[50; 10 000]	2140	248	9670
N8 (MON)	LU[50; 5 000]	69	56	2890
N21	LU[50; 10 000]	814	345	9560
N9 (LUS)	LU[50; 7 000]	212	92	5480
N22	LU[50; 10 000]	945	198	8970
N10 (NEO)	LU[50; 7 000]	521	134	5980
N23	LU[50; 10 000]	125	61	3900
N11 (LES)	LU[200; 12 000]	346	218	8010
N24	LU[200; 12 000]	1480	521	10780
N13	LU[300; 15 000]	428	314	13080
t1	LU[100; 3 000]	398	298	1780
t2	LU[100; 5 000]	998	547	2680
t3 (OSS)	LU[10; 3 000]	10	10	598
t4 (GEO)	LU[10; 3 000]	10	10	724
t5 (ARR)	LU[10; 3 000]	10	10	548
t6 (FAC)	LU[10; 3 000]	12	10	678
t7 (VIG)	LU[10; 3 000]	11	11	819
t8 (MUN)	LU[10; 3 000]	9	10	348
t9 (SAU)	LU[10; 3 000]	10	10	315
t14(PIA)	LU[10; 3 000]	9	10	327
t10 (MON)	LU[10; 3 000]	14	13	1890
t11 (LUS)	LU[10; 3 000]	13	11	689
t12 (NEO)	LU[10; 3 000]	11	10	657
t13 (LES)	LU[10; 3 000]	10	10	526
μ	U[0.0001; 0.01]	3.68e-04	1.87e-04	8.74e-04
P_{GSM}	LU[0.001; 0.3]	1.254e-03	1.17e-03	3.74e-02
P_{SNI}	LU[10 ⁻⁸ ; 10 ⁻⁴]	8.98e-09	8.56e-09	5.24e-07

All prior distributions for the demographic parameters are log-uniform distributions (LU).

For mutational parameters, hyper-priors distributions are uniform for the mean mutation rate over loci (μ) and log-uniform for the mean over loci of the geometric parameter of the GSM

mutation model (P_{GSM}) and for the mean probability over loci of single nucleotide insertions (P_{SNI}). Gamma distributions centred on the mean drawn from the hyperprior distributions are then used for each locus individually. Posterior distributions for all parameters, estimated under scenario 4, are summarised by their mode, and their first (q025) and last (q975) quantiles at a 0.025 level. Population sizes are expressed as the number of diploid individuals, times in generations and the mutation rate as the number of mutation per generation. See also the full scenario descriptions in the text, Fig. 2, and the DIYABC manual available at <http://www1.montpellier.inra.fr/CBGP/diyabc/download/file.php?id=197> for more details about those settings.

Figure 1

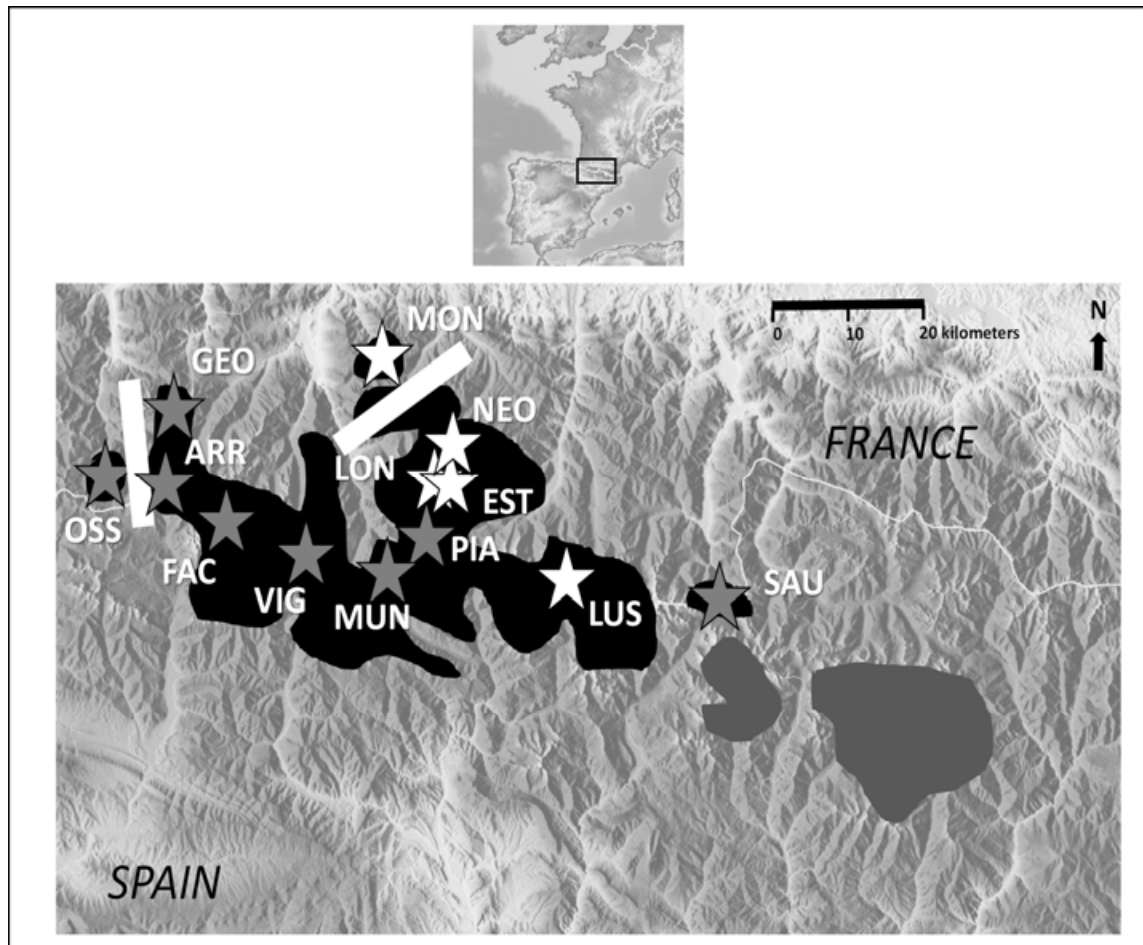


Figure 2

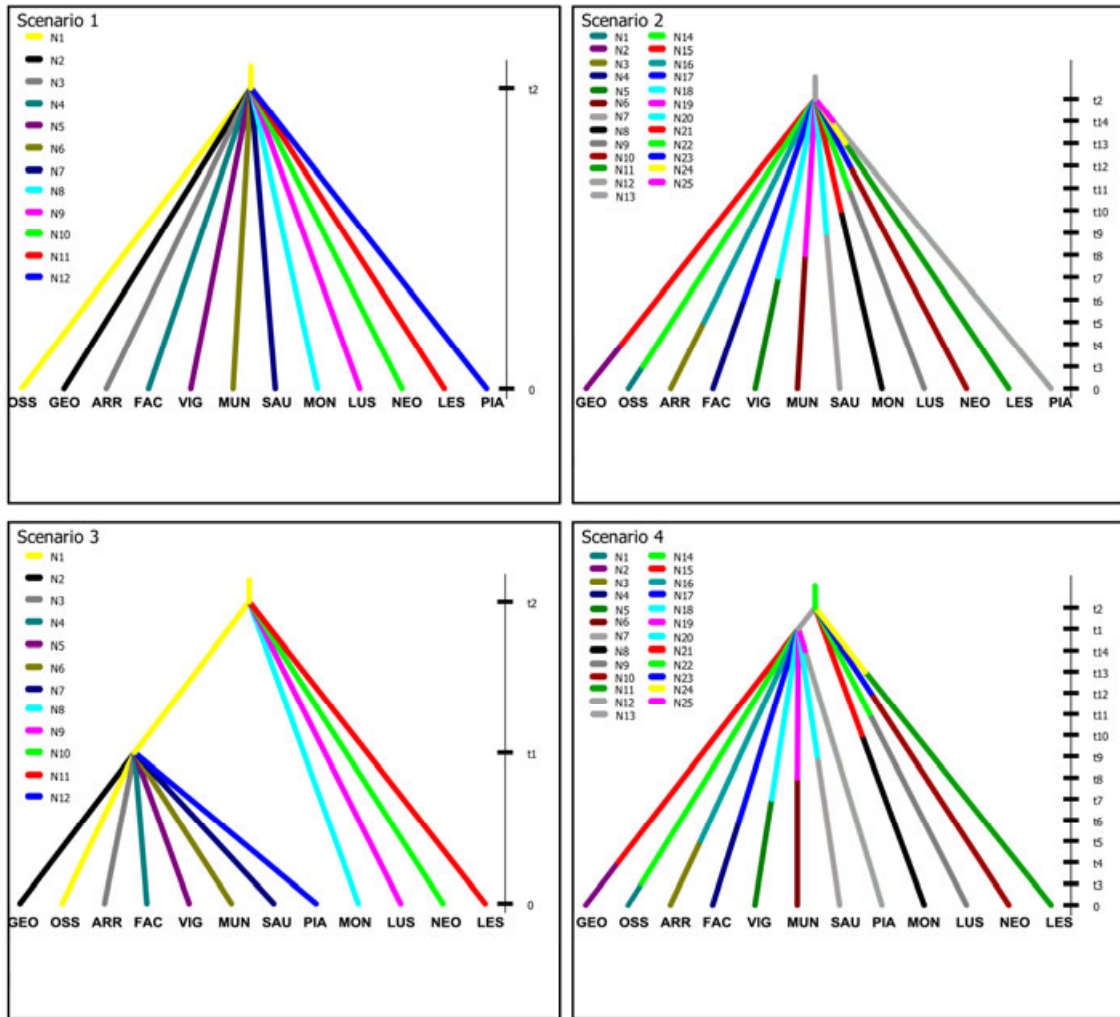


Figure 3:

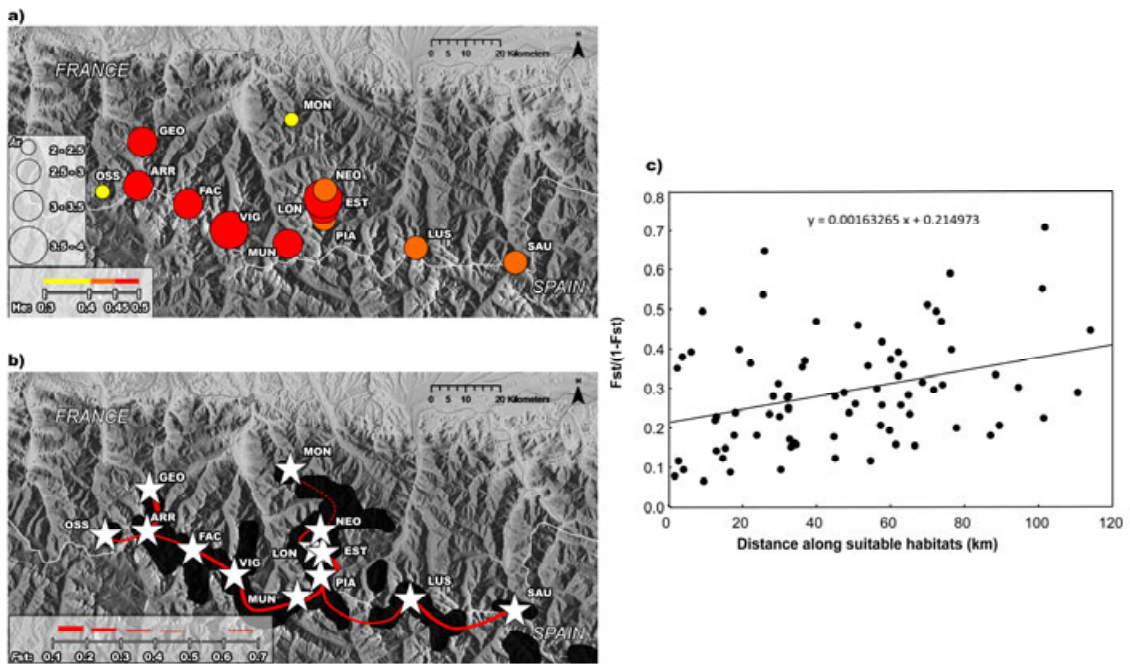


Figure 4:

