Phylogeographic breaks and how to find them: Separating vicariance from isolation by distance in a lizard with restricted dispersal

Running title: *Acanthodactylus erythrurus* phylogeography

Loïs Rancilhac[1,2,3*$], Aurélien Miralles[4*], Philippe Geniez[5], Daniel Mendez-Aranda[6], Menad Beddek[1,7], José Carlos Brito[8,9,10], Raphaël Leblois[11], Pierre-André Crochet[1]

1 CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France
2 Zoologisches Institut, Technische Universität Braunschweig, Mendelssohnstr. 4, 38106, Braunschweig, Germany
3 Animal Ecology, Institute of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden
4 Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France
5 CEFE, EPHE-PSL, CNRS, Université de Montpellier, Biogéographie et Ecologie des Vertébrés, Montpellier, France
6 Max Delbrueck Center for Molecular Medicine in the Helmholtz Association, Robert-Roessle-Str. 10, 13092, Berlin, Germany
7 Naturalia Environnement, Site Agroparc, Avignon, France
8 CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, Vairão, Portugal
9 BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, Vairão, Portugal
10 Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal
11 CBGP, INRAE, CIRAD, IRD, Institut Agro, Montpellier SupAgro, Univ. Montpellier, Montferrier-sur-Lez Cedex, France

***These authors contributed equally to this work**
$Corresponding author: Loïs Rancilhac, loisrancilhac@gmail.com

Abstract

Aim

Discontinuity in the distribution of genetic diversity (often based on mtDNA) is usually interpreted as evidence for phylogeographic breaks, underlying vicariant units. However, misleading signal of phylogeographic break can arise in the absence of barrier to gene flow, under mechanisms of isolation by distance (IBD). How and under which conditions phylogeographic breaks can be safely differentiated from populations evolving under IBD remain unclear. Here, we use multi-locus sequence data from a widely distributed lizard species to address these questions in an empirical setting.

Location

Morocco

Taxon

Spiny-footed lizard (*Acanthodactylus erythrurus*), Squamata: Lacertidae

Methods

Using 296 samples from 40 localities, we identified genetic discontinuities within *A. erythrurus* based on a mitochondrial fragment and nine nuclear markers independently. Using the nuclear markers, we then applied linear regression models to investigate whether genetic divergence could be explained by geographical distances alone, or barriers to gene flow (real phylogeographic breaks).

Result

A. erythrurus is characterized by an important mitochondrial diversity, with 11 strongly supported phylogeographic lineages with a crown age of 6 Mya. Nuclear markers, however, yielded weak phylogenetic support for these lineages. Using clustering methods based on genotypes at nine nuclear loci, we identified phylogeographic clusters that were partly discordant with the mtDNA lineages. Tests of IBD delimited at least four groups of populations separated by barriers to gene flow, but unambiguous separation of vicariance from IBD remained challenging in several cases.

Main conclusions

The genetic diversity of *A. erythrurus* originates from a mix of IBD and vicariance, which were difficult to distinguish, and resulted in similar levels of mitochondrial differentiation. These results highlight that phylogeographic breaks inferred from mitochondrial data should be further investigated using multi-locus data and explicit testing to rule out alternative processes generating discontinuities in mitochondrial diversity, including IBD. We identified four groups of populations within *A. erythrurus*, separated by barriers to gene flow, but even using nine independent nuclear makers the power of our approach was limited, and further investigation using genome-wide data will be required to resolve the phylogeographic history of this species.

Keywords: *Acanthodactylus*, Isolation by Distance, Morocco, Phylogeography, Vicariance, Admixture, Cytonuclear discordance, Phylogeographic breaks

Introduction

One of the main goals of phylogeography is to identify discontinuities in the distribution of genetic diversity across conspecific populations – phylogeographic breaks – and investigate the biotic and abiotic factors underlying them (Kumar & Kumar, 2018). Typically, phylogeographic breaks are interpreted as evidence for the existence of evolutionary lineages that diverged through vicariance events (Kuo & Avise, 2005). As such, they have been widely used to investigate biogeography and evolutionary biology questions, including the impact of geological events and landscape features on populations connectivity (e.g., Leaché et al., 2020), and their respective roles in promoting divergence and speciation (Huang, 2020). Furthermore, phylogeographic breaks are of interest for other fields of study, such as taxonomy and conservation biology (Braby et al., 2012; Shaffer et al., 2015). However, correct identification of phylogeographic breaks remains a non-trivial problem (Audzijonyte & Vrijenhoek, 2010).

Historically, phylogeographic studies have heavily relied on uniparentally inherited markers, such as mitochondrial DNA (mtDNA), to infer patterns of genetic diversity (Avise et al., 1987). However, while being characterized by a number of advantages (e.g., high mutation rate, lower effective population size, established standardized sequencing protocols), these markers often yield misleading signals of phylogeographic breaks (or the absence thereof) due to a number of processes (reviewed in Toews & Brelsford, 2012). Among them, different selective pressures (Irwin, 2012), introgression (e.g., Zieliński et al., 2013) and sex-biased dispersal (Palumbi & Baker, 1994) are well documented. Conversely, breaks in mitochondrial genetic diversity can also appear in continuous populations solely as a result of isolation by distance (IBD; Irwin, 2002), especially in the presence of sampling gaps (Audzijonyte & Vrijenhoek, 2010), and mtDNA is poorly suited to detect IBD patterns (Teske et al., 2018). Additionally, mtDNA gene trees often differ from the populations tree, mainly due to introgression and incomplete lineage sorting (ILS), hence phylogeographic breaks identified from mtDNA might not reflect actual populations splits (e.g., Dufresnes, Nicieza et al., 2020; Velo-Antón et al., 2018).

The limitations of mitochondrial phylogeography are usually overcome by considering multiple unlinked nuclear loci, which enable more accurate inferences (Brito & Edwards, 2009), particularly in the framework of the coalescent theory (Knowles, 2009). This approach reduces the likelihood of detecting spurious phylogeographic breaks, as independent loci are not expected to support concordant breaks in the absence of barriers to gene flow (Avise & Ball, 1990; Kuo & Avise, 2005). However, populations evolving along a single geographical gradient (i.e., under IBD) might still be apparently divided into distinct lineages, for example in the presence of large sampling gaps (Schwartz & McKelvey, 2009; Tolley et al., 2022; Wiemers & Fiedler, 2007). In that context, it appears crucial to test explicitly whether apparent phylogeographic breaks (e.g.,

reciprocally monophyletic groups in phylogenetic trees) delimit vicariant units separated by barriers to gene flow or have been generated by alternative processes such as IBD. A commonly used approach is to study the correlation between genetic and geographical distances among populations or individuals in a statistical framework (e.g., Ramachandran et al., 2005), to determine whether the null hypothesis that all populations differentiate along the same IBD gradient can be rejected, which would indicate the presence of a barrier to gene flow (e.g., Binks et al., 2019; Hausdorf & Hennig, 2020; Goudarzi et al., 2019). Other popular approaches include the estimation of migration rates over geographical surfaces, with low migration rates highlighting barriers to gene flow (Petkova et al., 2016). However, in practice, the implementation of such tests is often limited in phylogeographic studies by sparse sampling schemes and the low number of genetic markers available. Here, we investigate the extent to which vicariant units can safely be identified from populations evolving along IBD gradients, using multi-locus data of a genetically diverse North African lizard.

Compared to other regions of the Mediterranean Basin biodiversity hotspot, fewer phylogeographic studies have focused on North African communities (Beddek et al., 2018; Husemann et al., 2014). Among these, many studies have revealed important intra-specific genetic diversity in various taxa (see Husemann et al., 2014 for a review), in relation with the range of landscapes and bioclimatic regions and the complex biogeographic history that characterize this area. Consequently, important knowledge gaps still exist regarding the taxonomic diversity of North African taxa and the evolutionary processes underlying it (Beddek et al., 2018; Brito et al., 2014; Ficetola et al., 2013). Phylogeographic studies of North African taxa are of special interest, for two reasons. Firstly, they have the potential to significantly fill the aforementioned knowledge gaps, by providing extensive information on the extent and geographic distribution of intra-specific genetic diversity, and its underlying evolutionary and ecological factors, as well as undetected species boundaries. Secondly, biodiversity hotspots are prime areas to study the processes shaping the distribution of genetic diversity and underlying lineages diversification at a global scale. In that context, regionally distributed (as opposed to narrow-ranged) taxa are particularly valuable models, as they are likely to inform on a larger variety of processes.

In this study, we focus on the Red-tailed Spiny-footed lizard *Acanthodactylus erythrurus* (Squamata: Lacertidae), a species widely distributed across the Maghreb and the Iberian Peninsula. This species is characterized by pronounced morphological variations that led to the recognition of several morphological subspecies or species (Salvador, 1982; Bons & Geniez, 1995), which inhabit distinct habitat types (from coastal sand dunes to high altitude xerophytic vegetation through open forests and semi-arid steppes). Several molecular studies (Beddek et al., 2018; Fonseca et al., 2009; Miralles et al., 2020) have revealed an important and previously unsuspected intraspecific genetic

diversity, leading to the description of two new species, and emphasizing the need for a more in-depth taxonomic revision of this group. Notably, these studies highlight important disagreements between the accepted taxonomy and molecular clades. One of the main questions raised by these findings is whether the morphological and ecological diversification of this group has been produced by vicariance and lineage diversification, or by local adaptation in the face of historical gene flow. However, most studies (except Miralles et al., 2020) have relied only on mtDNA, offering an incomplete picture of the historical processes underlying *A. erythrurus*' genetic diversity. Here, we combine a comprehensive sampling with multi-locus data to clarify whether multiple vicariant units explain the genetic diversity in the *A. erythrurus* complex (referred to as *A. erythrurus* here).

The goals of this study are 1) to identify candidate phylogeographic breaks among Moroccan populations of *A. erythrurus* using both mtDNA and nuclear DNA (nDNA) markers; 2) to determine whether these candidate breaks delimit vicariant units or could results from IBD, and 3) integrate these processes into a phylogeographic scenario to explain *A. erythrurus*' present genetic diversity. Across the range of *A. erythrurus*, we decided to focus on Moroccan populations to mitigate the effect of sampling gaps, as we only had sparse sampling in other countries. We employed a two-step approach to untangle the processes at play: first, we analyzed mitochondrial and nuclear (nine gene sequence fragments) data independently to identify phylogeographic lineages. In both cases, phylogeographic lineages were defined as the smallest genetically homogeneous geographical units. Secondly, we used the nuclear dataset to test whether the described lineages corresponded to vicariant units or could result from IBD alone. Through this approach, we aim at providing a more comprehensive picture of the phylogeographic history of Moroccan *A. erythrurus*, but also empirically assess our ability to identify vicariant units using multi-locus data in a genetically variable species.

Material & Methods

1. Sampling

We used a previously published (Miralles et al., 2020) dataset of 392 *A. erythrurus* samples from the whole range of the complex (Algeria, Morocco, Portugal, Spain, Tunisia). The samples have been sequenced for DNA fragments from one mitochondrial gene (776 bp of the NADH-dehydrogenase subunit 4 [ND4] gene) and nine nuclear protein-coding genes: Recombination Activating Gene 2 (RAG2, 321 pb); Oocyte maturation factor (C-mos, 369 pb), Ornithine Decarboxylase (OD, 396 pb), Phosphogluconase Dehydrogenase intron 7 (PgD7, 420 pb), Melanocortin Receptor 1 (MC1R, 630 pb), Pyruvate Kinase Muscle (PKM, 351 pb), Nuclear

transcription factor Y gamma (NFYC, 666 pb), Acetylcholinergic receptor M4 (ACM4, 390 pb) and RNA fingerprint protein 35 (R35, 588 pb).

In the present study, we focus on the Moroccan populations, which offer the most complete sampling (296 samples). Details on the localities and number of samples are given in Figure 1 and Table S1. Samples from other lineages were also included to be used as outgroups in phylogenetic analyses, as described below. Finally, for the Moroccan samples, individual genotypes were inferred at each nuclear locus using the software PHASE (Stephens et al., 2001) implemented in DNAsp v. 5 (Librado & Rozas, 2009). The samples were grouped based on their mitochondrial lineages (cf. Results), and PHASE was run on each group separately, with 5,000 sampled steps and 10,000 burnin-steps in the MCMC. Uncertain phases (i.e., Posterior Probability < 0.90) were treated as unresolved sites using PHASE's default settings.

2. Mitochondrial phylogeography

As a first step to identify the lineages used as basic units for the subsequent analyses, we conducted phylogenetic inference based on 776 bp of the ND4 mitochondrial gene. In addition to all our Moroccan samples, two representatives of each of the following lineages (for the lineages nomenclature, we follow Miralles et al., 2020) were included as outgroups: *A. savignyi*, *A. lacrymae*, the AT lineage from eastern Algeria and Tunisia (including *A. blanci*), *A. montanus* and *A. cf. erythrurus* from central Algeria (CA lineage). Finally, two samples from Portugal and Spain were also included to allow estimation of the divergence times based on the opening of the Gibraltar straight (more details below). In total, 306 samples were included in the final alignment.

Maximum-likelihood (ML) phylogenetic inference was performed using IQTREE v. 1.6.8 (Chernomor et al., 2016; Nguyen et al., 2015). The best substitution model was selected using ModelFinder (Kalyaanamoorthy et al., 2017) as implemented in IQTREE, and the branches support was assessed using the SH-like approximate likelihood test (aLRT) with 5,000 pseudoreplicates. Based on this tree, we defined phylogeographic lineages corresponding to the smallest geographical units supported as monophyletic groups (in practice, with an aLRT support > 90). Finally, we estimated divergence times through two approaches. First, we calibrated the divergence between the Iberian populations and their closest Moroccan relatives based on the opening of the Strait of Gibraltar (between 5.3 and 5.5 Million years ago [Mya]). However, as more recent colonization events across the Strait of Gibraltar have been found in reptiles and amphibians (e.g., Carranza et al., 2006; Gutiérrez-Rodríguez et al., 2017), we also calibrated the tree based on a fixed mutation rate for the ND4 gene. Details on the procedure can be found in Supplementary materials S1.

3. Nuclear phylogeography

Preliminary analyses revealed that the phylogeographic signal in the nuclear data was much weaker compared to that of the mitochondrial dataset, with much higher levels of allele sharing among populations and lower nodes support in phylogenies. Our initial approach to infer genetic clusters and admixture among them was to use the software STRUCTURE (Pritchard et al., 2000). However, when conducting independent analyses of the same data, the MCMC chains struggled to reach convergence, and yielded very different, yet equally likely, clustering solutions. Likewise, phylogenetic reconstructions of the concatenated data failed to converge toward a consistent topology. In that context, it was difficult to apply a strict criterion to delimit phylogeographic lineages from the nuclear dataset. Thus, we relied on congruent results from several analyses, especially multivariate analyses of the genotypes at the nine nuclear exons, as described below.

3.1 Phylogenetic analyses

We first performed ML phylogenetic inference using the concatenation of the nine nuclear exons. Best substitution models and partition schemes were selected on a per-exon partition, and the analyses were run in IQTREE v. 1.6.8 with the same settings as for the mitochondrial analyses. Five independent ML searches were performed.

In addition, we also perform Neighbour-Joining (NJ) inference based on between-populations (i.e., sampling localities) pairwise distance matrices. This analysis was repeated twice, firstly using average p-distances calculated from the concatenation of the nine exons, secondly using Nei's distances computed from the genotypes at the nine exons. In the first case, both distances computation and tree inference were performed in MEGA 11 (Tamura et al., 2021); in the second case Nei's distances were calculated using adegenet v. 2.1.5 (Jombart, 2008) and NJ inference performed with ape v. 5.6 (Paradis & Schliep, 2019).

3.2 Multivariate analyses

As an alternative approach to detect phylogeographic breaks from the nuclear dataset, we performed Multiple Correspondence Analyses (MCA) from the genotypes at the 9 nuclear exons, using the software Genetix (Belkhir, 2004). We used an iterative approach to define phylogeographic lineages: the analysis was first run on a dataset including all Moroccan populations (same sampling as for the phylogenetic analyses but without outgroups and Iberian samples); the results were visually inspected and groups of populations that were clearly differentiated from the rest were identified as phylogeographic lineages. These populations were then removed from the dataset, and the analysis was run again on the resulting subset. This process

was repeated until no clear structure could be identified in the results (in practice, five iterations were needed, see Results).

4. Tests of Isolation by Distance and vicariance

Population structure analyses based on both mtDNA and nDNA data supported several genetic groups in Moroccan *A. erythrurus.* While the structure inferred from both types of markers was mostly concordant, there were a few notable differences. For further analyses, we considered the nDNA genetic groups (summarized in the following, cf. Table 1 and Figure 2 for more details), which minimized the number of clusters represented by a single locality. More in details, we define four inland groups: the Rif Mountains (Rif), the Middle-Atlas Mountains (Mid-Atl), the surroundings of Marrakech (Mara), and the population of Sidi-Yahia (S-Yah); as well as six groups distributed along the Atlantic coast, most of which consist in single localities, except for the northernmost (N-Atl) and southernmost (S-Atl) groups. To determine whether this pattern of genetic diversity reflects true phylogeographic breaks, we investigated whether these genetic groups differentiate along similar geographical gradients (i.e., IBD gradients) The populations were divided into two groups, analyzed separately for practical purposes: 1) the inland lineages and 2) the lineages distributed along the Atlantic coast. All scripts and distance matrices used to run these analyses are available at https://github.com/rancilhac/IBD_vs_Vicariance.

4.1 Differentiation between inland lineages

To assess whether the Rif, Mid-Atl and Mara groups are vicariant units, we analyzed the correlation between genetic and geographical distances in pairs of geographically adjacent lineages, using the statistical framework of Hausdorf & Hennig (2020). This approach relies on the null hypothesis that the two considered groups of population differentiate along the same IBD gradient. Testing this relies on a two-steps approach: firstly, we tested that the within-groups distances of both lineages followed similar regressions (hypothesis *H01* of Hausdorf & Hennig, 2020). In case this was verified, we then tested whether regressions fitted on between-groups distances only and all within-group distances are similar (hypothesis *H02* of Hausdorf & Hennig, 2020). A rejection of that hypothesis (i.e. within-groups and between-groups distances do not fit the same regression line) would support a barrier to gene flow between the tested groups. However, this test is not appropriate if *H01* is not verified. In such case, we tested whether between-groups distances follow the same regression as the within-group distances of one of the two lineages (hypothesis *H03* of Hausdorf & Hennig 2020). A verification of that hypothesis is incompatible with a barrier to gene flow. When testing *H03*, we performed two tests, *H031* and *H032*, in which the within-group

distances used for the comparison are respectively those of group 1 and group 2. A failure to reject both *H031* and *H032* would indicate inconclusive results deserving further investigation (Hausdorf & Hennig, 2020). In all cases, the significance of the difference between regression lines is assessed using a non-parametric jackknife test, as described in Hausdorf & Hennig (2020). We used this approach to test the following combinations of lineages: Rif + Atl-N, Rif + Mid-Atl, Mid-Atl + Mara, Mid-Atl + (S-Atl, O-Gha, S-Bou), Mara + S-Atl.

The case of S-Yah is more complicated to address, because this lineage is represented by a single locality, thus the tests described in the previous section are not applicable. Considering the results of previous analyses (cf. Results, Figure 2), we hypothesize that S-Yah is closely related to the Mid-Atl group. To test whether S-Yah differentiate within the same IBD gradient as Mid-Atl, we tested whether regressions within (Mid-Atl + S-Yah) and within Mid-Atl alone were similar. If this hypothesis is verified, we interpret it as an indication for an absence of barrier to gene flow between Mid-Atl and S-Yah.

All regression tests were performed using the prabclus R package (Hausdorf & Hennig, 2020) based on pairwise distances between individuals. In the specific case of S-Yah, the *regdistdiff* function had to be adapted in a custom R script to match our needs. Geographic distance matrices were calculated from coordinates using prabclus, and log transformed. Genetic distance matrices were computed from the genotype data at the nine nuclear loci using two metrics: the $\hat{a}$ statistic (an equivalent of $Fst/(1-Fst)$ calculated between individuals), calculated in Genepop (Rousset, 2008), and the shared allele distance (s), calculated in prabclus. Each analysis was repeated twice, using both measures of genetic distance.

4.2 Isolation by distance along the Atlantic coast

Because several of the genetic groups along the Atlantic coast are represented by single localities, testing for the presence of barriers to gene flow between each group is not straightforward, as explained in the case of S-Yah (above). Thus, rather than testing each putative phylogeographic breaks, we explored two hypotheses. First, considering that N-Atl is very distinct from other coastal populations on multivariate analyses and population trees (where it appears to be more related to the Rif, cf. Results), we hypothesize that there is a barrier to gene flow between N-Atl and the population adjacent to the south, M-Bou. Testing this hypothesis is relatively straightforward as the framework of Hausdorf & Hennig 2020 applies here, by testing (S-Atl, O-Gha, S-Bou, Blad-Mehd, M-Bou) and N-Atl. However, it is also possible that there are more barriers to gene flow along the coast, or none. To gain more insight into this, we first calculated a regression between genetic and logarithm of geographical distances including all coastal lineages

(to the exclusion of N-Atl in case the previous test showed that it does not evolve along the same IBD gradient as the southernmost populations). We then successively removed lineages and re-calculated the regression. In case removing populations did not significantly change the regression, we interpret this as an indication that all coastal lineages differentiate along a single IBD gradient. All tests were performed in the same set up as detailed in the previous section.

Results

1. Mitochondrial DNA supports deep phylogeographic lineages in Moroccan Acanthodactylus erythrurus

Using 776 bp of the ND4 mitochondrial gene, we recovered a relatively well supported phylogenetic tree (Figure 1, Figure S1), fully consistent with the analyses of the same data by Miralles et al. (2020). Particularly, the monophyly of the “Ibero-Moroccan” (IM) clade is fully supported (aLRT = 100), and this group is further divided into 11 phylogeographic lineages with aLRT support >90 (Figure 1, Table 1), namely (from north to south): 1) the northern Atlantic coast (N-Atl, from Souk-Tnine to Larache); 2) the northern part of the Rif mountains (N-Rif, from Martil to Ketama); 3) the southern part of the Rif (S-Rif, Tazzeka and the North of Ifrane); 4) the coastal population of Moulay-Bousselham (M-Bou); 5) the coastal populations of Blad-Dendoun and Mehdia (Blad-Mehd); 6) the inner-most parts of the Middle-Atlas mountains (Mid-Atl1, around Azrou and Ifrane); 7) the rest of the Middle-Atlas mountains (Mid-Atl2, from Oulmès to Ouaourioud and Boumia); 8) the population of Sidi-Yahia (S-Yah); 9) the coastal population of Sidi-Boumaine (S-Bou); 10) the populations surrounding Marrakech (Mara; including Sidi-Chiker and Timlite) and 11) the Southern Atlantic coast (S-Atl; from Ouled-Ghanem south to Essaouira). All lineages are geographically isolated by sampling gaps, except for lineages 3 and 6 which enter in contact in the north of Ifrane. Of the three samples from this locality, two belong to lineage 3 and one to lineage 6 (Figure S1).

The 11 lineages cluster into three main clades: i) Mara, S-Atl, S-Bou; ii) Rif N, Rif S (as well as the Iberian samples) and iii) N-Atl, M-Bou, Blad-Mehd, Mid-Atl1, Mid-Atl2. The aLRT support for these clades ranges from 77 to 100, but the relationships among them remain unresolved. Finally, S-Yah harbors a remarkably divergent mitochondrial lineage with a basal position within the Moroccan populations of the *A. erythrurus* group.

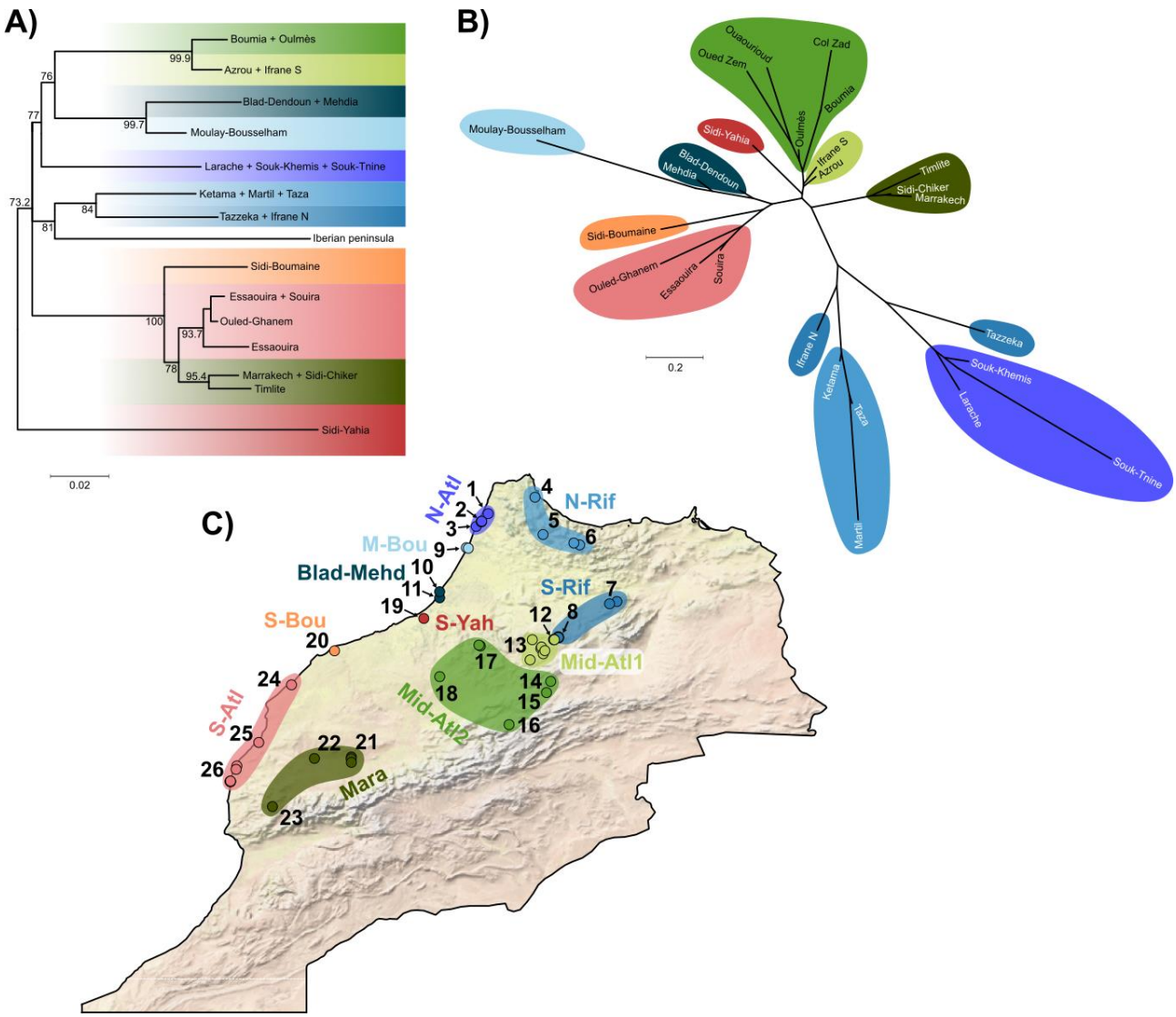

Figure 1. *A) Simplified phylogenetic tree obtained from the maximum likelihood analysis of 776 bp of the ND4 mitochondrial fragment. Outgroups were removed from the figure and terminal groups were pruned for improved graphical resolution (for full tree, see Figure S1). Numbers at the nodes show aLRT support. Color boxes show the 11 identified phylogeographic lineages. B) Neighbour-Joining population tree inferred from the pairwise Nei's distances between populations, calculated from the genotypes at nine nuclear exons, with colors referring to the mitochondrial lineages. C) Map of Morocco showing the sampled localities, with colors and texts showing the mitochondrial lineages they belong to. Localities are numbered, with details given in Table S1.*

The time-tree analyses based on the end of the Messinian crisis (Figure S2) inferred an origin of the "IM" clade at 5.99 Mya (95% credibility interval [5.36 – 7.07]). Inferences with a fixed mutation rate (Figures S3 & S4) yielded slightly older dates (6.05 and 7.12 Mya based on the mutation rate of *Podarcis* and *Gallotia* respectively), but the 95% confidence intervals were much wider. Overall, the three calibration approaches gave similar dates. However, these results should be taken with care since the analyses failed to reach proper convergence (see Supplementary materials S1 for more details).

Table 1. *Summary of the phylogeographic lineages delimited using 776 bp of the ND4 mitochondrial gene, and genotypes at nine nuclear exons. Numbers between brackets in the fourth column refer to the localities numbering on Figure 1C and Table S1.*

Mitochondrial lineages	**Nuclear lineages**	**Geographic range**	**Localities**
N-Atl	N-Atl	Northernmost Atlantic coast	Larache, Souk-Khemis, Souk-Tnine [1-3]
N-Rif	Rif	Northern part of the Rif mountains	Ketama, Martil, Bab-Taza [4-6]
S-Rif		Northern parts of Middle-Atlas	Tazzeka, Ifrane North [7-8]
M-Bou	M-Bou	Single locality on the North of the Atlantic coast south of Larache	Moulay-Bousselam [9]
Blad-Mehd	Blad-Mehd	Two nearby localities on the Atlantic coast just north of Rabat	Blad dendoun, Mehdia [10-11]
Mid-Atl1	Mid-Atl	North-eastern parts of the middle-Atlas	Ifrane South, Azrou [12-13]
Mid-Atl2		Southern and western parts of the middle-Atlas	Oulmès, Boumia, Col Zad, Ouaourioud, Oued Zem [14-19]
S-Yah	S-Yah	Single locality just south of Rabat	Sidi-Yahia [20]
S-Bou	S-Bou	Single locality on the Atlantic coast south of Casablanca	Sidi-Boumaine [21]
Mara	Mara	Populations surrounding Marrakech, in southern Morocco	Marakech, Sidi Chiker, Timlite [22-24]
S-Atl	O-Gha	Single locality on the Atlantic coast north of Essaouira	Ouled Ghanem [25]
	S-Atl	Populations of the southernmost Atlantic coast	Essaouira, Souira [26-27]

2. Multi-locus nuclear phylogeography

The ML phylogenetic analysis of the concatenation of the nine nuclear exons yielded a much less resolved topology compared to that of the mtDNA data (Figure S5). While the monophyly of the IM clade is strongly supported (aLRT = 99), most of the mitochondrial lineages are mixed and do not form reciprocally monophyletic groups. However, the following mitochondrial lineages are also supported as monophyletic in the nuclear phylogeny: Mara (aLRT = 97); S-Bou (aLRT = 81); and a clade composed of M-Bou and N-Atl (aLRT = 90) within which the former is paraphyletic but later is monophyletic (aLRT = 91). Finally, a clade comprising all N-Rif and S-Rif samples to the exclusion of one from the contact zone with Mid-Atl1 is relatively well supported (aLRT = 75), but also includes one sample from Mid-Atl1.

Multiple Correspondence Analyses (MCA) of the genotypes at the nine nuclear exons recovered a phylogeographic structure partially concordant with that of the mitochondrial data (Figure 2). With this approach, we could identify 10 phylogeographic lineages (from north to south, the names refer to those used for the mitochondrial lineages, see also Table 1): 1) N-Atl; 2) the Rif mountains (Rif, mitochondrial lineages N-Rif & S-Rif; including all three samples from IFRA-N); 3) M-Bou; 4) Blad-Mehd; 5) all the Middle Atlas populations (Mid-Atl, mitochondrial lineages Mid-Atl1 & Mid-Atl2); 6) S-Yah; 7) S-Bou; 8) Mara; 9) the coastal locality of Ouled-Ghanem (O-Gha, grouped within the S-Atl lineage in the mitochondrial data); and 10) S-Atl (to the exception of O-Gha). Overall, these groups are mostly overlapping with the mitochondrial lineages, but less straightforward to delimit due to the weak phylogenetic resolution of the nuclear markers. The locality O-Gha appears as an exception, being clustered with the other southern Atlantic populations in the mitochondrial data but quite divergent in the nuclear data (Figure 2C). Interestingly, S-Yah appears here closely related to Mid-Atl populations, from which it could be distinguished only at the finer scale (Figure 2E), contrasting with its high mitochondrial divergence.

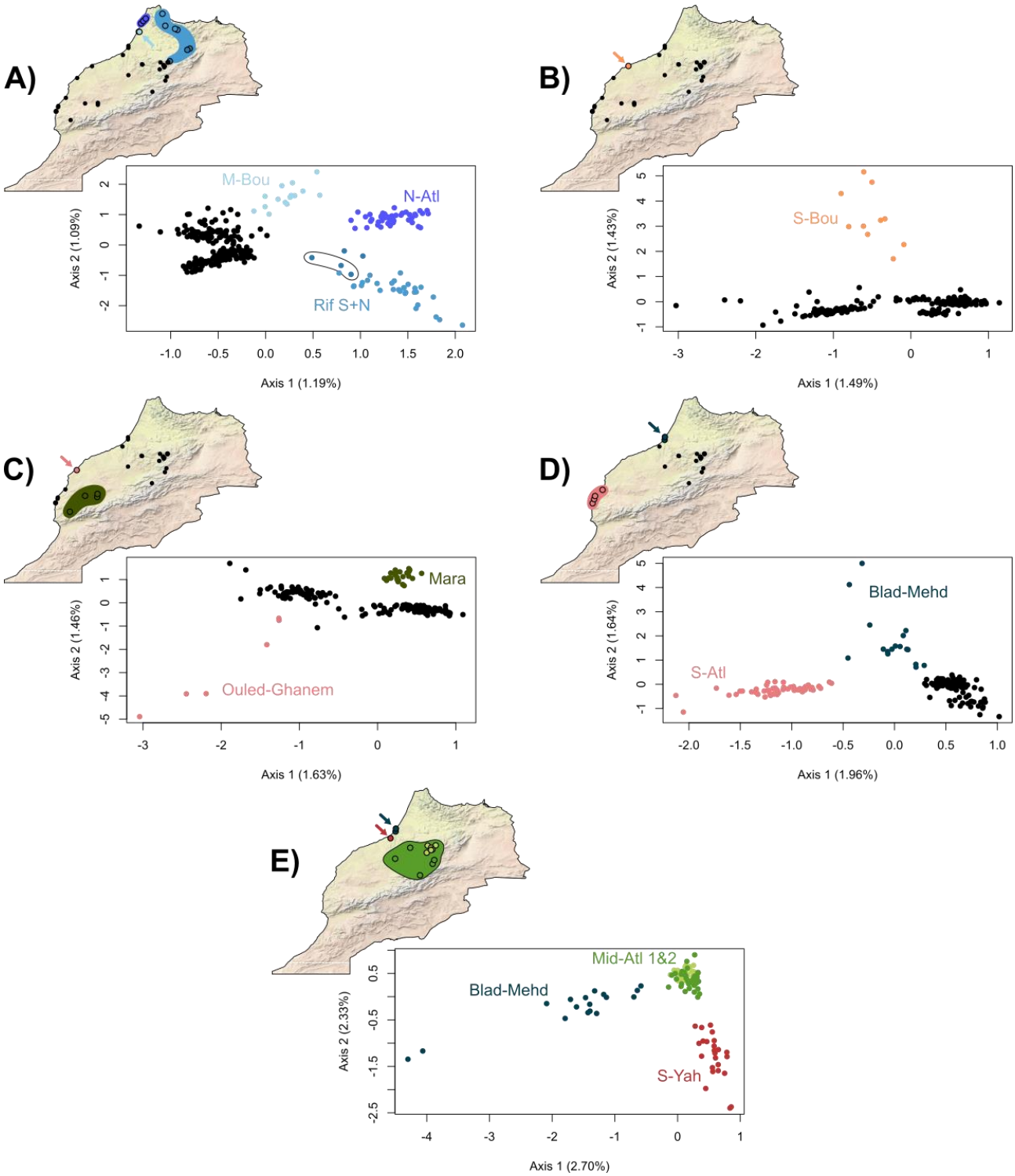

Figure 2. *Results of Multiple Correspondence Analyses calculated from the genotypes at the nine nuclear exons. The analysis was repeated five times iteratively with subsets of populations to detect finer phylogeographic structure. Colors and names correspond to the mitochondrial lineages. In A, the black ellipse highlights three samples from the contact zone between Rif and Mid-Atl. Blad-Mehd is highlighted twice in D and E to illustrate its intermediate position between the Southern Atlantic and Middle-Atlas populations.*

Population NJ trees inferred from both p-distances (Figure S4) and Nei distances (Figure 1B) supported a structure similar to that of the MCA. The Rif, N-Atl and Mara groups are separated

from the others by relatively long branches, while populations from the southern Atlantic coast, Mid-Atl and S-Yah appear closely related. Both genetic distances did not yield completely concordant topologies, especially regarding the position of M-Bou, which is alternatively placed close to N-Atl or S-Atl. On both trees, Blad-Mehd has an intermediate position, suggesting an effect of admixture between several genetic groups.

3. Isolation by Distance tests

IBD tests for pairs of inland lineages based on both $\hat{a}$ and *s* genetic distances were overall congruent (Tables 2 & 3). Figures 3 & 4 display the regression lines based on $\hat{a}$, while those based on *s* are provided in Figures S7 & S8.

Starting from the north, all tests congruently supported a barrier to gene flow between N-Atl and the Rif (Fig. 3A). This was also the case when comparing Rif and Mid-Atl (Fig. 3B), to the exception of one test: based on $\hat{a}$ genetic distances, we couldn't reject the null hypothesis *H032* that the between-groups regression differed from the within-group regression of Rif, although this difference was significative when considering the within-group regression of Mid-Atl (*H031*). The results for Mid-Atl and Mara (Fig. 3C) were also ambiguous: based on the *s* genetic distance, we found that the between-groups distances were different from the within-group distances of Mara, but other tests (i.e., *H02* based on $\hat{a}$ and *H031* based on *s*) were non-significant. On the contrary, tests between Mara and S-Atl (Fig. 3D) clearly supported a barrier to gene flow. Finally, tests based on both $\hat{a}$ and *s* unambiguously supported that S-Atl and Mid-Atl (Fig. 3E) differentiate along the same IBD gradient. While the differentiation between S-Yah and Mid-Atl could not be tested in the same framework because S-Yah is represented by a single locality, we were not able to distinguish regression lines calculated on the distances within Mid-Atl alone or Mid-Atl + S-Yah (Fig 3F), which we interpret as an indication that they evolve along the same IBD gradient.

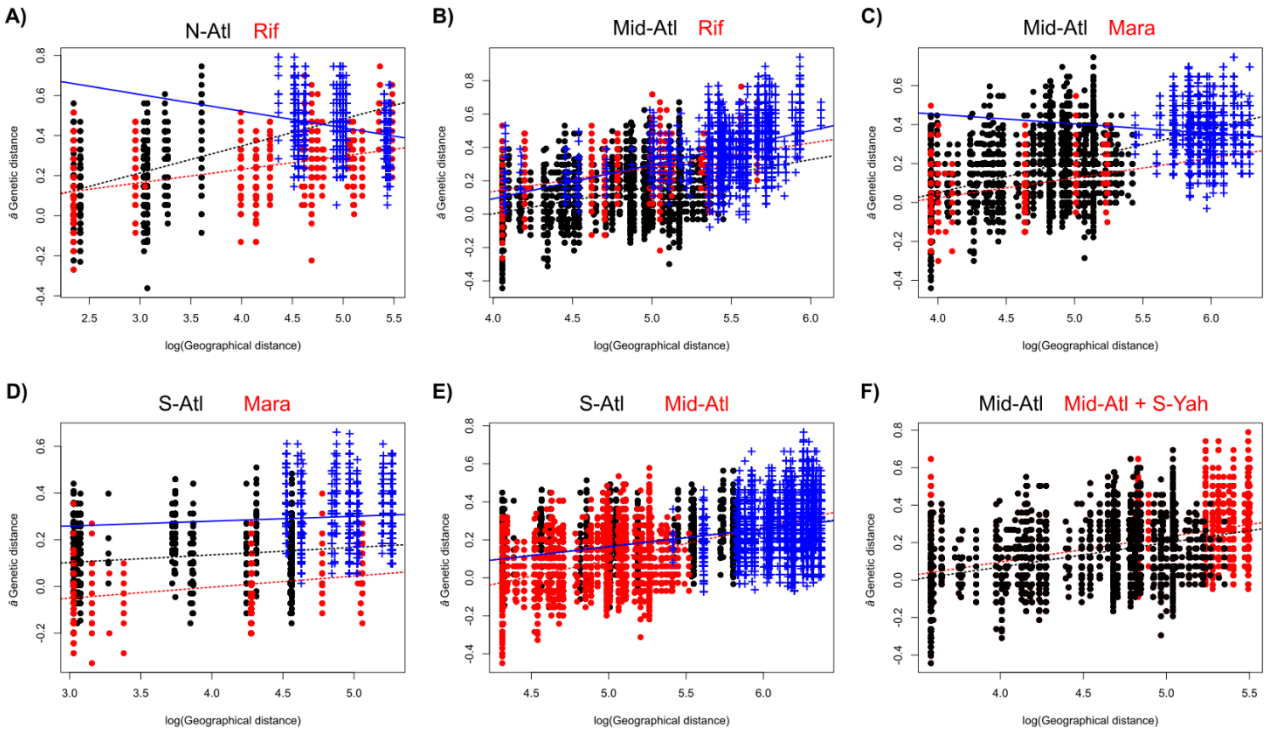

Figure 3. *Linear regressions between genetic (â) and geographical distances in pairs of geographically adjacent inland lineages. Black and red dots show within-group distances for both lineages, and corresponding dashed lines represent linear regression fitted to them. Blue crosses and solid lines show between-group distances. Between-group distances are not represented on the last graph as the test was adapted to account for zero-only geographical distances in S-Yah.*

On the Atlantic coast, tests based on both *â* and *s* genetic distances suggested a barrier to gene flow between N-Atl and the other coastal populations (Figure 4A, Table 2). However, these results remain inconclusive, as the regression on between-groups distances was significantly different from that on the within-group distances of N-Atl, but not the within-group distances of the rest of the coastal populations. Our tests further suggested that all coastal populations south of N-Atl differentiate along a single IBD gradient (Figure 4B).

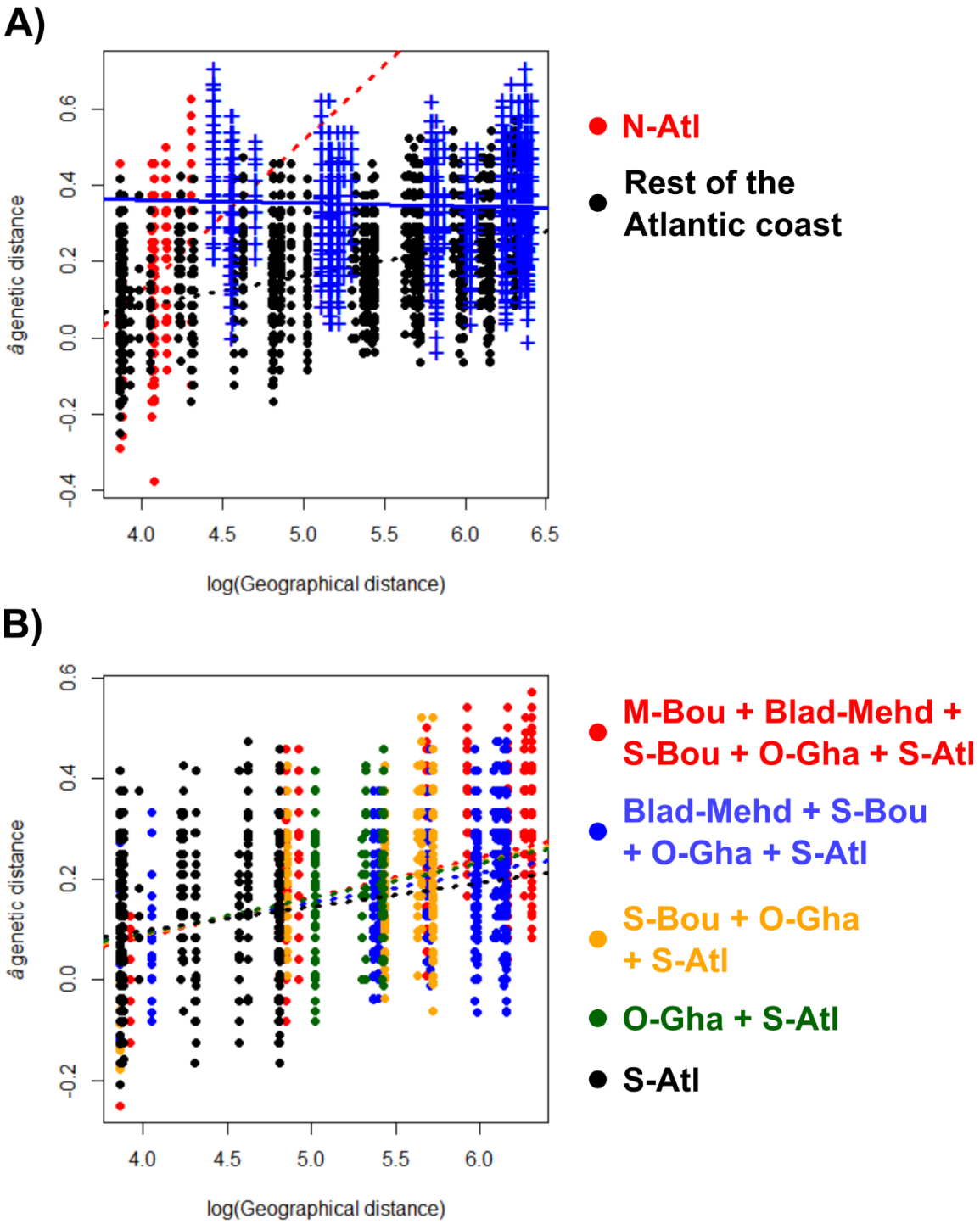

Figure 4. *Isolation by distance tests along the Atlantic coast based on the* â *genetic distance. A) Comparison between N-Atl (red) and the rest of the coastal populations (black). Blue cross show between-groups distances and lines represent the respective regressions. B) Regressions on within-group distances of several combinations of coastal populations South of N-Atl.*

Table 2. *Results of isolation by distance tests based on the comparisons of between-groups and within-groups distances. H01, H02, H031 and H032 refer to the different hypotheses tested, as explained in the text. â = â genetic distance; s = shared allele distance.*

		H01		***H02***		***H031***		***H032***	
Group 1	**Group 2**	***â***	***s***	***â***	***s***	***â***	***s***	***â***	***s***
Inland lineages									
N-Atl	Rif	0.35	0.40	5.69e-6	6.58e-6	-	-	-	-
Rif	Mid-Atl	2.96e-4	0.08	-	3.93e-7	2.17e-6	-	0.19	-
Mid-Atl	Mara	0.07	3.94e-5	0.55	-	-	0.91	-	0.025
Mara	S-Atl	1.25e-5	3.6e-13	-	-	4.14e-8	0.006	6.20e-8	6.48e-14
S-Atl	Mid-Atl	0.0005	0.0006	-	-	0.50	0.49	0.73	0.74
Coastal lineages									
N-Atl	(M-Bou + Bad-Mehd + S-Bou + O-Gha + S-Atl)	6.05e-20	1.33e-15	-	-	7.94e-13	3.02e-5	1.00	1.00

Table 3*. Results of simplified isolation by distance tests based on within-groups distances only, for the cases where the method of Henning & Hausdorf (2020) could not be applied.*

		p-value	
Group 1	**Group2**	***â***	***s***
Inland lineages			
Mid-Atl	S-Yah	0.39	0.71
Coastal lineages			
(M-Bou + Blad-Mehd + S-Bou + O-Gha + S-Atl)	(Blad-Mehd + S-Bou + O-Gha + S-Atl)	0.12	0.28

(M-Bou + Blad-Mehd + S-Bou + O-Gha + S-Atl)	(S-Bou + O-Gha + S-Atl)	1.00	0.95
(M-Bou + Blad-Mehd + S-Bou + O-Gha + S-Atl)	(O-Gha + S-Atl)	1.00	0.38
(M-Bou + Blad-Mehd + S-Bou + O-Gha + S-Atl)	S-Atl	0.14	0.80
(Blad-Mehd + S-Bou + O-Gha + S-Atl)	(S-Bou + O-Gha + S-Atl)	0.77	1.00
(Blad-Mehd + S-Bou + O-Gha + S-Atl)	(O-Gha + S-Atl)	1.00	0.32
(Blad-Mehd + S-Bou + O-Gha + S-Atl)	S-Atl	0.88	0.54
(S-Bou + O-Gha + S-Atl)	(O-Gha + S-Atl)	1.00	0.91
(S-Bou + O-Gha + S-Atl)	S-Atl	0.41	1.00
(O-Gha + S-Atl)	S-Atl	0.63	1.00

Discussion

1. Mitochondrial and nuclear markers yield contrasting phylogeographic signals in Moroccan **Acanthodactylus erythrurus**

As already highlighted by Miralles et al. (2020), our analyses of 776 bp of the ND4 mtDNA gene supported an important genetic diversity in the Moroccan populations of *A. erythrurus*, which can be divided into several deep lineages. Our analyses recovered an origin of the Ibero-Moroccan clade ("IM" clade) around 6 Mya, a relatively old age for a species-level lineage in comparison with calibrations in *Podarcis* wall lizards (Salvi et al., 2021) and other Lacertids (Garcia-Porta et al., 2019; Joger et al., 2007; Tamar et al., 2016). Furthermore, the overall congruence between calibrations based on the end of the Messinian crisis and on fixed mutation rates seem to confirm that Iberian populations have been isolated since the re-opening of the Strait of Gibraltar, around 5.4

Mya. While the deepest nodes within the IM clade are unresolved, possibly as a result of rapid diversification, or the limited informativeness of the sequence used, we could identify 11 strongly supported (aLRT > 90) clades corresponding to geographical units (subsequently referred to as mtDNA lineages) among the Moroccan populations. While some of these mtDNA lineages include several populations distributed across relatively large areas, many well differentiated monophyletic groups corresponded to single localities, especially along the Atlantic coast. This pattern could suggest the existence of many phylogeographic breaks across Morocco, resulting from allopatric diversification following gene flow disruptions. Especially, in the Middle-Atlas, two highly differentiated mitochondrial lineages (Rif and M-Atl) enter in contact with limited apparent admixture, a pattern consistent with allopatric divergence followed by secondary contact (Avise et al., 1987).

Contrastingly, our analyses of nine nuclear loci provided little support to the mtDNA lineages. Several “state-of-the-art” approaches, including phylogeny and STRUCTURE clustering, failed to recover consistent lineages, seemingly because of high levels of allele sharing among populations. This result was surprising, as we expected that populations evolving allopatrically for several millions of years would accumulate genome-wide mutations reflecting the mitochondrial structure. The high levels of allele sharing could thus indicate ongoing nuclear gene flow between lineages, or an important retention of ancestral polymorphisms due to large effective population sizes (Pamilo & Nei, 1988).

Contrary to other approaches based on more constrained models (i.e., phylogenetic trees, STRUCTURE), multivariate analyses of individual genotypes at the nine nuclear loci supported a phylogeographic structure somewhat consistent with the mtDNA lineages. One important difference was the position of the individuals from Sidi-Yahia (S-Yah, just south of Rabat), which harbors a very distinctive mitochondrial lineage, branching as sister to all the other populations. However, nuclear data suggest that these individuals are closely related to populations from the middle-Atlas and the Atlantic coast. Two scenarios can be considered to explain this discrepancy: firstly, the population of Sidi-Yahia could be closely related to its neighbours (i.e., Atlantic coast, Middle-Atlas), with a mitochondrial genome introgressed from a distinct genetic pool, absent from our analyses. Alternatively, Sidi-Yahia could be the remnant of an older lineage whose nuclear genome got swamped by that of invading populations from the middle-Atlas or Atlantic coast. Either way, it seems that this pattern is the legacy of an unsampled, possibly extinct *Acanthodactylus* lineage, whose mitochondrial genome persisted to that day. The existence of “ghost mitochondrial lineages” is well-documented in animals (for examples in reptiles, see Renoult et al., 2009; Salvi et al., 2017; Schultze et al., 2020), and their persistence can result from various mechanisms, including local

adaptation (Carpio et al., 2021). Future studies of *A. erythrurus* based on genome-wide data might be useful to understand the processes underlying such discordance.

At a larger scale, in the absence of a well resolved nDNA tree, it is difficult to assess the extent of cytonuclear discordances in *A. erythrurus*. However, genetic groups identified from mtDNA and nDNA are mostly concordant, the main difference residing in the lower phylogenetic resolution yielded by nDNA markers. We hypothesize that this pattern could result from two scenarios: 1) the mitochondrial lineages reflect phylogeographic breaks, and the poor resolution of nDNA markers is the result of ancestral polymorphism retention; or 2) the mitochondrial lineages are largely connected by nuclear gene flow. We tried to untangle these processes by studying the distribution of genetic differentiation along geographical gradients.

2. The conundrum of separating vicariance from isolation by distance

Geographical and genetic distances were overall strongly correlated in our data, highlighting the role of IBD in generating genetic diversity in *A. erythrurus*. However, IBD alone did not explain the observed phylogeographic patterns. Using linear regressions, we could identify at least four geographic groups separated by barriers to gene flow (summarized in Figure 5, discussed in-depth in the next section), within which we could not reject the null hypothesis that populations differentiate through IBD. It is interesting to note that our results suggest that similar levels of mitochondrial divergence can arise from both IBD and allopatric divergence (and even allopatric speciation, as discussed below). In that context, interpreting mitochondrial splits as evidence for phylogeographic breaks is potentially misleading and biased toward an over-estimation of the number of phylogeographic lineages. While contrasting mitochondrial and nuclear phylogeographic patterns can help determining the extant of mitochondrial diversity that can be attributed to vicariance, this issue remains very challenging in the absence of clear signal in nuclear markers. This problem is especially relevant as mitochondrial markers are still widely used to infer biogeographic patterns and species boundaries. In many cases, multi-locus studies employ only a single or few nuclear markers, whose signal gets eclipsed by that of the more variable mtDNA fragments. Hence, we argue that a thorough interpretation of mtDNA phylogeographic patterns should rely on a comparison with independently analyzed nuclear markers.

As we show here, employing statistical approaches specifically designed to test IBD patterns can yield valuable insights on the issue, compared to simpler approaches such as phylogenetic analyses. However, our analyses often yielded inconclusive results, even in a case where the signal for vicariance was apparently clear (i.e., the contact zone between Rif and Middle-Atlas, see further discussion in the next section), highlighting the sensitivity of our approach to false-negatives. Our

results highlight a large variance in genetic distances, making our tests more prone to stochastic errors. Employing more informative data (typically a larger number of markers) would likely reduce this variance, yielding more power to distinguish regression lines. More importantly, we identify the sampling scheme as a significant factor driving the results of IBD tests. For example, it seems that the inconclusive results for the Rif/Middle-Atlas tests could be due to the asymmetric sampling around the contact zone, with only a few samples available in the former group, and a large sampling gap between the contact zone and the core area. In-depth analyses of this transition by Doniol-Valcroze et al. (*in prep*) also identified sampling gaps as a limiting factor, even when considering genome-wide data. Likewise, our IBD tests for coastal populations were very limited because of "lineages" represented by single localities, which made thorough comparisons of within-groups and between-groups distances impossible. This issue could be mitigated by sampling many localities, especially to fill gaps between identified genetic groups. Finally, it is important to acknowledge that the IBD model used assumes that genetic and geographic distances follow a linear relationship (Hausdorf & Hennig, 2020). Because most genetic distance metrics are comprised between 0 and 1, this assumption is likely to be broken when genetic distances are high. Especially, intra- and between-groups regressions can be difficult to distinguish when intra-group distances are close to 1 (Hausdorf & Hennig, 2020). However, the linear relationship is expected at least in one case, when considering log-transformed geographic distances and $Fst/(1-Fst)$ as an estimate of genetic distance (and, by extension, the $\hat{a}$ genetic distance for individuals; Rousset, 1997). In our case, intra-group distances were very variable and sometimes very large, especially in S-Atl and Mid-Atl, which might result in false negatives or inconclusive results. Considering more markers (e.g., genome-wide SNPs) might improve this issue by reducing the variability across individuals and avoiding very high distances. Furthermore, violations of the linearity assumption might explain the differences in results between the $\hat{a}$ and s genetic distances.

To conclude, our results demonstrate that identifying phylogeographic breaks is far from a trivial issue and should not rely on mitochondrial markers alone. It would be interesting to further study the factors underlying the concordance of mitochondrial breaks with vicariance events, especially the effects of dispersal abilities and population sizes.

3. Mechanisms of diversification in the **Acanthodactylus erythrurus** *group*

Small terrestrial vertebrates, including lizards, often have low dispersal abilities, which sometimes result in strong phylogeographic patterns (e.g., Alfonso Silva et al., 2017; Potter et al., 2019). Hence, it is not surprising to find a high genetic diversity associated with strong IBD gradients in *A. erythrurus*. Especially, some Lacertidae species exhibit male-biased dispersal (e.g., Vignoli et al.,

2011), which could exacerbate the mitochondrial structure. However, low dispersal alone could not explain the phylogeographic pattern observed in *A. erythrurus* across Morocco, and we found at least four groups separated by barriers to gene flow (Figure 5): the northern Atlantic Coast, the Rif, the surroundings of Marrakech, and a large group including the remaining the Atlantic coast, the Middle-Atlas and Sidi-Yahia. However, several of the inferred barriers to gene flow were not unambiguously supported, highlighting a need for further investigations. Surprisingly, this was the case for Rif and Middle-Atlas, two groups that show very little to no traces of admixture at their contact zone, near Ifrane, where individuals belonging to either group occur 1 km apart. An in-depth analysis of this contact zone based on genome-wide data by Doniol-Valcroze et al. (*in prep*) showed a very steep transition, demonstrating that the Middle-Atlas and Rif are strongly differentiated groups with restricted admixture, consistent with the existence of intrinsic barriers to gene flow (Dufresnes et al., 2020; see Doniol-Valcroze et al. (*in prep*) for a more in-depth discussion of the Rif/Middle-Atlas transition). This result supports the hypothesis that ancestral polymorphism retention in our nine loci dataset, rather than widespread gene flow, blurred the phylogeographic signal, to the point that two lineages at an advanced stage of speciation were sometimes difficult to separate. Both the northern Atlantic and Marrakech plains populations showed patterns of divergence akin that of the Rif (i.e., monophyly in the nDNA tree, clear divergence in the MCAs), and might be candidates for deep intra-specific - or even species-level - lineages. Future studies employing a larger number of molecular markers will hopefully help clarifying how these groups are related, and whether they are linked by extensive gene flow or not.

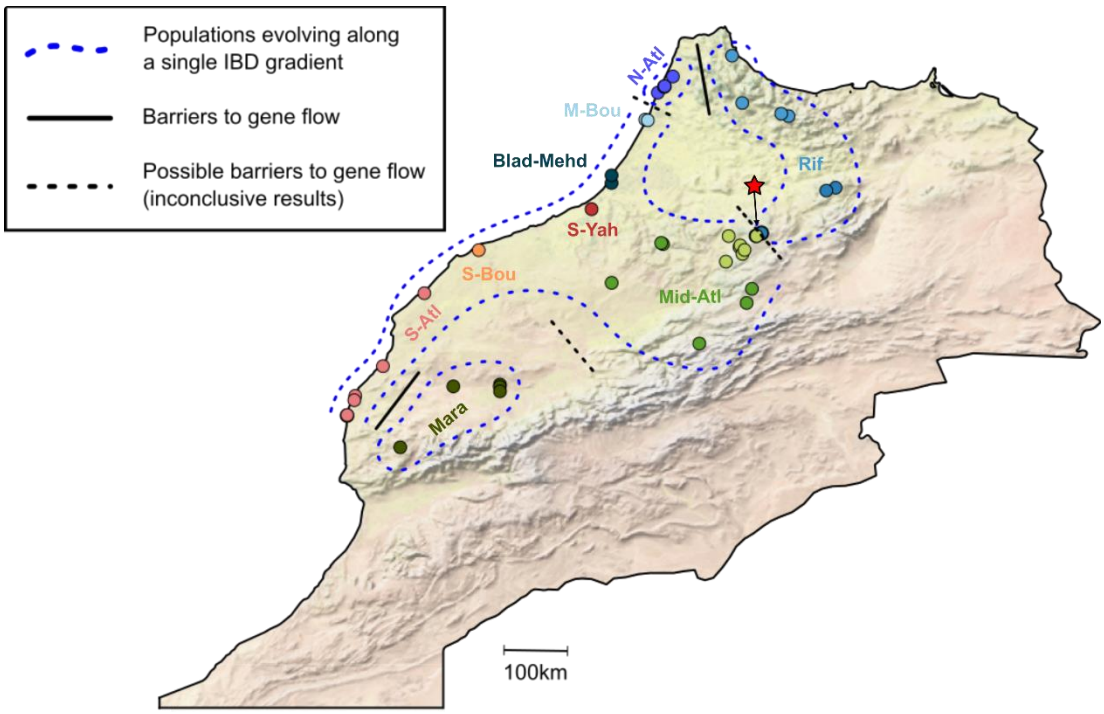

Figure 5. *Summary of the processes underlying populations differentiation in Moroccan* Acanthodactylus erythrurus. *The red star highlights the contact zone between the Rif and Middle-Atlas lineages, that is inconclusive in our analysis but is identified as a species-level barrier in Doniol-Valcroze et al. (*in prep*).*

Although the application of our approach was limited by lineages represented by single localities, our results support that populations from the Middle-Atlas, Atlantic coast to the exclusion of the three northernmost localities, and Sidi-Yahia are all part of the same IBD gradient, without barriers to gene flow detected. Alternatively, it is also possible that the Middle-Atlas and Atlantic coast represent two formerly allopatric lineages that resumed high levels of gene flow upon secondary contact, as supported by the intermediate position of some coastal populations (especially Blad-Dendoun and Mehdia) in both MCAs and population trees. In the absence of explicit measures of admixture, such scenario would be difficult to distinguish from strict IBD. Interestingly, these populations are separated into two morphologically well differentiated subspecies, *A. e. atlanticus* (Middle-Atlas) and *A. e. lineomaculatus* (Atlantic coast), the latter having even be proposed as a distinct species (Bons & Geniez, 1995). Our results suggest that this morphological variation could have arisen, or been maintained, in the face of gene flow, maybe as a result of local adaptation. Contrary to what we observed in *A. erythrurus*, phylogeographic studies of reptiles widespread in Western and Northern Morocco mostly report low levels of genetic diversity within Morocco (e.g., El-Farhati et al., 2021; Faria & Harris, 2020; Nicolas et al., 2018; Rosado et al., 2017; but see Klesser et al., 2021 and Veríssimo et al., 2016 for more complex patterns). It is interesting to note that the breaks we detected among Moroccan populations of *A. erythrurus* do not coincide with obvious present-day landscape features that could act as barriers to gene flow (e.g., large water bodies, mountain ranges). It is possible that this pattern is the legacy of populations fragmentation caused by climatic variations during the Pleistocene, when periods of high aridity occurred (Le Houérou, 1997). *Acanthodactylus erythrurus* populations, being associated with Mediterranean and mesic habitats (Fonseca et al., 2009), might have repeatedly retracted to cooler and more humid places during arid periods, such as mountain ranges or coasts (similar to “refugia-within-refugia”, Gómez & Lunt, 2007). Why the pattern of genetic diversity in *A. erythrurus* differs from other reptiles with similar distribution remains unclear but could be due to a more ancient expansion through Morocco. Our results highlight that the recent biogeographic history of taxa within Morocco, and the Maghreb in general, is still poorly understood, and call for more comparative phylogeography studies in this area, especially relying on multi-locus data sets.

Acknowledgements

The authors are grateful to Raouaa Fathalla and Patricia Sourrouille for their help with laboratory work, and to Christian Hennig for his helpful input on linear models. Part of the analyses were carried out by L.R. using resources in project 2022/22-653 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. Part of this work was carried out by using the resources of the national INRAe MIGALE (Migale bioinformatics Facility, https://doi.org/10.15454/1.5572390655343293E12) and GENOTOUL (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) bioinformatics HPC platforms, as well as the local Montpellier Bioinformatics Biodiversity (MBB, supported by the LabEx CeMEB ANR-10-LABX-04-01) and CBGP HPC platform services. R.L. was supported by the Agence Nationale de la Recherche (GENOSPACE ANR-16-CE02-0008 and INTROSPEC ANR-19-CE02-0011), and by recurrent funding from INRAE. This research was conducted in the scope of the LIA "Biodiversity and Evolution".

Conflict of interest statement

The authors state that there are no conflicts of interest regarding this research.

Authors contributions

A.M., J.C.B., M.B., P.A.C., P.G. and R.L. collected samples. P.A.C. and P.G. managed the samples. A.M., M.B. and D.M.A. generated the genetic data. A.M. performed preliminary explorations of the dataset. P.A.C, L.R. and R.L. conceived the research idea. L.R. performed the analyses presented in the manuscript with contributions from A.M., P.A.C and R.L. L.R. wrote the manuscript with contributions from P.A.C. All authors revised the manuscript.

References

Afonso Silva, A. C., Bragg, J. G., Potter, S., Fernandes, C., Coelho, M. M., & Moritz, C. (2017). Tropical specialist vs. climate generalist: diversification and demographic history of sister species of *Carlia* skinks from northwestern Australia. *Molecular Ecology*, *26*(15), 4045–4058.

Audzijonyte, A., & Vrijenhoek, R. C. (2010). When gaps really are gaps: statistical phylogeography of hydrothermal vent invertebrates. *Evolution 64*(8), 2369–2384.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, A. C. & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics*, *18*(1), 489–522.

Avise J. C., Ball R. M. Jr (1990). Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford surveys in evolutionary biology*, *7*, 45–67

Beddek, M., Zenboudji-Beddek, S., Geniez, P., Fathalla, R., Sourouille, P., Arnal, V., Dellaoui, B., Koudache, F., Telailia, S., Peyre, O., & Crochet, P. A. (2018). Comparative phylogeography of amphibians and reptiles in Algeria suggests common causes for the east-west phylogeographic breaks in the Maghreb. *PloS one*, *13*(8), e0201218.

Binks, R. M., Byrne, M., McMahon, K., Pitt, G., Murray, K., & Evans, R. D. (2019). Habitat discontinuities form strong barriers to gene flow among mangrove populations, despite the capacity for long-distance dispersal. *Diversity and Distributions*, *25*(2), 298–309.

Bons, J., & Geniez, P. (1995). Contribution to the systematics of the lizard *Acanthodactylus erythrurus* (Sauria, Lacertidae) in Morocco. *Herpetological Journal*, *5*, 271–271.

Braby, M. F., Eastwood, R., & Murray, N. (2012). The subspecies concept in butterflies: has its application in taxonomy and conservation biology outlived its usefulness? *Biological Journal of the Linnean Society*, *106*(4), 699–716.

Brito, P. H., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, *135*(3), 439–455.

Brito, J. C., Godinho, R., Martínez-Freiría, F., Pleguezuelos, J. M., Rebelo, H., Santos, X., ... & Carranza, S. (2014). Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. *Biological Reviews*, *89*(1), 215–231.

Carpio, A. J., Melo-Ferreira, J., Martinez-Haro, M., Farelo, L., Campos, R., Alves, P. C., & Acevedo, P. (2021). Iberian hares with anciently introgressed mitochondrial DNA express a marginal environmental niche. *Journal of Biogeography*, *48*(9), 2328–2336.

Dávalos, L. M., & Russell, A. L. (2014). Sex-biased dispersal produces high error rates in mitochondrial distance-based and tree-based species delimitation. *Journal of Mammalogy*, *95*(4), 781–791.

Dufresnes, C., Nicieza, A. G., Litvinchuk, S. N., Rodrigues, N., Jeffries, D. L., Vences, M., Perrin, N., & Martínez-Solano, Í. (2020). Are glacial refugia hotspots of speciation and cytonuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology*, *29*(5), 986–1000.

Dufresnes, C., Pribille, M., Alard, B., Gonçalves, H., Amat, F., Crochet, P. A., Dubey, S., Perrin, N., Fumagalli, L., Vences, M., & Martínez-Solano, I. (2020). Integrating hybrid zone analyses in species delimitation: lessons from two anuran radiations of the Western Mediterranean. *Heredity*, *124*(3), 423–438.

El-Farhati, H., Khaldi, M., Ribas, A., Hizem, M. W., Nouira, S., & Nicolas, V. (2021). Evolutionary history of the two North African hedgehogs (Mammalia: Erinaceidae) *Atelerix algirus* and *Paraechinus aethiopicus* based on phylogeography and species distribution modelling. *Vertebrate Zoology*, *71*, 799.

Faria, J. F., & Harris, D. J. (2020). Phylogeography of *Psammodromus algirus* (Linnaeus, 1758) in Morocco. *Herpetology Notes*, *13*, 1055–1061.

Fonseca, M. M., Brito, J. C., Paulo, O. S., Carretero, M. A., & Harris, D. J. (2009). Systematic and phylogeographical assessment of the *Acanthodactylus erythrurus* group (Reptilia: Lacertidae) based on phylogenetic analyses of mitochondrial and nuclear DNA. *Molecular Phylogenetics and Evolution*, *51*(2), 131–142.

Ficetola, G. F., Bonardi, A., Sindaco, R., & Padoa-Schioppa, E. (2013). Estimating patterns of reptile biodiversity in remote regions. *Journal of Biogeography*, *40*(6), 1202–1211.

Garcia-Porta, J., Irisarri, I., Kirchner, M., Rodríguez, A., Kirchhof, S., Brown, J. L., MacLeod, A., Turner, A. P., Ahmadzadeh, F., Albaladejo, G., Crnobrnja-Isailovic, J., De la Riva, I., Fawzi, A., Galan, P., Göçmen, B., Harris, D. J., Jiménez-Robles, O., Joger, U., Jovanović Glavaš, O., Karış,

M., Koziel, G., Künzel, S., Lyra, M., Miles, D., Nogales, M., Anıl Oğuz, M., Pafilis, P., Rancilhac, L., Rodriguez Concepcion, B., Sanchez, E., Salvi, D., Slimani, T., S'khifa, A., Turk Qashqaei, A., Žagar, A., Lemmon, A., Lemmon, E. M., Carretero, M. A., Carranza, S., Philippe, H., Sinervo, B., Müller, J., Vences, M., & Valero, K. C. W. (2019). Environmental temperatures shape thermal physiology as well as diversification and genome-wide substitution rates in lizards. *Nature communications*, *10*(1), 1–12.

Gómez, A., & Lunt, D. H. (2007). Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In *Phylogeography of southern European refugia* (pp. 155–188). Springer, Dordrecht.

Goudarzi, F., Hemami, M. R., Rancilhac, L., Malekian, M., Fakheran, S., Elmer, K. R., & Steinfartz, S. (2019). Geographic separation and genetic differentiation of populations are not coupled with niche differentiation in threatened Kaiser's spotted newt (*Neurergus kaiseri*). *Scientific Reports*, *9*(1), 1–12.

Hausdorf, B., & Hennig, C. (2020). Species delimitation and geography. *Molecular Ecology Resources*, *20*(4), 950–960.

Huang, J. P. (2020). Is population subdivision different from speciation? From phylogeography to species delimitation. *Ecology and Evolution*, *10*(14), 6890–6896.

Husemann, M., Schmitt, T., Zachos, F. E., Ulrich, W., & Habel, J. C. (2014). Palaearctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaearctic biota. *Journal of Biogeography*, *41*(1), 81–94.

Irwin, D. E. (2002). Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, *56*(12), 2383–2394.

Irwin, D. E. (2012). Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *The American Naturalist*, *180*(1), 35–49.

Joger, U., Fritz, U., Guicking, D., Kalyabina-Hauf, S., Nagy, Z. T., & Wink, M. (2007). Phylogeography of western Palaearctic reptiles–Spatial and temporal speciation patterns. *Zoologischer Anzeiger-A Journal of Comparative Zoology*, *246*(4), 293–313.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Klesser, R., Husemann, M., Schmitt, T., Sousa, P., Moussi, A., & Habel, J. C. (2021). Molecular biogeography of the Mediterranean *Buthus* species complex (Scorpiones: Buthidae) at its southern Palaearctic margin. *Biological Journal of the Linnean Society*, *133*(1), 166–178.

Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 593–612.

Kumar, R., & Kumar, V. (2018). A review of phylogeography: biotic and abiotic factors. *Geology, Ecology, and Landscapes*, *2*(4), 268–274.

Kuo, C. H., & Avise, J. C. (2005). Phylogeographic breaks in low-dispersal species: the emergence of concordance across gene trees. *Genetica*, *124*(2), 179–186.

Leaché, A. D., Oaks, J. R., Ofori-Boateng, C., & Fujita, M. K. (2020). Comparative phylogeography of West African amphibians and reptiles. *Evolution*, *74*(4), 716–724.

Le Houérou, H. N. (1997). Climate, flora and fauna changes in the Sahara over the past 500 million years. *Journal of Arid Environments*, *37*(4), 619–647.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452.

Miralles, A., Geniez, P., Beddek, M., Aranda, D. M., Brito, J. C., Leblois, R., & Crochet, P. A. (2020). Morphology and multilocus phylogeny of the Spiny-footed Lizard (*Acanthodactylus erythrurus*) complex reveal two new mountain species from the Moroccan Atlas. *Zootaxa*, *4747*(2).

Muir, G., & Schloetterer, C. (2005). Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular ecology*, *14*(2), 549–561.

Nicolas, V., Mataame, A., Crochet, P. A., Geniez, P., Fahd, S., & Ohler, A. (2018). Phylogeography and ecological niche modeling unravel the evolutionary history of the African green toad, *Bufotes boulengeri boulengeri* (Amphibia: Bufonidae), through the Quaternary. *Journal of Zoological Systematics and Evolutionary Research*, *56*(1), 102–116.

Palumbi, S. R., & Baker, C. S. (1994). Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, *11*(3), 426–435.

Pamilo, P., & Nei, M. (1988). Relationships between gene trees and species trees. *Molecular biology and evolution*, *5*(5), 568-583.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528.

Peters, J. L., Bolender, K. A., & Pearce, J. M. (2012). Behavioural vs. molecular sources of conflict between nuclear and mitochondrial DNA: the role of male-biased dispersal in a Holarctic sea duck. *Molecular Ecology*, *21*(14), 3562–3575.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature genetics*, *48*(1), 94–100.

Potter, S., Afonso Silva, A. C., Bragg, J. G., Catalano, S. R., Donnellan, S., Doughty, P., Scott, M. L., & Moritz, C. (2019). Contrasting scales of local persistence between monsoonal and arid biomes in closely related, low-dispersal vertebrates. *Journal of Biogeography*, *46*(11), 2506–2519.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences*, *102*(44), 15942–15947.

Renoult, J. P., Geniez, P., Bacquet, P., Benoit, L., & Crochet, P. A. (2009). Morphology and nuclear markers reveal extensive mitochondrial introgressions in the Iberian Wall Lizard species complex. *Molecular ecology*, *18*(20), 4298–4315.

Rosado, D., Rato, C., Salvi, D., & Harris, D. J. (2017). Evolutionary history of the Morocco lizard-fingered geckos of the *Saurodactylus brosseti* complex. *Evolutionary Biology*, *44*(3), 386–400.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219–1228.

Salvador, A. (1982) A revision of the lizards of the genus *Acanthodactylus* (Sauria: Lacertidae). Bonner Zoologisches Monographien, Nr. 16, *Zoologisches Forschungsinstitut und Museum Alexander Koenig*, 1–167.

Salvi, D., Pinho, C., & Harris, D. J. (2017). Digging up the roots of an insular hotspot of genetic diversity: decoupled mito-nuclear histories in the evolution of the Corsican-Sardinian endemic lizard *Podarcis tiliguerta*. *BMC evolutionary biology*, *17*(1), 1–22.

Salvi, D., Pinho, C., Mendes, J., & Harris, D. J. (2021). Fossil-calibrated time tree of *Podarcis* wall lizards provides limited support for biogeographic calibration models. *Molecular Phylogenetics and Evolution*, *161*, 107169.

Shaffer, H. B., Gidiş, M., McCartney-Melstad, E., Neal, K. M., Oyamaguchi, H. M., Tellez, M., & Toffelmier, E. M. (2015). Conservation genetics and genomics of amphibians and reptiles. *Annual Review of Animal Biosciences*, *3*(1), 113–138.

Schultze, N., Spitzweg, C., Corti, C., Delaugerre, M., Di Nicola, M. R., Geniez, P., Lapini, L., Liuzzi, C., Lunghi, E., Novarini, N., Picariello, O., Razzetti, E., Sperone, E., Stellati, L., Vignoli, L., Asztalos, M., Kindler, C., Vamberger, M., & Fritz, U. (2020). Mitochondrial ghost lineages blur phylogeography and taxonomy of *Natrix helvetica* and *N. natrix* in Italy and Corsica. *Zoologica Scripta*, *49*(4), 395–411.

Schwartz, M. K., & McKelvey, K. S. (2009). Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, *10*(2), 441.

Tamar, K., Carranza, S., Sindaco, R., Moravec, J., Trape, J. F., & Meiri, S. (2016). Out of Africa: phylogeny and biogeography of the widespread genus *Acanthodactylus* (Reptilia: Lacertidae). *Molecular Phylogenetics and evolution*, *103*, 6–18.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022–3027.

Teske, P. R., Golla, T. R., Sandoval-Castillo, J., Emami-Khoyi, A., Van der Lingen, C. D., Von der Heyden, S., ... & Beheregaray, L. B. (2018). Mitochondrial DNA is unsuitable to test for isolation by distance. *Scientific reports*, *8*(1), 1–9.

Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular ecology*, *21*(16), 3907–3930.

Tolley, K. A., Telford, N. S., Taft, J. M., Bates, M. F., Conradie, W., Makhubo, B. G., & Alexander, G. J. (2022). Taxonomic inflation due to inadequate sampling: are girdled lizards (*Cordylus minor* species complex) from the Great Karoo one and the same? *Biological Journal of the Linnean Society*, *135*(1), 1–24.

Velo-Antón, G., Martínez-Freiría, F., Pereira, P., Crochet, P. A., & Brito, J. C. (2018). Living on the edge: Ecological and genetic connectivity of the spiny-footed lizard, *Acanthodactylus aureus*, confirms the Atlantic Sahara desert as a biogeographic corridor and centre of lineage diversification. *Journal of Biogeography*, *45*(5), 1031–1042.

Veríssimo, J., Znari, M., Stuckas, H., Fritz, U., Pereira, P., Teixeira, J., Arculeo, M., Marrone, F., Sacco, F., Naimi, M., Kehlmaier, C., & Velo-Antón, G. (2016). Pleistocene diversification in Morocco and recent demographic expansion in the Mediterranean pond turtle *Mauremys leprosa*. *Biological Journal of the Linnean Society*, *119*(4), 943–959.

Vignoli, L., Vuerich, V., & Bologna, M. A. (2012). Experimental study of dispersal behaviour in a wall lizard species (*Podarcis sicula*) (Sauria Lacertidae). *Ethology Ecology & Evolution*, *24*(3), 244–256.

Wiemers, M., & Fiedler, K. (2007). Does the DNA barcoding gap exist? A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in zoology*, *4*(1), 1–16.

Zieliński, P., Nadachowska-Brzyska, K., Wielstra, B., Szkotak, R., Covaciu-Marcov, S. D., Cogălniceanu, D., & Babik, W. (2013). No evidence for nuclear introgression despite complete mt DNA replacement in the Carpathian newt (*Lissotriton montandoni*). *Molecular Ecology*, *22*(7), 1884–1903.