Received: 2 May 2023 | Revised: 18 July 2023 | Accepted: 10 August 2023

DOI: 10.1111/zsc.12628

Zoologica Scripta | WILEY

ORIGINAL ARTICLE

Combining RADseq and contact zone analysis to decipher cryptic diversification in reptiles: Insights from the Spiny-footed Lizard (Reptilia, Lacertidae)

Paul Doniol-Valcroze[1] | **Loïs Rancilhac**[2,3] | **José-Carlos Brito**[4,5,6] | **Aurélien Miralles**[1,7] | **Philippe Geniez**[8] | **Laure Benoit**[9] | **Anne Loiseau**[9] | **Raphaël Leblois**[9] | **Christophe Dufresnes**[10] | **Pierre-André Crochet**[1]

[1]CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France

[2]Zoological Institute, Technische Universität Braunschweig, Braunschweig, Germany

[3]Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

[4]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, Vairão, Portugal

[5]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, Vairão, Portugal

[6]Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal

[7]Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France

[8]CEFE, EPHE-PSL, CNRS, Université de Montpellier, Biogéographie et Ecologie des Vertébrés, Montpellier, France

[9]CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, University of Montpellier, Montpellier, France

[10]LASER, College of Biology and the Environment, Nanjing Forestry University, Nanjing, People's Republic of China

Correspondence
Paul Doniol-Valcroze, CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France.
Email: pauldoniol-valcroze@orange.fr

Funding information
Agence Nationale de la Recherche, Grant/Award Number: GENOSPACE ANR-16-CE02-0008, INTROSPEC ANR-19-CE02-0011, LabEx CeMEB ANR-10-LABX-04-01 and MOHMIE ANR-09-PEXT-0004

Abstract

Uncertainties on species taxonomy and distribution are major factors hampering efficient conservation planning in the current context of biodiversity erosion, even concerning widespread and abundant species in relatively well-studied regions. Species delimitation have long been based on phylogenetic analyses of a small number of standard markers, but accurate lineage identification through this approach can be hampered by incomplete lineage sorting, introgression or isolation by distance. In that context, analyses of introgression patterns at secondary contact zones offer an interesting alternative by allowing a direct estimation of reproductive isolation, especially when using genome-wide markers. Here, we investigated a contact zone between two genetic groups of the Spiny-footed Lizard *Acanthodactylus erythrurus* (Schinz, 1833) in Morocco, whose status as separate lineages remained disputed in previous multilocus studies. Based on thousands of genome-wide markers obtained through a RADseq approach, we confirmed that they represent distinct evolutionary lineages. Furthermore, the transition at their contact zone was very steep, with spatially restricted gene flow, highlighting

Paul Doniol-Valcroze and Loïs Rancilhac contributed equally to this study.

Christophe Dufresnes and Pierre-André Crochet contributed equally to this study.

levels of reproductive isolation consistent with species-level lineages. Our study further illustrates the power of RADseq-based studies of contact zones to understand cryptic diversity in non-model organisms.

KEYWORDS

Acanthodactylus erythrurus, admixture, cryptic speciation, hybrid zone

1 | INTRODUCTION

Uncertainties on species taxonomy and distribution (Linnaean and Wallacean shortfalls, respectively, Lomolino, 2004; Hortal et al., 2015) are major factors hampering efficient conservation planning in the current context of biodiversity erosion (e.g. Bini et al., 2006; Cardoso et al., 2011). Linnaean and Wallacean shortfalls do not only affect poorly studies organisms or remote biomes but are still widespread in conspicuous and abundant species in relatively well-studied regions. For example, in the Mediterranean biodiversity hotspot, a large fraction of the species diversity of small terrestrial reptiles remains (or remained until recently) unrecognized (e.g. Kiourtsoglou et al., 2021; Kornilios et al., 2020; Kotsakiozi et al., 2018; Psonis et al., 2018), especially but not only in North Africa and in the Middle East (e.g. Kapli et al., 2015; Kyriazi et al., 2008; Liz et al., 2021; Miralles et al., 2020; Montgelard et al., 2020; Pizzigalli et al., 2021). Efficiently addressing Linnaean and Wallacean biodiversity shortfalls relies on using accurate methods to identify independent evolutionary lineages and assess their taxonomic ranks (species or subspecies).

Phylogenetic analyses applied to standard molecular markers have significantly improved the accuracy and efficiency of species discovery and delimitation, but their results can be obscured by various processes such as incomplete lineage sorting, introgression and Isolation by Distance (IBD), especially when using limited genetic data (small number of loci and/or samples). Species delimitation models based on the multispecies coalescent have been proposed as an accurate way to objectively delimit species from multilocus data, but they can interpret strong population structure as species limits (Sukumaran & Knowles, 2017), making the use of complementary source of information necessary to assess the systematic ranks of the units identified by these models. In this context, two nonexclusive approaches can be used to resolve species boundaries: (1) using an integrative approach to assess the concordance and level of the divergences with other sets of characters (e.g. Miralles & Vences, 2013) or (2) assess reproductive isolation (RI) directly. This can be done in experimental settings where premating or postmating isolation can be evaluated or by quantifying admixture at secondary contact zones and determining how far introgression can be detected when moving towards the core range of the lineages, allowing for a direct application of the biological species concept (and by extension the unified species concept, de Queiroz, 2007, see Hillis, 2019 for a discussion in herpetology).

Analyses of introgression patterns at contact zones, based on genome-wide markers, have recently been popularized to investigate species boundaries in many taxa (e.g. Dufresnes et al., 2020; Dufresnes, Brelsford, et al., 2021; Dufresnes & Martínez-Solano, 2020; Lucek et al., 2020; Shipham et al., 2019), including squamates (e.g. Caeiro-Dias, Rocha, et al., 2021; Schield et al., 2017). The main factors limiting the applicability of this approach to poorly studied species are the need for the existence of contact zones (i.e. it does not work on allopatric taxa), its reliance on comprehensive sampling schemes and on multilocus datasets. However, recent developments of genome reduction approaches and reductions of sequencing costs now allow the generation of genome-wide phylogeographic datasets even in non-model organisms. Restriction-site Associated DNA sequencing (RADseq) and derived approaches have proven especially relevant for these studies, as they require little a priori genomic information, yield genotypes at many independent loci and at relatively low costs. Consequently, RADseq-based investigations of contact zones have been conducted in many systems to refine species boundaries (Caeiro-Dias, Brelsford, et al., 2021; Caeiro-Dias, Rocha, et al., 2021; Dufresnes et al., 2019; Dufresnes, Brelsford, et al., 2021; Dufresnes & Martínez-Solano, 2020). In this study, we use this approach to investigate a secondary contact zone of a Moroccan lizard, where previous multilocus analyses suggested low levels of admixture (Rancilhac et al., 2022).

The Spiny-footed Lizard *Acanthodactylus erythrurus* (Schinz, 1833) is widespread in the Iberian Peninsula and the Maghreb, from Morocco to Tunisia (Fonseca et al., 2009), where it occupies a wide diversity of Mediterranean and semi-arid habitats with open grounds from the Atlantic coastal plains to the High Atlas Mountains (Miralles et al., 2020). Marked variations in morphology and ecology (from coastal sands to open forests through steppes) have led to the recognition of the coastal Moroccan populations as a distinct species (*A. lineomaculatus*)

and of several subspecies within *A. erythrurus* (Bons & Geniez, 1995). However, phylogenetic reconstructions based on mitochondrial DNA (mtDNA) did not support this taxonomy and suggested considerable cryptic diversity within the group (Beddek et al., 2018; Fonseca et al., 2008, 2009; Harris et al., 2004). More recently, the analyses of one mtDNA and nine nuclear (nDNA) markers by Miralles et al. (2020) highlighted the existence of five major phylogeographic lineages: an Ibero-Moroccan clade (Iberian Peninsula and most of Morocco), a Central Algerian clade, an Algero-Tunisian clade (Tunisia and coastal populations of Eastern Algeria, including populations described under the name *A. blanci*) and two clades from the Eastern and Western High Atlas newly described as *A. lacrymae* and *A. montanus* by these authors. The Ibero-Moroccan, Central Algerian and Algero-Tunisian lineages are deeply diverged and paraphyletic, granting further investigations of their taxonomic status, although they remain aggregated under the name *A. erythrurus*, including *A. lineomaculatus*. Furthermore, the Ibero-Moroccan clade is subdivided into several evolutionary lineages, as revealed by both mtDNA and nuclear markers, although both types of markers do not entirely agree on the number and limits of these lineages: nine mitochondrial lineages were identified by Miralles et al. (2020), of which seven occur in Morocco, but only four lineages were recovered in Morocco by analyses of nine nuclear loci (Rancilhac et al., 2022). Part of this discrepancy surely stems from the high levels of allele sharing between lineages at these nuclear loci, coupled with sampling gaps and low phylogenetic resolution of these markers, but Rancilhac et al. (2022) also demonstrated that at least part of the phylogeographic structure in the Ibero-Moroccan clade is imputable to IBD. We thus still do not know how many independent lineages (evolutionary units) exist within the Ibero-Moroccan clade of *A. erythrurus*, which hampers our understanding of the species evolutionary history and precludes taxonomic revisions to formally recognize this diversity.

Highly relevant in this context, a contact zone between two highly divergent mitochondrial lineages has been identified near Ifrane, in the north of the Middle-Atlas (Miralles et al., 2020; Rancilhac et al., 2022). There, a lineage widely distributed across the Middle-Atlas (Figure 1; subsequently referred to as Mid-Atl) occurs at close range from individuals belonging to a lineage occurring in the Rif Mountain range (subsequently referred to as Rif). This contact zone coincides with a steep transition in several morphological characters (Bons & Geniez, 1995), including a fully diagnostic scalation feature (the presence or absence of a small scale inserted between the subocular plate and the lip) of unknown ecological relevance. Consequently, the two lineages have been placed into different subspecies *A. e. belli* (Rif) and *A. e. atlanticus* (Mid-Atl). They are characterized by deep mitochondrial divergences (c. 7.5 million years ago [Mya]), whereas nDNA analysis failed to recover them as reciprocally monophyletic (Rancilhac et al., 2022). Furthermore, statistical testing of IBD did not unambiguously support that these two groups are separated by a barrier to gene flow (Rancilhac et al., 2022). While these analyses are limited by the restricted dataset employed and sampling gaps, the contrast between mitochondrial and nuclear phylogeographic patterns is puzzling, and the existence of barriers to gene flow between the two groups remains to be confirmed. In case the barrier is confirmed, accurately quantifying admixture levels at the contact zone would allow to detect intrinsic barriers to gene flow (i.e. selection against hybrids) and determine whether the two lineages are different species.

In this study, we generated genome-wide data through a RADseq approach from 35 individuals along the transition zone between the Rif and Mid-Atl lineages. Based on these data, we investigate overall population structure and genetic transition at the contact zone to clarify (1) whether the Middle-Atlas and Rif populations represent two vicariant lineages as opposed to resulting from IBD processes and (2) whether these lineages are isolated enough to deserve species status.

2 | MATERIALS AND METHODS

2.1 | Sampling

A total of 35 individuals from the *A. erythrurus* complex, kept in the collection of the «Biogeography and Ecology of Vertebrates» team in Montpellier (CNRS & Ecole Pratique des Hautes Etudes, BEV-EPHE, UMR 5175-CEFE Montpellier) and already included in Rancilhac et al. (2022), were selected for this study (Table 1). These individuals were sampled during dedicated field trips in Morocco between 2004 and 2011 at 16 separate locations (Figure 1). Sampling locations are places where we stopped the car and searched for lizards on foot around the car; the maximum extent of a population is a circle with a radius of <300 m, while samples collected more than 500 m apart were allocated to different populations. Sampled locations cover two previously identified genetically and morphologically distinct lineages, Rif (Rif Massif and Northern Middle Atlas to Lake Ifrah) and Mid-Atl (Middle-Atlas south of Lake Ifrah; Bons & Geniez, 1995; Miralles et al., 2020). However, it should be noted that uncertainty remains about the ranges of these lineages since the identity of many other Moroccan and Algerian populations is unclear (Beddek et al., 2018; Miralles et al., 2020; Rancilhac et al., 2022).

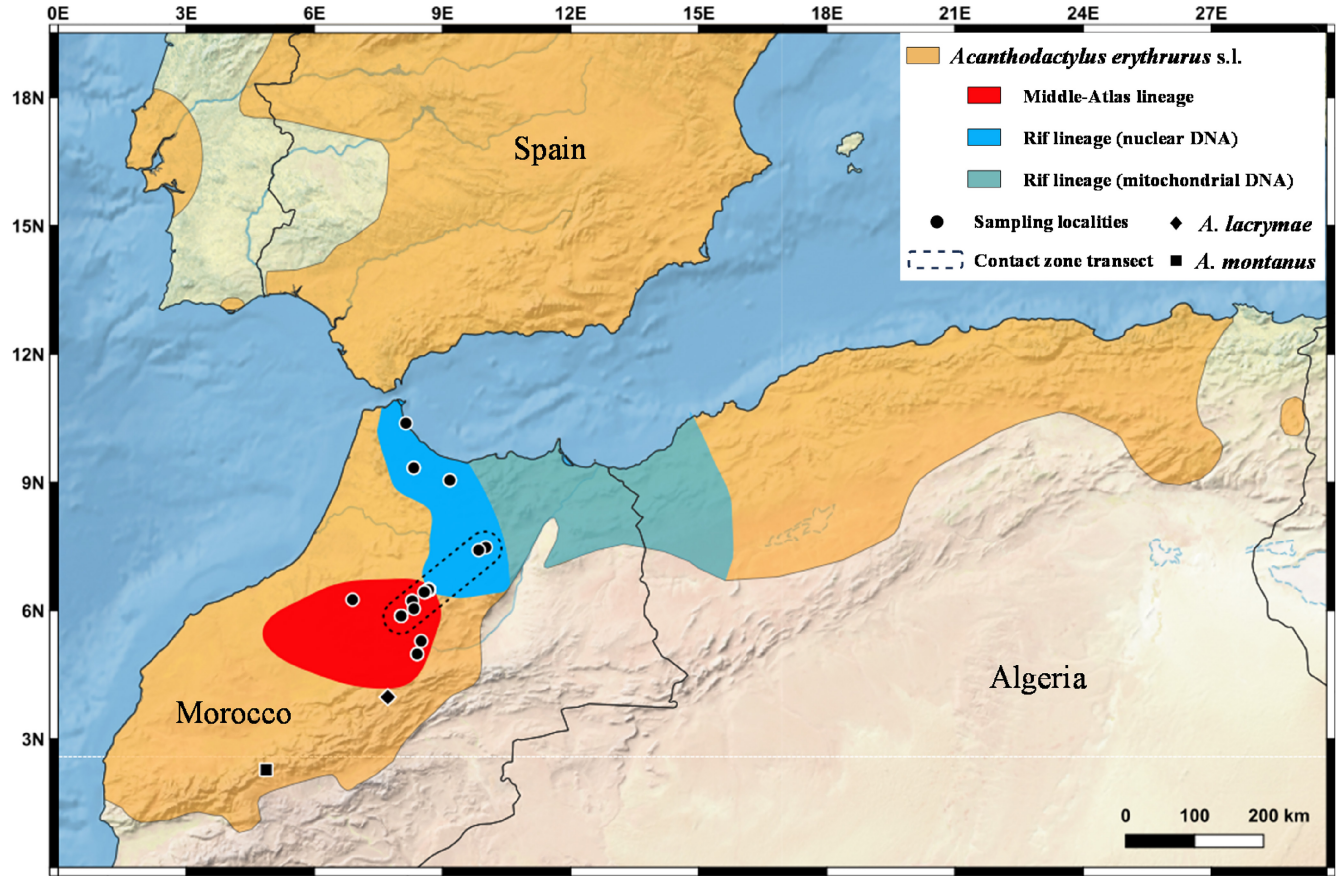

FIGURE 1 Distribution of the samples used for this study, with the dotted line showing the position of the 'transect' populations across the contact zone. The other populations are 'core' populations. The coloured areas denote the distribution range of *Acanthodactylus erythrurus* s.l. in orange (IUCN range polygon), the known range of the Middle-Atlas lineage in red (Miralles et al., 2020; Rancilhac et al., 2022) and the known range of the Rif lineage based on mitochondrial DNA haplotypes (light blue) and on nuclear DNA (dark blue; Beddek et al., 2018; Miralles et al., 2020; Rancilhac et al., 2022). The square and diamond symbols correspond to the outgroup *A. montanus* and *A. lacrymae*, respectively. Solid black lines represent country borders.

For each lineage, sampled localities are divided into two categories: localities in the core range ('core localities', 80–200 km from the identified contact zone, five samples in Rif and Mid-Atl each), where individuals are likely less affected by introgression, and localities along a transect through the contact zone ('transect localities', five samples from Rif and 20 from Mid-Atl, see Figure 1 for core and transect localities) with a special sampling effort at the contact zone itself (depicted in Figure 4d). The contact zone is located just north of Lake Ifrah (Dayet Ifrah, approximately 15 km ENE of the town of Ifrane, in the Middle Atlas (see Figure 1). In this contact zone, *A. erythrurus* is ubiquitous and in high density, and no ecological or habitat break is observable. Therefore, the somewhat limited sampling in the core contact zone (around Lake Ifrah) is only due to fieldwork and capturability constrains. One individual of each of the recently described *A. lacrymae* and *A. montanus*, sampled in the Hight-Atlas were included as outgroups (Table 1; Figure 1).

2.2 | Molecular laboratory procedures

Total genomic DNA was extracted using the Dneasy Blood and Tissue Kit (QIAGEN) following the manufacturer's recommended procedures. RADseq libraries were prepared at the Centre de Biologie pour la Gestion des Populations (CBGP) following the methods described by Baird et al. (2008) and Etter et al. (2011), with some modifications as detailed below. For each sample, 180 ng of genomic DNA was digested with the frequent cutter restriction enzyme *SbfI*. P1 adapters (synthesized by IDT, USA) barcodes (Multiplex IDentifiers: MIDs) were then ligated to the digested fragments. Low sample diversity is an intrinsic problem for restriction enzyme-based libraries (Krueger et al., 2011). To increase diversity within the library and reduce the number of phasing errors, which can be responsible for reduced sequence quality (Elshire et al., 2011), a set of 32 MIDs of 6 and 7 bp was used. The set was constructed to maximize the balance between nucleotides at each position and the MIDs differed by at least three nucleotides in order to reduce

TABLE 1 Origin and geographic position of samples along the transect and probability of STRUCTURE assignment to the Rif and Mid-Atl (MA) lineages for $K = 2$.

Sample number	Individual identifier	Lineage	Latitude	Longitude	Probability of Rif assignment	Probability of mid-Atl assignment
BEV.11873	MART3	RIF	35.6379	−5.2776	1	0
BEV.11898	KETA2	RIF	34.9373	−4.6158	1	0
BEV.11899	KETA3	RIF	34.9373	−4.6158	1	0
BEV.11830	TAZA1	RIF	35.0899	−5.1595	1	0
BEV.11831	TAZA2	RIF	35.0899	−5.1595	1	0
BEV.T11837	TZKA11	RIF	34.0708	−4.1806	1	0
BEV.T11836	TZKA10	RIF	34.1042	−4.0724	1	0
BEV.11925	**IFRA1**	RIF	33.5828	−4.9352	0.897	0.103
BEV.11926	**IFRA2**	RIF	33.5828	−4.9352	0.933	0.067
BEV.11927	**IFRA3**	RIF	33.5828	−4.9352	0.939	0.061
BEV.11235	OULM2	MA	33.4497	−6.0812	0.001	0.999
BEV.11238	OULM5	MA	33.4497	−6.0812	0	1
BEV.11972	BOUM4	MA	32.7655	−5.1081	0	1
BEV.11971	BOUM5	MA	32.7655	−5.1081	0	1
BEV.11949	CZAD1	MA	32.9297	−5.0465	0	1
BEV.11932	**IFRA4**	MA	33.5732	−4.9315	0.012	0.988
BEV.11933	**IFRA5**	MA	33.5732	−4.9315	0.004	0.996
BEV.11934	**IFRA6**	MA	33.5732	−4.9315	0.016	0.984
BEV.11935	**IFRA7**	MA	33.5732	−4.9315	0.035	0.965
BEV.11936	**IFRA8**	MA	33.5732	−4.9315	0.015	0.985
BEV.11937	**IFRA9**	MA	33.5732	−4.9315	0.024	0.976
BEV.11939	**IFRA11**	MA	33.5732	−4.9315	0.022	0.978
BEV.11940	**IFRA12**	MA	33.5732	−4.9315	0.001	0.999
BEV.11942	**IFRA13**	MA	33.5732	−4.9315	0.024	0.976
BEV.11944	**IFRA14**	MA	33.5732	−4.9315	0.005	0.996
BEV.11930	**IFRA15**	MA	33.577	−4.9349	0.025	0.975
BEV.11931	**IFRA16**	MA	33.577	−4.9349	0.076	0.924
BEV.T11809	**IFRA17**	MA	33.5449	−5.0004	0.001	0.999
BEV.11945	**IFRA18**	MA	33.5433	−4.9927	0.001	0.999
BEV.T11778	AZRU4	MA	33.4351	−5.1818	0	1
BEV.T11779	AZRU5	MA	33.4351	−5.1818	0	1
BEV.T11798	AZRU21	MA	33.2458	−5.3491	0.001	0.999
BEV.T11800	AZRU23	MA	33.2458	−5.3491	0	1
BEV.T11802	AZRU25	MA	33.2458	−5.3491	0	1
BEV.T11797	AZRU28	MA	33.3358	−5.1570	0.001	0.999
BEV.12001	*A. lacrymae*	-	32.2182	−5.5501	-	-
BEV.11794	*A. montanus*	-	31.2879	−7.3824	-	-

Note: Coordinates are in decimal degrees (WGS84). Samples in bold are from the contact zone near Lake Ifrah (Figure 4d).

misassignments due to sequencing errors. Ligated products were pooled in four different pools and sheared using an Ultrasonicator SS220 (Covaris, Inc.; 10% duty cycle, intensity equal to 5, cycle/burst equal to 200, duration of 70 s) and fragments of 300–600 bp were size-selected and purified with AMPure XP beads (Agencourt), for concentration and removal of small fragments (e.g. unligated adaptors). After end-repair and 3′ A-tailing, P2 adapters (including Illumina

primers for pair-end sequencing and 4 different 6–7 nucleotides barcodes [MIDs]) were ligated. Final enrichment PCR amplification was performed in 5 independent 25 µL wells. The initial amount of DNA was reduced to 30 ng and the number of PCR cycles was reduced to 17 to limit the amount of PCR duplicates. PCR products were pooled and homogenized, and the obtained library was quantified by qPCR using the Library Quantification Kit-Illumina/Universal (KAPA; KK4824) for sequencing preparation. It was then denatured using NaOH and diluted to 7 pM prior to clustering. Clustering and 100 nt pair-end read sequencing were performed on a single lane of a HiSeq 3000 flowcell at Toulouse GeT (Genotoul) following the manufacturer's instructions. A low-concentration spike-in (1%) of PhiX Control v3 was used as an in-lane positive control for alignment calculations and quantification efficiency. Image analyses and base calling were performed using the HiSeq Control Software (HCS) and Real-Time Analysis (RTA) software (Illumina).

2.3 | RADseq data processing and assembly

All raw reads were demultiplexed and trimmed to 80 bp using *process_radtags* in Stacks (Catchen et al., 2013), after quality checking with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). As de novo assembly of Pair-Ended RADseq data are challenging due to variable insert sizes (Etter et al., 2011), de novo assembly of the RAD loci was only performed from the R1 reads, in Ipyrad v.0.9.50 (Eaton & Overcast, 2020). Single-end sequences assembling does not allow PCR duplicates filtering, potentially impacting downstream analyses (Andrews et al., 2014, 2016), but recent evidence suggest that this issue may not be as severe as often thought (Euclide et al., 2020). The minimum depth for base calling was set to 8, and the optimal clustering threshold (CT, the minimal percentage of similarity for two sequences to be considered orthologous) was determined empirically to reduce the risk of introducing paralogs in the dataset by using an inappropriate value. To do so, the assembly was performed with CTs ranging from 0.85 to 0.99, and we observed the effect of varying the CT on (1) the number of recovered loci shared across ≥80% of the samples, (2) the number of Parsimony Informative Sites (PIS) in the assembly and (3) the proportion of missing data in the assembly. Both intrasamples CT (i.e. threshold used to cluster reads within samples) and between-samples CT (i.e. threshold used to cluster loci across individuals) were treated as the same value, as we are working at the intraspecific level, where we do not expect alleles from separate populations to be much more divergent than alleles within individuals. Ultimately, we selected a CT value that would maximize the number of PIS and loci shared by ≥80% of the samples, while minimizing the proportion of missing data. The other assembly parameters were left to default.

Two assemblies were performed: (1) Assembly 1 includes all 35 *A. erythrurus* individuals plus one sample each of *A. montanus* and *A. lacrymae* as outgroups for phylogenetic inference; (2) Assembly 2 does not include outgroups in order to mitigate the effects of loci dropout, and all loci with missing individuals were filtered out for population genetics analyses. Both assemblies were formatted to obtain complete loci sequences and unlinked phased genotypes corresponding to the SNP with highest coverage at each locus (or randomly drawn in case of equality).

2.4 | Phylogenetic reconstruction

The concatenation matrix from Assembly 1 was used to perform maximum likelihood (ML) phylogenetic inference. The analyses were conducted in IQTREE v. 1.6.8 (Nguyen et al., 2015) under a GTR+Γ substitution model, and branch support was assessed using the SH approximate likelihood ratio test (aLRT) with 1000 pseudoreplicates. The obtained topology was rooted using *A. montanus* and *A. lacrymae* as hierarchical outgroups.

2.5 | Population structure and contact zone analyses

The phylogeographic structure of our focal species was further investigated in two ways. First, the relative genetic distances among individuals were inferred using a Principal Component Analysis (PCA) computed on SNPs genotypes. This analysis was performed with the *adegenet 2.1.7* R package (Jombart, 2008) based on the unlinked SNPs from Assembly 2. Secondly, the Bayesian clustering algorithm STRUCTURE 2.3.4 (Pritchard et al., 2000) was used to infer the number of genetic groups in the data, and individual ancestries in these groups. Analyses were run with the number of groups (*K*) from 1 to 5, with four replicates for each. Most of the parameters were set to default values as advised by STRUTURE's manual (Pritchard et al., 2000). We applied the admixture model, the correlated allele frequencies option and allowed the degree of admixture alpha to be inferred from the data as recommended by Falush et al. (2003). For each run, 10,000 burn-in steps and 100,000 iterations were performed. We used Evanno et al. (2005) method implemented in STRUCTURE HARVESTER (Earl & von Holdt, 2012) to select the *K* value which maximized the DeltaK. We plotted the

distribution of STRUCTURE assignment probabilities (*Q*) of the contact zone samples to identify the modality of the hybrid zone, as bimodal hybrid zones are strongly suggestive of RI (Gay et al., 2008).

In order to quantify the geographic transition between the Rif and Mid-Atl lineages, we fitted sigmoid clines to assignment probabilities (*Q*) inferred by STRUCTURE, as well as to the allele frequencies of diagnostic SNPs (i.e. differentially fixed in the 'core populations' of the two lineages, see Figure 1) in 'transect populations', using the R package *hzar 0.2–5* (Derryberry et al., 2014). Allele frequency clines are defined by their width (*w*) and centre (*c*) and can be complemented by exponential tails to model long-range introgression ('stepped clines'). We fitted clines with width *w* and centre *c* parameters for each diagnostic SNP and the Akaike Information Criterion (AIC) was computed from the likelihood estimates to select the one most appropriate to the hybrid zone. Cline parameters result from a balance between selection against hybrids (which reduces cline width) and dispersal (which increases cline width) and can thus be informative of the processes mediating patterns of genetic introgression between hybridizing species at their secondary contact zone (Barton, 1979; Szymura & Barton, 1986). For instance, narrow clines flanked by large introgression tails can reflect strong selection against hybrids yet with the diffusion of neutral alleles far away from the contact zone (e.g. Dufresnes, Brelsford, et al., 2021). Asymmetric clines, such as stronger introgression towards one species than the other, can be signatures of hybrid zone movements (Wielstra, 2019).

3 | RESULTS

3.1 | RADseq data assembly and phylogenetic inference

After empirical optimization, the final assemblies were performed with a Clustering Threshold set to 95% (Sup Mat 1). The Assembly 1 (with outgroups and missing data included) yielded 97,765 loci (totalling 7,353,579 bp), containing 477,106 SNPs of which 87,828 are considered independent (one per locus). The Assembly 2 (no outgroups, no missing data) included 13,395 loci (totalling 1,007,162 bp), containing 69,542 SNPs of which 13,007 are considered independent. The ML phylogenetic inference of Assembly 1 yielded a robust tree (Figure 2),

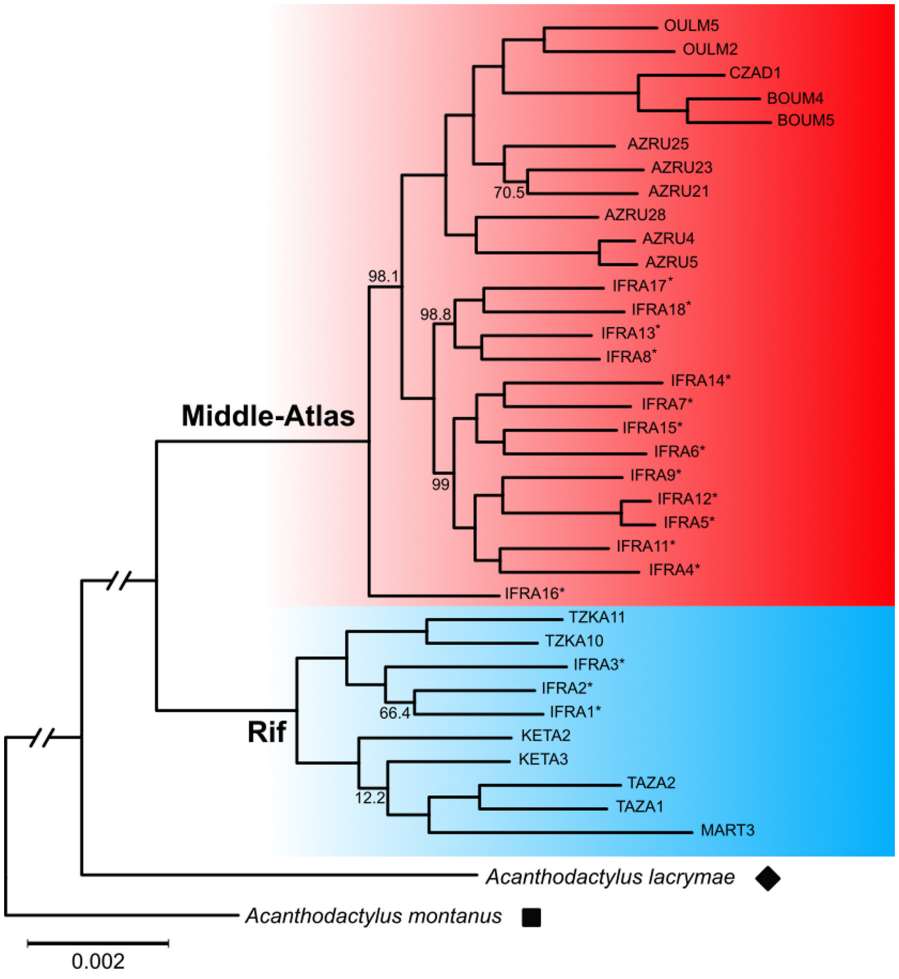

FIGURE 2 Maximum likelihood phylogenetic tree inferred from the concatenated RAD loci (7,353,579 bp). Approximate likelihood ratio test (aLRT) supports are shown at the nodes when <100. Samples originating from the contact zone are indicated by asterisks.

fully supporting the reciprocal monophyly of the Rif and Middle-Atlas populations.

3.2 | Analysis of the Rif/Middle-Atlas contact zone

The first two axes of the PCA explain respectively 11.24% and 5.10% of the total genetic variation. The Rif and Mid-Atl populations are clustered separately by PC1. PC2 shows additional structure within the Rif lineage, distinguishing individuals from the northern and southern parts (Figure 3, corresponding to the Rif-N and Rif-S mitochondrial lineages in Rancilhac et al., 2022). In both lineages, some individuals from Ifrane (near the contact zone) show a somewhat intermediate position on PC1, consistent with low but detectable levels of introgression.

STRUCTURE analyses identified $K=2$ as the best fit (Figure 4a), corresponding to the Rif and Mid-Atl lineages. In this clustering solution, all samples had an assignment probability >0.95 to either group (Figure 4b–d), except for several samples from the contact zone, which showed variable levels of genetic admixture (Table 1; Figure 4b–d). In Ifrane, the three northernmost individuals (IFRA 1–2–3) were assigned to the Rif lineage with weak introgression from Mid-Atl ($Q=.897$, .933 and .939 respectively) while the 14 remaining individuals clustered with the Mid-Atl lineage and were either pure ($Q>.95$, IFRA4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 17, 18) or weakly introgressed (IFRA16; Table 1; Figure 4b).

The best sigmoid cline model fitted on the average genomic ancestry (STRUCTURE's Q) for the transect populations did not include introgression tails. This cline modelled a very sharp transition between the Rif and Mid-Atl lineages, with a width estimated to 1260 m, and the centre located to the Lake Ifrah locality (Figure 5a). Both geographic clines (Figure 5a) and the distribution of the assignation scores within the contact zone (Figure 5b) reveal a bimodal hybrid zone with highly restricted introgression, and a lack of intermediate genotypes, that is, lack of hybrids or recently admixed individuals.

Locus-by-locus clines, fitted to 1820 diagnostic SNPs (with alleles frequencies of 0 or 1 in the core populations) mostly followed the genome average estimates (Figure 5c), with 95% of widths spanning 600–42,000 m (median: 7300 m), roughly following a Poisson distribution (Figure 5d). Most clines were asymmetric, with their centres shifted on the Rif side of the contact (negative distance positions on Figure 5c), where we lacked a proper transect sampling (Figure 4). Accordingly, 95% of the cline centres were located −19.4 to 2.5 km from the median centre.

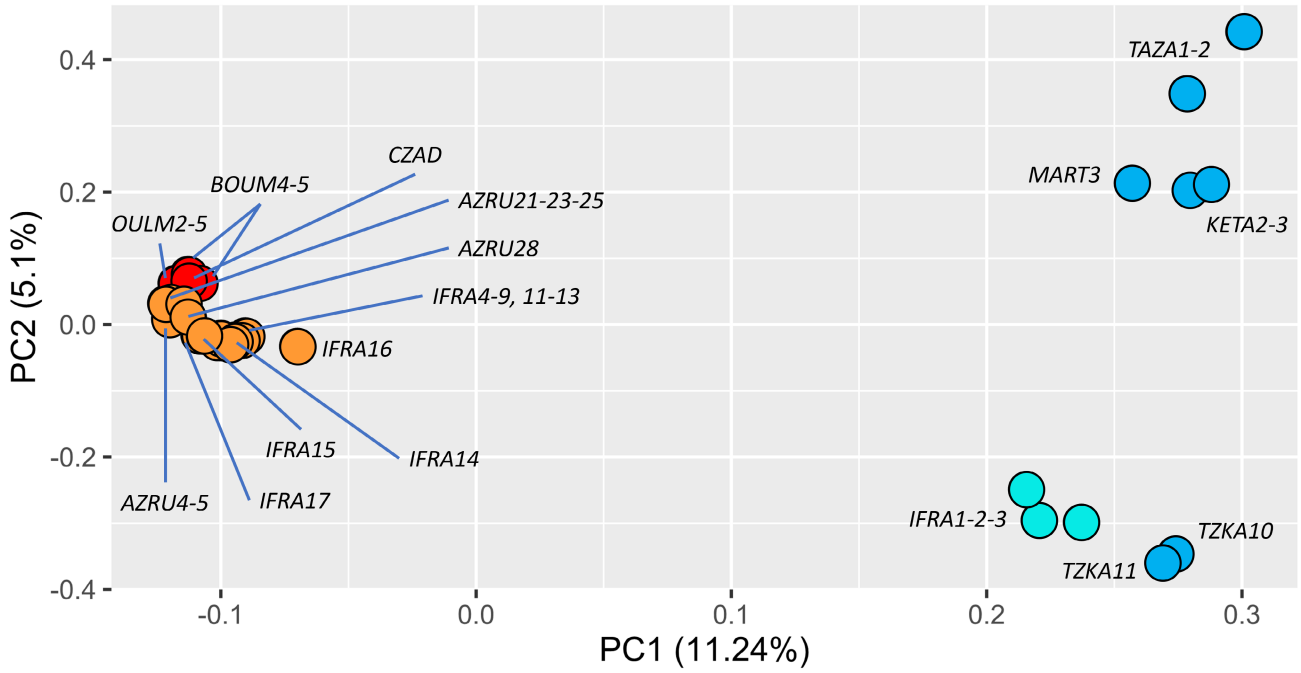

FIGURE 3 PCA performed on the independent SNPs matrix of Assembly 2, containing only *A. erythrurus* samples and no missing data. Red and orange circles: individuals from 'core' (red) and 'contact zone' (orange) populations of the Mid-Atl lineage; blue and light-blue circles: individuals from 'core' (blue) and 'contact zone' (light blue) population of the Rif lineage.

4 | DISCUSSION

4.1 | Genomics confirm a secondary contact between two species-level lineages within Moroccan populations of the *Acanthodactylus erythrurus* complex

Previous studies using multilocus data already highlighted the presence of several divergent phylogeographic lineages of *A. erythrurus* in Morocco, although the exact number of independent lineages has remained uncertain (Miralles et al., 2020; Rancilhac et al., 2022). In particular, tests of vicariance versus Isolation by Distance (IBD) in Rancilhac et al. (2022) did not unambiguously recover a phylogeographic break between the Rif and Middle-Atlas, probably because high levels of allele sharing made it difficult to estimate their levels of differentiation. Here, using thousands of genome-wide markers, we confirmed that the Rif and Mid-Atl lineages form two distinct clusters without intermediate individuals away from their contact zones, as expected with vicariant units. This is also in line with their morphology, as they differ by a diagnostic scalation feature leading to formal recognition of the Rif lineage as the subspecies *A. e. belli* in Bons and Geniez (1995). This pattern is most likely explained by an allopatric divergence scenario, where both lineages evolved separately long enough to accumulate mutations, ultimately leading to reproductive incompatibilities (Caeiro-Dias, Brelsford, et al., 2021; Caeiro-Dias, Rocha, et al., 2021; Pinho et al., 2009; Turelli et al., 2001). An alternative scenario of primary contact (evolution of divergence within a single population evolving along an IBD gradient) is theoretically less likely and seems empirically excluded by the lack of IBD within each lineage along the axis of divergence between the lineages, as shown by the PCA where the populations within each lineage do not segregate along axis 1 (separating the lineages) according to their proximity to the contact zone.

Within the contact zone, individuals from both lineages were sampled as close as 600 m apart and although introgressive hybridization occurs, as revealed by the presence of weakly admixed individuals, it occurs only at restricted geographic and genomic scales, consistent with the existence of a narrow hybrid zone which is strongly indicative of substantial selection against hybrids (Barton & Gale, 1993). Most importantly, the distribution of the hybrid index (here *Q*-values) is strongly bimodal, typical of a bimodal hybrid zone with restricted introgression (see Gay et al., 2008, figure S3). We interpret the level of divergence and reproductive isolation between the Rif and Middle-Atlas populations, as indicated by the strong genome-wide population structure and low levels of introgression, respectively, as evidence for species rank of these two lineages under the biological species concept (and by extension the unified species concept, de Queiroz, 1998, 2007; Mayr, 2000), leading us to consider that the Ibero-Moroccan clade of *A. erythrurus* comprises at least two distinct species (see also Helbig et al., 2002 for a discussion on how to interpret hybrid zones to assess species limits). Alternative explanations would require that we did not sample admixed individuals (sampling effects) or that the two lineages do not have opportunities to interbreed (lack of actual contact). We will examine these two hypotheses below.

Even if our sample size from the actual contact zone is limited to 14 individuals, their genomic background is weakly affected by introgression, excluding that we sampled an unimodal hybrid zone, in which all individuals have intermediate genotypes. We could have missed intermediate genotypes (hybrids) coexisting with parental genotypes (trimodal hybrid zone sensu Gay et al., 2008) but the probability of sampling only parental genotypes from 14 individuals in a trimodal hybrid zone implies that the hybrids would be rare. Furthermore, even if we missed hybrids, a trimodal hybrid zone with restricted introgression, as shown here, implies near complete reproductive isolation through low fitness of hybrids (Gay et al., 2008). In other words, the fact that we sampled only near-typical parental genotypes in the contact zone, even if our sample is not large, can only be compatible with a lack of admixture. Let us now examine if this lack of admixture could be due to an absence of opportunity for interbreeding.

An alternative explanation to explain the lack of admixture would be that the secondary contact is very recent, so that even without reproductive isolation, introgression could still be inconspicuous across the 600 m that currently separates the two lineages in our sample. This seems very unlikely given that no recent landscape changes that could have triggered a range expansion were recorded recently in the area and that *A. erythrurus* populations exhibiting the morphological features of the two lineages have been known in high density around the Dayet Ifrah for at least 30 years (PG pers. obs.). While we do not know of any estimate of dispersal distance in the genus *Acanthodactylus*, these are relatively large, very active animals that can run very fast and typically travel over several tens of meters during their daily activity (pers. obs.). We can thus safely exclude that restricted dispersal and recent contact prevented admixture within our contact zone samples. Another alternative explanation would be strong ecological isolation leading to low density or even the absence of lizards between our sampling points within the contact zone. This can also be safely excluded, as (i) the contact zone runs through homogenous habitats of clear oak forest with extensive grassy clearings and edges and

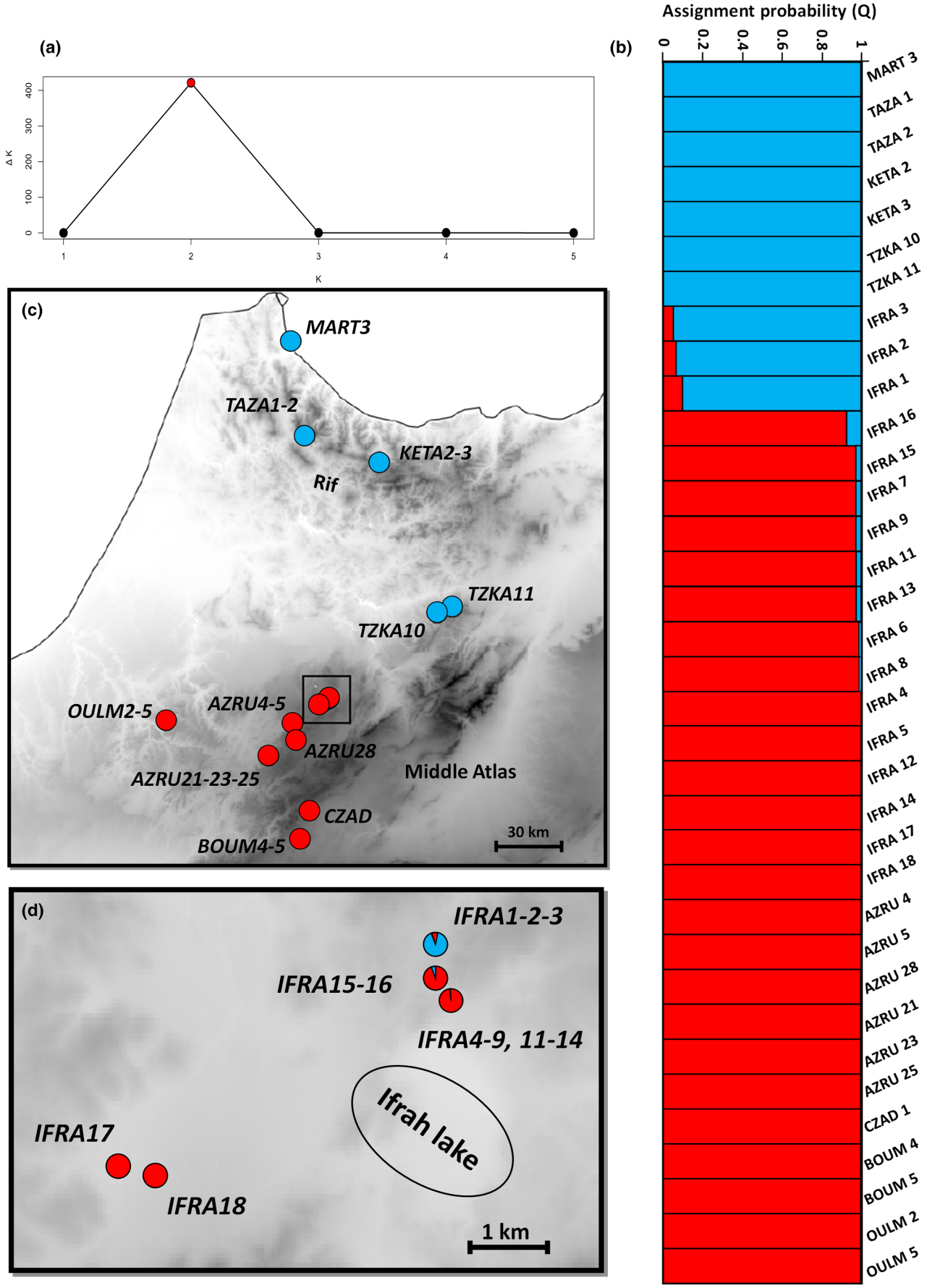

FIGURE 4 (a) Variation of DeltaK for STRUCTURE runs with *K* ranging 1–5, calculated as described by Evanno et al. (2005). (b) Assignment probabilities to the Rif (blue) and Mid-Atl (red) lineages for all individuals, obtained from a STRUCTURE run with $K=2$. Each bar represents an individual. (c) Map of the study area showing the position of the sampling sites. The pie charts represent the average assignment probabilities of the sampled populations to each lineage (Rif in blue and Middle Atlas in red) obtained from a STRUCTURE analysis with $K=2$. (d) Zoom on the contact zone around Lake Ifrah corresponding to the black square on (c).

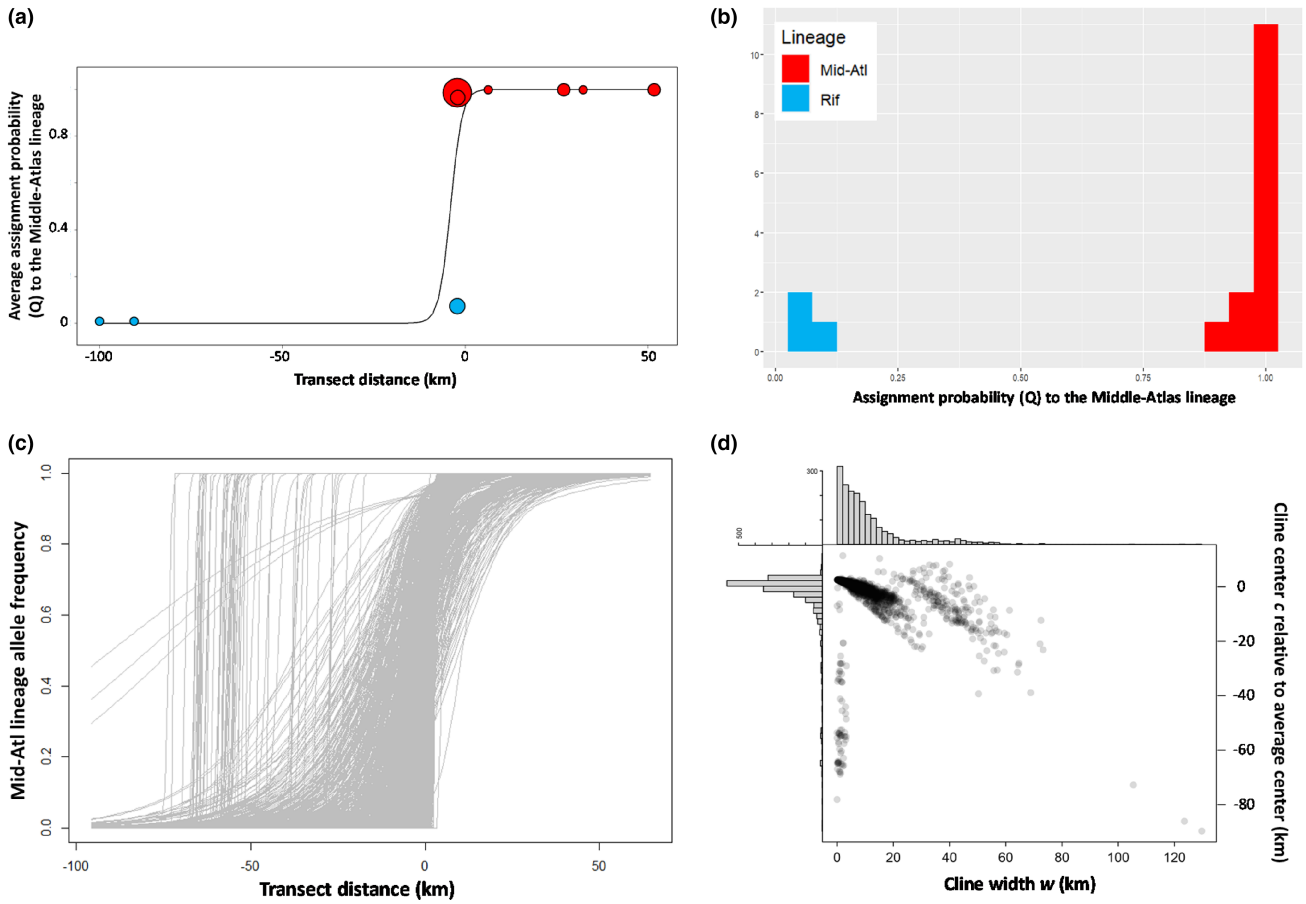

FIGURE 5 Geographic clines fitted across the Rif/Mid-Atl hybrid zone, including 10 populations along a broad north–south transect using the *hzar* package from and distribution of proportions of admixture in the hybrid zone: (a) Genomic average as the average *Q* assignment probabilities from a STRUCTURE run with $K=2$. Red circles = the Mid-Atl lineage, blue circles = Rif lineage, their diameters indicating the number of individuals per population. (b) Distribution of STRUCTURE probability of assignment to the Mid-Atl Lineage in the contact zone (samples in bold in Table 1). (c) Individual clines for 1820 diagnostic SNPs exhibiting alternative fixed alleles in core populations. (d) Width *w* and centre *c* parameters for the individual clines.

(ii) a subsequent visit of the contact zone to search for lizards between our first sampling points (after the samples used here were genotyped unfortunately) revealed that they are continuously distributed in high density all across the contact zone; we even caught individuals exhibiting the scalation features of both lineages together in the same spots. To conclude, neither sampling effect, nor recent secondary contact or lack of opportunity for interbreeding due to ecological segregation can explain our results. Only strong intrinsic reproductive isolation, due to efficient pre- or postmating barriers, can explain the distribution of the hybrid index and steep geographic clines that we observed in this contact zone.

This implies that *A. erythrurus* represents a species complex, as first illustrated by the discovery and description of *A. lacrymae* and *A. montanus* from the High Atlas by Miralles et al. (2020). The Ibero-Moroccan populations of *A. erythrurus* are further divided into several genetic groups with divergence levels seemingly comparable to that between the Mid-Atl and Rif lineages (Rancilhac et al., 2022). The generation of a large genome wide dataset covering an extensive sampling across the Maghreb is needed to accurately delimit species within the complex and determine their distributions. Although species delimitation will be more challenging for some of these lineages, given an anticipated absence of contact zones

because of their patchy distributions, analysing in parallel morphological and molecular data should allow a complete taxonomic revision of the complex. We refrain from formally naming the Rif and Middle Atlas lineages here, as we still do not know which of the various nomina available in the *A. erythrurus* complex can apply to these lineages; clarifying this will require genetic data on type specimens or on specimens from type localities of several nomina. For example, the nomen currently applied to the Rif population (*belli*) has been based on a specimen from 'Algiers', but we still do not know if the area around Algiers is inhabited by the Ibero-Moroccan clade or the Central Algerian clade of the complex (see figure 1 in Miralles et al., 2020).

4.2 | RADseq analyses of secondary contact zones as a tool for species delimitation in non-model organisms

Our study supports the relevance of genome wide data in hybrid zones to infer species boundaries, even when previous multilocus studies failed to unambiguously resolve them (here, Miralles et al., 2020; Rancilhac et al., 2022). Most of the geographic clines fitted to the genome hybrid index and to thousands of species-diagnostic SNPs all displayed a very steep transition. Moreover, the distribution of locus-specific width estimates roughly followed a Poisson distribution abutting zero, indicating that barrier loci, for example, loci potentially involved in postzygotic isolation, are scattered through the genome, a pattern characteristic of species boundaries in amphibians' hybrid zones (Dufresnes, Brelsford, et al., 2021).

However, it should be noted that many clines appeared asymmetric, with their centres being shifted towards the Rif side of the transect (Figure 5c,d). Shifts of cline centres can reflect differential introgression at specific loci caused not only by variation in selection parameters along loci (Barton, 1979; Szymura & Barton, 1986) but also by hybrid zone movement (Wielstra, 2019) or drift (Jofre & Rosenthal, 2021). It might also reflect unequal sampling along the transect: contrary to our dense and even sampling of Mid-Atl, a 100-km gap exists between the contact zone and the closest 'core' Rif populations (TZKA 10 & 11). Precision of parameter estimations in cline analyses is sensitive to sampling gaps, so part of the variation in cline centre positions on the Rif side, where we have no sample close to the contact zone, might result from such inflation of confidence intervals. Some species-diagnostic loci could truly introgress further away than our sampled 'contact zone' populations, but not as far as the reference populations, due to genuine differential introgression. This would result in cline centres shifted north more than their real positions because the precision of cline centres estimates north of the contact zone is limited by our sparse sampling. In addition, some of the markers considered as species diagnostic—because they featured allele frequency differences of 1 between our reference samples (located 100 km away from the contact zone for the Rif lineage)—may in fact reflect intraspecific structure between the Rif edge populations of the contact and the Rif reference populations. This explanation would account for the fact that several outlier loci showed cline centres located halfway between these populations, for example, ~50 km from the actual centre (Figure 5c,d). Another limitation of our study is that we only used diagnostic loci to model cline parameters. This affects the inference of variation in introgression levels among loci as loci that introgress more than the genomic average are also more likely to introgress further away from the contact zone, resulting in shared alleles in core populations that will exclude them from the set of diagnostic loci. In other words, using only diagnostic loci underestimate the amount and geographic extent of genomic introgression and historical gene flow across the contact zone. However, it does not affect the inference of current patterns of reproductive isolation, which is based on genotypic composition of individuals within the contact zones.

Note that these limitations do not affect the conclusions of our study: most of the single-locus clines were similar to that fitted to STRUCTURE's assignment probabilities, and even if it remains unclear whether some loci truly introgress further away from the contact, this would not jeopardize the specific status of the Rif and Mid-Atl lineages. Even between mostly incompatible genomes, some neutral loci are expected to diffuse far within species ranges once recombination has broken down linkage between them and the barrier loci (Barton, 1979; Polechová & Barton, 2011). Accordingly, speciation is often achieved before interspecies gene flow entirely stops, so sharp clines may be paralleled by far-ranging introgression without questioning the taxonomic status of lineages (as shown in *Bombina* toads, Dufresnes, Suchan, et al., 2021). However, the potential unreliability of the outlier clines reported here and the use of diagnostic loci only, prevents the interpretation of the genomic landscape of introgression. These shortcomings illustrate the need for careful design of hybrid zone studies, such as the inclusion of large numbers of reference specimens from different populations (in order to increase the chance to only flag lineage-diagnostic loci in analyses), dense yet even sampling along geographic transects (in order to improve the reliability of cline

estimates) and inclusion of nondiagnostic loci if the genomic landscape if introgression is to be interpreted.

ACKNOWLEDGEMENTS

Research conducted in the scope of the LIA 'Biodiversity and Evolution'. R.L. and P.-A.C. were supported by the Agence Nationale de la Recherche (MOHMIE ANR-09-PEXT-0004 and INTROSPEC ANR-19-CE02-0011) and the LabEx CeMEB (ANR-10-LABX-04-01). R.L. was also supported by the project GENOSPACE ANR-16-CE02-0008 and by recurrent funding from INRAE. A.M. was supported by the Agence Nationale de la Recherche (MOHMIE ANR-09-PEXT-0004). 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) and CBGP HPC platform services.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

ORCID

Paul Doniol-Valcroze https://orcid.org/0000-0002-2523-5979
Christophe Dufresnes https://orcid.org/0000-0002-8497-8908

REFERENCES

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Andrews, K. R., Hohenlohe, P. A., Miller, M. R., Hand, B. K., Seeb, J. E., & Luikart, G. (2014). Trade-offs and utility of alternative RADseq methods: reply to Puritz et al. *Molecular Ecology*, *23*(24), 5943–5946. https://doi.org/10.1111/mec.12964

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Barton, N. H. (1979). Gene flow past a cline. *Heredity*, *43*(3), 333–339.

Barton, N. H., & Gale, K. S. (1993). Genetic analysis of hybrid zones. In R. G. Harrison (Ed.), *Hybrid zones and the evolutionary process* (pp. 13–45). Oxford University Press.

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.

Bini, L. M., Diniz-Filho, J. A. F., Rangel, T. F., Bastos, R. P., & Pinto, M. P. (2006). Challenging Wallacean and Linnean shortfalls: Knowledge gradients and conservation planning in a biodiversity hotspot. *Diversity and Distributions*, *12*(5), 475–482. https://doi.org/10.1111/j.1366-9516.2006.00286.x

Bons, J., & Geniez, P. (1995). Contribution to the systematics of the lizard *Acanthodactylus erythrurus* (Sauria, Lacertidae) in Morocco. *Herpetological Journal*, *5*, 271–280.

Caeiro-Dias, G., Brelsford, A., Kaliontzopoulou, A., Meneses-Ribeiro, M., Crochet, P. A., & Pinho, C. (2021). Variable levels of introgression between the endangered *Podarcis carbonelli* and highly divergent congeneric species. *Heredity*, *126*(3), 463–476. https://doi.org/10.1038/s41437-020-00386-6

Caeiro-Dias, G., Rocha, S., Couto, A., Pereira, C., Brelsford, A., Crochet, P. A., & Pinho, C. (2021). Nuclear phylogenies and genomics of a contact zone establish the species rank of *Podarcis lusitanicus* (Squamata, Lacertidae). *Molecular Phylogenetics and Evolution*, *164*, 107270. https://doi.org/10.1016/j.ympev.2021.107270

Cardoso, P., Erwin, T. L., Borges, P. A., & New, T. R. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, *144*(11), 2647–2655. https://doi.org/10.1016/j.biocon.2011.07.024

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

de Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. In D. J. Howard & S. H. Berlocher (Eds.), *Endless Forms: Species and Speciation*. Oxford University Press.

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886.

Derryberry, E. P., Derryberry, G. E., Maley, J. M., & Brumfield, R. B. (2014). Hzar: Hybrid zone analysis using an R software package. *Molecular Ecology Resources*, *14*, 652–663. https://doi.org/10.1111/1755-0998.12209

Dufresnes, C., Brelsford, A., Jeffries, D. L., Mazepa, G., Suchan, T., Canestrelli, D., Nicieza, A., Fumagalli, L., Dubey, S., Martinez-Solano, I., Litvinchuck, S. N., Vences, M., Perrin, N., & Crochet, P. A. (2021). Mass of genes rather than master genes underlie the genomic architecture of amphibian speciation. *Proceedings of the National Academy of Sciences*, *118*(36), e2103963118. https://doi.org/10.1073/pnas.2103963118

Dufresnes, C., & Martínez-Solano, Í. (2020). Hybrid zone genomics supports candidate species in Iberian *Alytes obstetricans*. *Amphibia-Reptilia*, *41*(1), 105–112. https://doi.org/10.1163/15685381-20191312

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. https://doi.org/10.1038/s41437-020-0294-z

Dufresnes, C., Strachinis, I., Suriadna, N., Mykytynets, G., Cogălniceanu, D., Székely, P., Vukov, T., Arntzen, J. W., Wielstra, B., Lymberakis, P., Geffen, E., Gafny, S., Kumlutaş, Y., Ilgaz, Ç., Candan, K., Mizsei, E., Szabolcs, M., Kolenda, K., Smirnov, N.,

... Denoël, M. (2019). Phylogeography of a cryptic speciation continuum in Eurasian spadefoot toads (*Pelobates*). *Molecular Ecology*, *28*(13), 3257–3270. https://doi.org/10.1111/mec.15133

Dufresnes, C., Suchan, T., Smirnov, N. A., Denoël, M., Rosanov, J. M., & Litvinchuk, S. N. (2021). Revisiting a speciation classic: Comparative analyses support sharp but leaky transitions between Bombina toads. *Journal of Biogeography*, *48*(3), 548–560. https://doi.org/10.1111/jbi.14018

Earl, D. A., & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A. R., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011). Local de novo assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE*, *6*(4), e18561. https://doi.org/10.1371/journal.pone.0018561

Euclide, P. T., McKinney, G. J., Bootsma, M., Tarsa, C., Meek, M. H., & Larson, W. A. (2020). Attack of the PCR clones: Rates of clonality have little effect on RAD-seq genotype calls. *Molecular Ecology Resources*, *20*(1), 66–78. https://doi.org/10.1111/1755-0998.13087

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

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. https://doi.org/10.1016/j.ympev.2008.11.021

Fonseca, M. M., Brito, J. C., Rebelo, H., Kalboussi, M., Larbes, S., Carretero, M. A., & Harris, D. J. (2008). Genetic variation among spiny-footed lizards in the *Acanthodactylus pardalis* group from North Africa. *African Zoology*, *43*, 8–15. https://doi.org/10.1080/15627020.2008.11407401

Gay, L., Crochet, P. A., Bell, D. A., & Lenormand, T. (2008). Comparing clines on molecular and phenotypic traits in hybrid zones: A window on tension zone models. *Evolution: International Journal of Organic Evolution*, *62*(11), 2789–2806. https://doi.org/10.1111/j.1558-5646.2008.00491.x

Harris, D. J., Batista, V., & Carretero, M. A. (2004). Assessment of genetic diversity within *Acanthodactylus erythrurus* (Reptilia: Lacertidae) in Morocco and the Iberian Peninsula using mitochondrial DNA sequence data. *Amphibia-Reptilia*, *25*, 227–232.

Helbig, A. J., Knox, A. G., Parkin, D. T., Sangster, G., & Collinson, M. (2002). Guidelines for assigning species rank. *Ibis*, *144*(3), 518–525. https://doi.org/10.1046/j.1474-919X.2002.00091.x

Hillis, D. M. (2019). Species delimitation in herpetology. *Journal of Herpetology*, *53*(1), 3–12. https://doi.org/10.1670/18-123

Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Jofre, G. I., & Rosenthal, G. G. (2021). A narrow window for geographic cline analysis using genomic data: Effects of age, drift, and migration on error rates. *Molecular Ecology Resources*, *21*(7), 2278–2287. https://doi.org/10.1111/1755-0998.13428

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kapli, P., Lymberakis, P., Crochet, P. A., Geniez, P., Brito, J. C., Almutairi, M., Ahmadzadeh, F., Schmitz, A., Wilms, T., Pouyani, N. R., & Poulakakis, N. (2015). Historical biogeography of the lacertid lizard Mesalina in North Africa and the Middle East. *Journal of Biogeography*, *42*(2), 267–279. https://doi.org/10.1111/jbi.12420

Kiourtsoglou, A., Kaliontzopoulou, A., Poursanidis, D., Jablonski, D., Lymberakis, P., & Poulakakis, N. (2021). Evidence of cryptic diversity in *Podarcis peloponnesiacus* and re-evaluation of its current taxonomy; insights from genetic, morphological, and ecological data. *Journal of Zoological Systematics and Evolutionary Research*, *59*(8), 2350–2370. https://doi.org/10.1111/jzs.12540

Kornilios, P., Thanou, E., Lymberakis, P., Ilgaz, Ç., Kumlutaş, Y., & Leaché, A. (2020). A phylogenomic resolution for the taxonomy of Aegean green lizards. *Zoologica Scripta*, *49*(1), 14–27. https://doi.org/10.1111/zsc.12385

Kotsakiozi, P., Jablonski, D., Ilgaz, Ç., Kumlutaş, Y., Avcı, A., Meiri, S., Itescu, Y., Kukushkin, O., Gvoždík, V., Scillitani, G., Roussos, S. A., Jandzik, D., Kasapidis, P., Lymerakis, P., & Poulakakis, N. (2018). Multilocus phylogeny and coalescent species delimitation in Kotschy's gecko, *Mediodactylus kotschyi*: Hidden diversity and cryptic species. *Molecular Phylogenetics and Evolution*, *125*, 177–187. https://doi.org/10.1016/j.ympev.2018.03.022

Krueger, F., Andrews, S. R., & Osborne, C. S. (2011). Large scale loss of data in low-diversity illumina sequencing libraries can be recovered by deferred cluster calling. *PLoS ONE*, *6*(1), e16607. https://doi.org/10.1371/journal.pone.0016607

Kyriazi, P., Poulakakis, N., Parmakelis, A., Crochet, P. A., Moravec, J., Rastegar-Pouyani, N., Tsigenopoulos, C. S., Magoulas, A., Mylonas, M., & Lymberakis, P. (2008). Mitochondrial DNA reveals the genealogical history of the snake-eyed lizards (*Ophisops elegans* and *O. occidentalis*) (Sauria: Lacertidae). *Molecular Phylogenetics and Evolution*, *49*(3), 795–805. https://doi.org/10.1016/j.ympev.2008.08.021

Liz, A. V., Rödder, D., Gonçalves, D. V., Velo-Antón, G., Fonseca, M. M., Geniez, P., Crochet, P. A., & Brito, J. C. (2021). The role of Sahara highlands in the diversification and desert colonization of the Bosc's fringe-toed lizard. *Journal of Biogeography*, *48*(11), 2891–2906. https://doi.org/10.1111/jbi.14250

Lomolino, M. V. (2004). Conservation biogeography. In M. V. Lomolino & L. R. Heaney (Eds.), *Frontiers of biogeography: New directions in the geography of nature* (p. 293). Sinauer Associates, Inc.

Lucek, K., Butlin, R. K., & Patsiou, T. (2020). Secondary contact zones of closely-related Erebia butterflies overlap with narrow

phenotypic and parasitic clines. *Journal of Evolutionary Biology*, *33*(9), 1152–1163. https://doi.org/10.1111/jeb.13669

Mayr, E. (2000). The biological species concept. In Q. D. Wheeler & R. Meier (Eds.), *Species concepts and phylogenetic theory: A debate* (pp. 17–29). Columbia University Press.

Miralles, A., Geniez, P., Beddek, M., Mendez Aranda, D., 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*, 302–326. https://doi.org/10.11646/zootaxa.4747.2.4

Miralles, A., & Vences, M. (2013). New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS ONE*, *8*(7), e68242.

Montgelard, C., Behrooz, R., Arnal, V., Asadi, A., Geniez, P., & Kaboli, M. (2020). Diversification and cryptic diversity of *Ophisops elegans* (Sauria, Lacertidae). *Journal of Zoological Systematics and Evolutionary Research*, *58*(4), 1276–1289. https://doi.org/10.1111/jzs.12369

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Pinho, C., Kaliontzopoulou, A., Carretero, M. A., Harris, D. J., & Ferrand, N. (2009). Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *Podarcis carbonelli*: Evidence for limited natural hybridization and a bimodal hybrid zone. *Journal of Zoological Systematics and Evolutionary Research*, *47*(4), 368–377. https://doi.org/10.1111/j.1439-0469.2009.00532.x

Pizzigalli, C., Crochet, P. A., Geniez, P., Martínez-Freiría, F., Velo-Antón, G., & Brito, J. C. (2021). Phylogeographic diversification of the *Mesalina olivieri* species complex (Squamata: Lacertidae) with the description of a new species and a new subspecies endemic from north West Africa. *Journal of Zoological Systematics and Evolutionary Research*, *59*(8), 2321–2349. https://doi.org/10.1111/jzs.12516

Polechová, J., & Barton, N. (2011). Genetic drift widens the expected cline but narrows the expected cline width. *Genetics*, *189*(1), 227–235. https://doi.org/10.1534/genetics.111.129817

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. https://doi.org/10.1093/genetics/155.2.945

Psonis, N., Antoniou, A., Karameta, E., Leaché, A. D., Kotsakiozi, P., Darriba, D., Kozlov, A., Stamatakis, A., Poursanidis, D., Kukushkin, O., Jablonski, D., Crnibrnja-Isailović, J., Gherghel, I., Lymberakis, P., & Poulakakis, N. (2018). Resolving complex phylogeographic patterns in the Balkan Peninsula using closely related wall-lizard species as a model system. *Molecular Phylogenetics and Evolution*, *125*, 100–115. https://doi.org/10.1016/j.ympev.2018.03.021

Rancilhac, L., Miralles, A., Geniez, P., Mendez-Aranda, D., Beddek, M., Brito, J. C., Leblois, R., & Crochet, P.-A. (2022). Phylogeographic breaks and how to find them: An empirical attempt at separating vicariance from isolation by distance in a lizard with restricted dispersal. *bioRxiv*. https://doi.org/10.1101/2022.09.30.510256

Schield, D. R., Adams, R. H., Card, D. C., Perry, B. W., Pasquesi, G. M., Jezkova, T., Portik, D. M., Andrews, A. L., Spencer, C. L., Sanchez, E. E., Fujita, M. K., Mackessy, S. P., & Castoe, T. A. (2017). Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecology and Evolution*, *7*(11), 3951–3966. https://doi.org/10.1002/ece3.2996

Shipham, A., Joseph, L., Schmidt, D. J., Drew, A., Mason, I., & Hughes, J. M. (2019). Dissection by genomic and plumage variation of a geographically complex hybrid zone between two Australian non-sister parrot species, *Platycercus adscitus* and *Platycercus eximius*. *Heredity*, *122*(4), 402–416. https://doi.org/10.1038/s41437-018-0127-5

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Szymura, J. M., & Barton, N. H. (1986). Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution*, *40*(6), 1141–1159. https://doi.org/10.1111/j.1558-5646.1986.tb05740.x

Turelli, M., Barton, N. H., & Coyne, J. A. (2001). Theory and speciation. *Trends in Ecology & Evolution*, *16*(7), 330–343. https://doi.org/10.1016/S0169-5347(01)02177-2

Wielstra, B. (2019). Historical hybrid zone movement: More pervasive than appreciated. *Journal of Biogeography*, *46*(7), 1300–1305. https://doi.org/10.1111/jbi.13600

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Doniol-Valcroze, P., Rancilhac, L., Brito, J.-C., Miralles, A., Geniez, P., Benoit, L., Loiseau, A., Leblois, R., Dufresnes, C., & Crochet, P.-A. (2023). Combining RADseq and contact zone analysis to decipher cryptic diversification in reptiles: Insights from the Spiny-footed Lizard (Reptilia, Lacertidae). *Zoologica Scripta*, *00*, 1–15. https://doi.org/10.1111/zsc.12628