Assessing the role of aridity-induced vicariance and ecological divergence in species diversification in North-West Africa using Agama lizards

DUARTE V. GONÇALVES1,2,3*, PAULO PEREIRA1,2, GUILLERMO VELO-ANTÓN1, D. JAMES HARRIS1,2, SALVADOR CARRANZA3 and JOSÉ C. BRITO1,2

1CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
2Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre, 4169-007 Porto, Portugal
3Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37–49, E-08003 Barcelona, Spain

Received 24 January 2018; revised 13 April 2018; accepted for publication 13 April 2018

Diversification events in the Sahara–Sahel have mostly been attributed to regional aridification and subsequent arid–humid fluctuations, through vicariance or adaptation. However, no study has attempted to test these contrasting hypotheses. Here, we assess the importance of aridity-induced vicariance (as opposed to adaptation to new conditions) on diversification processes in North-West African Agama lizards. To test the hypothesis of vicariance as the main driver of diversification, we assessed the occurrence of the following three patterns expected to occur under the proposed scenario: (1) prevalent allopatric or parapatric distributions; (2) allopatric climatic refugia coincident with current distributions; and (3) niche similarity decreasing with increasing phylogenetic distance. We also reconstructed the centre of origin and range expansion dynamics for the Sahelian species to verify the congruence of the genetic signal with the vicariance scenario. All patterns expected from a neutral, non-adaptive niche divergence scenario were present. The diffusion models for the Sahelian species identified similar points of origin, corresponding to the areas of highest genetic diversity, topographic heterogeneity and climatic stability. Other patterns, such as mountain-isolated lineages, also indicate isolation by aridity. Our results support vicariance as the main driver of diversification in NW African Agama at both large and local scales. The importance of southern Mauritania for the conservation of biodiversity and the evolutionary process is highlighted.


INTRODUCTION

Global cooling and increased aridity in the late Miocene caused a significant increase in global coverage of deserts (Herbert et al., 2016) and created the opportunity and selective pressures for diversification and the consequent major changes in flora and fauna (Cerling et al., 1997; Badgley et al., 2008; Arakaki et al., 2011). Diversification in arid environments is commonly related to adaptation through key innovations [e.g. in cacti (Arakaki et al., 2011) or ice plants (Evans et al., 2009)] and other behavioural, morphological and physiological traits (e.g. Costa, 1995; Degen, 1997) that allow species to occupy different climatic niches and thus diversify or even radiate in arid environments. However, diversification does not occur solely through major adaptations. Some mesic taxa with higher levels of niche conservatism (the tendency to keep the ancestral niche; Wiens et al., 2010) can also diversify in arid regions. Examples range from ecological opportunity, with slow niche divergence and the progressive occupation of newly available (drier) neighbouring conditions (Evans et al., 2009), to allopatric speciation.

*Corresponding author. E-mail: duartenvg@gmail.com
attributable to the vicariant effect of aridification (Pepper et al., 2011).

In the Sahara–Sahel ecoregions of North Africa, diversification events have been attributed to the aridification of the region and the subsequent arid–humid fluctuations (Brito et al., 2014). This process started ~15 Mya and culminated with the appearance of the Sahara ~7 Mya (Zachos et al., 2001; Zhang et al., 2014), being followed by shifts between desert- and savannah-like states at regular intervals of 20,000–100,000 years (Le Houérou, 1997; Brito et al., 2014), which greatly affected species ranges (Brito et al., 2014). Vicariance and ecological adaptation have been suggested as the main processes responsible for diversification, and examples include the Sahara acting as a vicariant agent for mammals (Douady et al., 2003) and reptiles (Gonçalves et al., 2012), the adaptation to newly available arid habitats in geckos (Carranza et al., 2002), species radiations in skinks (Carranza et al., 2008), or a combination of processes (Metallinou et al., 2012). However, in most cases, biogeographical scenarios were proposed based only on the coincidence of divergence times and the periods of climatic cycles (Pliocene and Pleistocene) or empirical correlations (Brito et al., 2014). Using palaeontological data to verify past distributions and suitable climatic conditions would be an ideal solution but unrealistic because obtaining precise data on geological and palaeoecological events or fossils for the region is greatly hindered by stratigraphic discontinuities caused by the erosion associated with cyclic climatic changes (Swezey, 2003). Some studies have implemented projections of past distributions to support the inference of biogeographical scenarios concerning mesic faunal exchange across the desert (Nyári, Peterson & Rathbun, 2010; Gonçalves et al., 2018; Velo-Antón et al., 2018) or Pleistocene climatic refugia (Martínez-Freiría et al., 2017), but they are exceptions. The integration of phylogeographical data and quantitative niche comparisons allows the assessment of niche conservatism and niche divergence as a proxy for the role of climate in diversification (Hua & Wiens, 2013), but, to our knowledge, no study has attempted to test the vicariant hypothesis in the Sahara–Sahel ecoregions.

Here we combine dated phylogenies, palaeomodels and ecological niche comparisons to assess the importance of aridity-induced vicariance, as opposed to adaptation to new conditions, for species diversification in North-West Africa, using Agama lizards as a model. Previous studies have suggested aridity-induced vicariance as a major force behind the diversification of this genus in North Africa (Gonçalves et al., 2012). Although several species are distributed in the Mediterranean, Sahel and central Saharan mountains, the North African species do not present obvious interspecific morphological variation that can be related to adaptation to more or less arid habitats, thereby making them a good biogeographical model to assess the influence of humid–arid cycles and the changing landscape in shaping biodiversity patterns in the region (Gonçalves et al., 2012). Agamas are common lizards, found in arid and semi-arid habitats that include rocky outcrops, deserts and forests (Le Berre, 1989; Schleich, Kästle & Kabisch, 1996). Three major branches of the genus occur in North-West Africa, two of which have most of the species within the Sahara–Sahel (Leaché et al., 2014). One branch includes only Agama boulenegeri Lataste, 1886, and the other includes most species of the ‘Northern Africa radiation’ (Leaché et al., 2014): Agama impalearis Boettger, 1874; Agama boueti Chabanaud, 1917; and Agama tassiliensis Geniez, Padial & Crochet, 2011. The third branch includes the Agama agama species group (Leaché et al., 2017), distributed along the southern fringes of the Sahel, but most of its species occur south of the study region.

If aridity-induced vicariance was the main driver of diversification of mesic taxa, three main patterns are expected. The first pattern is prevalent allopatric or parapatric distributions for species and lineages (Wiley, 1988) attributable to recurrent historical constraints to dispersal linked to humid–arid cycles. If vicariant events were non-existent and ecological divergence and adaptation were predominant, phylogenetically close groups should occur more readily in sympathy (Wiens & Graham, 2005). The second pattern is that of allopatric climatic refugia. Stable climatic areas (potential refugia) should be mostly allopatric and coherent with present distributions (Avise, 2000). Conversely, if allopatry/parapatry was a product of adaptation to different climates (niche divergence), coincidence of climate stability with genetic structure should not be prevalent. The third pattern is niche similarity. If aridity-induced vicariance is prevalent, closely related clades should have similar climatic niches, as opposed to ecological adaptation to different conditions, where closer clades are expected to have more distinct niches (Wiens & Graham, 2005). To test these hypotheses, we compared niches at the intraspecific and interspecific levels. Intraspecific comparisons were focused on A. boulenegeri and A. boueti, both occurring at similar latitudes along the West Sahel and the species for which the sampling was more thorough.

Considering that the humid–arid cycles mostly translated into north–south movements of the climatic regions, particularly in the southern regions of the desert, it would also be expected that under a climate-driven scenario the species ranges would shift accordingly. To assess the congruence of the genetic signal with this pattern, we estimate the geographical ancestral origin of the lineages and reconstruct range dynamics for the species present in the Sahel–Sahara fringe (A. boueti and A. boulenegeri), using genetic
diversity measures and continuous diffusion models that integrate genetic and spatial data.

MATERIAL AND METHODS

PHYLOGENETIC ANALYSES

Sampling and study area
A total of 718 samples of *Agama* were available for this study (Fig. 1). Given the uneven spatial distribution of samples, an initial selection of samples best representing the geographical distribution of each species was carried out. For *A. impalearis*, as the distribution of genetic lineages had already been assessed previously (Brown, Suárez & Pestano, 2002), only a few additional samples were sequenced. A total of 72 were selected for *A. boueti*, 233 for *A. boulengeri*, seven for *A. impalearis* and 11 for *A. tassiliensis*. The limited sample size of *A. tassiliensis* was attributable to the difficulty of acquiring samples from the area of occurrence (see Brito et al., 2014).

DNA extraction and amplification
DNA was extracted from ethanol-preserved tissue using a commercially available kit (Easy-Spin). Amplifications were performed in 10 μL of 2× MyTaq Mix and 0.5 μM of each primer. The PCR conditions were as follows: pre-denaturation at 95 °C (15 min); 40 cycles with denaturing at 95 °C (30 s), annealing range of 48–52 °C (40 s), and extension at 72 °C (45 s); and final extension at 60 °C for 12 min. Some samples required minor adjustments to conditions. Four genes were amplified: ND4 (with tRNAs), 16S rRNA, NTF3 and c-mos. For the first three, we used primers from Arevalo, Davis & Sites (1994), Palumbi et al. (1991) and Wiens et al. (2008), respectively. For c-mos, new primers were designed based on S77/S78 of Lawson et al. (2005) and named A77 (5′-AATAGACTGGAAACAGTTGTG-3′) and A78 (5′-CCTTAGGTTAATTCTCTCACCT-3′). The PCR products were sequenced using cycle sequencing on an automated sequencer.

Phylogenetic analyses
Additional sequences of *Agama* species and outgroups were selected based mostly on Leaché et al. (2014) and retrieved from GenBank (Supporting Information, Table S1.1). DNA sequence alignments were inferred with MAFFT v7 (Katoh & Standley, 2013), with default parameters and the Q-INS-i option, then proofread and edited by eye. Coding genes (ND4, NTF3 and c-mos) were translated, and no unexpected stop codons were found. Independent maximum likelihood (ML) trees were inferred for each marker using RAxML v8.1.21 (Stamatakis, 2014), and no topological incongruences were found. Concatenated duplicate haplotypes (i.e. identical across all markers) were removed from the alignment when calculating the phylogenetic trees, in order to decrease computational load. The most appropriate models of molecular evolution and best-fit partitioning scheme were selected using PartitionFinder v.1.1.1 (Lanfear et al., 2012). Settings were: linked branch lengths, BEAST models, Bayesian information criterion model selection, and all partition schemes searched. An initial partition scheme by gene (ND4 and tRNAs separated) was used.

Nuclear haplotypes were inferred using PHASE 2.1 (Stephens, Smith & Donnelly, 2001), implemented in DnaSP. PHASE ran for 104 iterations, with a burn-in value of 1000 and a thinning interval of five. Haplotype networks were produced using TCS v1.21 (Clement, Posada & Crandall, 2000), with gaps treated as missing data and otherwise default parameters. Graphical representations were obtained using tcsBU (Santos et al., 2015). Uncorrected p-distances (proportion of nucleotide sites at which two sequences being compared are different) within and among species and lineages were calculated in MEGA6 (Tamura et al., 2013) for each mitochondrial marker.

Phylogenies were inferred with Bayesian inference (BI) and ML methods using MrBayes v3.2.6 (Ronquist et al., 2012) and RAXML v8.1.21 through RaxmlGUI 1.5b1 (Silvestro & Michalak, 2013), respectively. Gene partitions were applied according to PartitionFinder results. MrBayes ran for 2 × 107 generations in two independent runs, sampling every 1000 generations. Parameters of sequence evolution (statefreq, revmat, shape and pinvar) were unlinked for all partitions, and the overall rate (ratepr) variable among them. Burn-in was determined using Tracer v1.6 (Rambaut et al., 2014), upon stabilization of log-likelihood, average standard deviation of split frequencies, and ESS for all the parameters. RAxML used the same partition scheme and the GTR + G model (General Time Reversible model with Gamma distributed rate variation among sites), with ten random addition replicates and 1000 thorough bootstrapping replicates. To support lineage delimitation for further analyses, we used the species delimitation method bPTP (species.h-its.org), an update of PTP (Zhang et al., 2013). This was not intended to delimit potential species, but rather to avoid an ad hoc intraspecific lineage delimitation with lineages at different phylogenetic depths. bPTP ran with both ML and BI trees, using mitochondrial markers and removing outgroups.

Time calibration
In order to perform the time calibration, representatives of each species and supported lineage were selected (considering sequence length) based on
the results of the ML and BI phylogenetic analyses (Supporting Information, Fig. S2.2), resulting in a total of 118 sequences, including outgroups (Supporting Information, Table S1.1). We used the same calibration scheme as Leaché et al. (2014), who used the 62.5 Mya Calotes–Phrynocephalus divergence, calculated using 11 fossil calibrations of squamates (Wiens, Brandley & Reeder, 2006), and the 16.4–19.6 Mya Xenagama–Pseudotrachelus divergence, inferred from pairwise sequence divergence in an agamid species (Macey et al., 2006). Analyses were run using BEAST v1.8.3 (Drummond et al., 2012) in the CIPRES gateway (Miller, Pfeiffer & Schwartz, 2011). We performed three independent runs of $5 \times 10^7$ generations, sampling every 5000, using unlinked substitution and clock models, and an uncorrelated lognormal relaxed clock (Drummond et al., 2006). We used a Yule speciation tree prior (Yule, 1925; Gernhard, 2008), and treated ambiguities in nuclear sequences as informative sites (setting the option useAmbiguities as ‘true’ in the XML file). Burn-in was determined using Tracer v1.6 (Rambaut et al., 2014). Runs were combined with LogCombiner, and a maximum credibility tree was generated with TreeAnnotator (both in the BEAST package).

Genetic diversity and spatial diffusion models
Sequence and nucleotide diversity measures and demographic statistics were calculated in DnaSP v.5.10.1 (Librado & Rozas, 2009) for all markers. Spatially explicit representations of genetic diversity were produced using a predefined radius search around each sample in order to create pseudo-populations, from which diversity was estimated. The method was described by Veríssimo et al. (2016). The resulting diversity scores could then be spatially interpolated, in this case using the Kriging function in ArcMap. We used a radius of ~100 km, in order to include the isolated samples that would otherwise be ignored for having no neighbours.

Continuous diffusion phylogeographical models were produced using BEAST. These models use geographical coordinates of samples as continuous traits to reconstruct the geographical origin and spatially explicit expansion of organisms across a continuous landscape over time, and have already been implemented for predicting the origin of lineages’ ancestors (Veríssimo et al., 2016; Gutiérrez-Rodríguez, Barbosa & Martínez-Solano, 2017; Leaché et al., 2017). Two independent models were generated for A. boueti and A. boulengeiri, using all the unique concatenated haplotypes from each species. A Cauchy relaxed random walk (RRW) model (Lemey et al., 2010) and a coalescent constant population size were used as priors. Markov chain Monte Carlo chains were run for $5 \times 10^7$ generations, sampling every 5000. Runs were evaluated and processed as above. Output trees were fed into SPREAD (Bielejec et al., 2011) to create a spatial representation of the spread of lineages through time.

ECOLOGICAL NICHE ANALYSES
Presence data
To develop ecological niche-based models, a total of 1063 observations (169 A. boueti, 542 A. boulengeiri, 228 A. impalearis and 124 A. tassiliensis) with ≤1 km resolution (World Geodetic System - WGS, 1984 datum) were collected from fieldwork ($N = 889$) and museum databases ($N = 174$). A 50 km buffer around a minimal convex polygon including all samples was used to delimit the study area. All spatial analyses were conducted in ESRI ArcGIS 10. In order to reduce spatial bias in the ecological models attributable to uneven sampling (Merow, Smith & Silander, 2013), localities were removed at random from clusters of species occurrence, forcing a point-free minimal radius of 5 km around each retained presence. The final dataset included 96 presence points for A. boueti, 259 for A. boulengeiri, 152 for A. impalearis and 66 for A. tassiliensis.

Climatic variables
Nineteen variables representing present climatic conditions were downloaded from WorldClim (www.worldclim.org; Hijmans et al., 2005) at 30 arc-s resolution (~1 km × 1 km at the equator). Variables were then clipped to the study area and upscaled to 2.5 arc-min (~5 km × 5 km at the equator). Given that these layers were derived through interpolations and represent macroclimatic conditions, this pixel size allows a reduction in computation time without affecting inference power. Five variables were excluded because of the presence of obvious spatial artefacts. To identify the significantly correlated variables, pairwise correlation among the remaining 14 was calculated using the Band Collection Statistics in ArcGIS and a threshold of $R = 0.7$. BIO1 + BIO6 and BIO2 + BIO5 were all retained despite slightly higher correlation ($R = 0.83$ for both pairs), owing to the potential relevance of those variables for restricting the distribution of ectotherms. The final set contained BIO1, BIO2, BIO5, BIO6, BIO7, BIO12 and BIO14 (Table 1).

The corresponding variables were downloaded from WorldClim for the middle Holocene (mid-Hol; 6000 years before present (BP); Palaeoclimate Modelling Intercomparison Project Phase III - Coupled Modelling Intercomparison Project Phase V (PMIP3-CMIP5), Last Glacial Maximum (LGM; ~21 000 years BP; PMIP3-CMIP5) and Last Interglacial (LIG; ~120 000–140 000 years BP; Otto-Bliesner et al., 2006). The LGM variables were at 2.5 arc-min resolution
Table 1. Climatic variables used for ecological models and projections, global and regional variation

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Units</th>
<th>Global</th>
<th>Regional</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIO1</td>
<td>Annual mean temperature</td>
<td>Degrees Celsius</td>
<td>−15.5 to 31.9</td>
<td>3.5–30.8</td>
</tr>
<tr>
<td>BIO2</td>
<td>Mean diurnal range</td>
<td>Degrees Celsius</td>
<td>5.8–20.7</td>
<td>4.5–18.5</td>
</tr>
<tr>
<td>BIO5</td>
<td>Maximal temperature of warmest month</td>
<td>Degrees Celsius</td>
<td>2.9–48.9</td>
<td>22.6–48.9</td>
</tr>
<tr>
<td>BIO6</td>
<td>Minimal temperature of coldest month</td>
<td>Degrees Celsius</td>
<td>−33.5 to 23</td>
<td>−12.4 to 18.8</td>
</tr>
<tr>
<td>BIO7</td>
<td>Temperature annual range</td>
<td>Degrees Celsius</td>
<td>11–47.8</td>
<td>10.9–42.8</td>
</tr>
<tr>
<td>BIO12</td>
<td>Annual precipitation</td>
<td>Millimetres</td>
<td>0–2769</td>
<td>2–1401</td>
</tr>
<tr>
<td>BIO14</td>
<td>Precipitation of driest month</td>
<td>Millimetres</td>
<td>0–42</td>
<td>0–24</td>
</tr>
</tbody>
</table>

Table 2. Summary of genetic diversity and demographic statistics for mitochondrial markers, based on all available samples (including duplicate haplotypes). Uncorrected p-distances among species can be found in Table S1.6.

Group 16S

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>P</th>
<th>N</th>
<th>h</th>
<th>Hd</th>
<th>(\pi)</th>
<th>(R^2)</th>
<th>D</th>
<th>(F_{ST})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. boueti</td>
<td>501/387</td>
<td>26</td>
<td>96</td>
<td>31</td>
<td>0.94 ± 0.011</td>
<td>0.01059 ± 0.00037</td>
<td>0.0744</td>
<td>−0.83499</td>
<td>−14.824</td>
</tr>
<tr>
<td>boueti C</td>
<td>500/473</td>
<td>12</td>
<td>46</td>
<td>15</td>
<td>0.916 ± 0.018</td>
<td>0.00414 ± 0.00030</td>
<td>0.0771</td>
<td>−1.0279</td>
<td>−7.694</td>
</tr>
<tr>
<td>boueti E</td>
<td>501/390</td>
<td>13</td>
<td>30</td>
<td>11</td>
<td>0.816 ± 0.053</td>
<td>0.00527 ± 0.00527</td>
<td>0.0747</td>
<td>−1.23408</td>
<td>−4.116</td>
</tr>
<tr>
<td>boueti W</td>
<td>499/471</td>
<td>8</td>
<td>20</td>
<td>6</td>
<td>0.516 ± 0.132</td>
<td>0.00189 ± 0.00078</td>
<td>0.1143</td>
<td>−2.04091</td>
<td>−7.627</td>
</tr>
<tr>
<td>A. boulengeri</td>
<td>504/367</td>
<td>51</td>
<td>243</td>
<td>30</td>
<td>0.876 ± 0.011</td>
<td>0.03426 ± 0.00093</td>
<td>0.1226</td>
<td>1.21511</td>
<td>3.254</td>
</tr>
<tr>
<td>boulengeri E</td>
<td>499/393</td>
<td>6</td>
<td>60</td>
<td>6</td>
<td>0.513 ± 0.057</td>
<td>0.00272 ± 0.00031</td>
<td>0.0886</td>
<td>−0.41465</td>
<td>−0.644</td>
</tr>
<tr>
<td>boulengeri F</td>
<td>500/389</td>
<td>12</td>
<td>110</td>
<td>14</td>
<td>0.654 ± 0.046</td>
<td>0.00276 ± 0.00072</td>
<td>0.0443</td>
<td>−1.38591</td>
<td>−8.225</td>
</tr>
<tr>
<td>boulengeri W</td>
<td>499/474</td>
<td>17</td>
<td>16</td>
<td>12</td>
<td>0.723 ± 0.035</td>
<td>0.00441 ± 0.00073</td>
<td>0.0676</td>
<td>−0.96446</td>
<td>−1.903</td>
</tr>
<tr>
<td>A. impalearis</td>
<td>499/462</td>
<td>22</td>
<td>27</td>
<td>18</td>
<td>0.915 ± 0.047</td>
<td>0.01122 ± 0.00000</td>
<td>0.104</td>
<td>−0.47252</td>
<td>−7.527</td>
</tr>
<tr>
<td>A. tassiliensis</td>
<td>502/408</td>
<td>17</td>
<td>21</td>
<td>5</td>
<td>0.767 ± 0.067</td>
<td>0.0158 ± 0.00223</td>
<td>0.1896</td>
<td>1.34757</td>
<td>5.535</td>
</tr>
</tbody>
</table>

Group ND4 + tRNAs

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>P</th>
<th>N</th>
<th>h</th>
<th>Hd</th>
<th>(\pi)</th>
<th>(R^2)</th>
<th>D</th>
<th>(F_{ST})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. boueti</td>
<td>869/640</td>
<td>92</td>
<td>70</td>
<td>41</td>
<td>0.971 ± 0.010</td>
<td>0.02785 ± 0.00167</td>
<td>0.0923</td>
<td>−0.2609</td>
<td>−6.797</td>
</tr>
<tr>
<td>boueti C</td>
<td>869/663</td>
<td>39</td>
<td>38</td>
<td>20</td>
<td>0.927 ± 0.028</td>
<td>0.00726 ± 0.00085</td>
<td>0.0565</td>
<td>−1.70358</td>
<td>−7.428</td>
</tr>
<tr>
<td>boueti E</td>
<td>867/742</td>
<td>46</td>
<td>16</td>
<td>12</td>
<td>0.958 ± 0.036</td>
<td>0.01613 ± 0.00322</td>
<td>0.1183</td>
<td>−0.5775</td>
<td>−0.932</td>
</tr>
<tr>
<td>boueti W</td>
<td>867/811</td>
<td>10</td>
<td>16</td>
<td>9</td>
<td>0.892 ± 0.060</td>
<td>0.00294 ± 0.00042</td>
<td>0.1036</td>
<td>−0.78006</td>
<td>−3.437</td>
</tr>
<tr>
<td>A. boulengeri</td>
<td>869/663</td>
<td>175</td>
<td>189</td>
<td>87</td>
<td>0.969 ± 0.006</td>
<td>0.07312 ± 0.00195</td>
<td>0.1377</td>
<td>1.48748</td>
<td>−2.553</td>
</tr>
<tr>
<td>boulengeri E</td>
<td>867/751</td>
<td>42</td>
<td>57</td>
<td>25</td>
<td>0.790 ± 0.057</td>
<td>0.01078 ± 0.00092</td>
<td>0.0895</td>
<td>−0.59023</td>
<td>−4.359</td>
</tr>
<tr>
<td>boulengeri N</td>
<td>865/666</td>
<td>36</td>
<td>73</td>
<td>30</td>
<td>0.943 ± 0.012</td>
<td>0.00727 ± 0.00045</td>
<td>0.0633</td>
<td>−1.17293</td>
<td>−13.678</td>
</tr>
<tr>
<td>boulengeri S</td>
<td>868/765</td>
<td>76</td>
<td>55</td>
<td>30</td>
<td>0.958 ± 0.015</td>
<td>0.01051 ± 0.00191</td>
<td>0.0486</td>
<td>−1.81807</td>
<td>−9.789</td>
</tr>
<tr>
<td>A. impalearis</td>
<td>867/785</td>
<td>40</td>
<td>8</td>
<td>6</td>
<td>0.929 ± 0.084</td>
<td>0.02106 ± 0.00470</td>
<td>0.1818</td>
<td>0.3845</td>
<td>1.978</td>
</tr>
<tr>
<td>A. tassiliensis</td>
<td>872/658</td>
<td>64</td>
<td>16</td>
<td>8</td>
<td>0.808 ± 0.093</td>
<td>0.03894 ± 0.00615</td>
<td>0.1993</td>
<td>1.31617</td>
<td>6.66</td>
</tr>
</tbody>
</table>

\(D\), Tajima’s \(D\) (significant values in bold font); \(F_{ST}\), Fu’s \(F_{ST}\) statistic; \(h\), number of unique haplotypes; \(Hd\), haplotype diversity; \(L\), minimal length, excluding sites with gaps and missing data; \(N\), number of samples or phased sequences; \(P\), number of polymorphic sites, excluding sites with gaps and missing data; \(\pi\) = nucleotide diversity; \(R^2\), Ramos-Onsins and Rozas \(R^2\) statistic.

(~5 km × 5 km) and were retrieved for all three global circulation models (GCM) available: CCSM4, MIROC-ESM and MPI-ESM-P. The LIG and midHol variables were at 30 arc-s resolution. For LIG, the GCM was National Center for Atmospheric Research, Community Climate System Model (NCAR-CCSM); for midHol, we obtained the same GCM as for LGM. The LIG and midHol layers were also upsampled to 2.5 arc-min resolution.

Although the palaeoclimatic data accounts for only the past 120 000 years, and the duration and intensity of the cycles was variable, the last cycle is likely to be representative of previous cycles throughout the whole Pleistocene, because the direction of temperature changes was the same, and the later cycles had the largest amplitudes (Snyder, 2016). During the Pliocene, geological and palaeontological records indicate that, albeit with varying intensity, similar cycles
also took place globally (Lisiecki & Raymo, 2005) and in North Africa (Rohling, Marino & Grant, 2015).

**Palaeoclimatic modelling**

Ecological modelling was performed in BIOMOD2 (Thuiller et al., 2016), using two machine learning [artificial neural networks (ANNs) and maximal entropy (MaxEnt)] and two regression-based techniques [generalized additive models (GAMs) and generalized linear models (GLMs)]. This approach aims at reducing uncertainties that may affect a given modelling technique (Wiens et al., 2009).

The study area was as described in above (Sampling and study area), which also encompasses the expected variation in Sahara–Sahel extension. Given that the remoteness of the study area precludes extensive sampling, we used random pseudo-absences. To control for potential bias introduced in this step, ten independent sets of 10⁴ pseudo-absences were generated using the ‘disk’ (buffer) function, with 50 km around presence points. Given the large average distances between the presence points used for modelling, the relatively large buffer size should ensure that pseudo-absences correspond to real absences.

Ten replicates were run for each technique and pseudo-absence set, in a total of 400 replicates. Presence data for training and testing models were randomly selected for each replicate using a ratio of 70% training to 30% testing. The performance of replicates was evaluated using the true skill statistic (TSS) metric (Allouche, Tsoar & Kadmon, 2006) and a threshold of 0.75. The threshold was selected after visualizing the output, as a compromise between performance, geographical coherence with species distribution and the representativeness of modelling techniques. The consensus model was generated by averaging individual replicates (Marmion et al., 2009), and agreement among replicates was assessed using the standard deviation (Thuiller et al., 2009). Individual model replicates were projected to past climatic conditions (mid-Hol, LGM and LIG). Projections were assessed using clamping masks delimiting areas with environmental conditions outside the current range of climatic conditions (Elith, Kearney & Phillips, 2010). Consensus models and standard deviations were calculated as described above. The consensus models for present and past conditions were then averaged in order to identify climatically stable areas, i.e. potentially persistent areas of occurrence that could serve as refugia through time (Carnaval et al., 2009).

**Ecological niche comparisons**

Ecological niches were compared at three different phylogenetic levels, in order to evaluate whether niche divergence followed a Brownian (neutral) motion: intrageneric branch (A. boulengeri vs. the boueti-impalearis-tassiliensis group), species and intraspecific lineage. The intraspecific lineage comparisons were focused on A. boueti and A. boulengeri. We used the same climatic layers as those used in the niche models and the PCA-env approach developed by Broennimann et al. (2012) and updated with functions from the ‘ecospat’ R package (Broennimann, Cola & Guisan, 2016). This method uses a principal component analysis (PCA) to create a two-dimensional representation of climatic space, on which it performs comparisons between pairs of entities, in this case defined by minimal convex polygons encompassing lineage and species distributions. Overlap was measured using the D metric (Warren, Glor & Turelli, 2008), following Broennimann et al. (2012). Both equivalency and similarity tests (Warren et al., 2008) were run with 500 replicates. However, it should be noted that equivalency tests are more restrictive and affected by allopatric ranges and are thus typically less adequate than similarity tests when addressing biogeographical questions (Peterson, 2011).

## RESULTS

### Phylogeography

A total of 296, 246, 301 and 246 samples were successfully sequenced for 16S (522 bp, aligned), ND4 (699 bp coding portion plus 195 bp of tRNAs), c-mos (570 bp) and NTF3 (669 bp), respectively (Supporting Information, Table S1.1). From those and the previously available data, 341 unique concatenated sequences were kept for the concatenated tree (Supporting Information, Fig. S2.2). PartitionFinder model selection is summarized in the Supporting Information (Table S1.2).

The monophyly of all species was confirmed, and no nuclear haplotype sharing was detected among them (Fig. 2; Supporting Information, Fig. S2.1), supporting the phylogenetic relationships described in previous studies. Estimated species crown ages spanned the Pleistocene (A. boueti and A. impalearis), the Pliocene (A. tassiliensis) and the Miocene (A. boulengeri). The bPTP lineage delimitation (Supporting Information, Fig. S2.3) was consistent between the ML and BI trees, recovering two lineages within A. impalearis (‘NW’ and ‘SE’, for intercardinal directions), five within A. boueti (‘C’ (central), ‘W’ and three grouped under ‘boueti E’), three within A. tassiliensis (‘A’, ‘H’ and ‘T’, the initials of Air, Hoggar and Tassili mountains) and six within A. boulengeri (‘N’, ‘S’, ‘E’ and ‘M’ in Mali; the two other lineages include three samples in the northern distribution of ‘S’ and are sister lineages to it with a relatively recent split, thus were included in ‘S’).

The intraspecific lineages share nuclear haplotypes at varying degrees, with the exception of boueti W and tassiliensis A and T for NTF3 (Fig. 2). All intraspecific

© 2018 The Linnean Society of London, Biological Journal of the Linnean Society, 2018, XX, 1–18
lineages are seemingly parapatric or allopatric (Fig. 2), in agreement with the first expected scenario. The lineages of *A. boulengeri* occur in separate mountain systems in Mauritania and nearby Mali. A similar pattern is suggested for *A. tassiliensis*, with one lineage in each of three central Saharan mountain systems (Tassili N’Ajjér and Hoggar in Algeria, and Aïr in Niger). *Agama boueti* lineages have a predominantly east–west distribution along the Sahel, although the *boueti* C lineage reaches from the Senegal River in the south up to Adrar Soutouf in Morocco in the north. All major Sahelian species’ lineages occur in the Assaba–Tagant–Afollé region in Mauritania.

Diffusion models identified the same area around Tagant, Assaba and Afollé in Mauritania as the origin of dispersal of the extant diversity of both *A. boueti* and *A. boulengeri*. A second, younger possible ancestral area was depicted in Niger for *A. boueti*. Both species began marked range expansions out of the ancestral area in Mauritania at ~320 thousand years ago, particularly towards the northern parts of their current distribution (Fig. 3). *Agama boueti* also showed another, more recent expansion towards the West, the current range of the lineage *boueti* W. The genetic diversity of *boueti* C and *boulengeri* N lineages shows signs of recent demographic expansion, although values were not significant (Table 2; Supporting Information, Table S1.3). The spatial interpolation of diversity showed a concordant pattern of steady decrease from the centre towards the borders of both species’ distributions (Fig. 3).

**Palaeomodelling**

The climatically stable areas (Fig. 4) were generally coincident with the current species’ distributions (Fig. 1), except for *A. boulengeri*, with the southern and eastern parts of most of its current area of occurrence, the mountains of Mauritania, appearing as less suitable than the neighbouring areas. For intraspecific variability, climate stability is roughly concordant with the present distribution of lineages in the *boueti–impalearis–tassiliensis* group. Two major highly stable areas can be observed with the same NW–SE disposition as the *A. impalearis* lineages. For *A. boueti*, the identified stable areas in SW Mauritania, Senegal River mouth and SE Mauritania are roughly concordant with the C, W and E (in Mauritania) lineages, respectively. In the case of *A. tassiliensis*, the southern stable area corresponds to the distribution of lineage A, whereas the northern stable area is coincident with lineages T and H, with the cores of Tassili N’Ajjér (T) and Hoggar (H) also having higher stability than the surrounding area. The LIG was predicted to be the least favourable period for all species. The
Figure 2. Phylogeographic relationships among species and lineages. Left panels show dated phylogenetic trees (condensed in top panel, with two relevant subsections below; for the complete version, see Supporting Information, Fig. S2.1) of combined mitochondrial DNA (ND4 and 16S) and nuclear DNA (c-mos and NT3) data, calculated with BEAST; Bayesian posterior probability (bpp) support values (percentages) are indicated next to the nodes, with asterisks representing 100% support; some support values were omitted to improve clarity; species and lineages (capital letters) are indicated next to the sample codes. Right panels show maps and nuclear haplotype networks representing the distribution and relationships among species (top) and lineages (second and third rows); colours represent species and lineages according to the respective tree. Dashed lines in the maps represent the species’ extent of occurrence. Each circle in the networks represents a different haplotype, and size is proportional to the number of samples sharing that haplotype; the smallest circles along the lines represent mutated positions; the networks in the second and third rows were scaled up from the relevant section of the first row.
LGM also seems to have been overall less suitable for *A. boulengeri* (Supporting Information, Fig. S2.4). Model evaluation scores and environmental variable contributions are summarized in the Supporting Information (Table S1.4).

**ECOLOGICAL NICHE COMPARISONS**

Niche overlap (Table 3) was mostly > 10% for intra-specific lineages (average 11% in *A. boueti* and 30% in *A. boulengeri*) and decreased as phylogenetic distance increased (average 3.1% among species, 1% between intrageneric branches). All values were generally low, especially considering that they were weighted by the density of occurrence of each entity within the climatic space, which is likely to be attributable to the allopatry among compared entities. Similarity tests were significant among lineages of *A. boulengeri* and, although not significant, showed the same tendency for *A. boueti* lineages, except the E vs. W comparison. The remaining tests were non-significant. Using an extended climatic layer dataset, only comparisons against *boueti* W were non-significant (Supporting Information, Table S1.5). Tests for significant dissimilarity were all negative (Supporting Information, Table S1.5).

<table>
<thead>
<tr>
<th>Level</th>
<th>Comparison (A vs. B)</th>
<th>D</th>
<th>Equiv.</th>
<th>B → A</th>
<th>A → B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lineage</td>
<td>boue_C vs. boue_E</td>
<td>0.155</td>
<td>0.968</td>
<td>0.082</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>boue_C vs. boue_W</td>
<td>0.111</td>
<td>1.000</td>
<td>0.082</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>boue_E vs. boue_W</td>
<td>0.058</td>
<td>0.998</td>
<td>0.236</td>
<td>0.230</td>
</tr>
<tr>
<td></td>
<td>boul_E vs. boul_N</td>
<td>0.125</td>
<td>1.000</td>
<td>0.046*</td>
<td>0.030*</td>
</tr>
<tr>
<td></td>
<td>boul_E vs. boul_S</td>
<td>0.579</td>
<td>0.002</td>
<td>0.002*</td>
<td>0.004*</td>
</tr>
<tr>
<td></td>
<td>boul_N vs. boul_S</td>
<td>0.207</td>
<td>0.002</td>
<td>0.006*</td>
<td>0.008*</td>
</tr>
<tr>
<td>Species</td>
<td>boue vs. boul</td>
<td>0.036</td>
<td>1.000</td>
<td>0.365</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>boue vs. imp</td>
<td>0.035</td>
<td>1.000</td>
<td>0.419</td>
<td>0.365</td>
</tr>
<tr>
<td></td>
<td>boue vs. taas</td>
<td>0.032</td>
<td>1.000</td>
<td>0.573</td>
<td>0.561</td>
</tr>
<tr>
<td></td>
<td>boul vs. imp</td>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>boul vs. taas</td>
<td>0.029</td>
<td>1.000</td>
<td>0.234</td>
<td>0.214</td>
</tr>
<tr>
<td></td>
<td>imp vs. taas</td>
<td>0.055</td>
<td>1.000</td>
<td>0.305</td>
<td>0.355</td>
</tr>
<tr>
<td>Branch</td>
<td>bit† vs. boul</td>
<td>0.01</td>
<td>1.000</td>
<td>0.070</td>
<td>0.060</td>
</tr>
</tbody>
</table>

B → A and A → B, similarity tests; D, measured niche overlap; Equiv = equivalency test; boue = *A. boueti*; boul = *A. boulengeri*; imp = *A. impalearis*; taas = *A. tassiliensis*.* Significant (P < 0.05) similarity tests. Equivalency tests are not directly comparable with those calculated using the old methodology (Broennimann et al., 2012; see Discussion). †*boueti–impalearis–tassiliensis* group.

**DISCUSSION**

All the patterns expected in an aridity-induced vicariance scenario were present, supporting the role of climate in maintaining and probably giving rise to the diversity of the North-African radiation group and *A. boulengeri*, which most probably occurred as a result of increased and cyclic aridity for the Saharan–Sahelian populations. Diversification in the genus *Agama* has previously been related to the expansion of savannah habitats in Africa in the late Miocene (Leaché et al., 2014), and phylogeographical patterns in southern Africa have also suggested aridity-induced vicariance as a lead driver of diversification (Matthee & Flemming, 2002; Swart, Tolley & Matthee, 2009). However, adaptation to new habitats or niches is also found in *Agama*; for instance, the sand-burrowing behaviour of *Agama etoshae* (Arnold, 1995) or the tail retention as a possible climbing adaptation in *Agama lionotus* (Loumbourdis, 1986). The mountains in southern Mauritania are identified as a diversity hotspot and predicted climatic refugium for the local *Agama* species and therefore possibly for other species with similar climatic requirements, which highlights the importance of the region in terms of biodiversity conservation. The apparent ability of *Agama* from the Sahel–Sahara fringe to survive in that area during major climatic fluctuations emphasizes its pivotal role in the future survival of mesic species in the face of ongoing global warming.

**SPATIAL STRUCTURE OF GENETIC VARIABILITY**

The Agaminae subfamily is most likely to have colonized Africa from the Arabian Peninsula through semi-arid corridors, the Sahel region being among them (Kissling et al., 2016) and, as expected from an Afro-tropical group, most of the agama diversity in North Africa is found in the Sahel. The geographical coverage of available genetic information and the number of major intraspecific genetic lineages identified have greatly increased from previous studies (Gonçalves et al., 2012; Mediannikov, Trape & Trape, 2012; Leaché et al., 2014). Species crown ages and lineage split times were all younger than the
oldest age of the Sahara (7 Mya; Schuster et al., 2006) and mostly fall within the Pliocene–Pleistocene (A. boulenegeri crown age is 6.4 Mya), a pattern shared with other taxa in the region, including reptiles, birds and mammals (Brito et al., 2014) and usually attributed to the aridification and/or climatic cycles.

Species and lineages are almost exclusively parapatric or allopatric, which verifies the first expected pattern under the hypothesis of climate-induced vicariance and allopatric diversification. Aridity-induced isolation in sky islands, a pattern also present, for instance, in Myrtus shrubs (Migliore et al., 2013) or Ptyodactylus geckos (Metallinou et al., 2015), is apparent through the one-per-mountain-range distributions of the lineages of A. tassiliensis and A. boulenegeri. Evidence of north–south vicariance along the Atlantic coast is found in the north–south separation of sister species A. impalearis and A. boueti, in agreement with previous studies (Gonçalves et al., 2012) and matching the generally recognized role of the Sahara in separating Mediterranean and Sahelian populations (Douady et al., 2003; Brito et al., 2014). Still, the role of the Atlantic coast as a corridor allowing species dispersal across the desert (Brito et al., 2014; Gonçalves

Figure 3. Top four lines: continuous spatial diffusion phylogenetic models for A. boulenegeri and A. boueti. Only the polygons representing the 80% occurrence confidence interval are shown. Polygon colour represents relative age, the darker being more recent; the estimated origins for each polygon can be found in the raw output, which can be visualized in Google Earth (Supporting Information, Appendix S4). Bottom: distribution of genetic diversity. The colour represent the nucleotide diversity calculated for A. boueti (left) and A. boulenegeri (right) based on the mitochondrial markers.
et al., 2018) is illustrated by *A. impalearis* being the only species north of the Sahara and the fact that at the present time the distributions of these two species are very close, almost parapatric.

Although genetic analyses revealed no signs of contact between *A. impalearis* and *A. boueti*, the difficulty of sampling in the geographical area between both species prevents ruling out possible secondary contacts. A past secondary contact between *A. boueti* and *A. tassiliensis* in the Hoggar Mountains, previously hypothesized based on signs of nuclear introgression (Gonçalves et al., 2012), was supported by the palaeo-models (Fig. 4; Supporting Information, Fig. S3). If that was the case, it would contrast with the pattern found in other terrestrial vertebrates, such as the snake *Psammophis schokari* (Gonçalves et al., 2018) or the anurans *Pelophylax saharicus* (Nicolas et al., 2015) and *Bufotes boulengeri* (Nicolas et al., 2018), which seem to have reached the Hoggar Mountains from the north. The phylogeographical pattern recovered for *A. boulengeri* of deep allopatric mitochondrial lineages with low nuclear diversity and extensively shared haplotypes may indicate climate-mediated range fluctuations and secondary-contact events, but further research is needed to clarify this (Supporting Information, Appendix S3).

**Climatic refugia**

The climatically stable areas and areas of predicted present occurrence for species were separated, with large extensions of highly unsuitable regions between them (Fig. 4; Supporting Information, Fig. S2.4), suggesting that climatic conditions did not favour contact. The general correspondence of climatically stable areas to species and lineage distribution in the *boueti–impalearis–tassiliensis* group is concordant with the vicariance hypothesis, because they indicate potential refugia (Carnaval et al., 2009) whose geographical isolation can lead to allopatric speciation (Avise, 2000). However, in the particular cases of *impalearis* NW and SE it is more likely to be attributable to a colder and more humid climate in the Atlas Mountains (Fig. 4; Supporting Information, Fig. S2.4) rather than
increased aridity, in a similar manner to other taxa in that region, such as *Pelophylax* frogs, *Mauremys* pond turtles or *Daboia* snakes (*Lansari et al., 2015; Verissimo et al., 2016; Martínez-Freiría et al., 2017*). For *A. boulengeri*, the stable climatic areas include sandy areas where the species does not occur at present, which could indicate that only the outskirts of the mountains, rather than the whole massifs, acted as microrefugia not detected by the models, or that there was more exposed rock in the past. It could also be an artefact from a climatic layer dataset that does not include habitat and is therefore insufficient to predict the stable areas for a species occurring almost exclusively on rocky outcrops [see Vale et al., (2012) for topoclimatic models for the present].

The only two other available studies applying multiple climatic phase palaeodistribution modelling in north Africa found contrasting patterns: Martínez-Freiría et al. (2017) found a similar pattern of allopatric potential refugia broadly coherent with lineage distributions of *Daboia* snakes in Northern Maghreb, whereas Gonçalves et al. (2018) found signs of a continuous coastal climatic refugium along the Atlantic for *Psammophis schohaki*. Although, in the latter case, the stable areas were continuous in terms of climate, habitat suitability was probably interrupted by the basin of the large intermittent (flowing in humid phases) Tamanrasset palaeo-river, which depicts a different example of the impact of Pleistocene climatic cycles on species diversification in the region. Although not using palaeo-modelling, refugia in arid North Africa have been suggested in multiple studies in the last decades. Most of the focus has been on mesic species, with refugia including the mountain systems of central Sahara (Metallinou et al., 2015), inland water bodies of a large scale such as Lake Chad (Granjon & Dobigny, 2003) and a micro-scale such as oases (Shaibi & Moritz, 2010), or the peripheral regions of Sahel and the Mediterranean (Douady et al., 2003). Refugia for aridity-adapted species, which should have suffered range contractions during wet periods (Carranza et al., 2008; Pook et al., 2009; Metallinou et al., 2012; Tamar et al., 2016), are comparatively lacking. In fact, climatic shifts can be so fast and pronounced (DeMenocal et al., 2000) that many species probably need refugia for both humid and arid periods. That is exemplified by the aridity-adapted reptile genus *Mesalina*, which may survive on the outskirts of the central Saharan mountains in periods of extreme aridity (Kapli et al., 2015). The pattern of refugia is still not well understood in North Africa, but based on examples from other regions, such as the reptile community in south-western Australia, the interaction of climate, physical barriers and range shifts can be complex (Edwards, Keogh & Knowles, 2012). Techniques such as demographic simulations could help to address this subject with more detail and accuracy, as illustrated by a related example concerning the lizard *Lerista lineopunctulata* (He, Edwards & Knowles, 2013), but proper population sampling and a higher number of loci are needed to pursue such an approach.

**Range dynamics**

The mountain regions of southern Mauritania seem to have been the origin of dispersal of extant diversity of *A. boueti* and *A. boulengeri*, according to evidence from the diffusion models, interpolation of genetic diversity, and climatic stability. A second, younger possible ancestral area is depicted in Niger for *A. boueti*, which could indicate quick dispersal or a more ancient widespread distribution in the Sahel, but the sampling gap between Mauritania and Niger precludes further inferences. The diffusion models predicted a recent expansion to the fringes of distribution in Mauritania, a pattern also reflected by lower genetic diversity in those areas (Fig. 3). The marked difference in the expansion signal (Fig. 3, Table 2; Supporting Information, Table S1.3) recovered in the northernmost lineages of *A. boueti* and *A. boulengeri* (C and N, respectively), when compared with the southern ones from less arid regions, indicates a relatively recent colonization, starting from the climatically more stable central-southern areas. The niche similarity among lineages suggests that the expansion was not attributable to adaptation to different ecological conditions, and the fact that ‘backwards’ migrations in the diffusion models were predicted only for recent times (osf.io/pdthw/; doi:10.17605/OSF.IO/PDTHW) suggests repeated local extinctions and loss of signal from previous range expansions. Climatic changes in the region can be sudden and pronounced enough to allow it (DeMenocal et al., 2000). The synchronous marked expansions to the north and east in *A. boulengeri* and *A. boueti* and north and west in *A. boueti*, at ~320 thousand years ago (Fig. 1), followed a glacial termination (Lisiecki & Raymo, 2005), suggesting that these expansions took place in a more humid phase. The stable climatic suitability areas around the southern Mauritanian mountains also stress the importance of the opportunity for elevational displacement for the survival of mesic species as climate fluctuates (Dobrowski, 2011; Velo-Antón et al., 2013). This pattern of apparent lack of adaptation to fluctuating climate indicates niche conservatism (Kozak & Wiens, 2010) and is another indication of the role of climate in the dispersal and subsequent diversification of mesic species.

**Ecological niche comparisons**

Observed niche overlap decreased as phylogenetic distance increased, as expected from a Brownian motion of niche diversification and low degrees of ecological adaptation (Wiens & Graham, 2005). It could happen
that niche overlap is biased by geographical proximity (Warren et al., 2008), so overlap alone would not be highly significant because we are comparing two Sahelian lineages. However, the randomized niche similarity tests displayed a similar trend; similarity among species was much lower than among lineages. Given the wide geographical scale, the fact that no dissimilarity test was significant (all \( P > 0.4 \)) again shows support for some conservatism of niche and susceptibility to vicariance. The higher similarity among intra-generic branches (A. boulenieri vs. the rest), even if not significant, does not contradict the vicariance hypothesis, as the divergence between both groups took place well before the significant increase in aridity (Fig. 2; Supporting Information, Fig. S2.1) and is also coherent with a general niche conservatism within the genus and subfamily (Kissing et al., 2016).

The significant similarity for all comparisons among lineages of A. boulenieri (Table 3; Supporting Information, Table S1.5) also follows the expected pattern under the hypothesis of vicariance. The A. boueti lineages C and E follow the same pattern, but it is clear that boueti W has a less similar climatic niche compared with the rest. Given no allele sharing in NTF3, this could indicate some level of adaptation or divergence and is compatible with the existence of an undescribed species and the proposal of a ‘boueti’ species complex by other authors based on the morphological and ecological variation within A. boueti (Mediannikov et al., 2012; Leaché et al., 2014).

It is worth noting that niche distinction between A. boueti and A. impalearis is much greater than among A. boulenieri lineages of around the same age. Together with the case of boueti W, this could indicate a lower degree of niche conservatism in the boueti–impalearis–tassiliensis group, which could translate into higher adaptation potential to different climates and habitat (Wiens et al., 2010) and help to explain why only one of the groups was able to cross and colonize the Sahara. However, the rock habitat specialization of A. boulenieri is the most likely cause for its geographically restricted occurrence; species of the boueti–impalearis–tassiliensis group occur in soft to hard soil with sparse vegetation and therefore do not require rock connectivity to disperse to different areas.

Equivalency tests were mostly non-significant, meaning that the areas where the compared entities occur are more different than expected by chance (except among A. boulenieri lineages). This can be explained by all of the compared areas being allopatric and, in general, having low ecological niche overlap. In addition, equivalency tests are not weighted by the distribution density, meaning that they include all the area within the minimal convex polygon without correcting for actual occupancy (as the similarity tests and niche overlap measurement do); and given

the characteristics of the landscape, a significant portion of the area within the minimal convex polygons can be unsuitable for both entities, possibly rendering comparisons meaningless.

Previous studies have reported low \( P \)-values for equivalency tests (Rato et al., 2015; Ahmadzadeh et al., 2016; Martínez-Freiría et al., 2017; Gonçalves et al., 2018), but the ones obtained here cannot be compared directly with those results, given differences in the way the \( P \)-value for equivalency tests is calculated in ‘ecospat’ and in previous scripts (Broennimann et al., 2012). If subject to the previous methodology, the equivalency values obtained here would be close to zero, thus following the same pattern as other published works (Supporting Information, Appendix S4).

Most studies addressing niche divergence and conservatism in North Africa are focused in the Mediterranean region (e.g. Anadón et al., 2015; Rato et al., 2015), with examples relating varying levels of niche divergence to patterns of temperature seasonality in the Mediterranean and Atlantic climates (Ahmadzadeh et al., 2016) or linking niche conservatism and allopatric diversification to Pleistocene climatic oscillations (Martínez-Freiría et al., 2017). Diversification in a scenario of climatic niche conservatism has also been described for other arid regions (e.g. Loera, Sosa & Ickert-Bond, 2012), and it might be a common pattern. These examples show that adaptation or niche specialization is not a requirement for diversification in deserts. A similar inference has been reached by Wien, Kozak & Silva (2013) when comparing niche breaths of phrynosomatid lizards and concluding that high diversity in arid regions was probably attributable to longer time evolving in those habitats.

CONCLUSIONS

This study highlights the importance of climate-induced vicariance in the maintenance of diversity and probably in the diversification of the Agama group and is, to our knowledge, the first study critically evaluating the role of this environmental process in species diversification across the Sahara–Sahel ecoregions. Using Agama species as a model, we have verified the occurrence of the expected patterns for species with a conserved niche under a climatically induced vicariance scenario. Based on the role of the southern Mauritanian mountains as a diversity hotspot and predicted climatic refugia for the local Agama species, we highlight the probable importance of this region for the conservation of local biodiversity and evolutionary processes.

ACKNOWLEDGEMENTS

The authors acknowledge BIODESERTS group members for their assistance during fieldwork. P.A. Crochet,
P. Geniez and J. M. Patial provided access to samples. This work was funded by the National Geographic Society (CRE 7629-04/8412-08), Mohamed bin Zayed Species Conservation Fund (11052709, 11052707 and 11052499), FCT - Fundação para a Ciência e a Tecnologia (PTDC/BIA-BEC/099934/2008 and PTDC/BIA-BIC/2903/2012), Ministério de Economia y Competitividad (CGL2015-70390-P) and by European Regional Development Fund through the Operational Programme for Competitiveness Factors, COMPETE (FCOMP-01-0124-FEDER-008917/028276). D.V.G., P.P., G.V.-A. and J.C.B. were supported by FCT (SRH/BD/78402/2011, PD/BD/128492/2017, IF/01425/2014 and IF/00459/2013, respectively) within QREN-POPH-T4.1 funded by European Science Foundation and Portuguese Ministério da Educação e Ciência. Logistic support for fieldwork was given by Pedro Santos Lda (Trimble GPS), Off Road Power Shop, P. N. Banc d’Arguin (Mauritania), Ministry of Environment and Faculty of Sciences and Technologies, University of Nouakchott. The authors also thank the two anonymous referees whose helpful comments helped to improve the manuscript.

REFERENCES


continuous space and time. Molecular Biology and Evolution 27: 1877–1885.


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supplementary tables.
Appendix S2. Supplementary figures.
Appendix S3. The case of Agama boulengeri.
Appendix S4. Notes on equivalency tests.

SHARED DATA

Supplementary data (DOI: 10.17605/OSF.IO/PDTHW) for the diffusion models can be found at the website of the Open Science Framework (www.osf.io/pdthw).