Evolutionary history of the two North African hedgehogs (Mammalia: Erinaceidae) *Atelerix algirus* and *Paraechinus aethiopicus* based on phylogeography and species distribution modelling

Haithem El-Farhati¹, Mourad Khaldi², Alexis Ribas³, Mohamed Wassim Hizem¹, Saïd Nouira¹,*, Violaine Nicolas⁴,*

¹ Ecologie de la Faune Sauvage UR17-ES44, Faculté des Sciences de Tunis, Université de Tunis Manar, 2092 Tunis, Tunisia
² Département des Sciences Agronomiques, Faculté des Sciences, Université Mohamed Boudiaf de M’sila, M’sila 28000, Algérie
³ Section of Parasitology, Department of Biology, Healthcare and the Environment, IRBio (Research Institute of Biodiversity), Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain
⁴ Institut de Systématique, Evolution, Biodiversité (ISTEB), Muséum national d’Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP31, 57 rue Cuvier, 75005 Paris, France

* Shared last authorship, contributed equally

http://zoobank.org/6C6B1A34-1BA6-4DDA-959B-114991DA0968

Corresponding author: Violaine Nicolas (violaine.colin@mnhn.fr)

Academic editor: Clara Stefen | Received: 2 July 2021 | Accepted: 10 November 2021 | Published: 6 December 2021


Abstract

Two species of hedgehogs are known to occur in northern part of Africa: the Algerian hedgehog *Atelerix algirus* and the Ethiopian hedgehog *Paraechinus aethiopicus*. Within each species several subspecies were described based on morphometrical data and pelage coloration, but all these subspecies have enigmatic and unclear definitions. We investigated the phylogeographical history and taxonomy of these two species based on mitochondrial DNA data covering the entire geographical distribution of *A. algirus* and the North African distribution of *P. aethiopicus*. We also used climatic niche modelling to make inferences about their evolutionary history. Low genetic diversity was recovered in both species. While no phylogeographic pattern was found in *P. aethiopicus*, two haplogroups were identified within *A. algirus*. This could be explained by the fact that continuous high or moderate climatic suitability occurred throughout most of the Saharan desert since the LGM (Last Glacial Maximum) for the first species, while during the LGM there were several disconnected areas of high climatic suitability for *A. algirus*: one in South-West Morocco, one at the coastal Moroccan-Algerian border and one in Tunisia-coastal Libya. Our genetic results confirm that *A. algirus* recently colonized Spain, Balearic and Canary Islands, and that this colonization was probably mediated by humans. Suitable climatic conditions occurred throughout most of the Southern and Eastern Iberian Peninsula during the last 6,000 years which could have favored the spatial expansion of the Algerian hedgehog after its arrival in Europe. According to our molecular results subspecific recognition within North Africa is unwarranted for both species.

Key words

Climatic niche modelling; haplotype network; mammal; mitochondrial DNA; North Africa; phylogeny; taxonomy
Introduction

Two species of hedgehogs are known to occur in northern part of Africa (Fig. 1): the North African or Algerian hedgehog *Atelerix algirus* (Lereboullet, 1842) and the Ethiopian hedgehog *Paraechinus aethiopicus* (Ehrenberg, 1832), which can be distinguished based on morphological characteristics (Happold and Happold 2013) and genetic data (He et al. 2012; Bannikova et al. 2014; Velo-Antón et al. 2019). The Algerian hedgehog is recorded from North Africa (Morocco, Algeria, Tunisia and Libya), in parts of the Mediterranean coastal regions of Spain, and the Balearic, Canary and Maltese islands (Best 2018). It occurs in coastal dry and warm Mediterranean scrublands and grasslands. It does not inhabit arid deserts, but it has been documented at oases in the Sahara desert (Happold and Happold 2013; Best 2018). The Ethiopian hedgehog occurs in arid and semi-arid habitats in Africa, from Senegal and Mauritania through scattered locations in the Sahara to Egypt, Sudan, and Somalia, and in the Middle East, from Syria and Iraq to the Southern Arabian Peninsula (Corbet 1988; Hutterer 2016; Best 2018).

Within each species several subspecies and forms were described based on morphometrical data and pelage coloration. According to Hutterer (2005) the Algerian hedgehog comprises three subspecies: *A. algirus algirus* (Lereboullet, 1842) from North Africa (with four synonyms: *caniculus* (Thomas, 1915), *diadematus* (Dobson, 1882), *fallax* (Dobson, 1882) and *lavaudeni* (Cabrera, 1928)), *A. a. girbaensis* (Vesmanis, 1981) from Djerba Island (Tunisia), *A. a. vagans* (Thomas, 1901) from Spain and Minorca and other islands (with one synonym: *krugi* (Peters, 1877)). According to the same author the Ethiopian hedgehog comprises five subspecies: *P. aethiopicus aethiopicus* (Ehrenberg, 1832) from Soudan and Ethiopia, *P. a. albatus* Thomas, 1922 from Djerba island (with three synonyms: *albior* Pocock, 1934, *blancalis* (Thomas, 1921) and *brachydaactylus* (Wagner, 1841)), *P. a. deserti* (Loche, 1858) from North Africa, *P. a. ludlowi* (Thomas, 1919) from Iraq and a part of Iran (with two synonyms: *onisicus* (Thomas, 1922) and *pallidus* (Fitzinger, 1867)) and *P. a. pectoralis* (Heuglin, 1861) from Arabian Peninsula (with three synonyms: *dorsalis* (Anderson and de Winton, 1901), *sennaariensis* (Hedenborg, 1839) and *wassifi* (Setzer, 1957)). However, as stated by Hutterer (2005) and by Best (2018) all these subspecies have enigmatic and unclear definitions, and subspecific taxonomy required reassessment.

Until recently, virtually nothing was known on the genetic variability within these two species. However, during the last five years, several studies were published. Based on one mitochondrial (Cytochrome b) and one nuclear (BFIBR) genes Khaldi et al. (2016) showed low genetic diversity across the geographical range of *A. algirus* and a recent colonization of Spain, Balearic and Canary Islands by this species, probably mediated by humans. Following these results, Best (2018) proposed to consider *A. a. algirus* and *A. a. vagans* as a single subspecies: *A. a. algirus*. Derouiche et al. (2016) also found low genetic diversity within this species in Algeria based
on a fragment of the control region. However, with the same DNA fragment Velo-Antón et al. (2019) found three distinct genetic lineages within this species: two lineages endemic to Morocco and one lineage widely distributed from Eastern Morocco to Libya. This result suggests that the control region could be a suitable marker to study the phylogeographical history and evolutionary history of this species across its distribution range, owing to its fast evolutionary rate and presumed selective neutrality as a non-coding region. Molecular studies of *P. aethiopicus* found no phylogeographic structure within North Africa, whatever the DNA fragment used (Derouiche et al. 2016; Velo-Antón et al. 2019), and genetic differentiation between North African and Arabian populations (Derouiche et al. 2017).

The aim of this study is to investigate the phylogeographical history and taxonomy of these two species based on improved sampling covering the entire geographical distribution of *A. algirus* and the North African distribution of *P. aethiopicus*. New samples from Morocco, Algeria, Tunisia, peninsular Spain, and Balearic and Canary Islands were included. We also used ecological niche modelling to make inferences about the evolutionary history of these lineages and the potential impact of paleoclimatic events on the current genetic variability and distribution of each species.

**Material and methods**

**Sampling**

A total of 126 new samples (See supplementary file 1) were collected in the field, either from road-killed animals or live-trapping or from museum collections (Museum of Natural Science of Barcelona) and stored in 96% ethanol. These specimens come from Tunisia (72 specimens), Algeria (32), Morocco (2 specimens), Spain (15), Balearic (2) and Canary Islands (3). Geographical coordinates for the collected samples were recorded with a Global Positioning System. Species identification in the field was based on morphological characteristics, namely ventral, leg and facial color patterns, as well as ear length (Aulagnier et al. 2008).

**DNA extraction, amplification and sequencing**

Whole genomic DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN), following the manufacturers’ protocols. We amplified a fragment of 456–465 bp of the mitochondrial control region (CR) using the primers developed by Derouiche et al. (2016): CRF: 5′-CATCA-ACACCCCAAAGTTG-3′ and Erinaceinae CRR: 5′-TGA-AGAAAGAACCAGATG-3′. The amplification conditions consisted of an initial denaturation phase at 94 °C for 3 min, followed by 40 cycles each of denaturation at 94 °C for 40 s, hybridization at 53 °C for 40 s and elongation at 72 °C for 60 s; then a final elongation phase ran at 72 °C for 5 min. The quality and size of the PCR products were then verified by electrophoresis on 1% agarose and ethidium bromide-supplemented gel. The double-stranded PCR products were sent for purification and Sanger sequencing at EUROFINS. All newly determined sequences were submitted to GenBank (accession numbers OL692499–OL692624).

**Phylogenetic analyses and genetic diversity**

Chromatograms were checked and sequences manually corrected in CodonCode Aligner 8.0.2. In our phylogenetic analyses, we included all the newly sequenced specimens and all specimens of *Atelerix* and *Paraechinus* available in the GenBank database for which the control region was sequenced. We also added several references sequences for the hedgehog species *Erinaceus europaeus*, *E. roumanicus* and *E. concolor*. Sequences were aligned with ClustalW implemented in Geneious 4.6.4., and the alignment was corrected by eyes. The final alignment comprised 283 sequences (see supplementary file 1 for specimens list, and see supplementary file 2 for sequence alignment). Sequences of 423 bp were retained for the final analyses as they were available for most specimens.

We estimated evolutionary relationships among sequences by constructing phylogenetic trees using maximum-likelihood (ML) and Bayesian inference (BA). The computer program MrModeltest ver.2 (Nylander 2004) was used to evaluate the fit of 24 nested models of nucleotide substitution to the data. The model (HKY+I+G) chosen by MrModeltest according to the Akaike Information Criterion was then used in ML and BA analyses. The ML analysis was performed with PhyML ver. 3.0 (Guindon et al. 2010) on the PhyML online web server (Guindon et al. 2005). In the PhyML procedure, we used the BIONJ distance-based tree as the starting tree and 100 bootstrap replicates. Bayesian inference was performed with MrBayes ver. 3.2.6 (Ronquist et al. 2012) using 7,000,000 generations (burnin = 25%), sampling every 1,000 steps. Convergence between chains was checked by manual inspection using the ‘sump’ command and by using the mean SD of split frequencies, which was below 0.01 in our runs, as an indicator of convergence.

A haplotype network for each species (See supplementary file 3 for *A. algirus* sequences alignment, and see supplementary file 4 for *P. aethiopicus* sequences alignment) was constructed in TCS 1.2.0 (Clement et al. 2000) using statistical parsimony to calculate the minimum number of mutational steps among haplotypes with connections based on 95% confidence interval. Gaps were treated as 5th state. The web-based program tcsBU was used to draw the network (Murias dos Santos et al. 2016).

Nucleotide diversity, haplotype diversity (Nei 1987) and the mean number of pairwise differences were calculated using DnaSP 6.12 (Librado and Rozas 2009).
Inferences of population expansion were made using two different methods: the Fs statistic (Fu 1997) and the R2 statistic (Ramos-Onsins and Rozas 2002). Ramos-Onsins and Rozas (2002) demonstrated that these statistics have the greatest power to detect population expansion for nonrecombining regions of the genome under a variety of different circumstances, especially when population sample sizes are large or when sample sizes are small. The significance of Fu’s Fs and R2 was obtained by examining the null distribution of 1,000 coalescent simulations of these statistics using DnaSP 6.12 (Rozas et al. 2017). Significantly large negative Fu’s Fs values and significantly positive R2 values were taken as evidence of a population expansion.

Species distribution modelling

To evaluate current and past habitat suitability for the two species of hedgehogs (A. algirus and P. aethiopicus) we produced distribution models under three different climatic conditions (current, mid-Holocene and Last Glacial Maximum [LGM]) using a distribution modeling approach. Species occurrence data were assembled from the literature (Derouiche et al. 2016; Aulagnier et al. 2017; Deef 2020; El-Farhati et al. 2020; de Smet et al. 2021), museum collections (Global Biodiversity Information Facility, GBIF.org (2 December 2020) GBIF Occurrence Download https://doi.org/10.15468/dlf.gkz4ya and https://doi.org/10.15468/dlf.m8y5zz) and our own records (See supplementary file 5). Datasets of this type often exhibit sampling bias, which can lead to unreliable ecological niche modeling results (Fourcade et al. 2014). To correct sampling bias, a subsample of records regularly distributed in the geographical space was selected. The resolution of the reference grid was set to 2.5 arc min; i.e. the same as that used for the environmental layers. This coarse resolution was chosen to better match the coordinate uncertainties associated with georeferenced, textual localities of museum specimens. After filtering, the final dataset used for modeling consisted of 791 distribution records for A. algirus and 481 distribution records for P. aethiopicus.

As environmental layers, we used available climatic data from the WorldClim database (Hijmans et al. 2005). We removed from our analyses variables BIO2, BIO3, BIO5, BIO7 and BIO15 because those variables show a high level of discrepancy between general circulation models (GCMs) for North Africa (Varela et al. 2015). Collinearity of the initial variables was measured by Pearson’s correlation coefficient. We retained six variables that are considered ecologically important for hedgehogs and for which pairwise correlations were mostly smaller than 0.75: annual mean temperature (BIO1), seasonality of temperature (BIO4), minimal temperature for coldest period (BIO6), mean temperature of the warmest quarter of the year (BIO10), mean temperature of the coldest quarter of the year (BIO11), annual precipitation (BIO12). These climatic variables are ecologically important for hedgehogs because temperature affects hedgehogs’ prey activity, and hedgehogs are known to hibernate in cold winters (Mouhoub-Sayah 2009; Mouhoub Sayah et al. 2009). Climatic variables were downloaded for present conditions, for mid-Holocene (a wetter period, about 6,000 years ago) and for the LGM (a cold and dry period, about 22,000 years ago). Mid-Holocene and LGM climate data were drawn from General Circulation Model (GCM) simulations from two climate models: the Community Climate System Model (CCSM ver. 3) (Collins et al. 2006) and the Model for Interdisciplinary Research on Climate (MIROC ver. 3.2) (Hasumi and Emori 2004). These climate models differ in temperature and precipitation patterns. For a discussion of the uncertainties associated with the climatic data see Schorr et al. (2012) and Varela et al. (2015). The use of these two different climate models enabled us to assess and account for modeling uncertainty due to LGM climate data. To predict the potential distribution of the species in current conditions and in the past, we used Maxent ver. 3.3.3 (Phillips et al. 2006), which outputs a model with relative occurrence probability of a species within the grid cells of the study area. To ensure consistency of model predictions among repeated runs, we performed a 50-fold cross-validation with random seed. To determine whether the predictions for current conditions generated by Maxent were better than random predictions we used the area under the receiver operating characteristic curve (AUC), a commonly used measurement for comparison of model performance (Elith et al. 2006). The AUC ranges from 0 to 1, with greater scores indicating better discrimination ability; an AUC greater than 0.5 indicates that the model discriminates better than random.

Results

Phylogenetic analyses and genetic diversity

The final alignment contained 283 sequences of 423 pb long with 32 sites with alignment gaps. ML and BA analyses resulted in a well resolved tree (Fig. 2) supporting the monophyly of the three included genera (Atelerix, Erinaceus and Paraechinus). All P. aethiopicus specimens cluster in a well-supported lineage. Specimens of A. algirus cluster in two distinct well supported lineages (I and II), which are not sister lineages. The first lineage includes most specimens and is distributed throughout the geographical range of the species. The second lineage includes only two Moroccan specimens. The average number of nucleotide differences between these two lineages is 27.574 with 22 fixed differences. As a matter of comparison, it is 27.000 with 27 fixed differences between A. algirus lineage II and A. albiventris lineage 1, and it is 27.667 with 26 fixed differences between A. algirus lineage II and A. albiventris lineage 2. The two sequences of A. algirus lineage II were previously published by Velo-Antón et al. (2019). Considering the high
genetic divergence between these two lineages and the topology of the tree, doubts on the species identification and/or validity of these sequences are raised (see discussion part). Thus, they are not considered as belonging to *A. algirus* in subsequent analyses.

The 204 specimens of *A. algirus* lineage I resulted in 58 haplotypes (Table 1). The nucleotide diversity is 0.00745 ± 0.00037, haplotype diversity is 0.924 ± 0.008 and the average number of nucleotide differences is 3.010. No strong phylogeographic structure is recorded within this species (Fig. 3). One abundant haplotype is shared between specimens from Tunisia, Algeria and Morocco, and several haplotypes are shared between specimens from Algeria and Morocco, or from Algeria and Tunisia. One haplotype is shared between specimens from Libya and Tunisia. High genetic diversity is observed within Morocco (51 specimens, 23 haplotypes) with numerous endemic haplotypes. Haplotypes from Morocco can be divided in two groups: the first haplogroup (haplotypes A01 to A13) is composed exclusively of haplotypes en-
Table 1. Diversity and neutrality estimates for the two species and main haplogroups identified in the Statistical parsimony haplotype networks. Significant values are indicated by asterix (* for P < 0.05, ** for P < 0.01, *** for P < 0.001). Number of sequences (N), number of distinct haplotypes (Nh), haplotype diversity (Hd), nucleotide diversity (Pi), average number of nucleotide differences (k) and results of the Fu’s Fs and R2 tests are given.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Nh</th>
<th>Hd</th>
<th>π</th>
<th>k</th>
<th>Fu’s Fs</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>All <em>A. algirus</em> lineages</td>
<td>204</td>
<td>58</td>
<td>0.924 ± 0.008</td>
<td>0.00745 ± 0.00037</td>
<td>3.010</td>
<td>-28,426***</td>
<td>0.057</td>
</tr>
<tr>
<td>Haplogroup 1</td>
<td>25</td>
<td>13</td>
<td>0.827 ± 0.067</td>
<td>0.00514 ± 0.000069</td>
<td>2.077</td>
<td>-3.555**</td>
<td>0.110</td>
</tr>
<tr>
<td>Haplogroup 2</td>
<td>179</td>
<td>43</td>
<td>0.904 ± 0.010</td>
<td>0.00592 ± 0.000030</td>
<td>2.399</td>
<td>-19.395***</td>
<td>0.053</td>
</tr>
<tr>
<td>All <em>P. aethiopicus</em></td>
<td>45</td>
<td>19</td>
<td>0.905 ± 0.026</td>
<td>0.01086 ± 0.000050</td>
<td>4.386</td>
<td>-3.846</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Figure 3. Statistical parsimony haplotype networks for *Atelerix algirus* (top) and *Paraechinus aethiopicus* (bottom) based on the control region. Codes refer to haplotype numbers (see Appendix S1 for the correspondance between haplotypes numbers, specimens numbers and geographic origin).
demic to Morocco that differ from haplotypes from other countries by more than three mutations (except one haplotype shared between Morocco and Canary Islands), and the second haplogroup is composed of haplotypes shared (or closely related: one mutation) with specimens from Algeria and Tunisia. The first Moroccan haplogroup is mainly distributed in Atlantic plains of Morocco, and the second one is mainly distributed in southern and eastern Morocco. Both haplogroups occur in Atlas and Rif mountains (Fig. 1). Three haplotypes are found in Spain, and one of them is shared with Balearic Islands. We obtained significant negative values of Fu’s Fs for all haplogroups (Table 1). Positive values of R2 were obtained for all haplogroups, but they were not significant.

The 45 specimens of *P. aethiopicus* resulted in 19 haplotypes. Nucleotide diversity is 0.01086 ± 0.00050, haplotype diversity is 0.905 ± 0.026 and the average number of nucleotide differences is 4.386. A higher number of haplotypes is found in Algeria, but this is probably due to the greater sample size in this country (Fig. 3). No phylogeographic structure is apparent in the network: specimens from Morocco, Mauritania and Senegal can share the same haplotype, and this is also true for specimens from Algeria, Morocco and Mauritania, from Tunisia, Algeria and Mauritania, and from Algeria, Morocco and Tunisia. We obtained a negative value of Fu’s Fs and a positive value of R2, but they were not significant (Table 1).

**Figure 4.** Species distribution modeling of *Atelerix algirus* as estimated by Maxent for present-day conditions (current), for Holocene and the Last Glacial Maximum (LGM) based on the Model for Interdisciplinary Research on Climate (MIROC) and on the Community Climate System Model (CCSM) paleoclimatic models. Warmer colors show areas with higher probability of presence.

**Figure 5.** Species distribution modeling of *Paraechinus aethiopicus* as estimated by Maxent for present-day conditions (current), for Holocene and the Last Glacial Maximum (LGM) based on the Model for Interdisciplinary Research on Climate (MIROC) and on the Community Climate System Model (CCSM) paleoclimatic models. Warmer colors show areas with higher probability of presence.

**Species distribution modelling**

We built species distribution models for the two African hedgehog species based on the known presence localities. The AUC values were 0.935 and 0.808 for *A. algirus* and *P. aethiopicus*, respectively. Thus, they are considered to correspond to a useful predictive model.

Under present-day climatic conditions the model revealed climatic suitability for *A. algirus* across all the Mediterranean Basin (Fig 4). This area is larger than the currently known distribution of this species, as it is not
Discussion

Phylogeography

One of the most remarkable findings from this study is the discrepancy between our results and those published by Velo-Antón et al. (2019) for *A. algirus*. These authors found this species to be monophyletic and to include three distinct genetic lineages (lineages 1, 2 and 3), while we did not recover the monophyly of this species and found two distantly related lineages within it (lineages I and II). Our lineage I corresponds to the two sister lineages *A. algirus* 1 and *A. algirus* 2 of Velo-Antón et al. (2019). In their study, haplotypes from lineages 1 and 2 differ by at least 22 mutations. In our study these two lineages are not recovered in the phylogenetic tree, and they can only be identified in the haplotype network because the minimum number of mutations observed between haplotypes of these two lineages is only of three mutations (between haplotypes A13 and A14, or haplotypes A09 and A49). This discrepancy between the results of these two studies is striking. It is not only due to our greater sample size (allowing the recovery of more haplotypes), but may also reflect alignment differences and/or differences in the way of treating gaps and missing data (our network analyses are based on complete sequences only, the alignment included 7 gaps and gaps were treated as a fifth state; see Appendix S3).

To sum up, our results are congruent with the study of Khalidi et al. (Khalidi et al. 2016) in showing low genetic diversity in *A. algirus* across its geographic range. However, our more extensive sampling scheme and the use of a more rapidly evolving genetic marker allows us to highlight mitochondrial phylogeographic structure within this species: one haplogroup is endemic to Morocco and Canary Islands, and another haplogroup is widely distributed from Morocco to Libya, and through Spain and Balearic Islands. These results suggest that the mitochondrial phylogeographic structure of the species has a recent origin. This low mitochondrial structure at continental scale contrasts with the marked east–west divergence that has been commonly found in previous phylogeographic studies of North African organisms and which identified two common disjunction areas: one around the Algeria-Morocco border and the other one in Kabylia (central Algeria) (Nicolas et al. 2015; Beddek et al. 2018). It is interesting to note that the two haplogroups co-occur in Atlas and Rif mountains. Two hypotheses can be proposed to explain this pattern: (i) vicariance of an eastern and a western haplogroup followed by secondary expansion or (ii) retention of deep mitochondrial DNA diversity in Morocco followed by loss of diversity during eastward expansion of the species out of Morocco. Note that both scenarios imply a population expansion of haplogroup 2, for which, we find evidence only with Fu’s Fs test. Our climatic niche modeling results support the first hypothesis since they show that during the LGM there were several disconnected areas of high climatic suitability for *A. algirus*: one in South-West Morocco, one at the coastal Moroccan-Algerian border and one in Tunisia-coastal Libya.

Our genetic results confirm that the Algerian hedgehog recently colonized Spain, Balearic and Canary Islands,
and that this colonization was probably mediated by humans, as already suggested by several authors (Hutterer 1983; Borg 2005; Morales and Rofes 2008; Khalidi et al. 2016). Our ecological niche modeling analyses suggest that suitable climatic conditions occurred throughout most Southern and Eastern lowland Iberian Peninsula during the last 6000 years. This would have favored the spatial expansion of the Algerian hedgehog after its arrival in Europe. It is interesting to note that the current distribution of this species in Spain is much more reduced than expected based on climatic niche modelling. Species interactions (e.g. mutual competitive exclusion, predation, parasitism, and mutualism) are also important in determining species distribution (Pearman et al. 2008; Sinclair et al. 2010; Giannini et al. 2013). Competition with the European hedgehog, Erinaceus europaeus, a species already well established in Spain before the introduction of the Algerian hedgehog, may explain partially its restricted distribution in Spain.

Our study shows no phylogeographic structure in North African P. aethiopicus. This is consistent with our climatic niche modeling results showing a more or less continuous high or moderate climatic suitability for this species throughout most of the Saharan desert since the LGM (except in the most central parts). This study significantly increases our knowledge of the distribution of P. aethiopicus expanding currently recognized ranges (IUCN, 2015; Best, 2018). An extension of its distribution range in Atlantic Sahara was already proposed by Aulagnier et al. (2017). Velo-Antón et al. (2019), based on molecular data, proposed its extension in Mauritania, Senegal, Atlantic Sahara and Libya. We show that its distribution range should also be extended in Algeria since we captured it in the M’sila region, up to 35.748°N, confirming, with molecular data, the previous observation of Sellami et al. (1989).

**Taxonomy**

Several subspecies have been described for Atelerix algirus, either from the mainland or islands, Atelerix algirus vagans from Spain, Minorca and Balearic Islands and A. algirus girbaensis from Djerba Island (Tunisia). Given that the differential diagnosis is mainly based on pelage variations (hairs, spines), the validity of these subspecies has been regularly questioned (Corbet 1988; Hutterer 2005; Happold 2013). Our results do not show any strong genetic discontinuity between the Algerian, Moroccan, Tunisian, Libyan, Spanish, Balearic and Canarian populations, suggesting that subspecific recognition may be unwarranted. The status of the subspecies from Djerba Island (A. a. girbaensis) still needs to be assessed by genetic data. Several subspecies were also described within P. aethiopicus: P. a. aethiopicus from Soudan and Ethiopia, P. a. pectoralis from Arabian Peninsula, P. a. deserti from North Africa, P. a. albus from Djerba Island and P. a. ludlowi from Iraq. Our genetic data do not show any genetic discontinuity throughout North Africa, from Sene-}

**Conclusion**

In the two studied hedgehog species low genetic diversity was found in North Africa. However, while no phylogeographic pattern was found in P. aethiopicus, two haplogroups were identified within A. algirus. This could be explained by the fact that a more or less continuous high or moderate climatic suitability occurred throughout most of the Saharan desert since the LGM for the first species, while during the LGM there were several disconnected areas of high climatic suitability for A. algirus: one in South-West Morocco, one in coastal Moroccan-Algerian border and one in Tunisia-coastal Libya. Our genetic results confirm that the Algerian hedgehog colonized recently Spain, Balearic and Canary Islands, and that this colonization was probably mediated by humans. Suitable climatic conditions occurred throughout most lowland Southern and Eastern Iberian Peninsula during the last 6000 years which could have favored the spatial expansion of the Algerian hedgehog after its arrival in Europe.

Several subspecies have been described within the two African hedgehog species. Our results do not show any strong genetic discontinuity within either species in North Africa, suggesting that subspecific recognition within this region may be unwarranted. Additional sampling in Djerba Island (for both species), Soudan, Ethiopia and in the Middle East (for P. aethiopicus) is needed to test the genetic differentiation of other proposed subspecies.

Our phylogeographic and taxonomic conclusions are only based on short mitochondrial fragment, which could explain the low bootstrap support of several important nodes in our phylogeny. Increasing the amount of genetic data frequently allows resolution of even the most problematic phylogenetic relationships (DeFilippis and Moore 2000; Rokas and Carroll 2005; Fabre et al. 2017). One such approach is based on sequencing of complete mitochondrial genomes (“mitogenomics”). However, due to its maternal inheritance, it only estimates the matrilineal histories of individuals and populations. Because of the absence of recombination, the mitochondrial DNA should be still considered as a single locus and reconstructed phylogenies represent only “single gene” trees. There are many reasons why historical inferences derived directly from a gene genealogy may be inaccurate or misleading (Knowles 2004, 2009a, b). Each gene has its own history,
and gene trees can be different from the species tree. To address this problem, recent phylogenomic approaches target multiple nuclear fragments, which allow to infer a phylogenetic tree that accounts for discord among hundreds of independent loci at nuclear DNA (Lemmon et al. 2012). To better investigate the evolutionary history of the African hedgehogs such multi-locus data are needed. Moreover, integrative taxonomic studies combing these multilocus data and morphometric, behavioral and ecological data are required to test our taxonomic conclusions.

Acknowledgments

Molecular analyses were performed at the Service de Systématique Moléculaire (UMS 2700 Acquisition et Analyse de Données pour l’Histoire naturelle, 2AD, Paris, France). We are indebted to many local authorities for providing permits to carry out the research, especially ‘Direction Générale des forêts’ (Tunisia), ‘Haut Commissariat aux Eaux et forêts et lutte contre la desertification’ (Morocco). For help in the field and with logistics, we acknowledge A. Lalis, C. Denys, M. Benyettou, G. Barech, P. Foronda and M. Alcover. We thank the curator of the Museum of Natural Science of Barcelona for access to the tissue samples from Spain.

References


Supplementary material 1

List of specimens

Authors: Haithem El-Farhati, Mourad Khaldi, Alexis Ribas, Mohamed Wassim Hizem, Saïd Nouira, Violaine Nicolas (2021)

Data type: .xlsx

Explanation note: List of specimens included in the mitochondrial DNA phylogenetic.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/vz.71.e70989.suppl1

Supplementary material 2

Sequence alignment – phylogenetic tree analyses

Authors: Haithem El-Farhati, Mourad Khaldi, Alexis Ribas, Mohamed Wassim Hizem, Saïd Nouira, Violaine Nicolas (2021)

Data type: .csv

Explanation note: Sequence alignment used for the phylogenetic tree analyses.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/vz.71.e70989.suppl2

Supplementary material 3

Sequence alignment used for haplotype network analysis of A. algirus

Authors: Haithem El-Farhati, Mourad Khaldi, Alexis Ribas, Mohamed Wassim Hizem, Saïd Nouira, Violaine Nicolas (2021)

Data type: .csv

Explanation note: Sequence alignment used for haplotype network analysis of A. algirus.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/vz.71.e70989.suppl3
Supplementary material 4

Sequence alignment used for haplotype network analysis of *P. aethiopicus*

**Authors:** Haithem El-Farhati, Mourad Khaldi, Alexis Ribas, Mohamed Wassim Hizem, Saïd Nouira, Violaine Nicolas (2021)

**Data type:** .csv

**Explanation note:** Sequence alignment used for haplotype network analysis of *P. aethiopicus*.

**Copyright notice:** This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** https://doi.org/10.3897/vz.71.e70989.suppl4

Supplementary material 5

Species occurrence data

**Authors:** Haithem El-Farhati, Mourad Khaldi, Alexis Ribas, Mohamed Wassim Hizem, Saïd Nouira, Violaine Nicolas (2021)

**Data type:** .xlsx

**Explanation note:** Species occurrence data used in our climatic niche modelling analyses.

**Copyright notice:** This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

**Link:** https://doi.org/10.3897/vz.71.e70989.suppl5