Molecular Phylogenetics and Evolution 58 (2011) 365–373

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

When selection deceives phylogeographic interpretation: The case of the Mediterranean house gecko, *Hemidactylus turcicus* (Linnaeus, 1758)

C. Rato [a,b,c,*], S. Carranza [c], D.J. Harris [a,b]

[a] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[b] Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal
[c] Institute of Evolutionary Biology (CSIC-UPF) – Passeig Marítim de la Barceloneta, 37-49, E-08003 Barcelona, Spain

ARTICLE INFO

Article history:
Received 30 September 2010
Revised 8 December 2010
Accepted 10 December 2010
Available online 16 December 2010

Keywords:
Hemidactylus turcicus
Phylogeography
Selective sweep
mtDNA
nDNA

ABSTRACT

A previous study on *Hemidactylus turcicus* based on mtDNA makers indicated that this gecko has a Middle-East origin, and that the current phylogeographic pattern is the result of a very rapid spread from the east to the west of the species' range. The same study identified two distinct mitochondrial lineages with low differentiation and genetic diversity. Since *H. turcicus* is known to be closely associated to humanized environments, its present distribution range and phylogeography is frequently interpreted to be the result of recurrent human-mediated introductions. These conclusions used to be the same as those used to interpret the results obtained for the European populations of another gecko, *Tarentola mauritanica*. However, a recent study has revealed that the phylogeographic pattern of *T. mauritanica* is not solely the result of a recent colonization, but also of a mitochondrial selective sweep. Could the same be occurring in *H. turcicus*? To answer this question, two mitochondrial (12S rRNA and cytochrome *b*) and two nuclear genes (ACM4 and Rag2) were used in this study. From the mtDNA data we confirmed the existence of two distinct phylogeographic lineages; one occurring exclusively in the northern Mediterranean (Clade A), and another one more widespread that is the only lineage present in North Africa (Clade B). In light of these results, we could hypothesize that *H. turcicus* had its origin in Turkey, and from there Clade A moved to Europe and Clade B to North Africa spreading latter into Europe. However, Clade A presents significantly higher nucleotide diversity for the nuclear DNA compared to the mtDNA, and neutrality tests gave significant results for the mitochondrial data. These results suggest that the lack of mtDNA genetic diversity and structure in the European population of *H. turcicus* could also be due to a selective sweep, and not only because of a recent colonization. Together with the situation reported in *T. mauritanica*, the identification of a hitch-hiking process occurring in *H. turcicus*, represents two unprecedented cases of a selective sweep taking place in the same geographic area shaping the phylogeographic patterns of two unrelated genera of geckos.

1. Introduction

Hemidactylus geckos are one of the most species-rich genera of the Gekkonidae comprising over 85 species (Carranza and Arnold, 2006). Their distribution range includes tropical Asia and Africa, arid areas of northeast Africa and southwest Asia, and they also occur across the Mediterranean region. These geckos are very frequently associated to humanized habitats, living around or inside houses. Probably because of this, considerable evidence exists of possible human-mediated introductions of this genus, based on both direct observations and genetic markers (Carranza and Arnold, 2006; Jesus et al., 2005; Rocha et al., 2005; Vences et al., 2004).

Hemidactylus turcicus has a mainly circum-Mediterranean distribution including many islands and with populations extending to the south along the Nile River up to the border with Sudan (Mateus and Jacinto, 2008; Sindaco and Jeremcenko, 2008). They have also been introduced recently in the Canary islands (Geniez, 2002), Mexico, Cuba, Florida (Smith, 1946), and in other areas of the United States (Pianka and Vitt, 2003; White and Tumlison, 1999). Due to this rapid range expansion outside the Mediterranean region, Rödder and Lötters (2009) decided to assess the differences in climatic niches between the native and invaded ranges of *H. turcicus*, and to analyse which environmental variables are more conserved versus subject to niche shifts. Their results indicate that the degree of conservatism of niches in *H. turcicus* clearly depends on the predictors and variables considered *a priori*, and that these should always be tested with ecological niche models (ENMs). According to a recent phylogenetic study from Carranza and Arnold (2006), *H. turcicus* may have originated in the Middle East from

* Corresponding author at: CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.
E-mail address: catarina.rato@mail.icav.up.pt (C. Rato).

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.12.004

where it moved westwards across the whole Mediterranean, reaching later the Atlantic Ocean. In this same study, the authors reported two distinct mtDNA lineages of *H. turcicus* with little genetic divergence between them, suggesting that the phylogeographic pattern obtained was the result of a very rapid and recent spread.

Tarentola mauritanica is another gecko whose European populations present a similar mtDNA genealogy to *H. turcicus*, with low genetic diversity (Harris et al., 2004a,b). However, recent studies (Rato et al., 2010) have demonstrated that this mtDNA pattern is not solely the result of a recent and rapid spread (possibly human-mediated), but also due to a genetic hitch-hiking process (Maynard-Smith and Haigh, 1974). This process occurs when neutral alleles increase their frequency due to gametic disequilibrium with mutations that are favoured by natural selection. This will lead to a selective sweep, which results in a decrease of polymorphism at the affected markers. Since mitochondrial genomes do not experience recombination, genetic hitch-hiking has the potential to cause strong selective sweeps (Hamilton, 2009).

In the present study, we intend to investigate if a similar pattern is happening in *H. turcicus*. Therefore, we have increased the geographic sampling of *H. turcicus* covering approximately its entire native range, and sequenced two nuclear DNA markers and two mitochondrial regions in order to evaluate if the low genetic diversity in mtDNA lineages is reflected by limited variation in the nuclear markers. Tests were then applied to identify if a selective sweep may also have occurred in this species.

2. Material and methods

2.1. Samples, DNA extraction and amplification

The study was carried out using a total of 100 specimens of *H. turcicus* from 78 distinct localities and one *Hemidactylus lemurinus* to root the tree (see details in Table 1 and Fig. 1). Sixty-seven of these specimens correspond to new samples, and the remaining ones are from Carranza and Arnold (2006). From this previous study, we did not include the specimens of *H.t. turcicus* and *H.t. lavadeserticus* from Jordan, since their taxonomy is not clear (S. Carranza, pers. observ.).

Total genomic DNA was extracted from tail tissue, using the DNeasy Extraction Kit from Qiagen, following the manufacturer's protocol. Fragments of two mitochondrial (12S rRNA and cytochrome *b*) and two nuclear exons (ACM4, and Rag2) were amplified by PCR for 67 and 61 individuals, respectively. For amplification and sequencing of the 12S and cytochrome *b* (cytb) partial sequences we used the primers 12Sa/12Sb, and cytochrome b1/cytochrome b2, respectively, both from Kocher et al. (1989). PCR conditions were the same as those described in Harris et al. (1998). The ACM4 nuclear gene fragment was amplified and sequenced using the primers tg-F and tg-R published by Gamble et al. (2008b), and PCR reactions were carried out as in Rato et al. (2010). Amplification and sequencing of the Rag2 gene fragment was performed using two sets of primers; 31FN.Venk/Lung.460R (amplification) and Lung.35F/Lung.320R (amplification and sequencing) published by Hoegg et al. (2004). PCR conditions were the same as described in Chiari et al. (2004). Amplified fragments were sequenced on an ABI3730XL automated capillary DNA sequencer. Sequences obtained for this study have been deposited in GenBank.

2.2. Sequence alignment and phylogenetic analyses

Sequences of 12SrRNA and cytb from *H. turcicus* previously published elsewhere (Carranza and Arnold, 2006) were added to our dataset, and *H. lemurinus* was designated as the outgroup, since it has been shown to be the sister taxa of *H. turcicus* (Arnold and Carranza, unpublished data). Alignment of the obtained sequences was performed using MAFFT v.6.811 (Katoh, 2008) with default parameters (gap opening penalty = 1.53; gap extension penalty = 0.123; progressive method = FFT-NS-2), and exported to Bioedit (Hall, 1999) to be checked and adjusted by hand.

The appropriate model of nucleotide substitution for each mtDNA gene, and concatenated data was determined using jModeltest v.0.1.1 (Posada, 2008), under the Akaike Information Criterion (Akaike, 1974). Maximum likelihood (ML) and Bayesian phylogenetic analyses were conducted on the combined mtDNA dataset, considering the model of sequence evolution estimated earlier. ML analysis was performed using GARLI v.1.0 (Zwickl, 2006). Tree search was conducted using 5000–10,000 generations (*genthreshfortopoterm*) considering a stochastic algorithm, each resulting in a single best tree. Since no significant differences in the topology were observed when the number of generations was increased, bootstrap support was calculated from 1000 bootstrap replicates using 6000 *genthreshfortopoterm*. A 50% majority-rule consensus tree was generated using the software PAUP* v.4.0d10 (Swofford, 2002). Partitioned Bayesian phylogenetic analyses were conducted using MrBayes v.3.1.2 (Huelsenbeck and Ronquist, 2001). Both runs began with a random starting tree and ran for 2×10^6 generations, saving one tree in each 100 generations. Substitution-model parameters were always unlinked across partitions, and each subset was allowed to have its own rate (*prset* = variable). In both searches, the stationarity of the Markov chain was determined as the point when sampled log-likelihood values plotted against generation time, reached a stable mean equilibrium value 25% of the sampled trees were discarded (burnin = 5000). Tracer v1.4 (Rambaut and Drummond, 2007) was used to confirm that all parameters had an ESS > 100 after burnin. The remaining trees were combined and a 50% majority consensus tree was generated.

2.3. Population genetic analyses

In order to investigate the haplotype diversity and structure within each marker, we built two haplotype networks; one for the combined mtDNA and another one for each nDNA gene. A median-joining haplotype network was constructed using the Fluxus Phylogenetic Network Analysis software v.4.5.1.0 (Bandelt et al., 1999; http://www.fluxus-engineering.com) with the parameter epsilon set to 0. Prior to this analysis, for each nuclear gene an input file with identification of all heterozygous positions was created using SeqPHASE (Flot, 2010). This generated file was then used to run PHASE v.2.1.1 (Stephens and Donnelly, 2003; Stephens et al., 2001) in order to reconstruct the phase of each haplotype, using the default parameters (thresholds are $p = q = 90\%$). Definition of the number of populations of *H. turcicus* was carried out with the mtDNA dataset using BAPS v.5 (Bayesian Analysis of Population Structure) (Corander et al., 2008). Genetic mixture analysis was performed at the individual level considering a spatial model. The *Snn* test (Hudson, 2000) was applied to each nuclear gene fragment to test for genetic differentiation between the obtained mtDNA clades. In order to calculate the *p*-value of the *Snn* we performed a permutation test with 1000 replicates implemented in the software DnaSP v.5 (Librado and Rozas, 2009). Within each of the defined populations we calculated Tajima's *D* (Tajima, 1989) for both mitochondrial and nuclear DNA, since it is expected that under a hitch-hiking model the value of Tajima's *D* will be large and negative, indicating a skew toward rare variants (Braverman et al., 1995; Fay and Wu, 2000; Kim, 2006). In order to compare the number of mutations in internal and external branches of a genealogy with the expectation of selective neutrality, we applied