Gene 512 (2013) 166–173

Contents lists available at SciVerse ScienceDirect

Gene

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

Short Communication

Evolutionary patterns of the mitochondrial genome in the Moorish gecko, *Tarentola mauritanica*

C. Rato [a,b,c,*], S. Carranza [c,1], A. Perera [a,2], D.J. Harris [a,b,2]

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

ARTICLE INFO

Article history:
Accepted 5 September 2012
Available online 18 September 2012

Keywords:
Tarentola mauritanica
Selective sweep
Positive selection
Background selection
Polymorphism
Bottleneck

ABSTRACT

A previous study on the evolutionary patterns of *Tarentola mauritanica* demonstrated that low levels of mitochondrial diversity observed in the European populations relative to nuclear markers were consistent with a selective sweep hypothesis. In order to unravel the mitochondrial evolutionary history in this European population and two other lineages of *T. mauritanica* (Iberian and North African clades), variation within 22 nearly complete mitogenomes was analyzed. Surprisingly, each clade seems to have a distinct evolutionary history; with both the European and Iberian clades presenting a decrease of polymorphism, which in the former is consistent with departure from neutrality of the mtDNA (positive or background selection), but in the latter seems to be the result of a bottleneck after a population expansion. The pattern exhibited by the North African clade seems to be a consequence of adaptation to certain mtDNA variants by positive selection.

1. Introduction

The vertebrate mitochondrial genome is around 16–18 kb in size, circular, with double stranded molecules, and typically composed of 13 protein-coding genes, 22 transfer RNA genes, two ribosomal RNA genes, a non-coding control region, and some intergenic spacers. Over the last 30 years, mitochondrial DNA has been the most popular and commonly used molecular marker to estimate genetic diversity in a variety of organisms and its initial success was mainly because it was thought to be always maternally inherited, nearly neutral, and with an almost clock-like nature of its substitution rates. Because these organelles perform such important role in cellular respiration, and it is where most of the ATP is produced, any type of malfunction could be lethal or would seriously compromise an individual's fitness. Therefore, it was initially assumed that the mitochondrial genome would evolve primarily under constant purifying selection that would eliminate possible deleterious mutations. However, a number of exceptions have been reported showing evidences of positive selection acting on the mitochondria's evolution (Castoe et al., 2008, 2009; Grossman et al., 2004; Yu et al., 2011). Nevertheless, neutrality assumptions were previously challenged by Bazin et al. (2006), using a broad range of vertebrate and invertebrate animals, concluding that mtDNA genetic diversity is not affected by population size, which could be explained by recurrent selective sweeps in large populations. These recurrent sweeps are in agreement with the "genetic draft" model of Gillespie (2000). Considering the non-recombining nature of the mitochondria, this genetic hitch-hiking has the potential to cause strong selective sweeps (Hamilton, 2009). In fact, in the past few years several cases of possible selective sweeps have been recorded, namely in warblers (Bensch et al., 2006), birds from the genus *Emberiza* (Irwin et al., 2009), butterflies (Jiggins, 2003), and Hawaiian crickets (Shaw, 2002). More recently, it has been suggested that the mtDNA pattern presented by the European populations of two species of geckos, *Tarentola mauritanica* (Rato et al., 2010), and *Hemidactylus turcicus* (Rato et al., 2011) could have been shaped by a selective sweep leading to a decrease of the nucleotide diversity to levels below those presented by nuclear markers. However, at that time this hypothesis could not be verified at the genome level.

Abbreviations: Ala, Alanine; Asn, Asparagine; Asp, Aspartic acid; Arg, Arginine; ATP6, ATP synthase 6; ATP8, ATP synthase 8; bp, Base pair; COXI, Cytochrome c oxidase 1; COXII, Cytochrome c oxidase 2; COXIII, Cytochrome c oxidase 3; Cys, Cysteine; CYTB, Cytochrome b; dNTP, Deoxyribonucleoside triphosphate; Gln, Glutamine; Glu, Glutamic acid; Gly, Glycine; HKA test, Hudson, Kreitman and Aguadé test; His, Histidine; Ile, Isoleucine; Leu, Leucine; Lys, Lysine; Met, Methionine; MgCl2, Magnesium chloride; MLHKA, Maximum likelihood version of the HKA test; mtDNA, Mitochondrial DNA; ND1, NADH dehydrogenase 1; ND2, NADH dehydrogenase 2; ND3, NADH dehydrogenase 3; ND4, NADH dehydrogenase 4; ND4L, NADH-ubiquinone oxidoreductase chain 4L; ND5, NADH dehydrogenase 5; ND6, NADH dehydrogenase 6; Numt, Nuclear mitochondrial DNA; PCR, Polymerase chain reaction; Phe, Phenylalanine; Pro, Proline; rRNA, Ribosomal RNA; Ser, Serine; Thr, Threonine; tRNA, Transfer RNA genes; Trp, Tryptophan; Tyr, Tyrosine; Val, Valine.

* Corresponding author at: CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. Tel./fax: +351 252660411, +351 252 661 780.
E-mail addresses: catarina.rato@cibio.up.pt (C. Rato), salvador.carranza@ibe.upf-csic.es (S. Carranza), perera@cibio.up.pt (A. Perera), james@cibio.up.pt (D.J. Harris).
[1] Tel./fax: +34 93 230 96 44, +34 93 230 95 55.
[2] Tel./fax: +351 252660411, +351 252 661 780.

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2012.09.032

Nevertheless, genetic hitch-hiking can also take place when negative or purifying selection removes deleterious mutations from the population driving them to loss (Hamilton, 2009). This negative selection against deleterious mutations is also expected to reduce polymorphism through hitch-hiking in a process called background selection (Charlesworth et al., 1993).

In the present study, we have used 22 nearly complete mitochondrial genomes of *T. mauritanica* to test selective sweep hypothesis. This gecko species belonging to the Family Phyllodactylidae (Gamble et al., 2008), is apparently a complex of several highly divergent cryptic species (Harris et al., 2004a, 2004b, 2009; Perera and Harris, 2008), but evidences of a selective sweep were detected in only one lineage, geographically widespread across Europe and North Africa (Rato et al., 2010). Thus, used mitogenomes in this study correspond to three different lineages of *T. mauritanica*; one lineage previously identified as subject to a selective sweep, and two others where this process was not detected. Minimal genetic variation is expected in the European lineage, but also the detection of fixed mutations responsible for amino acid changes that could be the driving force behind the sweep. Populations from the Iberian Peninsula are also expected to present little mitochondrial genetic diversity, consistent with a post-bottleneck demographic expansion pattern, since they also present low values of variation for nuclear markers (Rato et al., 2010). We report on rates and patterns of mtDNA sequence variation, and perform specific tests to detect selection.

2. Material and methods

2.1. Sample collection and DNA extraction

The 22 specimens used in this study are representatives of three different lineages of *T. mauritanica* identified in previous studies (Harris et al., 2004a, 2004b, 2009; Perera and Harris, 2008; Rato et al., 2010); namely six specimens (DB11105, DB9107, DB9112, DB9113, DB9115, and DB9116) representing the widespread European lineage; five representatives (DB3832, DB3839, DB3843, DB3846, DB3853) from the Iberian Peninsula lineage; and 10 specimens (DB11003, DB11004, DB11007, DB11008, DB11009, DB11013, DB11022, DB11035, DB11042, DB2635) from the Central and Southern Morocco lineage. A complete mitochondrial genome from *T. mauritanica* already available in GenBank (EU443255.1 from Albert et al., 2009) was used in the analyses, and assigned to the European Clade according to the phylogenetic results from 12SrRNA and 16SrRNA fragments (data not shown).

Genomic DNA extraction from fresh tail muscle tissue was carried out using Qiagen's DNeasy Blood & Tissue Kit, following the manufacturer's protocol.

2.2. Primer design, amplification and sequencing

Since amplification of several mitochondrial regions was unsuccessful using the primers from Albert et al. (2009), new primers were developed for this study. The complete genome from these authors was used as a reference for primer design, as well as for the alignment of the obtained amplicons. Primers were designed by hand, and checked for quality using the software AmplifX v1.5.4 (http://ifrjr.nord.univ-mrs.fr/AmplifX-Home-page). The complete list of primers used in this study is represented in Table S1 from Additional file 1.

Mitochondrial fragments were amplified through a standard Polymerase Chain Reaction (PCR) protocol, using 25 μl volumes, containing 2.5 μl of 10× reaction buffer, 2.0 mM of $MgCl_2$, 0.5 μM of each dNTP, 0.2 mM of each primer, 1 U of Taq DNA polymerase (Invitrogen), and approximately 100 ng of template DNA, and under the following cycling conditions for all primer sets; 94 °C for 3 min of initial denaturation followed by 35 cycles of denaturation at 94 °C for 45 s, annealing at 48–54 °C for 30 s, and extension at 72 °C for 1 min and 30 s. Amplification ended with a 10 min at 72 °C final extension step. The amplicons were checked in a 2% agarose gel with a negative control. Sequencing of the PCR products was performed in Macrogen Inc. (Korea) using an ABI 3730XL DNA Analyzer (Applied Biosystems). All obtained genomes were deposited in GenBank with accession numbers from JQ425040 to JQ425060.

2.3. Genome annotation

The newly obtained sequences were assembled and aligned with the currently available mitochondrial genome of *T. mauritanica*, using the software package Geneious Pro v5.4.3 (Drummond et al., 2010). Alignment was carried out with MAFFT v6.814b (Katoh et al., 2002) considering the auto algorithm, gap open penalty = 1.53, offset value = 0.123, and the scoring matrix 200PAM/k = 2. The protein-coding gene sequences were identified and annotated in Geneious Pro v5.4.3 by the start and stop codons, and sequence similarity in the alignment. Inconsistencies between the obtained sequences in this study, and the mitogenome from Albert et al. (2009) were detected, regarding the orientation of two tRNAs (tRNA-Gln and tRNA-Glu). Using the program ARWEN for tRNA detection in metazoan mitochondrial sequences (Laslett and Canbäck, 2008), we identified an incorrect orientation of both tRNA fragments on the GenBank sequence. Since transfer or "transposition" of cytoplasmic mitochondrial DNA sequences into the separate nuclear genome of a eukaryotic organism (Numt) is a possible phenomenon (Bensasson et al., 2001), all amplicons were translated into amino acids in order to inspect for the presence of stop codons inside the protein-coding genes.

In order to identify the nucleotide fixed differences in the European clade compared to the remaining lineages, we used the software SITES (Hey and Wakeley, 1997), and classified these differences into synonymous and non-synonymous using the Geneious Pro v5.4.3 software package.

2.4. Population genetic analyses and tests for selection

Population genetic parameters, such as the nucleotide diversity (π), number of segregating sites, number of invariable sites and average number of nucleotide differences between lineages, were calculated using DnaSP v5.10 (Librado and Rozas, 2009) for each lineage and locus. In order to calculate the nucleotide diversity across the coding portion of the mitochondrial genome, we used a 500 steps sliding window with a step size of 100 bp. Calculations were performed for each lineage, and for the complete set of mtDNA genomes of *T. mauritanica*.

Within each lineage, all coding genes were tested for evidences of selection using Tajima's *D*, which compares the number of rare to intermediate-frequency variants (Tajima, 1989). However, a test sensitive to high-frequency variants such as Fay and Wu's H is particularly useful to detect hitch-hiking and therefore, positive selection (Fay and Wu, 2000). This test was implemented using as outgroups two mitogenome sequences of *Gekko gecko* (GenBank accession numbers: AY282753.1 and HM370130.1), which belongs to the Gekkonidae, the sister Family of Phyllodactylidae (Gamble et al., 2008). Both tests were implemented in the software package DnaSP v5.10 (Librado and Rozas, 2009), where also statistical significance was assessed through coalescence, considering 10,000 simulations.

Another neutrality test was implemented, namely a maximum likelihood version of the Hudson, Kreitman and Aguadé (HKA) test (Hudson et al., 1987) developed by Wright and Charlesworth (2004) and available from S.I. Wright (http://www.yorku.ca/stephenw/). Similar to the HKA test, the MLHKA compares sequence polymorphism within species to rates of divergence between species, in order to test for deviations from neutrality. Nevertheless, while the HKA test does not identify which loci deviate from expectations, MLHKA is able to test directly the specific loci for evidences of selection.

The MLHKA test was conducted for each locus and each lineage, using as the outgroup taxon the two mitogenomes of *Gekko gecko*