The complete mitochondrial genome of the parthenogenetic Caucasian rock lizard *Darevskia unisexualis* (Squamata: Lacertidae) contains long tandem repeat formed by 59 bp monomer

Aleksey Komissarov\(^a\), Vitaly Korchagin\(^b\), Sergei Kliver\(^a\), Pavel Dobrynin\(^a\), Serafima Semyenova\(^b\), Andrey Vergun\(^b\), Stephen O'Brien\(^a\) and Alexey Ryskov\(^a\)

\(^a\)Theodosius Dobzhansky Center for Genome Bioinformatics Saint Petersburg State University, St. Petersburg, Russia; \(^b\)Laboratory of Genome Organization, Institute of Gene Biology of the Russian Academy of Sciences, Moscow, Russia

**ABSTRACT**

The first complete mitochondrial genome sequence of parthenogenetic Caucasian rock lizard *Darevskia unisexualis* (Lacertidae family) is determined by hybrid assembly with Illumina HiSeq and PacBio RS II platforms. The circular 21.4 kbp mitogenome contains 13 protein-coding genes, 125 and 16S rRNA genes, 20 tRNAs, two pseudogenized tRNAs, and one long tandem repeats with 4.1 kbp length formed by 59 bp monomer repeated 17.6 times located before control region. This finding represents a new example of mitogenome variation in lizards of hybrid origin, providing fundamental data for following study of a unique hybridization system formed by parthenogenetic and bisexual species in the mountain steppe of central Armenia.

Parthenogenetic lizard species represent a unique model of organisms for study of genetic and ecological bases of hybridochroic speciation. *Darevskia unisexualis* is one of the seven parthenogenetic species originated from the interspecific hybridization of *D. raddei nairensis* (maternal species) and *D. valentini* (paternal species), which belong to different clades (Moritz et al. 1992; Murphy et al. 2000; Fujita & Moritz 2010).

Mitochondrial genomes of *Squamata* often have unusual and complicated structure (Amer and Kumazawa 2005; Dong & Kumazawa 2005; Amer & Kumazawa 2007; Fujita et al. 2007; Kumazawa et al. 2014). In this article we report a new kind of rearrangement in mitochondrial genome of assemblage lizard *Darevskia unisexualis*: a long tandem repeat with 59 bp monomer located before the control region (Daversky et al. 1985; Moritz et al. 1992; Murphy et al. 2000; Fujita & Moritz 2010).

DNA samples extracted from blood of *D. unisexualis* obtained previously between 1997 and 2006 from lizards of Armenian population Takyurlu (Artavaz, 40°37’20.2”N 44°34’51.4”E) were delivered to the Asex/Macrogen (Korea) for genomic library preparation and sequencing with Illumina HiSeq 2000 and PacBio RS II.

To overcome complexity of control region assembly we used approach based on reads extraction with following greedy assembly previously successfully adopted in *Selenodon paradoxus* mitogenome assembly (Brandt et al. 2016). A subset of Illumina reads having common kmers with *Podarcis siculus* mitogenome (NC_011609; Podnar et al. 2009) was extracted with Cookiecutter software (Starostina et al. 2015), mapped back to the reference mitogenome with bowtie2 (Langmead & Salzberg 2012), and assembled into consensuses sequence, that was used as query for alignment against PacBio reads with Blast aligner (Chaisson & Tesler 2012). Cookiecutter’s kmer database was updated from 215 successfully found PacBio reads. Updated kmer database was used for accurate read extraction from Illumina HiSeq raw data. After that Illumina reads were mapped back to PacBio reads and assembled to the final sequence. The consistency of assembly was verified: (1) by consistency with PacBio reads that were mapped to it with MAFFT tool (Katoh & Standley 2013); (2) by consistency with previously published restriction map (Moritz et al. 1992) that has a perfect match with assembled mitogenome except two inverted restriction sites; (3) consistency of Illumina reads coverage. The assembled genome was annotated with MITOS WebServer (Bert et al. 2013) with additional checking for predicted tRNA and protein-coding genes.

The circular assembled mitogenome size is 21,433 bp, with 13 predicted protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, and 20 tRNAs, two pseudogenized tRNAs, and one complex tandem repeats with 59 bp monomer repeated 17.6 times and located before control region. MITOS WebServer predict second ND4 in 17202–18491 regions inside found tandem repeat, however, verification by Blast alignment did not find any significant match in this region. The validation of phylogenetic position is shown in Figure 1.

**CONTACT** Aleksey Komissarov (ad3002@gmail.com) Theodosius Dobzhansky Center for Genome Bioinformatics Saint Petersburg State University, St. Petersburg, 199034, Russia

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Unusual mitochondrial genome of parthenogenetic lizards containing long tandem repeat raises questions about stability and evolution of mtDNA after recent hybridization events. Moreover, its features provide unique opportunities to study mtDNA evolution not only in Squamata taxon, but in all animals.

Data availability

_D. unisexualis_ mitogenome assembly is available from Genebank with accession number KX644918.

Disclosure statement

The authors declare that they have no conflicts of interest concerning this article.

Funding

This study was supported by the Russian Science Foundation, 10.13039/501100006769 [14-14-00832]

ORCID

Andrey Vorgun [http://orcid.org/0000-0001-5255-7802

References


