The determination and analysis of the complete mitochondrial genome of *Dario dario* (Anabantiformes: Badidae)

Qi Wang\textsuperscript{a,b}, Xinyu Zhao\textsuperscript{a,b}, Chuanchen Ren\textsuperscript{a,b}, Fang Meng\textsuperscript{a,b}, Hongqiang Wang\textsuperscript{c}, Yifan Liu\textsuperscript{a,b}, Bingjian Liu\textsuperscript{a,b} and Enshang Yang\textsuperscript{d}

\textsuperscript{a}National Engineering Research Center for Marine Aquaculture, Zhejiang Ocean University, Zhoushan, China; \textsuperscript{b}National Engineering Laboratory of Marine Germplasm Resources Exploration and Utilization, Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, China; \textsuperscript{c}Zouchuan Hospital of Zhejiang Province, Zhoushan, China; \textsuperscript{d}School of Marine Engineering Equipment, Zhejiang Ocean University, Zhoushan, China

**ABSTRACT**

The classification of Badidae family based on morphology has been revised several times, but data on complete mitogenome are scarce, the complete mitochondrial genome of the Badidae fish *Dario dario* was characterized for the first time in the present study. The whole mitogenome was 16,830 bp in size and consisted of 13 protein-coding genes, 22 tRNAs, two rRNAs genes, a control region and origin of light-strand replication. The proportion of coding sequences with a total length of 11,431 bp was 67.92%, which encoded 3800 amino acids. The genome composition was highly A+T biased (58.12%), and exhibited a negative AT-skew (−0.0045) and GC-skew (−0.2347). All protein-coding genes started with ATG except for GTG in CO1, while stopped with the standard TAN codons or a single T. The control region (D-loop) ranging from 15,658 bp to 16,830 bp was 1173 bp in size. Phylogenetic analysis showed that *D. dario* was most closely related to *Badis badis*. The complete mitochondrial genome sequence provided new insight into taxonomic classification, and a more complex picture of species diversity within the Anabantiformes.

**ARTICLE HISTORY**

Received 30 May 2021
Accepted 13 September 2021

**KEYWORDS**

*Dario dario*; Badidae; Anabantiformes; mitochondrial genome; phylogenetic relationship

**CONTACT**

Enshang Yang yangenshang@126.com School of Marine Engineering Equipment, Zhejiang Ocean University, No. 1, Haida South Road, Lincheng Changzhi Island, Zhoushan, Zhejiang 316022, China

Supplemental data for this article is available online at https://doi.org/10.1080/23802359.2021.1981164.
Thirteen PCGs were 11,431 bp (67.92%) and encoded 3800 amino acids. Moreover, the AT-skew (−0.0733) and GC-skew (−0.2918) for the PCGs in *D. dario* were negative. All the PCGs used the initiation codon ATG except for GTG in CO1. Besides, CO2, ND4, CytB ended by single T, ND3 ended by TAG, all the others ended by TAA. The base content of nucleotides differed in the sense strands of the PCGs (T, 31.05%; A, 26.80%; G, 14.92%; C, 27.22%). The overall A + T content in the sense strands of the PCGs (57.85%) showed the obvious bias in the AT nucleotide composition. The values of RSCU showed that Leu2; Val; Ser1; Pro; Thr; Ala; Arg; Gly were higher codon usage at the same level encoded by four synonymous codons, while the others were lower codon usage encoded by either three or two codons.

The lengths of 12S rRNA and 16S rRNA were 950 bp and 1694 bp, it showed a positive AT skew (0.2138) and negative GC skew (−0.0755). The total length of the 22 tRNAs in the *D. dario* mitochondrial genome was 1553 bp, and the overall A + T content of tRNAs was 56.60%. It had a positive AT skew (0.1081), but negative GC skew (−0.1246). The length of CR was 1173 bp, ranging from 15,658 bp to 16,830 bp, 420 nucleotides for A, 385 nucleotides for T, both of them accounting for 68.63% of the whole D-loop, the AT and GC skew values were 0.0435 and −0.1957.

Based on the Akaike information criterion (AIC), GTR + G + I + F was indicated as the best-fitting substitution model for the phylogenetic relationship analysis. In Figure 1, it is obvious that *D. dario* was most closely related to *B. badis*; these two species formed a monophyletic clade with high support value constituting a Badidae group. Besides, Anabantidae + Helostomatidae + Osphronemidae forms a monophyletic clade, and formed sister branches with Badidae + Pristolepididae + Channidae. Phylogenetic analysis was used to get a clear understanding of classification status, and here better clarification of the phylogenetic classification of *D. dario*. The more discovery of these species will further promote more research on Badidaes.

**Disclosure statement**

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

**Funding**

This research was supported by Zhejiang Provincial Natural Science Foundation of China [Nos. LY20C190008, LY19H160012], Science and Technology Project of Zhoushan [No. 2020C21016], Medical and Health Science and Technology Project of Zhejiang Province [No. 2019337264], Science and Technology Project of Traditional Chinese Medicine of Zhejiang Province [No. 2018ZA126], Open Foundation from Marine Sciences in the First-Class Subjects of Zhejiang [Nos. 20200201,
20200202), and Starting Research Fund from the Zhejiang Ocean University.

Data availability statement
The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/ under the accession no. MT344964.1.

References
Barnett R, Larson G. 2012. A phenol–chloroform protocol for extracting DNA from ancient samples. Methods Mol Biol. 840:13–19.
Basumatary S, Choudhury H, Baishya RA, Sarma D, Vishwanath W. 2016. Badis panchratnaensis, a new percid fish species from Brahmaputra River drainage, Assam, India (Teleostei: Badidae). Vertebr Zool. 66: 151–156.
Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27(8): 1767–1780.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. Mega X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549.
Lowe T, Eddy S. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25(5): 955–964.
Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol. 41(3):353–358.
Rüber L, Britz R, Kullander SO, Zardoya R. 2004. Evolutionary and biogeographic patterns of the Badidae (Teleostei: Perciformes) inferred from mitochondrial and nuclear DNA sequence data. Mol Phylogenet Evol. 32(3):1010–1022.