The complete mitochondrial genome of the endangered freshwater fish *Odontobutis obscurus* from Korea

Kang-Rae Kim\textsuperscript{a}, Yeong-Ho Kwak\textsuperscript{b}, Mu-Sung Sung\textsuperscript{a}, Gun Young Lee\textsuperscript{a}, Yong Hwi Kim\textsuperscript{a} and In-Chul Bang\textsuperscript{a}\textsuperscript{,} \textsuperscript{,}

\textsuperscript{a}Department of Life Science & Biotechnology, Soonchunhyang University, Asan, Republic of Korea; \textsuperscript{b}Inland Fisheries Research Institute, National Institute of Fisheries Science, Gapyeong, Republic of Korea

**ABSTRACT**

We report the first complete mitochondrial genome of *Odontobutis obscurus*, which consists of 17,038 bp harboring 13 protein-coding genes, two ribosomal RNA genes, 22 transfer RNA genes, and a control region (D-loop). The overall base composition of the complete genome is A (30.63%), C (28.72%), T (25.70%), G (14.95%). The complete mitogenome of *Odontobutis obscurus*, most closely related to congeners in the Bayesian inference and maximum likelihood tree, provides a better understanding of the phylogeny of the genus *Odontobutis*.

A DNA library was prepared to obtain the complete mitochondrial DNA sequence. DNA library preparation was conducted using the MGIEasy DNA Library Prep Kit (MGI Technology, Shenzhen, China) according to the manufacturer’s instructions, and sequencing was performed with 150 base pairs (bp) paired-ends using the MGISEQ-2000 platform (MGI Technology). The raw data obtained from the MGISEQ-2000 platform were assembled using Gendieus 11.0.3 (https://www.geneious.com) and annotated with the MITOS web server (Bernt et al. 2013). The complete mitogenome sequence of *O. obscurus* was deposited at the National Center for Biotechnology Information GenBank under accession number MW646297.

The mitochondrial genome of *O. obscurus* consists of 17,038 bp, harboring 13 protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, 22 transfer RNA (tRNA) genes, and a control region (D-loop). Two PCGs (CO1 and ND3) start with a GTG codon and the other 11 PCGs start with ATG. The overall base composition is A (30.63%), C (28.72%), T (25.70%), and G (14.95%), with a high AT content of 56.33%. The two rRNAs are 16S (1,675 bp) and 12S (951 bp).

The dataset was aligned using the ClustalW module of MEGA X and the TIM2 + I + G model was the best fit nucleotide substitution model as calculated by jModelTest 2.1.10 (Guindon and Gascuel 2003; Darriba et al. 2012). A phylogenetic tree was constructed based on 13 PCGs using the Bayesian inference (MrBayes 3.2.7) and maximum likelihood (PhyML 3.0) methods (Guindon et al. 2010; Ronquist et al. 2012).

**CONTACT** In-Chul Bang incbang@gmail.com Department of Life Science & Biotechnology, Soonchunhyang University, Asan 31538, Republic of Korea

© 2022 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.
2012), which grouped O. obscurus in the genus Odontobutis (Figure 1).

**Author contributions**

The first author (Kim Kang-Rae) analyzed the data and wrote the final thesis. The second author (Yeong-Ho Kwak) collected the data and wrote the thesis. The third author (Mu-Sung Sung) designed the study. The fourth author (Gun Young Lee) prepared the draft and collected the data. The fifth author (Yong Hwi Kim) prepared the draft and wrote the thesis. The corresponding author (Bang In-Chul) interpreted and approved the data while checking the content of final version of the paper.

**Disclosure statement**

The authors report no conflicts of interests. The authors alone are responsible for the content and writing of this article.

**Funding**

This study was supported by Soonchunhyang University Research Fund and Korea Native Animal Resources Utilization Convergence Research Institute [NRF-2021R1A6A1A03039503].

**ORCID**

Kang-Rae Kim [http://orcid.org/0000-0002-1006-3123](http://orcid.org/0000-0002-1006-3123)
Yong Hwi Kim [http://orcid.org/0000-0001-9901-5445](http://orcid.org/0000-0001-9901-5445)
In-Chul Bang [http://orcid.org/0000-0003-4584-5384](http://orcid.org/0000-0003-4584-5384)

**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/](https://www.ncbi.nlm.nih.gov/) under the accession no. MW646297. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA752481, SRX11659825 and SAMN20599215, respectively.

**References**

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 69(2): 313–319.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 9(8):772.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3): 307–321.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52(5): 696–704.

[ME] Ministry of Environment. 2012. Wildlife protection and management act (Law No. 10977). [http://www.law.go.kr/lsInfoP.do?lsiSeq=125808&viewCls=lsRvsDocInfoR#](http://www.law.go.kr/lsInfoP.do?lsiSeq=125808&viewCls=lsRvsDocInfoR#).

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61(3):539–542.