

The complete mitochondrial genome of *Sarotherodon galilaeus* (Linnaeus, 1758) (Perciformes: Cichlidae) and its phylogenetic placement

Congqiang Luo^{a,b,c,d}, Pinhong Yang^{a,b,c,d} and Suqin Wang^{a,b,c,d}

^aHunan Engineering Research Center of Aquatic Organism Resources and Environmental Ecology, Hunan University of Arts and Science, Changde, China; ^bHunan Provincial Key Laboratory for Health Aquaculture and Product Processing in Dongting Lake Area, Hunan University of Arts and Science, Changde, China; ^cHunan Provincial Collaborative Innovation Center for Efficient and Health Production of Fisheries, Hunan University of Arts and Science, Changde; ^dChangde Research Center for Agricultural Biomacromolecule, Hunan University of Arts and Science, Changde, China

ABSTRACT

Sarotherodon galilaeus (Linnaeus, 1758), a cichlid species that is naturally distributed in African and Eurasian waters, was introduced in many Asian countries for aquaculture. To date, rare genetic studies focused on this species have hindered our understanding of this species. Here, we reported the complete mitochondrial genome of *S. galilaeus* that was sequenced using next-generation sequencing technology. The resulting mitogenome of *S. galilaeus* was 16,630 in length and comprised 13 protein-coding genes (PCG), 22 transfer RNA (tRNA) genes, 2 ribosomal RNA genes (rRNA), and one control region (D-loop). Phylogenetic analysis indicated that Oreochromini species contained two lineages (I and II) and *S. galilaeus* clustered with *Oreochromis aureus* rather than other *Sarotherodon* species.

ARTICLE HISTORY

Received 7 October 2020
Accepted 5 February 2021

KEYWORDS

Sarotherodon galilaeus;
mitochondrial genome;
phylogenetic analyses



Sarotherodon galilaeus (Linnaeus, 1758), a cichlid species that is naturally distributed in Africa and Eurasia, was introduced in many Asian countries for aquaculture, including in China (www.fishbase.org). Given that this species can tolerate hypersaline conditions, and can reproduce even in the high level of salinity, *S. galilaeus* was used to hybridize with other Oreochromini species (i.e. *Oreochromis niloticus*) for improving adaptive capacity (Yan and Wang 2010). To date, studies largely focused on the physiology and biology of *S. galilaeus* (Zeilstra et al. 2003; Yan and Wang 2010), whereas sparse genetic analyses have been conducted for this species. Therefore, this study reported the whole mitochondrial genome of *S. galilaeus* for the first time using next-generation sequencing technology and attempted to infer its phylogenetic status.

We collected the sample of *S. galilaeus* in July 2019 from Hengxian county (22.677 N, 109.266E), Guangxi Province, China. The sample (Voucher number: JLVLF2019001) was stored in the fish collection of Hunan University of Arts and Science. Total genomic DNA was extracted from fin tissues using a Genomic DNA Isolation Kit (QiaGene, Germany). We used the Illumina MiSeq platform to sequence the complete mitochondrial genome (Illumina Inc, San Diego, CA, USA) and assembled the raw sequence reads into contigs using SPAdes 3.9.0 (Bankevich et al. 2012). We finally obtained the complete mitochondrial genome using the contigs in SOAPdenovo (Luo et al. 2012) and annotated the protein-

coding genes and rRNA genes using MITOS (Bernt et al. 2013).

The size of the *S. galilaeus* mitogenome was found to be 16,630 base pairs (bp) in length and comprised (*ND1*, *ND2*, *ND3*, *ND4*, *ND4L*, *ND5*, *ND6*, *COI*, *COII*, *COIII*, *ATP6*, *ATP8*, *Cyt b*), 2 rRNA genes (12S rRNA and 16S rRNA), 22 tRNA genes and a control region (D-loop) (GenBank nos: MW046257). The structural organization and gene order were in line with those in other typical teleosts. A total of 18 mitogenomes including 17 Oreochromini mitogenomes and one outgroup were aligned using MUSCLE (Edgar 2004). All 13 protein-coding genes were selected manually and then were combined into a single sequence for phylogenetic analyses. Maximum likelihood (ML) technique with nucleotide substitution model (GTR + I + G) was employed to construct trees. The ML tree was constructed in RAXML-VI-HPC (Stamatakis 2006) with 1000 nonparametric bootstrap replicates.

ML tree obtained two well supported major lineages with high supported values (Lineage I and II; Figure 1). Both clades I and II contained species from genus *Oreochromis*, implying taxonomic problems occurred in this genus. In addition, *S. galilaeus* was found to be grouped with *Oreochromis aureus*, rather than other *Sarotherodon* species (Figure 1). The result suggested genus *Sarotherodon* was not monophyletic taxa. Thus, more work on morphological features and phylogenetic inferences should stand together to resolve the complex relationships among Oreochromini species.

CONTACT Suqin Wang  suqinwanghn@outlook.com  Hunan University of Arts and Science, Changde, 415000, China

© 2021 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.

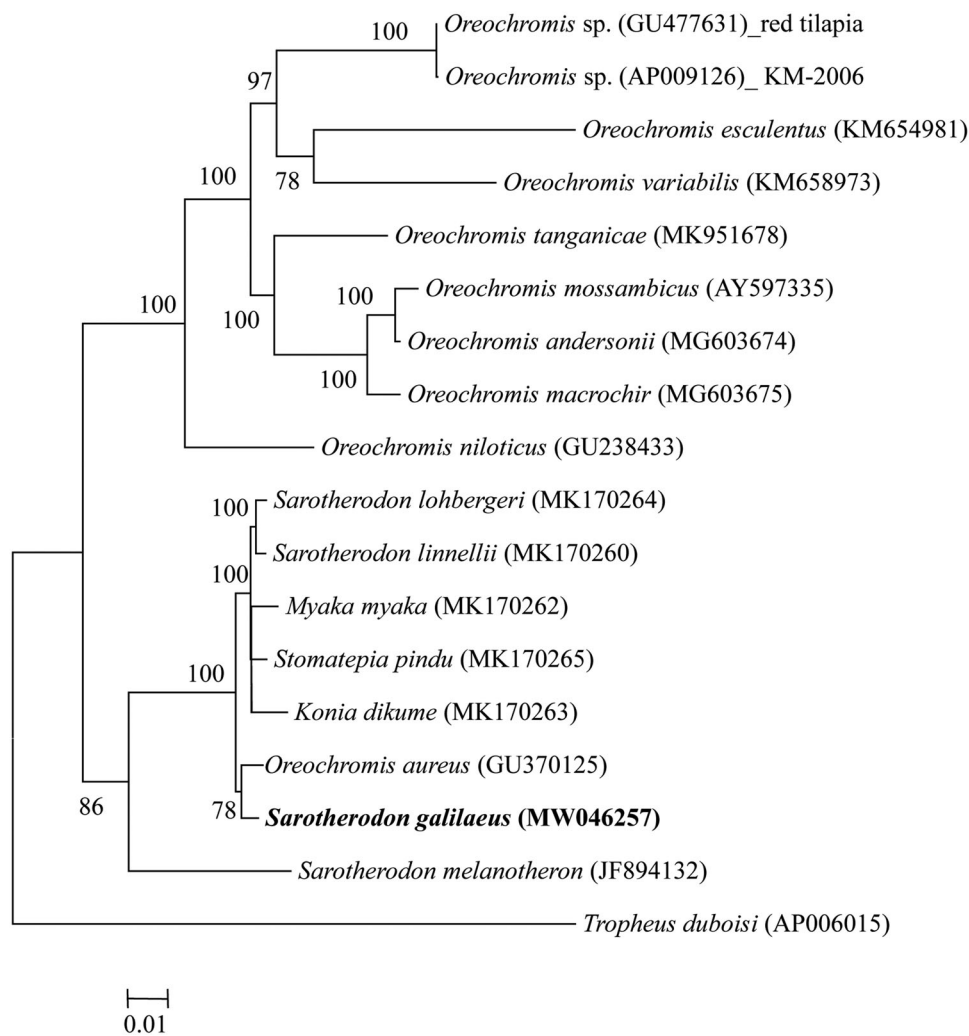


Figure 1. Phylogenetic trees based on maximum likelihood technique showing the phylogenetic relationships among 17 Oreochromini mitogenomes based on 13 protein-coding genes. Values on branches indicate bootstrap values from maximum likelihood analyses.

Ethical approval

Experiments were performed in accordance with the recommendations of the Ethics Committee of Hunan University of Arts and Science. These policies were enacted according to the Chinese Association for the Laboratory Animal Sciences and the Institutional Animal Care and Use Committee (IACUC) protocols.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Department of Education General program [19C1257] and Changde Research Center for Agricultural Biomacromolecule Open Project fund [2020AB08].

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. MW046257. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA695435, SRR13570802 and SAMN17614357, respectively.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzscher G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol.* 69(2):313–319.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience.* 1(1):18.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 22(21):2688–2690.
- Yan B, Wang Z. 2010. Growth, salinity tolerance and microsatellite analysis of the F2 reciprocal hybrids of *Oreochromis niloticus* × *Sarotherodon galilaeus* at different salinities. *Aquac Res.* 41(9):336–344.
- Zeilstra I, Ros A, Oliveira R. 2003. Mate choice in the Galilee St. Peter's fish, *Sarotherodon galilaeus*. *Behaviour.* 140(8-9):1173–1188.