Characterization of the complete mitochondrial genome and phylogenetic analysis of *Opsarius pulchellus* (Cypriniformes, Danionidae, Chedrinae)

Xiao Jiang Chen*, Lin Song*, Wen Zhao Liu* and Peng Gao*

*College of Fishery Science and Technology, Jiangsu Agri-Animal Husbandry Vocational College, Taizhou City, P. R. China; bCollege of Fisheries and Life Science, Dalian Ocean University, Dalian City, P. R. China

**ABSTRACT**

This study aimed to sequence and annotate the complete mitochondrial DNA genome sequence of *Opsarius pulchellus*. Complete mitochondrial genome is a circular molecule of 16,552 bp in length, including 13 protein-coding genes (PCGs), 22 tRNA genes, two rRNA genes, and a displacement loop region. Proportion of nucleotides in mitochondrial genome is 26.8% T, 26.9% C, 27.6% A, and 18.8% G, with an AT bias of 54.4%. Maximum-likelihood phylogenetic tree was reconstructed using the connected protein sequences of 13 PCGs of *O. pulchellus* and other 18 fishes. Result of phylogenetic analysis supports a close relationship between *O. pulchellus* and *Barilius bendelisis* (Hamilton, 1807). Fundamental genetic data of *O. pulchellus* would be beneficial for further studies on genetic diversity and phylogeny.

*Opsarius pulchellus* (Smith, 1931) belongs to the subfamily Chedrinae (family: Danionidae). It is widely distributed in Cambodia; the Yunnan Province of China (Lancang River system and Yuanjiang River system); Lao People’s Democratic Republic; Thailand and Viet Nam. Habitat degradation caused by deforestation and agricultural pollution is the main potential threat to *O. pulchellus* (Vidthayanon 2012). The synonyms of *O. pulchellus* such as *Barilius pulchellus*, *Barilius Pellegrini* and *Danoiops macropterus* are not valid. At present, the research on *O. pulchellus* is limited to morphological characteristics (Kottelat 1998; Beamish et al. 2006; Tejavej 2012). This study is the first to report the complete mitochondrial genome of *O. pulchellus* using the next-generation sequencing method, which will be beneficial for future studies on population genetics and evolution of the subfamily Chedrinae.

Samples were collected from Xishuangbanna, Yunnan province, China (21°98′75.2″N, 100°81′43.28″E). *Opsarius pulchellus* was identified by consulting FishBase Database, Fish Database in Inland Waters of China, and Native Ornamental Fish of China Illustrated Book (Chu and Chen 1989; Zheng et al. 1991; Zheng et al. 2016). Samples were anesthetized with MS-222 before the experiment. Some specimens were fixed with 10% formaldehyde and deposited at Aquatic Science and Technology Institution Herbarium (voucher ASTIH-21b0616d06, https://www.jshwcv.edu.cn/, XJ CHEN, cq_cxj@126.com). Genomic DNA was extracted from muscle tissues by Tguide cell/tissue genomic DNA Extraction Kit (Tiangen, Beijing, China) and then sequenced on the Illumina HiSeq platform. MetaSPAdes (Nurk et al. 2017) was used to assemble the sequence and MitoMaker (Bernt et al. 2013) for gene annotation. The animal experiments were approved by the Ethical Committee for Animal Experiments of Jiangsu Agri-animal Husbandry Vocational College and conducted following the Guidelines for Experimental Animals of the Ministry of Science and Technology (Beijing, China).

The complete mitochondrial genome of *O. pulchellus* is 16,552 bp in length, including 13 proteins, 22 tRNAs, two rRNAs, and one displacement loop region (D-loop). The D-loop is located between *tRNA^{Phe}* and *tRNA^{Pro}* with 932 bp in size. The proportion of nucleotides in mitochondrial genome is T (26.8%), C (26.9%), A (27.6%), G (18.8%), with an AT bias of 54.4%. Among the 37 genes, 27 genes are encoded on the H-strand while *tRNA^{Phe}* and *tRNA^{Pro}* with 932 bp in size. The overlap length of 11 genes is 42 bp, the interval length of 15 genes is 117 bp, ranging from 1 to 35 bp. The overlap length of 11 genes is 42 bp, with overlapping base numbers between 1 and 19 bp (Table 1).

Among 13 protein-coding genes (PCGs), the longest gene is *ND5* (1812 bp) and the shortest gene is *ATP8* (165 bp). Two genes (*CO I* and *ND3*) take GTG and ATA as the starting codon, respectively. The other 11 genes take ATG as the starting codon. Seven genes (*ND1, CO I, ATP8, ATP6, COIII, ND4L, and ND6*) take TAA as complete termination codon while three genes (*ND2, ND3*, and *ND5*) end with TAG, respectively. Three genes (*CO II, ND4*, and *Cyt b*) end with an...
incomplete termination codon (T). The length of tRNA is ranging from 66 to 76 bp. The length of 12S rRNA and 16S rRNA is 953 bp and 1661 bp, respectively.

The phylogenetic relationships of *O. pulchellus* (MZ853157.1) were reconstructed using the maximum-likelihood (ML) method with 1000 bootstrap replicates based on concatenated amino acids and nucleotides of 13 PCGs of 19 fishes, *Amblypharyngodon mola* and *Rasbora tornieri* were used as outgroups. Evolutionary analyses were conducted in MEGA X (Kumar et al. 2018), and the General Reversible Mitochondrial (mtREV) + Gamma distributed with invariant sites (G+I) were used as the best model (Adachi and Hasegawa 1996). The phylogenetic tree was analyzed based on amino acids and nucleotides sequences of 13 PCGs, both analyses were run with very similar topology.

The result of the phylogenetic analysis showed that *O. pulchellus* was close to *B. bendelisis* (Figure 1). Presently, the studies recording *O. pulchellus* were limited, and we believed that the fundamental genetic data could be beneficial for further studies on genetic diversity and phylogeny.

### Authors' contributions

Xiao Jiang Chen made substantial contributions to the conception or design of the work, and drafting the paper, and final approval of the version to be published; and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; Lin Song made substantial contributions to the design of the work, and drafting the work and analysis and final approval of the version to be published; and agreement to be accountable for all aspects of the work; Wen Zhao Liu and Peng Gao were involved in the acquisition, analysis and interpretation of the data; the drafting of the paper, and the final approval of the version to be published, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

### Acknowledgments

The authors thank Dr. Li Yanping of Neijiang Normal University, Dr. Wei Li and Yi Wen Gu for their help, and acknowledge the Shanghai Genesky Biotechnologies Inc., for their technical support.

### Disclosure statement

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

### Table 1. Characteristics of the mitochondrial genome of *Opsarius pulchellus*.

| Gene     | Position | Nucleotide size (bp) | Amino acid | Space (+) overlap (−) | Codon |
|----------|----------|----------------------|------------|-----------------------|-------|
| tRNAPhe  | 16,547   | 63                   | 68         | 0                     | L     |
| 12SrRNA  | 64       | 1016                 | 953        | 0                     | H     |
| tRNAVal  | 1018     | 1089                 | 72         | 1                     | H     |
| 16SrRNA  | 1108     | 2768                 | 1661       | 18                    | H     |
| tRNA Arg | 2768     | 2842                 | 75         | −1                    | H     |
| ND1      | 2847     | 3818                 | 972        | 323                   | ATG   |
| tRNA Leu | 3819     | 3890                 | 72         | 0                     | TAA   |
| tRNA Glu | 3889     | 3959                 | 71         | −2                    | H     |
| tRNA Met | 3962     | 4030                 | 69         | 2                     | H     |
| ND2      | 4031     | 5077                 | 1047       | 348                   | ATG   |
| tRNA Thr | 5076     | 5145                 | 70         | −2                    | TAG   |
| tRNA Pro | 5147     | 5214                 | 68         | 1                     | H     |
| tRNA Leu | 5215     | 5287                 | 73         | 0                     | L     |
| tRNA Cys | 5323     | 5388                 | 66         | 35                    | L     |
| tRNA Tyr | 5388     | 5457                 | 70         | −1                    | L     |
| CO I     | 5459     | 7009                 | 1551       | 516                   | GTG   |
| tRNA Arg | 7010     | 7080                 | 71         | 0                     | TAA   |
| tRNA Thr | 7084     | 7153                 | 70         | 3                     | ATG   |
| CO II    | 7168     | 7858                 | 691        | 230                   | 14    |
| tRNA Arg | 7859     | 7934                 | 76         | 0                     | ATG   |
| ATPase8  | 7936     | 8100                 | 165        | 54                    | T     |
| ATPase6  | 8082     | 8777                 | 696        | 231                   | −19   |
| CO II    | 8777     | 9562                 | 786        | 261                   | −1    |
| tRNA Arg | 9562     | 9632                 | 71         | −1                    | ATG   |
| ND3      | 9633     | 9983                 | 351        | 116                   | T     |
| tRNA Arg | 9982     | 10,051               | 70         | −2                    | L     |
| ND4      | 10,052   | 10,348               | 297        | 98                    | ATG   |
| ND4      | 10,342   | 11,719               | 1378       | 459                   | T     |
| tRNA Thr | 11,724   | 11,793               | 70         | 4                     | H     |
| tRNA Arg | 11,794   | 11,862               | 69         | 0                     | H     |
| tRNA Arg | 11,864   | 11,936               | 73         | 1                     | H     |
| NDS      | 11,938   | 13,749               | 1812       | 603                   | 1     |
| NDS      | 13,746   | 14,240               | 495        | 164                   | 4     |
| tRNA Glu | 14,268   | 14,335               | 68         | 27                    | ATG   |
| Cyt b    | 14,340   | 15,480               | 1141       | 380                   | T     |
| tRNA Thr | 15,481   | 15,552               | 72         | 0                     | T     |
| tRNA Thr | 15,551   | 15,620               | 70         | −2                    | L     |
| D-loop   | 15,621   | 16,552               | 932        | 0                     | –     |
Figure 1. Phylogenetic tree reconstructed using the connected protein sequences of 13 PCGs of *O. pulchellus* together with other 18 fishes based on maximum-likelihood (ML), tree topology was evaluated by 1000 bootstrap replicates. *Amblypharyngodon mola* and *Rasbora tornieri* were used as outgroups. The species names are followed by their GenBank accession numbers.

**Funding**

This work was supported by The 311 Talents Project in Taizhou City under Grant [2017III - 804]; The Research Project of Jiangsu Agri-Animal Husbandry Vocational College under Grants [NSF2021ZR14 and NSFHP202102].

**ORCID**

Xiao Jiang Chen http://orcid.org/0000-0002-0934-0330

**Data availability statement**

The genome sequence data that support the findings of this study are openly available in GenBank at https://www.ncbi.nlm.nih.gov/, accession number MZ853157.1.

The associated BioProject, SRA, and Bio-Sample numbers are PRJNA769989, SRR16282297, and SAMN22187389, respectively.

**References**

Adachi J, Hasegawa M. 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. J Mol Evol. 42(4):459–468.

Beamish F, Sa-Ardrit P, Tongnunui S. 2006. Habitat characteristics of the Cyprinidae in small rivers in Central Thailand. Environ Biol Fish. 76(2–4):237–253.

Bernt M, Donath A, Juhling 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.

Chu XL, Chen YR. 1989. [Ichthyology of Yunnan]. Beijing (China): Science Press; p. 19–23. Chinese.

Kottelat M. 1998. Fishes of the Nam Theun and Xe Bangfai basins, Laos, with diagnoses of twenty-two new species (Teleostei: Cyprinidae, Balitoridae, Cobitidae, Coiidae and Odontobutidae). Ichthyol Explor Freshwat. 9(1):1–128.

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.

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. MetaSPAdes: a new versatile metagenomic assembler. Genome Res. 27(5):824–834.

Tejavej A. 2012. *Barilius signicaudus*, a new species of cyprinid fish from Maeklong Basin, western Thailand (Cypriniformes: Cyprinidae). Zoo taxa. 3586(1):138–147.

Vidthayanon, C. 2012. *Opsarius pulchellus*. The IUCN Red List of Threatened Species; https://www.iucnredlist.org/species/181090/1697253. 2012: e.T181090A1697253

Zheng CY, Pan JH, Zhong L, Zheng CY, Wu HL, Liu JH. 1991. [The freshwater fishes of Guangdong Province]. Guangzhou (China): Guangdong Science and Technology Press; p. 68–72. Chinese.

Zheng SM, Wu Q, He LJ, Zhang ZG. 2016. [Native ornamental fish of China illustrated book]. Beijing (China): Science Press; p. 23. Chinese.