

The complete mitochondrial DNA sequence of Kashgarian loach (*Triplophysa yarkandensis*) from Bosten Lake

Xia Ning^a, Yan-Zhen Zhang^b, Zhi-Hai Sui^b, Xian-Qing Quan^b, Hai-Guang Zhang^b, Ling-Xiao Liu^c, Qing-Dian Han^b and Yun-Guo Liu^b

^aJinan Zhangqiu District Center for Disease Control and Prevention, Jinan, China; ^bCollege of Life Sciences, Linyi University, Linyi, China; ^cLinyi Academy of Agricultural Sciences, Linyi, China

ABSTRACT

Triplophysa yarkandensis is a specific cobitidae species that is endemic to Xinjiang Tarim River basin, China. The complete mitochondrial genome sequence of *T. yarkandensis* from Bosten Lake was determined in this study (Accession number MN821008). The mitogenome (16,552 bp) consists of 22 tRNA genes, 2 ribosomal RNA genes, 13 protein-coding genes, and 1 control region (D-loop region). The complete mitochondrial genome sequence of the *T. yarkandensis* provides an important data set for further study in genetic mechanism and classification.

ARTICLE HISTORY

Received 3 January 2020
Accepted 7 January 2020

KEYWORDS

Kashgarian loach;
Triplophysa yarkandensis;
mitochondrial genome

As a kind of small economic fish, which belongs to *Triplophysa*, Nemacheilidae, Cypriniformes, Actinopterygii, Osteichthyes, *Triplophysa yarkandensis*, also called Kashgarian loach, is a specific omnivorous Cobitidae species that is only distributed in Xinjiang Tarim River basin, China (Chen et al. 2016). Most of them are just finger thick with bearded and scaleless, which is a relatively large individual species in the genus of *Triplophysa*. Due to the gradual depletion of water resources in Tarim River and the backward cultivation technology, *T. yarkandensis* is facing the danger of extinction (Gong et al. 2016). Here, the complete mitochondrial genome of *T. yarkandensis* was sequenced and characterized in detail. Kashgarian loach sample was collected from Bosten Lake (41°73'N, 86°95'E), a relatively closed water area, Xinjiang, China in July 2019. The specimen of Kashgarian loach, named as TriyarBost-01 was stored in the College of Life Sciences, Linyi University, Linyi, China. Total genomic DNA was extracted from *T. yarkandensis* muscle according to Liu et al. (2012, 2016). Then, the complete mitochondrial genome was sequenced using a shotgun approach and assembly. Subsequently, DNA sequence was analyzed using MEGA 7 (Kumar et al. 2016) and protein-coding genes were analyzed by ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) using the invertebrate mitochondrial code. The tRNA genes were identified using ARWEN (Laslett and Canback 2008) and tRNA-scan SE (Lowe and Eddy 1997).

The complete *T. yarkandensis* mitochondrial genome (Accession number MH MN821008) was 16,552 nucleotides long, of which 15,557 nucleotides are coding DNA and 965 nucleotides are non-coding DNA. The mitochondrial genome of *T. yarkandensis* was found to contain 22 transfer RNA genes, 2 ribosomal RNA genes, 13 protein-coding genes, and 1 control region. The arrangement and composition of the mitochondrial genome are also comparable to other vertebrates (Liu and You 2015; Xu et al. 2015; Shan and Liu 2016; Li, Liu, Sui et al. 2019; Li, Liu, Zhang 2019). A phylogenetic tree was constructed based on the comparison of the complete mitochondrial genome sequences with other Cobitidae species using the neighbour-joining method (Figure 1). There were 13 intergenic spacers (total 73 bp) and 7 overlapping regions (total 27 bp) among the genes. The overall base composition was 27.21% A, 28.21% T, 19.03% C, and 25.55% G, with an AT content of 55.42%. To investigate the nucleotide bias, skew for a given strand was calculated as $(A-T)/(A+T)$ or $(G-C)/(G+C)$ (Perna and Kocher 1995). The AT and GC skews for the *T. yarkandensis* mitochondrial genome were -0.018 and 0.146 , respectively; this finding indicated that the strand that encoded genes contained more T and G than A and C and this skew was evidence of codon usage bias. All 13 protein-coding genes identified in the *T. yarkandensis* mitochondrial genome started with the common initiation codon ATG, except for *COX1* and *ND6*, which began with

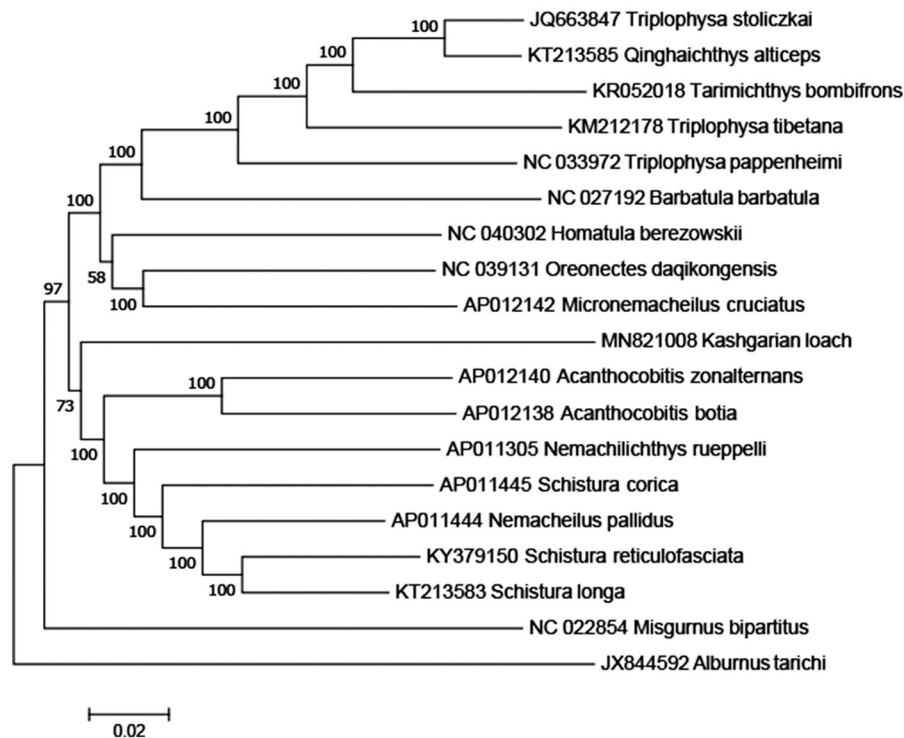


Figure 1. A phylogenetic tree constructed based on the comparison of complete mitochondrial genome sequences of the Nemacheilidae, *T. yarkandensis* (Kashgarian loach) and other 16 species of Nemacheilidae family. They are *Triplophysa stoliczkai*, *Qinghaichthys alticeps*, and *Triplophysa tibetana* (Tibetan stone loach), *Tarimichthys bombifrons*, *Barbatula barbatula*, *Oreonectes daqikongensis* and *Nemachilichthys rueppelli* (stone loach), *Triplophysa pappenheimi* (Yellow River stone loach), *Homatula berezowskii*, *Micronemacheilus cruciatus*, *Acanthocobitis zonalternans*, *Schistura longa* and *Schistura corica* (ray-finned loach), *Acanthocobitis botia* (Thailand mottled loach), *Nemacheilus pallidus* (undertone loach), *Schistura reticulofasciata* (stream loach). *Misgurnus bipartitus* and *Alburnus tarichi* are used as outgroup. Genbank accession numbers for all sequences are listed in the figure. The numbers at the nodes are bootstrap percent probability values based on 1000 replications.

GTG and TTA, respectively. For termination codons, *ND1*, *ND5*, and *COX1* terminate with TAG, whereas *ATP8*, *ATP 6*, and *ND4L* terminate with TAA, *ND6* terminate with CAT, respectively. In addition, the six remaining genes terminate with incomplete stop codon T-, that T- is the 5' terminal of the adjacent gene, which presumptively formed a complete stop codon by post-transcriptional polyadenylation (Anderson et al. 1981). All the mitogenome genes were encoded on the H strand except for *ND6* and eight tRNA genes (*tRNA-Gln*, *Ala*, *Asn*, *Cys*, *Tyr*, *Ser*, *Glu*, *Pro*). The 12S rRNA (951 bp) gene and 16S rRNA (1678 bp) gene were located between the *tRNA-Phe* and *tRNA-Leu* genes and separated by the *tRNA-Val* gene. The lengths of 22 tRNA coding genes range in size from 69 to 75 bp. In addition, the control region was located between *tRNA-Pro* and *tRNA-Phe* with a length of 892 bp. Thus, the data would facilitate further investigations of phylogenetic relationships within Nemacheilidae.

Disclosure statement

The authors report no conflicts of interest, as well as only the authors are responsible for the content and writing of the article.

Funding

This work was supported by the National Natural Science Foundation of China [31560719], the Key Research and Development Program in

Shandong Province [2019YYSPO26], and the High-level Talent Introduction Project of Linyi University [LYDX2018BS032].

References

- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature*. 290(5806): 457–464.
- Chen S, Ya N, Xie C, Wang S, Ren D. 2016. Complete mitochondrial genome of the *Triplophysa (Hedinichthys) yarkandensis* (day). *Mitochondr DNA B*. 1 (1):235–236.
- Gong X, Xu H, Xiang W, Fan Z, Zheng L, Wang C. 2016. Complete mitochondrial genome of Kashgarian loach, *Triplophysa Yarkandensis* (day, 1877) in the Tarim river. *DNA Sequence*. 27(5):2.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 33(7): 1870–1874.
- Laslett D, Canback B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*. 24(2): 172–175.
- Li YY, Liu LX, Sui ZH, Ma C, Liu YG. 2019. The complete mitochondrial DNA sequence of Mongolian grayling *Thymallus brevisrostris*. *Mitochondr DNA B*. 4(1):1204–1205.
- Li YY, Liu LX, Zhang RH, Ma C, Liu YG. 2019. The complete mitochondrial DNA sequence of ruffe *Acerina cernua*. *Mitochondr DNA B*. 4(1): 1590–1591.
- Liu YG, Guo YG, Hao J, Liu LX. 2012. Genetic diversity of swimming crab (*Portunus trituberculatus*) populations from Shandong peninsula as assessed by microsatellite markers. *Biochem Syst Ecol*. 41: 91–97.

- Liu YG, Li YY, Meng W, Liu LX. 2016. The complete mitochondrial DNA sequence of Xinjiang arctic grayling *Thymallus arcticus grubei*. *Mitochondr DNA B*. 1(1):724–725.
- Liu T, You P. 2015. The complete mitochondrial genome of *Triplophysa* sp. (Teleostei: Cypriniformes: Balitoridae). *Mitochondr DNA*. 27(6):1–2.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl Acids Res*. 25(5): 955–964.
- Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at four-fold degenerate sites of animal mitochondrial genomes. *J Mol Evol*. 41(3):353–358.
- Shan WJ, Liu YG. 2016. The complete mitochondrial DNA sequence of the cape hare *Lepus capensis pamirensis*. *Mitochondr DNA*. 27(6): 4572–4573.
- Xu D, He CQ, Li QH, He J, Ma HM. 2015. The complete mitochondrial genome of the Daweizi pig. *Mitochondr DNA*. 26(4):640–641.