

# The complete mitochondrial genome of a new species of the genus *Schizothorax* from Sichuan, China (Cypriniformes: Cyprinidae)

Ying Wang<sup>a</sup>, Pan Shang<sup>b</sup>, Yuanxing Dai<sup>c</sup>, Dayong Xu<sup>a</sup>, Yanzhen Dong<sup>a</sup> and Zhiqiu Huang<sup>a</sup>

<sup>a</sup>Key Laboratory of Application of Ecology and Environmental Protection in Plateau Wetland of Sichuan, Xichang University, Xichang, PR China; <sup>b</sup>Key Laboratory of Freshwater Aquatic Genetic Resources, Ministry of Agriculture, Shanghai Ocean University, Shanghai, PR China; <sup>c</sup>Gulin County Bureau of Agriculture and Rural Affairs, Luzhou, PR China

## ABSTRACT

*Schizothorax gulinensis* sp. nov. is a new species of the genus *Schizothorax* from Sichuan, China (Cypriniformes: Cyprinidae). In this study, we have first reported the complete mitochondrial genome of *S. gulinensis* with Illumina sequencing. There were 16,587 nucleotide pairs in the mitochondrial genome (mitogenome) of *S. gulinensis*, including 13 protein-coding genes (PCGs), two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs), as well as one non-coding control region (CR). The proportion of nucleotides in mitochondrial genome was 29.67% (A), 25.45% (T), 17.84% (G), 27.05% (C), and A + T content was 55.12%. All PCGs have the same start codon of the standard ATG, excepting for that of NADH dehydrogenase subunit 1 (*nad1*) which was the ATC, NADH dehydrogenase subunit 5 (*nad5*) which was the ATT and cytochrome c oxidase 1 (*cox1*) which was the ATC. Phylogenetic analysis results supported that *S. gulinensis* was closely related to *Schizothorax grahami*. The complete mitochondrial sequence of *S. gulinensis* will contribute to mitochondrial genome database and provide useful resources for population genetics and evolution analyses.

## ARTICLE HISTORY

Received 21 May 2023  
Accepted 21 November 2023

## KEYWORDS

Cyprinidae; *Schizothorax gulinensis*; mitochondrial genome; phylogenetic

## 1. Introduction

Schizothoracinae is an important subfamily of endemic fish on the Qinghai-Tibetan Plateau (QTP), which includes 12 genera and more than 70 species/subspecies (Qi et al. 2012; Xu et al. 2016). However, in 1964, a new *Schizothoracinae* fish was discovered in Sichuan province (Cao and Wu 1964). It has been identified as a new species of the *Schizothorax* genus in the Cyprinidae family, and has been named *Schizothorax gulinensis* sp. nov. Ding et al. 2022 (Cypriniformes: Cyprinidae) (Ding et al. 2022) (Figure 1). However, the comprehensive understanding of *S. gulinensis* ecology has been long hampered by the homoplasy of morphological characteristics due to the massive subspecies. Besides, the inadequate genetic information restrains the further exploration of evolutionary phylogenetic relationship of this species within *Schizothoracinae* family. In this sense, the complete mitochondrial genome of *S. gulinensis* was described in the present study through Illumina sequencing. Our results are expected to contribute to future research of the evolution and phylogenetic status of *S. gulinensis*.

## 2. Materials and methods

### 2.1. Sample collection and preservation



In this study, specimens of *S. gulinensis* were collected in live from a tributary (Baisha River) of the Chishui River, Gulin County,


Sichuan Province, China (27°48′51.03″ N, 105°45′33.05″ E) in August 2022. The specimens were morphologically identified (Figure 1) and confirmed refer to previous studies (Ding et al. 2022). The samples obtained were stored in the aquatic museum of Xichang University (accession number: DKXY-0101279, Dayong Xu, 512573607@qq.com).

### 2.2. DNA sequencing and genome assembly

Total genomic DNA was extracted from the muscle tissue using a DNeasy tissue kit (Qiagen, Beijing, China) following the manufacturer's protocols. After DNA isolation, 1 µg of purified DNA was fragmented into approximately 500 bp by Covaris E220 system (Covaris, Woburn, MA). Afterwards, the fragmented DNA was used to construct short-insert libraries according to the manufacturer's instructions (TruSeq™ Nano DNA Sample Prep Kit, Illumina), and then sequenced on an Illumina NovaSeq 6000 platform (BIOZERON Co., Ltd, Shanghai, China) with 150 bp paired-end reads length.

Prior to the assembly, the filtering step was performed through Trimmomatic 0.39 (Bolger et al. 2014), by which certain types of the raw reads were removed such as reads with adaptors, lower quality score than 20 ( $Q < 20$ ), higher uncalled bases ('N' characters) percentage than 10% and any duplicated sequences. The read coverage depth map was shown in Figure S1. In order to reconstruct the mitochondrial genome, a *de*

**CONTACT** Zhiqiu Huang  [xcchuangzhiqiu@hotmail.com](mailto:xcchuangzhiqiu@hotmail.com)  Key Laboratory of Application of Ecology and Environmental Protection in Plateau Wetland of Sichuan, Xichang University, Xichang, PR China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2023.2288437>.

© 2023 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-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

*novo* combination and reference-guided assemblies were applied. There were three following steps to assemble the mitogenome. First, by using MitoZ version 2.3 (<https://github.com/linzhi2013/MitoZ>; Meng et al. 2019) the filtered reads were assembled into contigs, from which the potential mitochondrial contigs were extracted by the alignment with the mitogenome database from NCBI. Second, after the alignment of potential mitochondrial contigs were aligned with the reference mitogenomes using BLAST version 2.8.1+ (<https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>; Camacho et al. 2009), the aligned contigs ( $\geq 80\%$  similarity and query coverage) were manually ordered and connected according to the reference mitogenomes. Third, whether these selected contigs were circular was confirmed by MUMmer version 3.23



**Figure 1.** The morphological characteristics of *S. gulinensis* (photographed by Yuanxing Dai).

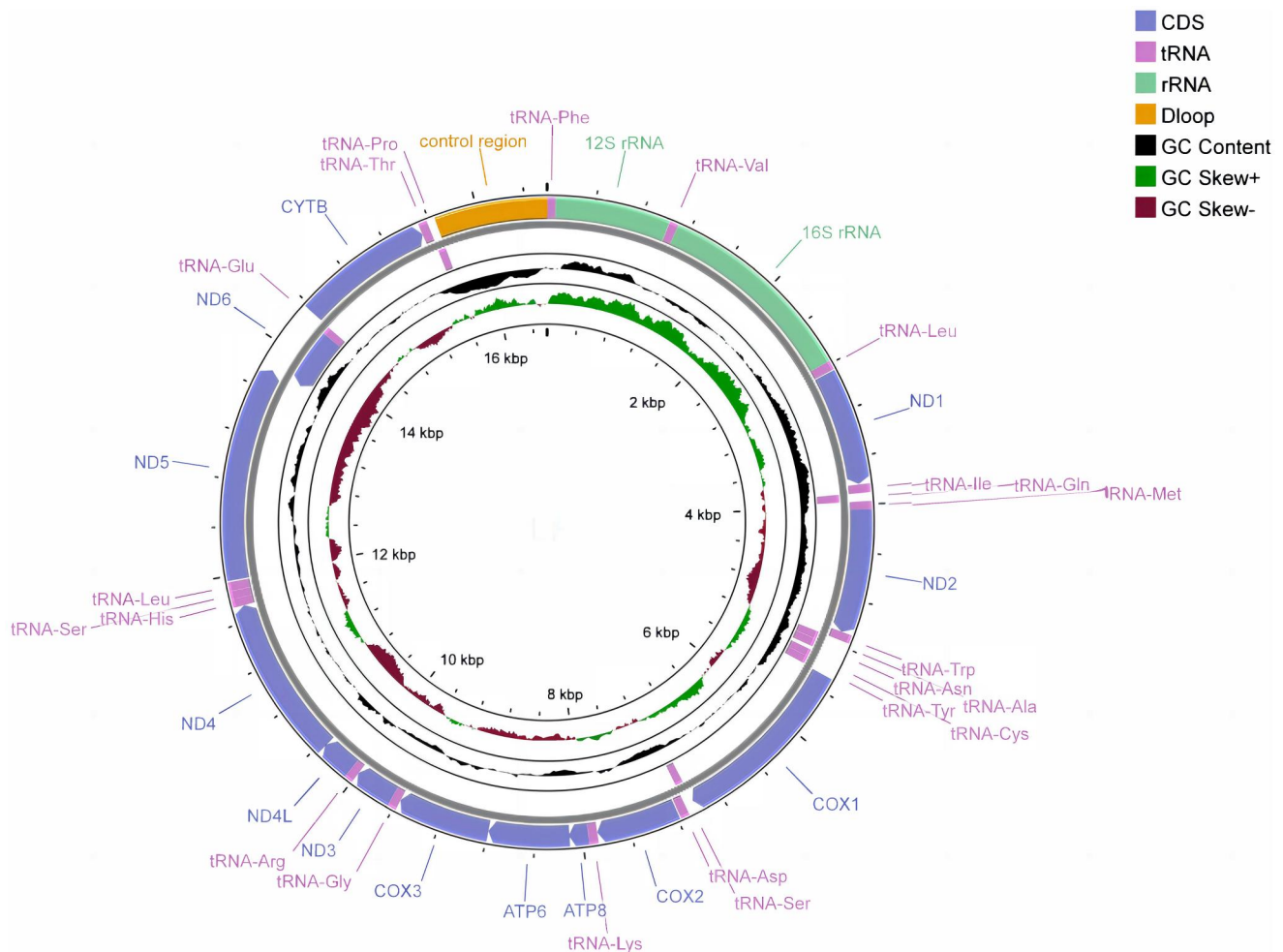
(<https://github.com/chienchi/MUMmer>; Kurtz et al. 2004). Eventually, the circle of the *S. gulinensis* mitogenome was obtained from the assembly steps above.

### 2.3. Annotation and analysis

The mitochondrion genes were annotated using the online MITOS tool (Bernt et al. 2013), using default parameters to predict protein coding genes, transfer RNA (tRNA) genes, and ribosome RNA (rRNA) genes. The position of each coding gene was determined using BLAST searches against ref mt genes. Manual corrections of genes for start/stop codons were performed in SnapGene Viewer by referencing the ref mt genome. The circular mt genome map of *S. gulinensis* was drawn using the CPGview tool (Figure S2) (Grant and Stothard 2008). Phylogenetic tree constructed using the IQ-TREE (Minh et al. 2020) and MrBayes (Ronquist et al. 2012) method based on the amino acid sequences of 13 proteins translated by the mitochondrial genome. And based on the BIC value, we selected mtVer + F + R3 as the best-fit model.

### 3. Results

The complete mitochondrial genome of *S. gulinensis* was 16,587 bp in length (GenBank accession number OQ730545)



**Figure 2.** Circular map of the mitochondrial genome of *S. gulinensis*.

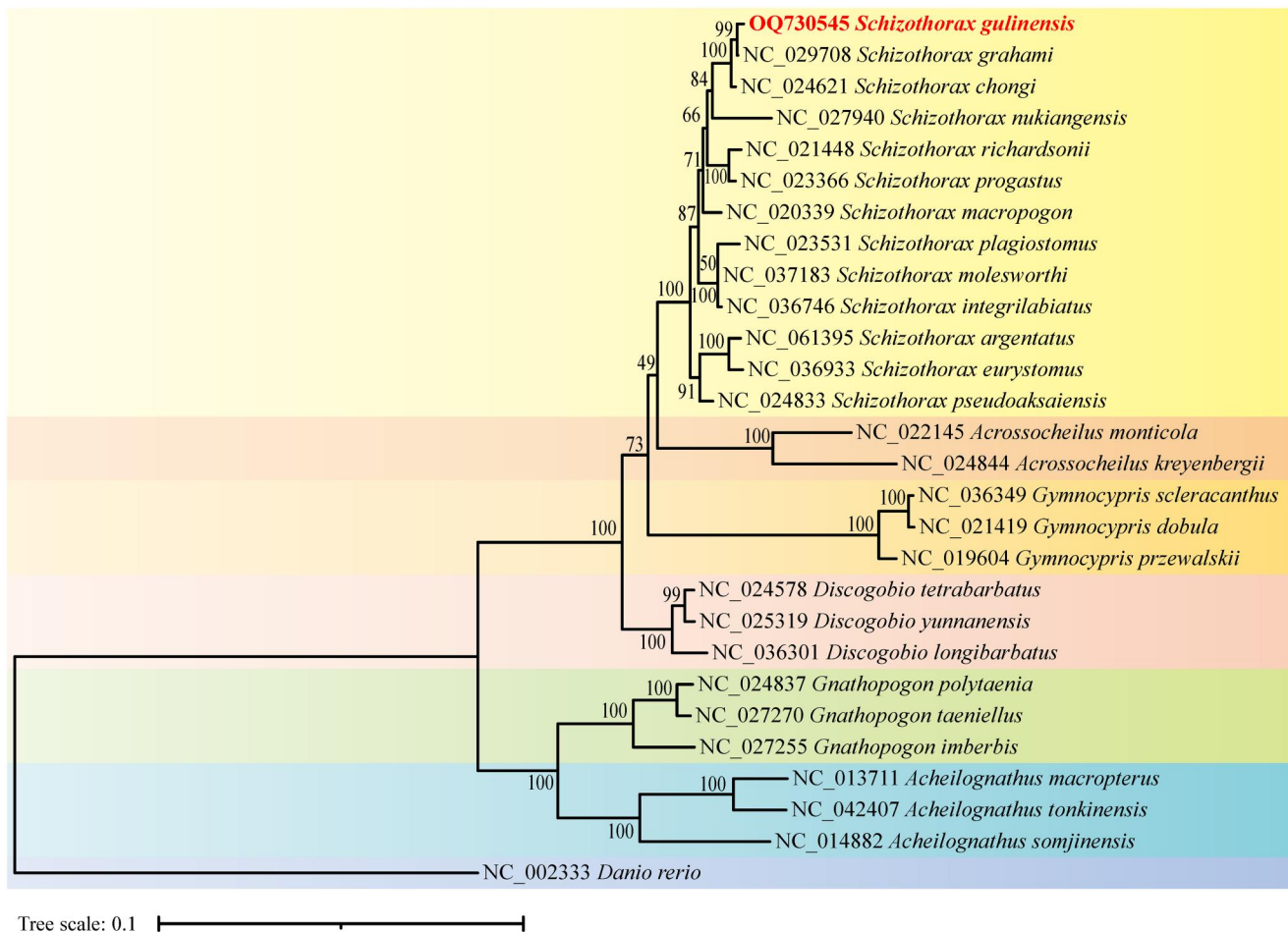
and the proportion of nucleotides in mitochondrial genome was A (29.67%), in which the proportion of each nucleotide was 29.67% (A), 25.45% (T), 17.84% (G), 27.05% (C) with the content of A + T as 55.12%. It consisted of 13 protein-coding genes (PCGs), 22 tRNA genes, and two rRNA genes (Figure 2, Figure S2). Twelve out of all 13 PCGs encoded on the heavy strand, while *nad6* was the only PCG located in the light-strand. Similar to the PCGs, 14 of the 22 tRNA genes were preferentially located in the heavy strand as well. The PCGs used three different start codons: ATT, ATC, or ATG and two stop codons: TAA or TAG, except that *cox2*, *nad4*, and *cob* ended with T. The conserved 13 PCGs varied in length from 165 bp (*atp8*) to 1815 bp (*nad5*). The length of 22 tRNAs ranges from 67 bp (tRNA-Cys) to 74 bp (tRNA-Leu). The length of 12S rRNA and 16S rRNA was 954 and 1677 bp, respectively (Table S1). The gene arrangement of *S. gulinensis* mitochondrial genome was identical to that of other *Schizothorax* genus.

Given the mitogenome of zebrafish as the outgroup data (Broughton et al. 2001), the maximum-likelihood phylogenetic tree was constructed to infer the phylogeny relationship using the protein sequences of the 13 PCGs of 28 species. In Figure 3, *S. gulinensis* was preferentially clustered into a clade with *S. grahami* (Zheng et al. 2016) then with the clade of other species in the *Schizothorax* genus, indicating that *S. gulinensis*

is most closely related to *S. grahami*. In this sense, the analysis of molecular phylogenetics in the present study supported the attribution of *S. gulinensis* to the *Schizothorax* genus (Ding et al. 2022). Considering that the nuclear genomes of the *Schizothorax* genus where *S. gulinensis* belongs have not been sequenced, our results of the complete mitochondrial genome sequence for *gulinensis* are expected to be contributive for the future phylogeographical studies of the *Schizothorax* genus.

#### 4. Discussion and conclusion

The uplift of the QTP plays an important role in the formation of the intricate water ecological environment, which also has a profound impact on the production of rich and diverse *Schizothorax* fish. The distribution of *Schizothorax* fish has a complex water system pattern, which may cause the migration of the same fish species between different water and hybridization between species within the same water (Li et al. 2015). This phenomenon leads to a significant similarity in the morphology of *Schizothorax* fish, which increased the difficulty in studying the morphology of this species. In addition, with the increasing demand for protection and development of fish resources, the requirements for accurate identification of *Schizothorax* species and optimal selection of species resources



**Figure 3.** The phylogenetic tree was constructed using the amino acid sequences of 13 proteins translated by the mitochondrial genome of 28 species based on the maximum likelihood with 1000 bootstrap replicates using IQ-TREE. The bootstrap support values were shown by the numbers on the branches, and the red font represented the target species in this study. GenBank accession numbers were provided near branch tips for all sequences used. References used for comparative analysis were provided in Table S2.

are also increasing. Conventional morphological data and detection methods are difficult to satisfy the identification and evaluation demand, therefore, we need to use molecular biology methods to provide more scientific basis for identification of this species. In recent years, mitochondrial DNA has gradually been applied to research on species evolution, population genetic diversity, phylogenetic relationships, and species classification (Mabuchi et al. 2006) due to its genetic characteristics, such as semi autonomy, maternal inheritance, polymorphism, large copy count, extremely high mutation rate, and unorganized specificity (Meland et al. 1991).

In this study, we reported the whole mitochondrial genome information of *S. gulinensis* and analyzed the genetic structure and identified its position in phylogenetic evolution with other subfamilies. The mitochondrial genome of *S. gulinensis* was 16,587 bp in length and expresses high AT bias. Phylogenetic analysis revealed the phylogenetic position of *S. gulinensis* and indicated that *S. gulinensis* belongs to the *schizothorax* genus (Ding et al. 2022). The complete *S. gulinensis* mitochondrial genome reported here adds to the number of *schizothoracine* fish mitochondrial genomes available for future work. Additionally, our contributed genome was helpful in the accurate identification of *S. gulinensis* through combined with a previously published analysis of *S. gulinensis* morphological. And this study will contribute to the future research on population genetics and evolution of *Schizothorax* genus.

## Acknowledgments

We appreciate anonymous reviewers for providing valuable comments on this article.

## Ethical approval

All applicable international, national, and/or institutional guidelines for the care and use of animals were strictly followed. All animal sample collection protocols complied with the current laws of China. All animal procedures performed in this research were in accordance with the ethical standards approved by the Ethics Committee for the Use of Animal Subjects of Xichang University (No. xcc2022030).

## Author contributions

Conceived and designed the experiments: Ying Wang and Pan Shang. Performed the experiments: Ying Wang. Sample collection: Yuanxing Dai and Ying Wang. Analyzed the data: Ying Wang and Pan Shang. Wrote the manuscript: Ying Wang. Helped to proofread the manuscript: Zhiqiu Huang, Yanzhen Dong, and Dayong Xu. All authors have read and agreed to the published version of the manuscript.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The study was supported by the Doctoral Research Initiation Project of Xichang University under Grant [No.117281676].

## 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. OQ730545. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA954193, SRR24138089, and SAMN34147823, respectively.

## References

- Bernt M, Donath A, J uhling F, Externbrink F, Florentz C, Fritsch G, P utz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol.* 69(2):313–319. doi: [10.1016/j.ympev.2012.08.023](https://doi.org/10.1016/j.ympev.2012.08.023).
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30(15):2114–2120. doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- Broughton RE, Milam JE, Roe BA. 2001. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genome Res.* 11(11):1958–1967. doi: [10.1101/gr.156801](https://doi.org/10.1101/gr.156801).
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics.* 10(1):421. doi: [10.1186/1471-2105-10-421](https://doi.org/10.1186/1471-2105-10-421).
- Cao WX, Wu XW. 1964. Fauna of cyprinidacin China (volume I). Shanghai, China: Shanghai Science and Technology Press; p. 137–197.
- Ding RH, Dai YX, Huang YY. 2022. A new species of the genus *Schizothorax* from Sichuan, China (Cypriniformes: cyprinidae). *Sichuan J Zool.* 41(3):300–303.
- Grant JR, Stothard P. 2008. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res.* 36:W181–4. doi: [10.1093/nar/gkn179](https://doi.org/10.1093/nar/gkn179).
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5(2):R12. doi: [10.1186/gb-2004-5-2-r12](https://doi.org/10.1186/gb-2004-5-2-r12).
- Li ZL, Chen YX, Hu SY, Zhao HT. 2015. Multivariate Analysis on the Morphological Differentiation of Kozlov's *Schizothorax kozlovi* and David's *Schizothorax (S. davidi)*. *Chin J Zool.* 50(4):547–554.
- Mabuchi K, Miya M, Senou H, Suzuki T, Nishida M. 2006. Complete mitochondrial DNA sequence of the Lake Biwa wild strain of common carp (*Cyprinus carpio* L.): further evidence for an ancient origin. *Aquaculture.* 257(1–4):68–77. doi: [10.1016/j.aquaculture.2006.03.040](https://doi.org/10.1016/j.aquaculture.2006.03.040).
- Meland S, Johansen S, Johansen T, Haugli K, Haugli F. 1991. Rapid disappearance of one parental mitochondrial genotype after isogamous mating in the myxomycete *Physarum polycephalum*. *Curr Genet.* 19(1):55–59. doi: [10.1007/BF00362088](https://doi.org/10.1007/BF00362088).
- Meng GL, Li YY, Yang CT, Liu SL. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Res.* 47(11):e63–e63. doi: [10.1093/nar/gkz173](https://doi.org/10.1093/nar/gkz173).
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 37(5):1530–1534. doi: [10.1093/molbev/msaa015](https://doi.org/10.1093/molbev/msaa015).
- Qi D, Chao Y, Guo S, Zhao L, Li T, Wei F, Zhao X. 2012. Convergent, parallel and correlated evolution of trophic morphologies in the subfamily *Schizothoracinae* from the Qinghai-Tibetan Plateau. *PLoS One.* 7(3): e34070. doi: [10.1371/journal.pone.0034070](https://doi.org/10.1371/journal.pone.0034070).
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, H hna 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. doi: [10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029).
- Xu QH, Zhang C, Zhang DS, Jiang HP, Peng SH, Liu Y, Zhao K, Wang CC, Chen LB. 2016. Analysis of the erythropoietin of a Tibetan Plateau *Schizothoracine* fish (*Gymnocypris dobula*) reveals enhanced cytoprotection function in hypoxic environments. *BMC Evol. Biol.* 16(1):11.
- Zheng Z, Yu E, Li ZF, Ou HX, Wang GJ. 2016. The complete mitochondrial genome of *Schizothorax grahmi* (Cypriniformes: cyprinidae). *Mitochondrial DNA B Resour.* 1(1):775–776. doi: [10.1080/23802359.2016.1238757](https://doi.org/10.1080/23802359.2016.1238757).