Yu et al. *Zool. Res.* 2022, 43(1): 30−32
https://doi.org/10.24272/j.issn.2095-8137.2021.243

ZR | Zoological Research

Letter to the editor Open Access

Comparative mitogenomic analyses unveil conserved and variable mitogenomic features and phylogeny of Chedrinae fish

Chedrinae fish, which belong to Danionidae, have important ornamental, economic, and scientific value. At present, however, their mitogenomic features are unclear and their phylogenetic relationships remain controversial. In this study, we presented five new Chedrinae mitochondrial genomes (mitogenomes) and analyzed the conserved and variable mitogenomic characteristics of 17 Chedrinae fish. The gene composition and arrangement and secondary structure of transfer RNAs (tRNAs) were highly conserved among the Chedrinae mitogenomes. However, the length of the control region and base composition were variable. Interestingly, the mitogenome of *Barilius barila* was unusual, with lower A+T content in the first codon of protein-coding genes (PCGs) (47.32% versus average of 54.47%) and distinct pattern of codons per thousand codons (CDspT). Three Chedrinae fish had a long tandem repeat (>291 bp) in the 5'-end of the control region, which may increase their adaptability. In addition, $tRNA^{Lys}$ had notably larger DHU and TΨC loops than other tRNAs. The phylogenetic trees of the Chedrinae fish suggested that the *Barilius* genus was not a monophyletic group but could be divided into two main groups based on significant differences in A+T content. This study provides insights into the mitogenomic features and phylogenetic implications of Chedrinae fish, which should benefit their systematics and conservation.

Chedrinae Bleeker 1863, also called Chedrina/Chedrini (Liao et al., 2011a), is one of the three subfamilies of Danionidae (Tang et al., 2010). Chedrinae fish usually inhabit turbulent rivers, mountain streams, and freshwater reservoirs. They are economically important in mountain regions and a valuable food source, as well as popular aquarium species due to their beautiful colors and patterns (Prabhu et al., 2020). Chedrinae fish are wildly distributed across Asia and Africa. The genus *Raiamas* is found on both continents, while other genera are found in Asia or Africa only, making them suitable models for studying the biogeographic history of freshwater systems. Due to habitat fragmentation, chemical fertilizer and pesticide use, and overfishing, Chedrinae and other mountain stream fish have become seriously threatened (Prabhu et al., 2020). Accurate species identification is essential for targeted protection; however, Chedrinae fish are cryptic, especially at the juvenile stage, and are thus difficult to distinguish morphologically.

Mitochondrial genomes (mitogenomes) are widely used in species identification and molecular phylogenetics (Jiang et al., 2021; Yu et al., 2021). The gene arrangements of fish mitogenomes are generally conserved, whereas genome sequence length, base composition bias, and control region (CR) can show considerable species-level diversity (Yu et al., 2021). Although some complete Chedrinae mitogenomes have been sequenced and are available in the NCBI GenBank database, the mitogenomic characteristics of Chedrinae fish remain unclear. In addition, the phylogeny of Chedrinae fish is still controversial. In the current study, we sequenced five complete Chedrini mitogenomes and compared them with 12 other Chedrini to clarify the features of the Chedrinae mitogenomes. We also constructed the phylogenetic relationship of Chedrinae fish based on a 11 383 bp sequence matrix of 13 PCGs.

The complete mitogenomes of five Chedrinae fish, including *B. bernatziki* (GenBank accession No. MW625809), *B. barila* (MW625806), *B. pulchellus* (MW625808), *B. ardens* (MW625805), and *B. canarensis* (MW625807), were first sequenced and annotated (Figure 1A; Supplementary Table S1), then compared with 12 previously reported Chedrinae mitogenomes (Supplementary Table S2). The A+T content,

Received: 31 August 2021; Accepted: 09 November 2021; Online: 10 November 2021

Foundation items: This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000), China Agriculture Research System of MOF and MARA, and Wuhan Branch, Supercomputing Centre, Chinese Academy of Sciences, China.

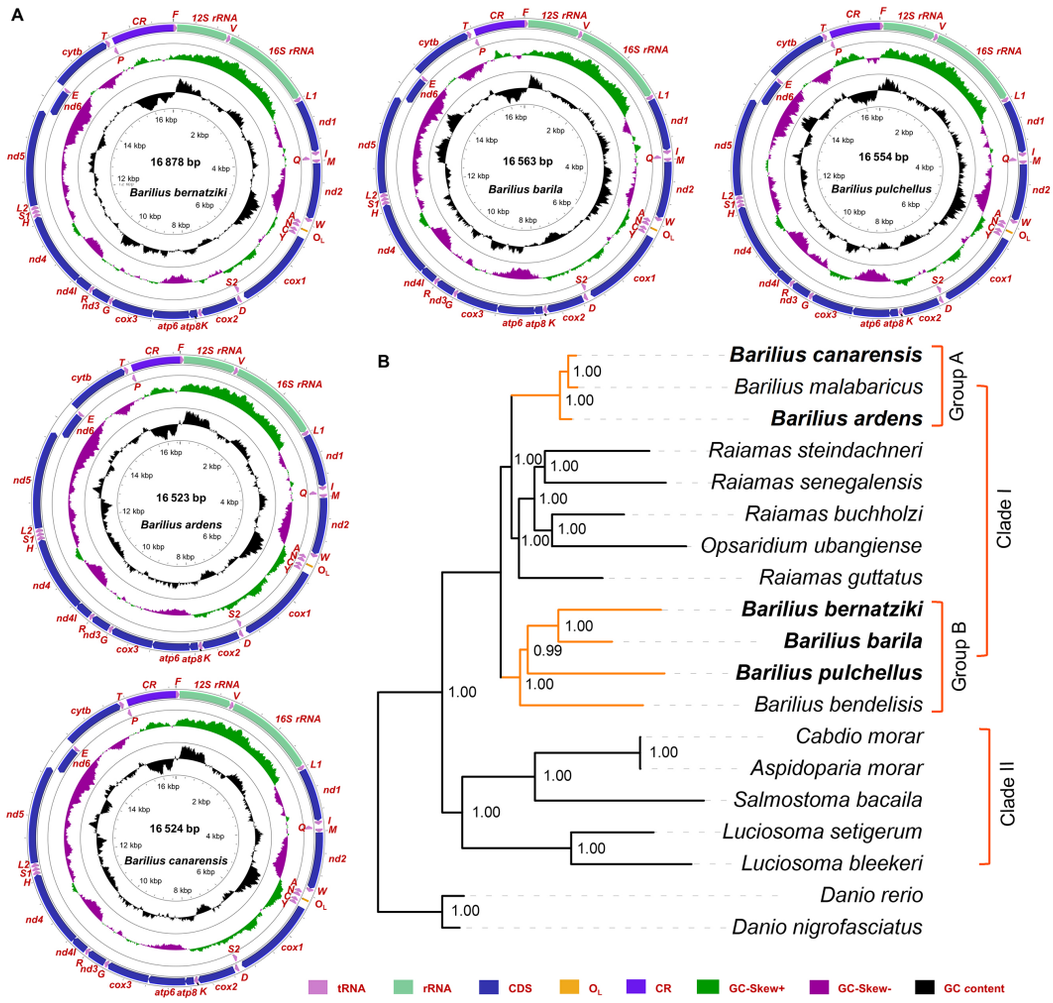

Figure 1 Mitogenome circular sketch map of five fish sequenced in this study (A) and phylogenetic tree constructed by BI methods based on 13 PCGs of 17 Chedrinae mitogenomes (B)

A: Different colors represent different gene blocks. B: *D. rerio* and *D. nigrofasciatus* were chosen as outgroups. Node numbers represent values of posterior probability.

AT-skew, and GC-skew of the PCGs, tRNAs, ribosomal RNAs (rRNAs), and CRs of the 17 Chedrinae mitogenomes were calculated (Supplementary Figure S1A–C). All counted units exhibited AT bias, and A+T content was highest in the CRs (63.00%±2.50%) (Supplementary Figure S1A, Supplementary Table S3). Of note, A+T content at the first position of *B. barila* was low (47.32% versus average of 54.47%).

We aligned 22 tRNA sequences of 17 Chedrinae fish, which were relatively conserved (Supplementary Figure S2). Although the secondary structure of each tRNA was highly conserved among the 17 Chedrinae mitogenomes, there were significant differences among the different tRNAs (Supplementary Figure S3). For example, among the 22 tRNAs, 21 showed a typical clover-leaf secondary structure, whereas $tRNA^{Ser(AGN)}$ *(S1)* lacked the DHU arm. In addition, the sizes of the DHU loops varied among the 21 tRNAs and the DHU loop of $tRNA^{Lys}$ (*K*) (15 bp) was much larger than that of the other tRNAs. Interestingly, *K* also had the largest TΨC loop (11 bp). Large DHU and TΨC loops in *K* have been reported in other fish (Yu et al., 2021), and thus may be a common characteristic of fish.

The CR was the most variable region in the Chedrinae mitogenomes, with a length ranging from 798 bp (*Aspidoparia morar*) to 1 245 bp (*Raiamas senegalensis*) (Supplementary Table S3). There was one short tandem repeat unit (TA) at the 3'-end of each Chedrinae CR, except for *Cabdio morar*, which had two independent TA repeat units (Supplementary Figure S4). In addition, we identified a long tandem repeat (>291 bp) at the 5'-end in three species (*B. bernatziki*, *Opsaridium ubangiense*, and *R. senegalensis*), which resulted in a longer CR (>1 000 bp). These results indicate that the variable CR lengths were primarily caused by tandem repeats. Previous studies have suggested that tandem repeats in mitogenome CRs can be transferred to the nuclear genome, in a somewhat similar way to transposons, which may help species gain a fitness advantage (Dover, 1982; Lin et al., 2021). Among the three Chedrinae fish with longer CRs, two are dominant or widely distributed in their natural habitats. For example, *O. ubangiense* is widespread in the Lowa Basin in Africa (Kisekelwa et al., 2020) and *R. senegalensis* is a dominant

species among the 316 recorded freshwater bony fish in Nigeria (Olanrewaju et al., 2017). Although the function of tandem repeats in the 5'-end of CRs in fish remains unclear, we speculate that the appearance of long tandem repeats could help fish become more adaptable as mitogenome repeats may function as a molecular driver via transposition (Dover, 1982; Lin et al., 2021). Although there were significant differences in length and tandem repeats, it was still possible to identify important and conserved domains in the CRs in the Chedrinae fish. For instance, two conserved domains (CSB) were recognized in the longest CR of *B. bernatziki* (Supplementary Figure S5).

The CDspT of the 17 Chedrinae mitogenomes were analyzed (Supplementary Figure S6). Results showed that the CDspT pattern was largely comparable among the mitogenomes, except for *B. barila*, which had more Leu$^{(CUN)}$ codons (137.99) and fewer Ile codons (66.23 vs. average of 77.13). Thus, given its lower A+T content and unusual CDspT pattern, *B. barila* differed from the other Chedrinae fish.

Based on the phylogenetic trees, the BI and ML analyses generated similar topologies (Figure 1B and Supplementary Figure S7). Chedrinae could be divided into two clades: i.e., *Barilius-Raiamas-Opsaridium* (clade I) and *Cabdio-Aspidoparia-Salmostoma-Luciosoma* (clade II). In clade I, the genus *Barilius* was not a monophyletic group, but could be divided into two groups (A and B). Group A contained three species (*B. canarensis*, *B. malabaricus*, and *B. ardens*) and showed the closest relationship with the *Raiamas-Opsaridium* group, and then clustered with Group B (*B. bernatziki*, *B. barila*, *B. pulchellus*, and *B. bendelisis*). These findings differ from Liao et al. (2011b) and Tang et al. (2010) but are similar to Prabhu et al. (2020). In addition, A+T content of the third codon in the Group A species was higher (>60%) than that in Group B (53.65%–57.84%) (Supplementary Table S3).

In conclusion, we reported on the complete mitogenomes of five Chedrinae fish and found some conserved and variable mitogenomic characteristics via comparative mitogenomic analyses, which should shed light on the architecture and evolution of fish mitogenomes. Additionally, we found several unusual features in the *B. barila* mitogenome and identified a long tandem repeat at the 5'-end in three species (*B. bernatziki*, *O. ubangiense*, and *R. senegalensis*). We also analyzed the phylogeny of Chedrinae fish, which indicated that the genus *Barilius* was not a monophyletic group but could be divided into two groups based on differences in A+T content. The current study provides comprehensive insight into the mitogenomic features and phylogenetic implications of Chedrinae fish, which should benefit their systematics and conservation.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

P.Y. designed the study, collected the samples, analyzed the data, prepared the figures, and wrote the original draft. J.F.G. and Y.W. revised the manuscript. L.Z. revised the manuscript and applied for funding. X.L.Y., L.L.L., and H.J. collected the samples and analyzed the data. Z.L., Z.W.W., and X.Y.L. collected the samples and identified the species. X.J.Z. submitted the manuscript. All authors read and approved the final version of the manuscript.

Peng Yu[1,2], Li Zhou[1,2,3], Xiao-Li Yang[1,3], Liang-Liang Li[1,3], Hui Jin[1,3], Zhi Li[1], Zhong-Wei Wang[1], Xi-Yin Li[1], Xiao-Juan Zhang[1], Yang Wang[1,*], Jian-Fang Gui[1,2,3,*]

[1] *State Key Laboratory of Freshwater Ecology and Biotechnology*, *Institute of Hydrobiology*, *Innovation Academy of Seed Design*, *Chinese Academy of Sciences*, *Wuhan*, *Hubei* 430072, *China*
[2] *Hubei Hongshan Laboratory*, *Wuhan*, *Hubei* 430070, *China*
[3] *University of Chinese Academy of Sciences*, *Beijing* 10049, *China*
*Corresponding authors, E-mail: wangyang@ihb.ac.cn; jfgui@ihb.ac.cn

REFERENCES

Dover G. 1982. Molecular drive: a cohesive mode of species evolution. *Nature*, **299**(5879): 111−117.

Jiang W, Zhou TT, Wang HZ, Yu P, Erséus C, Cui YD. 2021. Genetic and morphological analyses uncover a new record and a cryptic species in *Allonais* (clitellata: Naididae). *Biologia*, **76**(6): 1705−1714.

Kisekelwa T, Snoeks J, Vreven E. 2020. An annotated checklist of the fish fauna of the river systems draining the kahuzi-biega national park (upper congo: Eastern DR congo). *Journal of Fish Biology*, **96**(3): 700−721.

Liao TY, Kullander SO, Fang F. 2011a. Phylogenetic position of rasborin cyprinids and monophyly of major lineages among the danioninae, based on morphological characters (cypriniformes: Cyprinidae). *Journal of Zoological Systematics and Evolutionary Research*, **49**(3): 224−232.

Liao TY, Ünlü E, Kullander SO. 2011b. Western boundary of the subfamily danioninae in asia (teleostei, cyprinidae): derived from the systematic position of *Barilius mesopotamicus* based on molecular and morphological data. *Zootaxa*, **2880**(1): 31−40.

Lin ZJ, Wang XZ, Wang JB, Tan YT, Tang XM, Werren JH, et al. 2021. Comparative analysis reveals the expansion of mitochondrial DNA control region containing unusually high G-C tandem repeat arrays in nasonia vitripennis. *International Journal of Biological Macromolecules*, **166**: 1246−1257.

Olanrewaju AN, Kareem OK, Nyaku RE, Tubo MT. 2017. Length-weight and length-length relationships of heterotis niloticus (cuvier, 1829) and *Raiamas senegalensis* (steindachner, 1870). *Journal of Aquaculture Research & Development*, **S2**: 011.

Prabhu VR, Singha HS, Kumar RG, Gopalakrishnan A, Nagarajan M. 2020. Characterization of the complete mitochondrial genome of *Barilius malabaricus* and its phylogenetic implications. *Genomics*, **112**(3): 2154−2163.

Tang KL, Agnew MK, Hirt MV, Sado T, Schneider LM, Freyhof J, et al. 2010. Systematics of the subfamily danioninae (teleostei: Cypriniformes: Cyprinidae). *Molecular Phylogenetics and Evolution*, **57**(1): 189−214.

Yu P, Zhou L, Yang WT, Miao LJ, Li Z, Zhang XJ, et al. 2021. Comparative mitogenome analyses uncover mitogenome features and phylogenetic implications of the subfamily cobitinae. *BMC Genomics*, **22**(1): 50.