Pakistan J. Zool., vol. 58(2), pp 997-1000, 2026 **DOI**: https://dx.doi.org/10.17582/journal.pjz/20231220061408

Short Communication

Analysis of the Complete Mitochondrial Genome Sequences and Gene Organization of *Hyphessobrycon rosaceus* (Characiformes, Characidae)

Xiaoli Zhang*, Shuya Liu, Liping Chen and Zhenli Wang

College of Environment and Life Health, Anhui Vocational and Technical College, Hefei 230011, China.

Article Information
Received 20 December 2023
Revised 18 January 2024
Accepted 13 February 2024
Available online 08 July 2025 (early access)
Published 25 February 2026

Authors' Contribution
XZ, SL, LC and ZW designed the study. XZ and SL executed experimental work. XZ and LC analyzed the data. XZ wrote the paper. XZ provided the laboratory equipment. XZ supervised the research.

Key words
***Hyphessobrycon rosaceus*, Mitochondrial genome, Phylogeny**

ABSTRACT

Hyphessobrycon rosaceus is an important oviparous ornamental fish distributed in the Essequibo and Suriname River basins in South America. To understand the mitochondrial genome characteristics and phylogenetic status of *H. rosaceus*, Illumina HiSeq high-throughput sequencing technology was used to analyze its entire mitochondrial genome, and the mitochondrial genome sequences of related species were downloaded for a phylogenetic analysis. The results showed that the length of all mitochondrial genes of *H. rosaceus* comprised 16,870 bp, including 13 protein-coding genes, 22 tRNA genes, two rRNA genes, and one D-loop control region. A phylogenetic tree was constructed based on 13 protein-coding gene sequences. Here, the target species *H. rosaceus* and *Hyphessobrycon socolofi* clustered well into one branch, supporting the classification of *H. rosaceus* as belonging to *Hyphessobrycon*. However, the mixed distribution of *Hyphessobrycon* genera throughout the phylogenetic tree suggests that further research is needed to classify this genus. These results provide an important basis for further clarifying the classification status of *H. rosaceus*, which could help with associated resource protection.

Hyphessobrycon rosaceus (Durbin, 1909) is an important ornamental fish of the order Characiformes and family Characidae. Moreover, this species is distributed in the Essequibo and Suriname River basins in South America. *H. rosaceus* is an oviparous fish that feeds on small insects and plant debris. Further, it has a red and transparent background; all fins are red, and the edges of the dorsal fins are black. In addition, the artificially improved White Winged Rose Flag has a dynamic white edge on its dorsal fins (Castro Paz *et al*., 2014; Barik *et al*., 2018).

The mitochondrial genome of *H. rosaceus* is widely used in studies of fish systematics, population polymorphism, and germplasm resource conservation (Boore, 1999; Gray et al., 1999; Sun et al., 2021, 2022a). In recent years, the development of gene sequencing technology has provided a convenient method for studying the entire mitochondrial genome. Using second-generation sequencing technology, entire mitochondrial genome sequences were obtained. These were supplemented with mitochondrial whole genome data of the Characidae family and combined with existing related mitochondrial genome sequence phylogenetic analyses to provide a reference basis to further clarify the classification and status of the genus *Hyphessobrycon*, which will be important for resource protection.

Materials and methods

H. rosaceus was collected from Yufeng Flower, Bird, Insect and Fish Market in Hefei City, Anhui Province (31.853174° N, 117.305852° E). After morphological analysis, a small amount of muscle tissue was taken from the sample and used for extraction of genomic DNA using the high-salt method (Aljanabi and Martinez, 1997). DNA quality was determined via 10 g/L agarose gel electrophoresis, and the DNA purity and concentration were determined using a NanoDrop ND-2000 ultra-micro spectrophotometer. To ensure sample accuracy, DNA barcodes were identified using the mitochondrial COI gene sequence (accession number KU568886).

Using a Covaris ultrasound to break the genomic

* Corresponding author: gzzxl89@163.com
0030-9923/2026/0002-0997 $ 9.00/0

DNA of the sample into approximately 350 bp fragments, repair the end of the DNA fragment, and add an A base and sequencing adapter to the 3 end. PCR amplification was performed on the connected products, which were then recovered and purified using magnetic beads to construct a library. After quality inspection of the DNA library, the Illumina HiSeq high-throughput sequencing platform was used for paired-end sequencing (Caporaso *et al*., 2012), with a sequencing data volume of no less than 6 Gb. Sequencing was performed by Shenggong Biotechnology Co. Ltd. (Shanghai, China). Low-quality reads and splice sequences were filtered out from the sequencing data, and NOVOPlasty was used for de novo assembly (Dierckxsens *et al*., 2017), referring to the published mitochondrial genome sequence of *Hyphessobrycon amandae* (accession number MT484069, Sun *et al*., 2021). BioEdit v7.2.5 Software was used to proofread the spliced sequences (Hall *et al*., 2011), and MITOS2 (http://mitos.bioinf.unileipzig.de/index.py) online software was used for gene annotation. The complete mitochondrial genome sequence of *H. rosaceus* has been submitted to the NCBI (National Center for Biotechnology Information) database (accession number: OP595700).

For phylogenetic analysis Blast alignment (Johnson *et al*., 2008) was performed based on the NCBI database using the mitochondrial gene sequence of *H. rosaceus*. The mitochondrial genome sequences of the 41 fish species showed high homology with the mitochondrial gene sequence of *H. rosaceus*. *Corydoras aeneus* (accession number MZ571336, Sun *et al*., 2022b) was used as an outlier for the phylogenetic analysis. Using BioEdit v7.2.5, 13 protein-coding genes sequences were extracted from the mitochondrial genome of the sample and spliced into one sequence. A Bayesian inference phylogenetic tree was constructed using MrBayes v3.2.6 (Ronquist and Huelsenbeck, 2003), with four independent Markov chains and 5,000,000 generations run, sampling once every 1000 generations; then, the top 25% were discarded. When the average standard deviation of the column frequency was less than 0.001, it was believed that the analysis tended to be stable, and the posterior support rates for each branch were calculated. A maximum likelihood phylogenetic tree using IQ-TREE v1.6.8 (Nguyen *et al*., 2015), and the confidence of each branch was tested using bootstrapping (5000 repetitions).

Results and discussion

The total length of the mitochondrial genome of *H. rosaceus* was 16,870 bp (Fig. 1), with each containing 13 protein-coding genes, 22 tRNA genes, two rRNA genes, and a control region (D-loop). The base composition of the mitochondrial genome sequence was similar, with A, T, C, and G contents of 29.5, 28.8, 26, and 15.6%, respectively. The A+T content (58.3%) was higher than the C+G content (41.6%).

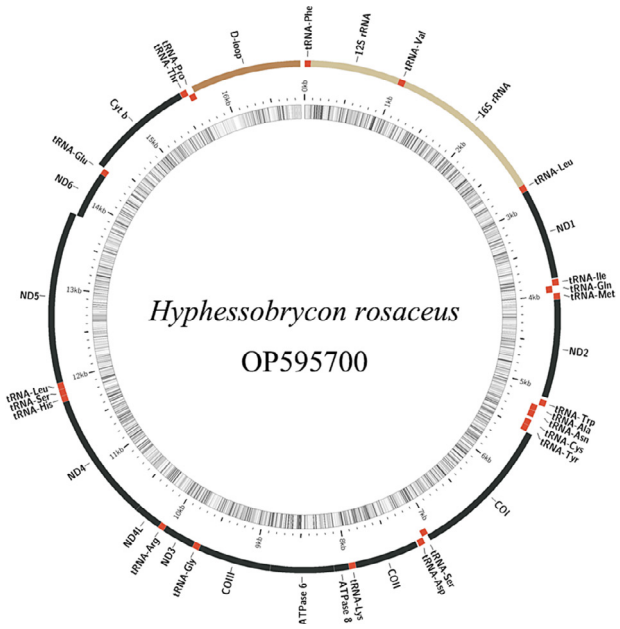

Fig. 1. Gene map of the mitochondrial genome of *Hyphessobrycon rosaceus.*

The 13 protein-coding gene combinations in the mitochondrial genome of *H. rosaceus* obtained in this study had a total length of 11,423 bp, with 3806 codons (excluding termination codons). Twelve genes (*ND1*, *ND2*, *COX1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *ND3*, *ND4L*, *ND4*, *ND5*, and *Cytb*) were encoded in the H chain, whereas only ND6 was encoded in the L chain. All 13 protein-coding genes had ATG as the starting codon. The termination codons included TAG, TAA, AGG, and an incomplete termination codon, T.

An analysis of the frequency of codon usage in protein-coding genes and the frequency of synonymous codon usage (Table I) showed that, among the 13 protein-coding genes, there were 36 preferred codons (RSCU codons), with the third codon C or T having a high usage frequency, whereas the codon ending in G had the lowest usage frequency. In codons ending in A or T at the third base, the RSCUs of AAU (N), AGU (S), CAU (H), CGU (R), GAU (D), GGU (G), UAU (Y), and UGU (C) codons were less than 1, whereas those of the other codons were greater than 1.

Figure 2 shows a phylogenetic tree of 43 mitochondrial genomes, including 39 from the Characidae family, two from Bryconidae, one from Callichthyidae (outer group), and one from Serrasalmidae. It shows that *Paracheirodon* and *Astyanax* genera are well supported. Among the three species of the *Psalidodon* genus, *Hyphessobrycon*

anisitsi was mixed. Other genera were mixed and not clustered according to the classification standards. The target species, *H. rosaceus* and *Hyphessobrycon socolofi*, clustered together well and were also clustered with *Hyphessobrycon megalopterus* and *Hyphessobrycon pulchripinnis*. However, 11 species of this genus did not form a single branch, resulting in a mixed distribution of species throughout the phylogenetic tree. The phylogenetic results indicate that further research is required to classify the *Hyphessobrycon* genus.

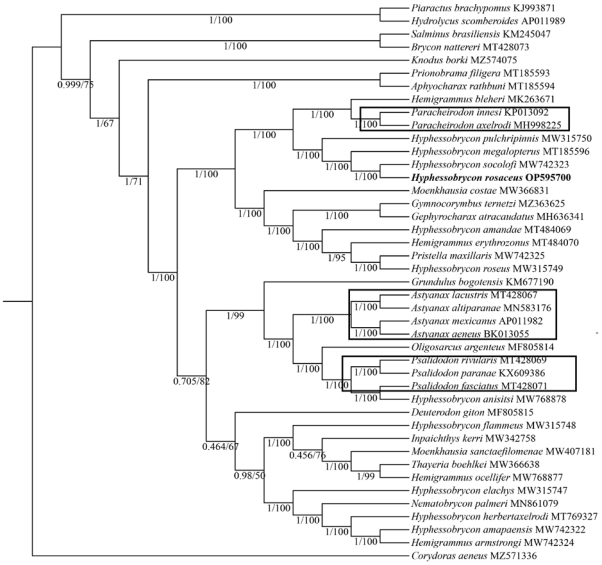

Fig. 2. Phylogenetic relationships among Characidae species based on concatenated nucleotide sequences of 13 mitochondrial protein-coding genes. Numbers at the nodes represent the posterior probability for the Bayesian analysis and bootstrap values for the maximum likelihood analysis.

Table I. Frequency of codon usage in the 13 protein-coding genes.

Codon	Count	RSCU
UUU(F)	**141**	**1.12**
UUC(F)	111	0.88
UUA(L)	**159**	**1.51**
UUG(L)	28	0.27
CUU(L)	**140**	**1.33**
CUC(L)	**107**	**1.02**
CUA(L)	**157**	**1.5**
CUG(L)	39	0.37
AUU(I)	**201**	**1.37**
AUC(I)	92	0.63
AUA(M)	**131**	**1.51**
AUG(M)	43	0.49
GUU(V)	**66**	**1.28**
GUC(V)	**57**	**1.11**
GUA(V)	**63**	**1.22**
GUG(V)	20	0.39
UCU(S)	**64**	**1.54**
UCC(S)	**52**	**1.25**
UCA(S)	**79**	**1.9**
UCG(S)	8	0.19
CCU(P)	**61**	**1.24**
CCC(P)	**53**	**1.08**
CCA(P)	**68**	**1.38**
CCG(P)	15	0.3
ACU(T)	**77**	**1.08**
ACC(T)	**96**	**1.35**
ACA(T)	**96**	**1.35**
ACG(T)	15	0.21
GCU(A)	**85**	**1.02**
GCC(A)	**126**	**1.52**
GCA(A)	**103**	**1.24**
GCG(A)	18	0.22
UAU(Y)	61	0.99
UAC(Y)	**62**	**1.01**
UAA(*)	**5**	**2.5**
UAG(*)	**2**	**1**
CAU(H)	42	0.82
CAC(H)	**60**	**1.18**
CAA(Q)	**80**	**1.55**
CAG(Q)	23	0.45
AAU(N)	55	0.84
AAC(N)	**76**	**1.16**
AAA(K)	**82**	**1.8**
AAG(K)	9	0.2
GAU(D)	32	0.82
GAC(D)	**46**	**1.18**
GAA(E)	**80**	**1.67**
GAG(E)	16	0.33
UGU(C)	14	0.9
UGC(C)	**17**	**1.1**
UGA(W)	**99**	**1.74**
UGG(W)	15	0.26
CGU(R)	16	0.89
CGC(R)	10	0.56
CGA(R)	**39**	**2.17**
CGG(R)	7	0.39
AGU(S)	17	0.41
AGC(S)	29	0.7
AGA(*)	0	0
AGG(*)	1	0.5
GGU(G)	42	0.7
GGC(G)	**76**	**1.27**
GGA(G)	**87**	**1.45**
GGG(G)	35	0.58

Table continued on next column.....

Note: bold represents preferred codons.

In this study, the mitochondrial whole genome structure of *H. rosaceus* was similar to that of other bony fish (Gutiérrez et al., 2015; Satoh et al., 2016) and showed a significant A+T preference in the base composition. The D-loop is the region with the most complex changes in mitochondrial DNA (Brown et al., 1986), and its evolutionary rate is 5–10 times that of other mtDNA segments. There were significant differences in the control regions between related species and even between different individuals of the same species. The size of the repeat sequences in the control region of teleost fish varies greatly, and the length of the control region used in this study was 1186 bp. However, the role of D-loop repeat sequences in evolution needs further exploration (Gvozdanović et al., 2019).

In the mitochondrial genome of *H. rosaceus*, protein-coding genes were determined to be present in both the H and L chains. Brown et al. (1982) reported that protein-coding genes on the H-chain are not easily transformed into protected double chains because most H-chains are hydrolyzed single chains. Most protein-coding genes in the mitochondrial genome of *H. rosaceus* were found to be located in the H chain and are prone to hydrolysis and oxidation. The *ND6* gene located in the L chain is much more stable, reflecting the differences in and importance of these genes.

Conclusions

The phylogenetic tree constructed showed that the target species, *H. rosaceus* and *H. socolofi*, clustered together well, supporting the classification of *H. rosaceus* as belonging to *Hyphessobrycon*.

Funding

Funded by Engineering Research Center for Tropical and Subtropical Aquatic Ecological Engineering, Ministry of Education, Jinan University, P.R. China (No. 2022A0201).

Statement of conflict of interest

The authors have declared no conflict of interest.

References

Aljanabi, S.M. and Martinez, I., 1997. *Nucl. Acids Res*., **25**: 4692-4693. https://doi.org/10.1093/nar/25.22.4692

Barik, M., Bhattacharjee, I., Ghosh, A. and Chandra, G., 2018. *BMC Res. Notes*, **11**: 1-5. https://doi.org/10.1186/s13104-018-3902-8

Boore, J.L., 1999. *Nucl. Acids Res*., **27**: 1767-1780. https://doi.org/10.1093/nar/27.8.1767

Brown, G.G., Gadaleta, G., Pepe, G., Saccone, C. and Sbisà, E., 1986. *J. mol. Biol*., **192**: 503-511. https://doi.org/10.1016/0022-2836(86)90272-X

Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C., 1982. *J. mol. Evol*., **18**: 225-239. https://doi.org/10.1007/BF01734101

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A. and Knight, R., 2012. *ISME J*., **6**: 1621-1624.https://doi.org/10.1038/ismej.2012.8

Castro-Paz, F.P., Batista, J.D.S. and Porto, J.I.R., 2014. *PLoS One*, **9**: e98603. https://doi.org/10.1371/journal.pone.0098603

Dierckxsens, N., Mardulyn, P. and Smits, G., 2017. *Nucl. Acids Res*., **45**: e18-e18. https://doi.org/10.1093/nar/gkw1314

Gray, M.W., Burger, G. and Lang, B.F., 1999. *Science*, **283**: 1476-1481. https://doi.org/10.1126/science.283.5407.1476

Gutiérrez, V., Rego, N., Naya, H. and García, G., 2015. *BMC Genomics*, **16**: 1-15. https://doi.org/10.1186/s12864-015-2090-3

Gvozdanović, K., Margeta, V., Margeta, P., Djurkin Kušec, I., Galović, D., Dovč, P. and Kušec, G., 2019. *Anim. Biotechnol*., **30**: 242-251. https://doi.org/10.1080/10495398.2018.1478847

Hall, T., Biosciences, I. and Carlsbad, C.J.G.B.B., 2011. *GERF Bull. Biosci*., **2**: 60-61.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T.L., 2008. *Nucl. Acids Res*., **36**: W5-W9. https://doi.org/10.1093/nar/gkn201

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A. and Minh, B.Q., 2015. *Mol. Biol. Evol*., **32**: 268-274. https://doi.org/10.1093/molbev/msu300

Ronquist, F. and Huelsenbeck, J.P., 2003. *Bioinformatics*, **19**: 1572-1574. https://doi.org/10.1093/bioinformatics/btg180

Satoh, T.P., Miya, M., Mabuchi, K. and Nishida, M., 2016. *BMC Genomics*, **17**: 1-20. https://doi.org/10.1186/s12864-016-3054-y

Sun, C.H., Zhang, Y.N., Zeng, X.S., Liu, D.W., Huang, Q., Zhang, X.L. and Zhang, Q., 2022a. *Mol. Biol. Rep*., **49**: 1741-1748.

Sun, C.H., Huang, Q., Zeng, X.S., Li, S., Zhang, X.L., Zhang, Y.N., Liao, J., Lu, C.H., Han, B.P. and Zhang, Q., 2022b. *Zookeys*, **1083**: 89. https://doi.org/10.3897/zookeys.1083.76887

Sun, C.H., Liu, H.Y., Xu, N., Zhang, X.L., Zhang, Q. and Han, B.P., 2021. *Front. Genet*., **12**: 627402. https://doi.org/10.3389/fgene.2021.627402