Pakistan J. Zool., pp 1-8, 2023. **DOI**: https://dx.doi.org/10.17582/journal.pjz/20230302100304

Structural Analysis and Phylogenetic Relationships of a Teleost Fish, *Pethia stoliczkana* Based on the Complete Mitochondrial Genome Sequence

Zhenglei Qiao*, Jie Xing and Fang Li

College of Life Science and Technology, Mudanjiang Normal University, Mudanjiang 157012, China

Article Information
Received 02 March 2023
Revised 21 March 2023
Accepted 03 April 2023
Available online 21 July 2023
(early access)

Authors' Contribution
ZQ collected and analysed the data and wrote the manuscript. JX and FL supervised the study, analysed the data, and edited the manuscript.

Key words
Mitogenome, mtDNA, Next-generation sequencing, Phylogeny

ABSTRACT

In this study, the whole mitochondrial genome sequence of *Pethia stoliczkana* was obtained using high-throughput sequencing technology, and its structure and characteristics were analyzed. The *P. stoliczkana* mitochondrial genome contained a total of 16,966 base pairs, including 13 protein-coding genes, 22 transport RNA genes, two ribosomal RNA genes, and one control region. The A+T content (59.7%) of the whole mitochondrial genome was greater than the G+C content (40.3%), indicating an obvious A+T preference. The mitochondrial genome of *P. stoliczkana* is similar to that of most teleost fish, and no gene rearrangements were detected. The phylogenetic relationship tree of Smiliogastrinae fish was constructed based on 13 protein-coding genes using the Bayesian inference and maximum likelihood methods. We found that *P. stoliczkana* was closely related to *Pethia ticto* and *Pethia padamya*. These results enrich the mitochondrial genome database of Smiliogastrinae fish and provide reference materials for systematic classification of this group of fish.

INTRODUCTION

Pethia stoliczkana (Day, 1871) belongs to the order Cypriniformes, family Cyprinidae, and subfamily Smiliogastrinae. This tropical benthic freshwater fish is mainly distributed in Laos, Thailand, Myanmar, and India (Nath *et al*., 2022). The main morphological characteristics of *P. stoliczkana* species are as follows: Flank behind gill opening with vertically elongated black blotch; caudal peduncle with vertically elongated black blotch; dorsal fin of sexually active male is red with black margin and two rows of black spots; no barbels; and last simple dorsal ray serrated posteriorly; this fish has important economic and ornamental value (Atkore *et al*., 2015; Nath *et al*., 2022).

Mitochondrial DNA (mtDNA) is the only genetic material outside of the cell nucleus in animals that can replicate and transcribe independently. In contrast to nuclear DNA, mtDNA is maternally inherited, has a simple molecular structure, undergoes rapid evolution, and exhibits unorganized specificity. mtDNA is a powerful tool for studying the origin and phylogeny of species, genetic differentiation between related species and intraspecific populations, species identification, and genetic diversity (Funk and Omland, 2003; Wolstenholme, 1992). Fish mtDNA is useful for studying evolutionary genetics; particularly, the mitochondrial genome sequence contains more information than a single gene and more comprehensively reflects the genetic characteristics of species and phylogenetic relationships at different taxonomic levels (Avise *et al*., 1987). In the past ten years, the mitochondrial genomes of fish have been widely studied using high-flux sequencing technology, leading to an increase in reports on completed mitochondrial genome sequences of fish.

In this study, we determined the full mitochondrial genome sequence of *P. stoliczkana* using high-throughput sequencing technology and analyzed its gene composition and structural characteristics. Combined with mitochondrial sequence information of related species, the phylogenetic relationships of Smiliogastrinae fish were determined using the protein-coding genome sequence. These results fill a knowledge gap in the molecular biology of *P. stoliczkana* and complement and improve the limited mitochondrial genome data on Smiliogastrinae fish. This sequence provides molecular evidence and a theoretical

* Corresponding author: qzlqzqjr@126.com
0030-9923/2023/0001-0001 $ 9.00/0

reference for classification and identification, germplasm resource evaluation and development, and utilization of this group of fish.

MATERIALS AND METHODS

Experimental materials, DNA extraction, and species identification

Samples were purchased from a flower, bird, fish, and insect Market in Mudanjiang, China in June 2022 and preliminarily identified on site based on their morphological characteristics. Genomic DNA was extracted from the fish fins using a noninvasive extraction method. The quality and concentration of the extracted DNA were determined using 1% agarose gel electrophoresis and a NanoDrop 2000 nucleic acid analyzer (Thermo Fisher Scientific, Waltham, MA, USA), respectively. DNA barcoding technology was performed to further identify the species.

Sequencing

DNA samples were sent to Wuhan Beina Biotechnology Co., Ltd. to construct a 350 bp small fragment sequencing library and for high-throughput sequencing. Using sequencing by synthesis technology and an Illumina HiSeq X sequencing platform (San Diego, CA, USA), the constructed sequencing library was sequenced by 150 bp at both ends, and the original sequencing data were filtered using NGS QC Toolkit 2.3.3 (Patel and Jain, 2012) to remove adapter sequences, low-quality terminals, reads with N >10%, and fragments of less than 25 bp.

Assembly, annotation, and feature analysis

Leverage SPAdes v3.11.1 (http://cab.spbu.ru/software/spades/) (Bankevich *et al*., 2012) was used to splice clean reads to build contigs. SSPACE (Boetzer *et al*., 2011) was used to extend the contigs and obtain the final complete mitochondrial genome sequence. MITOS (Bernt *et al*., 2013) was used to annotate the mitochondrial genome sequence. The results were verified by homology comparison with the mitochondrial genes of known Smiliogastrinae species. tRNAscan-SE software (http://lowelab.ucsc.edu/tRNAscan-SE/) (Lowe and Chan, 2016) was used to search for the tRNA gene. Mega 11 (Tamura *et al*., 2021) was used to calculate the base composition, codon usage frequency, AT-skew, and GC-skew of each coding gene in the mitochondrial genome of *P. stoliczkana*.

Table I. Origins of mitochondrial genomes of Cyprinidae fishes.

Taxon (Species)	Size (bp)	AT %	AT-Skew	GC-Skew	Accession number
Cyprinidae					
Smiliogastrinae					
Barbodes binotatus	16573	57	0.159	-0.281	KY305681
Barbodes semifasciolatus	16594	58.2	0.103	-0.256	KC113209
Dawkinsia denisonii	16899	58.6	0.126	-0.263	KF019637
Enteromius thysi	16688	60.5	0.085	-0.22	OP819561
Enteromius trimaculatus	16417	60.8	0.049	-0.188	AB239600
Hampala macrolepidota	16766	58.2	0.151	-0.285	KF670818
Hampala salweenensis	16913	58.9	0.14	-0.284	MW548258
Oliotius oligolepis	16636	58.4	0.102	-0.247	ON864407
Oreichthys crenuchoides	16596	60.2	0.087	-0.209	MK456608
Osteobrama belangeri	16602	60.7	0.091	-0.239	KY887473
Osteobrama belangeri	16594	60.7	0.091	-0.239	MK749691
Pethia padamya	16792	58.6	0.109	-0.225	ON864408
Pethia stoliczkana	16996	59.7	0.109	-0.247	OP785085
Pethia ticto	17302	60	0.109	-0.249	AB238969
Puntigrus tetrazona	16550	59.8	0.095	-0.242	EU287909
Puntius sachsii	16587	58.2	0.103	-0.256	MZ364158
Puntius snyderi	16578	59.3	0.097	-0.251	KC113210
Sahyadria chalakkudiensis	16989	59.9	0.112	-0.259	JX311437
Systomus sarana	16590	58.6	0.122	-0.256	KU886061
Cyprininae (outgroup)					
Sinocyclocheilus bicornutus	17426	57.9	0.119	-0.273	KX528071
Schizopygopsinae					
Gymnocypris eckloni	16686	56.1	0.024	-0.179	JQ004279

Phylogenetic analysis

To examine the phylogenetic status of *P. stoliczkana* in Smiliogastrinae, the nucleotide sequences of 13 protein-coding genes (PCGs) in the mitochondrial genome were used for phylogenetic analysis. The mitochondrial genomes from 18 species of Smiliogastrinae were selected as reference sequences, and a phylogenetic tree was constructed using the maximum likelihood (ML) and Bayesian (BI) methods, with *Sinocyclius bicornutus* and *Gymnocypris eckloni* as the outgroup (Table I). After multiple nucleotide sequence alignments using Cluster X 2.0 (Larkin *et al*., 2007), the results were filtered using Gblocks v0.91b (Castresana, 2000), and the alignment results for each gene were concatenated using SequenceMatrix v1.7 (Vaidya *et al*., 2011). Using SMS software (Lefort *et al*., 2017) and ModelFinder (Kalyaanamoorthy *et al*., 2017), the most suitable alternative model obtained from the evaluation of the tree-building dataset was GTR+I+G. The ML phylogenetic tree was built through 50,000 bootstrap operations using PhyML 3.0 (Guindon *et al*., 2010). MrBayes3 (Ronquist and Huelsenbeck, 2003) was used to calculate 20,000,000 generations, with the sequences sampled and saved every 100 generations; we discarded 25% of the aging samples and built a BI phylogenetic tree.

RESULTS

Gene structure and composition

The mitochondrial genome of *P. stoliczkana* obtained using high-throughput sequencing was 16,993 bp in length (Fig. 1) and contained 22 tRNA genes (tRNAs), 13PCGs, two ribosomal RNA genes (rRNAs), and one control region.

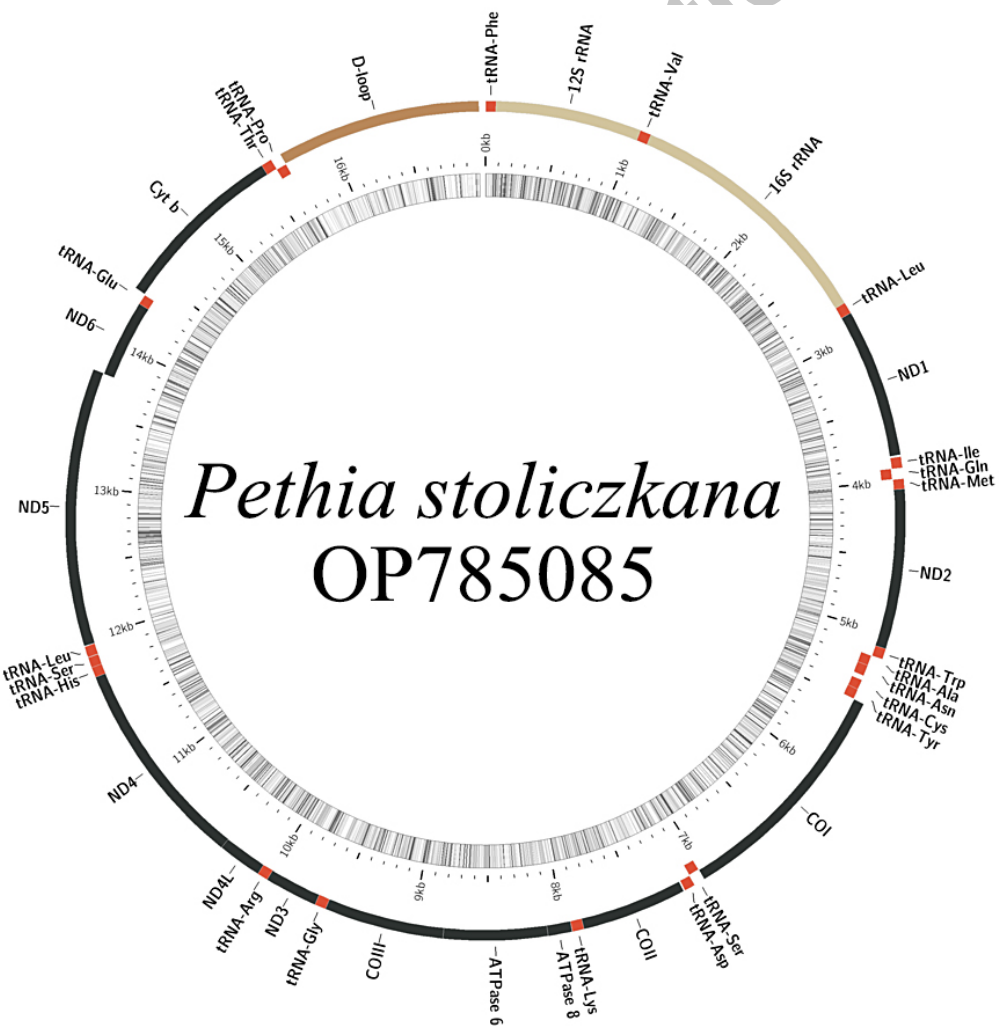

Fig. 1. Mitochondrial genome of *Pethia stoliczkana.*

Table II. Characteristics of the mitochondrial genome of *Pethia stoliczkana.*

Gene	Position		Size	Inter-genic nucle-otides	Codon		
	From	To			Start	Stop	Strand
tRNA-Phe	1	69	69	0			H
12S RNA	70	1025	956	0			H
tRNA-Val	1026	1097	72	0			H
16S RNA	1098	2780	1683	0			H
tRNA-Leu2	2781	2855	75	1			H
ND1	2857	3831	975	12	ATG	TAA	H
tRNA-Ile	3844	3915	72	-1			H
tRNA-Gln	3915	3985	71	1			L
tRNA-Met	3987	4055	69	0			H
ND2	4056	5100	1045	0	ATG	T	H
tRNA-Trp	5101	5171	71	1			H
tRNA-Ala	5173	5241	69	1			L
tRNA-Asn	5243	5315	73	31			L
tRNA-Cys	5347	5413	67	-1			L
tRNA-Tyr	5413	5482	70	1			L
COI	5484	7034	1551	0	GTG	TAA	H
tRNA-Ser2	7035	7105	71	1			L
tRNA-Asp	7107	7178	72	9			H
COII	7188	7878	691	0	ATG	T	H
tRNA-Lys	7879	7954	76	1			H
ATP8	7956	8120	165	-7	ATG	TAG	H
ATP6	8114	8796	683	0	ATG	TA	H
COIII	8797	9580	784	0	ATG	T	H
tRNA-Gly	9581	9653	73	0			H
ND3	9654	10002	349	0	ATG	T	H
tRNA-Arg	10003	10072	70	0			H
ND4L	10073	10369	297	-7	ATG	TAA	H
ND4	10363	11743	1381	0	ATG	T	H
tRNA-His	11744	11812	69	0			H
tRNA-Ser1	11813	11880	68	1			H
tRNA-Leu1	11882	11954	73	3			H
ND5	11958	13781	1824	-4	ATG	TAA	H
ND6	13778	14299	522	0	ATG	TAA	L
tRNA-Glu	14300	14368	69	6			L
Cyt b	14375	15515	1141	0	ATG	T	H
tRNA-Thr	15516	15587	72	-1			H
tRNA-Pro	15587	15656	70	0			L
D-loop	15657	16996	1340	0			H

In the control region, eight tRNAs and ND6 genes were in the light chain (L chain) and the remaining 28 genes were in the heavy chain (H chain) (Table II). There were six gene overlaps and 13 gene gaps in the whole mitochondrial genome of *P. stoliczkana* (Fig. 1, Table II). The total length of the gene interval was 69 bp, with a maximum interval of 31 bp between tRNA-Asn and tRNA-Cys. The total length of gene overlap was 21 bp. Large overlaps were observed between ATP8 and ATP6, ND4L, and ND4. The base number of the overlap was 7 bp. The A+T content (59.7%) was higher than the G+C content (40.3%) in the mitochondrial genome of *P. stoliczkana*, revealing a preference for A+T and base anti-G bias. These results are consistent with the preference for A+T bases in vertebrates (Sun *et al*., 2020, 2022, 2023).

Table III. Nucleotide composition of protein-coding genes and rRNA in *Pethia stoliczkana.*

Gene	length (bp)	T (%)	C (%)	A (%)	G (%)	AT (%)	GC (%)	AT skew	GC skew
ATP6	683	29.6	26.1	32.5	11.9	62.1	38	0.047	-0.375
ATP8	165	27.9	26.1	35.8	10.3	63.7	36.4	0.124	-0.433
COI	1551	30.3	23.8	28.4	17.5	58.7	41.3	-0.032	-0.153
COII	691	27.1	24.6	33.6	14.8	60.7	39.4	0.107	-0.25
COIII	784	27.4	26.9	30.1	15.6	57.5	42.5	0.047	-0.267
Cyt b	1141	29.4	26.6	30.8	13.2	60.2	39.8	0.023	-0.336
ND1	975	27.4	26.2	31.9	14.6	59.3	40.8	0.076	-0.285
ND2	1045	24.5	28.4	35.3	11.8	59.8	40.2	0.181	-0.414
ND3	349	32.7	24.4	30.4	12.6	63.1	37	-0.036	-0.318
ND4	1381	27.5	27.2	30.9	14.4	58.4	41.6	0.058	-0.307
ND4L	297	27.9	27.9	27.6	16.5	55.5	44.4	-0.006	-0.258
ND5	1824	27.2	25.3	35	12.5	62.2	37.8	0.125	-0.339
ND6	522	43.1	11.5	15.9	29.5	59	41	-0.461	0.439
16S RNA	1683	21.5	22.5	36.8	19.3	58.3	41.8	0.262	-0.077
12S RNA	956	19.8	26.6	32.5	21.1	52.3	47.7	0.244	-0.114

PCGs

The total length of the 13 PCGs in the mitochondrial genome of *P. stoliczkana* was 11,408 bp. Except for ND6, which is in the L chain, all genes were in the H chain. Among the 13PCGs, the start codon of the COI gene was GTG, and the remaining start codons were ATG. Deletion of the termination codon is typically thought to be caused by polyadenylation. We found that the termination codons of the ND2, CO II, ATP6, CO III, ND3, ND4, and Cyt b genes in the mitochondrial genome of *P. stoliczkana* had the incomplete codons T or TA (Table II), which is

common in the mitochondrial genomes of metazoa and similar to the termination codons of most mitochondrial PCGs in teleost fish. The uneven distribution of bases is one of the most characteristic features of coding regions. Although the base contents of the different gene fragments differed, they all presented a lower G content and higher A+T enrichment (Table III).

Codon usage and amino acid composition

The relative synonymous codon usage of the *P. stoliczkana* mitochondrial genome was analyzed using MEGA to determine the ratio of the expected frequency of amino acids using synonymous codons to their observed frequency (Table IV, Fig. 2). There were 25 preferred codons (relative synonymous codon usage ≥1) (Behura and Severson, 2013) in the 13 PCGs of *P. stoliczkana*. The 11,408-bp gene sequence encoded 3794 amino acids. The most common amino acid in the mitochondrial genome of *P. stoliczkana* was leucine (Leu), with a content of (11.12%), whereas the least used amino acid was cysteine (Cys), with a content of only 0.66%.

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

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	128	1.09	UCU(S)	34	0.89	UAU(Y)	56	0.96	UGU(C)	9	0.72
UUC(F)	106	0.91	UCC(S)	42	1.1	UAC(Y)	61	1.04	UGC(C)	16	1.28
UUA(L)	184	1.8	UCA(S)	105	2.74	UAA(*)	5	3.33	UGA(W)	105	1.78
UUG(L)	7	0.07	UCG(S)	3	0.08	UAG(*)	1	0.67	UGG(W)	13	0.22
CUU(L)	86	0.84	CCU(P)	22	0.43	CAU(H)	34	0.64	CGU(R)	9	0.49
CUC(L)	63	0.62	CCC(P)	32	0.62	CAC(H)	73	1.36	CGC(R)	4	0.22
CUA(L)	239	2.34	CCA(P)	144	2.81	CAA(Q)	96	1.88	CGA(R)	53	2.86
CUG(L)	34	0.33	CCG(P)	7	0.14	CAG(Q)	6	0.12	CGG(R)	8	0.43
AUU(I)	205	1.41	ACU(T)	52	0.66	AAU(N)	47	0.79	AGU(S)	10	0.26
AUC(I)	86	0.59	ACC(T)	101	1.28	AAC(N)	72	1.21	AGC(S)	36	0.94
AUA(M)	154	1.61	ACA(T)	157	1.99	AAA(K)	83	1.95	AGA(*)	0	0
AUG(M)	37	0.39	ACG(T)	5	0.06	AAG(K)	2	0.05	AGG(*)	0	0
GUU(V)	52	0.98	GCU(A)	58	0.69	GAU(D)	21	0.56	GGU(G)	29	0.48
GUC(V)	34	0.64	GCC(A)	127	1.52	GAC(D)	54	1.44	GGC(G)	36	0.6
GUA(V)	111	2.08	GCA(A)	142	1.7	GAA(E)	91	1.75	GGA(G)	125	2.07
GUG(V)	16	0.3	GCG(A)	8	0.1	GAG(E)	13	0.25	GGG(G)	51	0.85

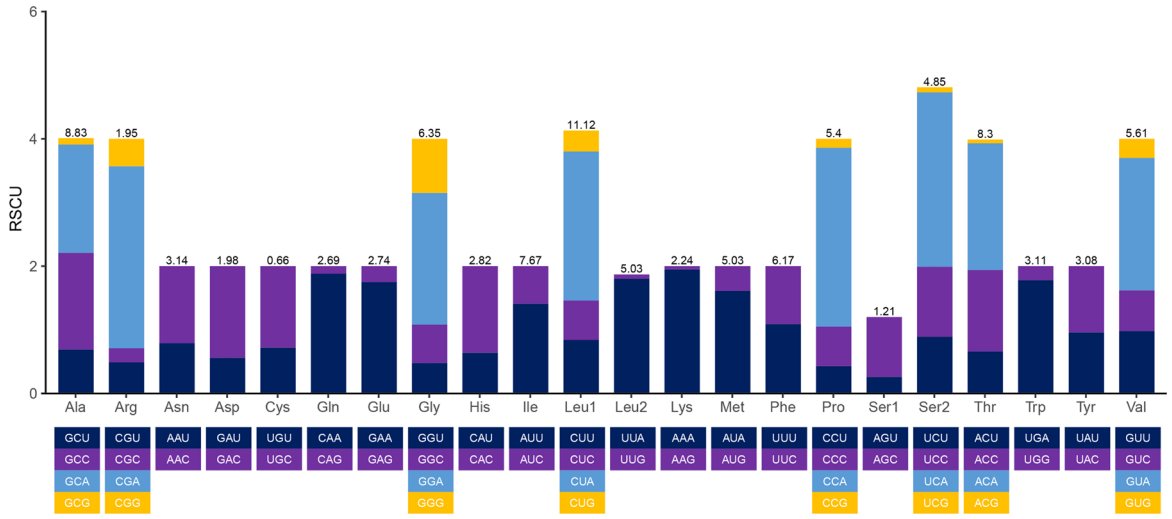

Fig. 2. Relative synonymous codon usage in mitochondrial protein-coding genes of *Pethia stoliczkana.*

rRNA, tRNA, and control region

Similar to those in common bony fish, the mitochondrial genome of *P. stoliczkana* contained 12S rRNA and 16S rRNA, which were between tRNA-Phe and tRNA-Leu2 on the H chain and separated from each other by tRNA-Val. The 12S rRNA sequence was 956 bp in length, its position in the mitochondrial sequence was 69–1025 bp, the length of the 16S rRNA sequence was 1683 bp, and its position in the mitochondrial sequence was 1098–2780 bp. The mitochondrial genome of *P. stoliczkana* was found to contain 22 tRNAs with a length of 67–76 bp. The 1340-bp control region was between tRNA-Pro and tRNA-Phe.

Phylogenetic relationships

ML and BI phylogenetic trees of Smiliogastrinae were constructed based on the nucleotide tandem sequences of the 13 PCGs. The two tree-building methods generated consistent topological structures (Figs. 3 and 4). *Pethia stoliczkana* and *Pethia ticto* were clustered into one branch together with *P. padamya*, with confidence values of 100%. Except for the genus *Puntius*, all genera were clustered into one branch with a high confidence value.

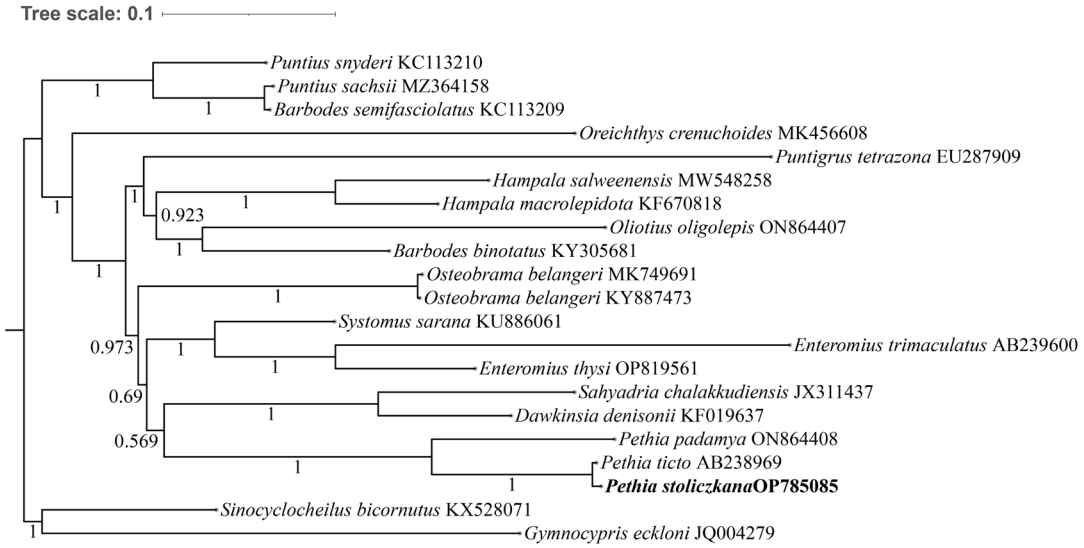

Fig. 3. Bayesian inference phylogenetic trees based on the nucleotide datasets of 13 protein-coding genes from the mitogenomes of 21 fishes. The numbers along the branches indicate the Bayesian posterior probability values.

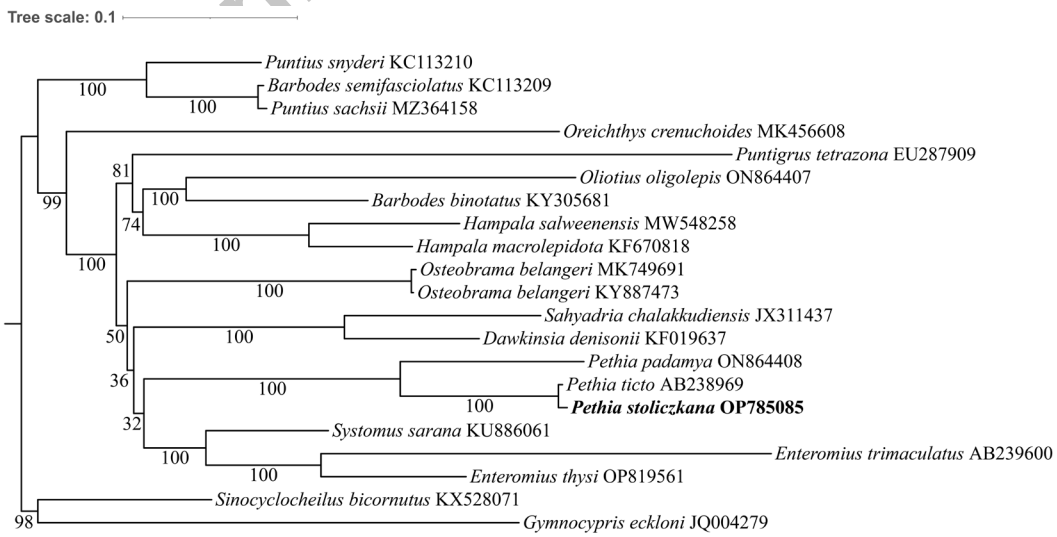

Fig. 4. Maximum likelihood phylogenetic trees based on the nucleotide datasets of 13 protein-coding genes from the mitogenomes of 21 fishes. The numbers along the branches indicate the ML bootstrap values.

DISCUSSION

With advancements in DNA sequencing technology and the rapid development of bioinformatics, fish mitochondrial genomes have been widely studied in the fields of fish germplasm protection, species identification, population polymorphism, and phylogenetic development. Previous studies showed that the mitochondrial genome of fish is typically 15–20 kb, often has a double-stranded closed circular structure, is closely arranged, and has a low molecular weight. The mitochondrial genomes of different species vary widely and contain tandem repeats, base insertions, and deletions (Peng *et al*., 2006). Each PCG has a different evolution rate. Zardoya and Meyer (1996) divided the evolution rates of 13 PCGs into good, medium, and poor groups, in which COI, ND2, ND4, Cytb, and ND5 genes were good, and COII, COIII, ND1, and ND6 were medium. ATP6, ATP8, ND3, and ND4L levels were poor. The evolution rate of most PCGs was between that in the control region and RNA, showing a moderate evolution rate. *Pethia stoliczkana* genes, such as CO I, Cyt b, and ND, which exhibit a rapid evolution rate, can be used as molecular markers to distinguish these fish from other Smiliogastrinae fishes and provide a reference for their germplasm resource protection. However, the 16S rRNA sequence in the mitochondrial genome is not a PCG and is not affected by codon selection pressure. Most mutations were neutral. In addition, the evolution rate of mtDNA is significantly higher than that of nuclear DNA. Therefore, the homology of mitochondrial 16S rRNA sequences can be compared to study phylogenetic relationships between species.

The system information contained in a single gene is too small to reflect the entire level of biological molecular evolution; thus, the results obtained by analyzing gene sequences encoded by multiple genomes are more reliable. In fish, the whole mitochondrial genome is widely used to study phylogenetic relationships at different stages. This study provides a basis for germplasm identification, phylogenetic evolution analysis, genetic diversity evaluation, and utilization of *P. stoliczkana*.

CONCLUSION

The whole mitochondrial genome of *P. stoliczkana* was obtained using second-generation sequencing. The arrangement pattern of genes in the mitochondrial genome was the same as that of *P. ticto* and *P. padamya* and was consistent with the ancestral pattern. Phylogenetic analysis supports the monophyly of the genus *Pethia*.

ACKNOWLEDGEMENTS

This research was funded by the Doctoral Research Fund of Mudanjiang Normal University [grant number 1002319042]. We would like to thank Editage (www.editage.cn) for English language editing.

Funding

This research was funded by the Doctoral Research Fund of Mudanjiang Normal University [grant number 1002319042].

IRB approval

All experiments were approved by the Animal Ethics Committee of the Mudanjiang Normal University and conducted in compliance with relevant animal welfare and protection laws.

Ethical statement

All experiments were conducted in accordance with Chinese laws.

Statement of conflict of interest

The authors have declared no conflict of interest.

REFERENCES

Atkore, V.M., Knight, J.M., Devi, K.R. and Krishnaswamy, J., 2015. A new species of Pethia from the western Ghats, India (Teleostei: Cyprinidae). *Copeia*, **103**: 290-296. https://doi.org/10.1643/OT-12-172

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. and Saunders, N.C., 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst*., **18**: 489-522. https://doi.org/10.1146/annurev.es.18.110187.002421

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A., 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol*., **19**: 455-477. https://doi.org/10.1089/cmb.2012.0021

Behura, S.K. and Severson, D.W., 2013. Codon usage bias: Causative factors, quantification methods and genome-wide patterns with emphasis on insect genomes. *Biol. Rev*., **88**: 49-61. https://doi.org/10.1111/j.1469-185X.2012.00242.x

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M. and Stadler, P.F., 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol*., **69**: 313-319. https://doi.org/10.1016/j.ympev.2012.08.023

Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D. and Pirovano, W., 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, **27**: 578-579. https://doi.org/10.1093/bioinformatics/btq683

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol*., **17**: 540-552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Funk, D.J. and Omland, K.E., 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annls Rev. Ecol. Evol. Syst*., **34**: 397-423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. and Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol*., **59**: 307-321. https://doi.org/10.1093/sysbio/syq010

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A. and Jermiin, L.S., 2017. Model finder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods*, **14**: 587-589. https://doi.org/10.1038/nmeth.4285

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**: 2947-2948. https://doi.org/10.1093/bioinformatics/btm404

Lefort, V., Longueville, J.E. and Gascuel, O., 2017. SMS: smart model selection in PhyML. *Mol. Biol. Evol*., **34**: 2422-2424. https://doi.org/10.1093/molbev/msx149

Lowe, T.M. and Chan, P., 2016. tRNAscan-SE On-line: Integrating search and context for analysis of transfer RNA genes. *Nucl. Acids Res*., **44**: W54-W57. https://doi.org/10.1093/nar/gkw413

Nath, D., Dutta, R., Sarmah, R., Ahmed, A.M., Pokhrel, H., Mudoi, L.P. and Bhagabati, S.K., 2022. First record of a barb, *Pethia stoliczkana* (Day) from Brahmaputra drainage, Assam, India. *J. appl. Ichthyol*., **38**: 609-614. https://doi.org/10.1111/jai.14363

Patel, R.K. and Jain, M., 2012. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS One*, **7**: e30619. https://doi.org/10.1371/journal.pone.0030619

Peng, Z., Wang, J. and He, S., 2006. The complete mitochondrial genome of the helmet catfish Cranoglanis bouderius (Siluriformes: Cranoglanididae) and the phylogeny of otophysan fishes. *Gene*, **376**: 290-297. https://doi.org/10.1016/j.gene.2006.04.014

Ronquist, F. and Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**: 1572-1574. https://doi.org/10.1093/bioinformatics/btg180

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., 2022. Comparative analysis of the mitogenomes of two Corydoras (Siluriformes, Loricarioidei) with nine known Corydoras, and a phylogenetic analysis of Loricarioidei. *Zookeys*, **1083**: 89–107. https://doi.org/10.3897/zookeys.1083.76887

Sun, C.H., Liu, H.Y. and Lu, C.H., 2020. Five new mitogenomes of Phylloscopus (Passeriformes, Phylloscopidae): Sequence, structure, and phylogenetic analyses. *Int. J. Biol. Macromol*., **146**: 638-647. https://doi.org/10.1016/j.ijbiomac.2019.12.253

Sun, C.H., Sun, P.Y., Lao, Y.L., Wu, T., Zhang, Y.N., Huang, Q. and Zhang, Q., 2023. Mitogenome of a monotypic genus, Oliotius Kottelat, 2013 (Cypriniformes: Cyprinidae): Genomic characterization and phylogenetic position. *Gene*, **851**: 147035. https://doi.org/10.1016/j.gene.2022.147035

Tamura, K., Stecher, G. and Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol*., **38**: 3022-3027. https://doi.org/10.1093/molbev/msab120

Vaidya, G., Lohman, D.J. and Meier, R., 2011. Sequence matrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**: 171-180. https://doi.org/10.1111/j.1096-0031.2010.00329.x

Wolstenholme, D.R., 1992. Animal mitochondrial DNA: Structure and evolution. *Int. Rev. Cytol*., **141**: 173-216. https://doi.org/10.1016/S0074-7696(08)62066-5

Zardoya, R. and Meyer, A., 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol. Biol. Evol*., **13**: 933-942. https://doi.org/10.1093/oxfordjournals.molbev.a025661