

POPULATION GENETIC DIVERSITY OF GOLDFISH
(*CARASSIUS AURATUS*) FROM ZHEJIANG RIVER BASIN
BASED ON MITOCHONDRIAL CYTOCHROME OXIDASE
SUBUNIT I (*COI*) GENE SEQUENCE

Jindian Liu¹, Xiaojun Xu¹, Quanfu Luo², Lianrong Wang³,
Pengcheng Sheng⁴, Hongjian Song¹, Shuai Liu¹, Bao Lou¹✉

Received on July 22, 2024

Presented by H. Najdenski, Member of BAS, on November 26, 2024

Abstract

Carassius auratus is an indigenous and economically important species in China. The genetic diversity of three *C. auratus* populations from Tiaoxi River (TXR), Oujiang River (OJR), and Qiantangjiang (QTJ) River basin, were studied. The genetic diversity of the three populations was evaluated by mitochondrial cytochrome oxidase subunit I (*COI*) sequences. Totally, 225 individuals were sampled from three populations. After the sequencing of 541 bp fragments of *COI* gene, 536 variable sites were found and no insertion or deletion was detected. The Hap_38 (haplotype 38) was the dominant haplotype shared by 225 individuals (11.11% of all individuals) and existed in all localities. The haplotype diversity (Hd) values ranged from 0.866 to 0.931, while the values of nucleotide diversity (Pi) varied from 0.388 to 0.549. All three populations showed high genetic diversity. The pair-wise fixation index (F_{ST}) value between TXR and QTJ populations was 0.1409, which indicated that they were in middle genetic differentiation ($0.05 < F_{ST} < 0.15$), while the F_{ST} value between TXR and OJR populations, and between OJR and QTJ populations were 0.1816 and 0.1775, respectively, which indicated that they were in high genetic differentiation ($F_{ST} > 0.15$). Based on the genetic structure, TXR and OJR populations were in the same cluster separated from

The study was supported by the Zhejiang Provincial Natural Science Foundation (LTGS23C040001).

<https://doi.org/10.7546/CRABS.2025.02.04>

QTJ population. In summary, by mitochondrial *COI* gene analysis, all three *C. auratus* populations had high genetic diversity. Populations TXR and OJR had a closer genetic distance, and they had a large genetic distance with QTJ population. This study provides a particularly important basis for *C. auratus* resource conservation and utilization.

Key words: *Carassius auratus*, mitochondrial *COI* gene, genetic diversity, population structure

Introduction. As an important part of biodiversity, genetic diversity is the basis of ecological diversity and species diversity. A better understanding of genetic diversity and population genetic structure is critical and can provide clues for the management and conservation of economically important species. It can also guide how to conserve resources and reproduce important species. So far, genetic analysis of genetic variation and population structure of the *C. auratus* population has been analyzed by mtDNA, D-loop gene, mitochondrial whole genome sequencing, and other molecular markers [1]. Most of these studies focused on individual *C. auratus* collected from some areas such as the Pingxiang district of Jiangxi Province, and Huaihe River Basin in China [1]. It is important to systematically study the genetic diversity and population structure of *C. auratus* for its germplasm resources protection and exploration [2]. Mitochondrial DNA (mtDNA) marker analysis is very effective in the assessment of genetic diversity and population structure of aquatic and economic animals, including *Portunus sanguinolentus*, *Nibea albiflora*, Chinese mitten crab (*E. sinensis*), silver carp, sheep (*Ovis aries*). Also mtDNA has certain advantages, including simple structure, maternal inheritance, rapid evolutionary speed, and almost no recombination [3].

At present, only one method, mtDNA gene sequence, is used to study the genetic diversity and population structure of *C. auratus*. To better elucidate the genetic relationship between different river populations and give full play to each other's advantages, mtDNA gene sequence analysis was used to infer the origin and evolutionary history of *C. auratus* in Zhejiang river basin. The results can guide the utilization of germplasm resources and the formulation of breeding strategies. In contrast, sampling and genetic analysis of larger *C. auratus* populations have been less reported. According to the existing studies, the genetic differentiation level of *C. auratus* in the Huaihe River Basin in China is low, and the genetic distance is small, and natural and artificial factors may play an important role in the genetic structure of the *C. auratus* population [4]. In addition, it is worth noting that overexploitation, habitat destruction, water pollution, and artificial breeding activities may greatly affect the natural resources of this species, the overall situation of genetic diversity and population structure of *C. auratus* may change in the future.

In this study, the mitochondrial cytochrome oxidase subunit I (*COI*) gene was used to analyze the genetic diversity and population structure of different goldfish

populations collected from the Qiantangjiang (QTJ), Oujiang River (OJR) and Tiaoxi River (TXR) basins. The research results can be used to guide the sustainable development and conservation of natural germplasm resources of goldfish and contribute to the conservation and management of wild resources of this species.

Materials and methods. Sample collection and DNA extraction. A total of 225 wild individuals of goldfish were collected by fishermen from three different rivers in the Tiaoxi River (Huzhou City, Zhejiang Province), Oujiang River (Qingtian City, Zhejiang Province), and Qiantangjiang (Hangzhou City, Zhejiang Province) River basins in Zhejiang Province. The geographic locations of goldfish populations are shown in Fig. 1. Out of the total 225, 50 individuals were collected in the Tiaoxi River, 75 individuals in the Oujiang River, and 100 individuals in the Qiantangjiang. A certain amount of muscle from each fish was taken, and stored in ethanol (99%) at -20°C . Total DNA was isolated from muscle using the DNA extraction Kit (Nuobai Biotec, SK8224, Shanghai, China) according to the instruction. The DNA was adjusted to 100 ng/ μl concentration and stored at -20°C .

COI gene amplification and sequencing. A pair of primers (Fish-COI-F: TCAACCAACCACAAAGACATTGGCAC and Fish-COI-R: TAGACTTCTGG GTGGCCAAAGAATCA) were used for *COI* gene amplification [5]. Polymerase



Fig. 1. Sampling locations of *C. auratus* populations. The sampling locations of rivers are Tiaoxi River (TXR), Oujiang River (OJR), and Qiantangjiang (QTJ), respectively

chain reaction (PCR) was performed on a PCR apparatus (Verity 96 well) in 50 μ l total volume (including Fish-COI-F 1 μ l, Fish-COI-R 1 μ l, dNTP 10 mM 1 μ l, Taq Buffer 5 μ l, 25 mM MgCl₂ 5 μ l, 5 U/ μ l Taq polymerase 0.5 μ l, mtDNA template 1 μ l and ddH₂O 35.5 μ l) under the following conditions: one cycle of denaturation at 95 °C for 3 min; 35 cycles of 30 s at 95 °C, 30 s at 55 °C, and 50 s at 72 °C. At last, the products were extended for 7 min at 72 °C. The PCR products were separated and recovered on 1.0% agarose gels. Then they were purified by the PCR product purification and recovery kit (GENEWIZ, SK1141, Suzhou, China). Then the products were sequenced by the dideoxy chain termination method using the ABI3770 automatic DNA sequencer (PE Company).

Mitochondrial COI gene sequence analysis. Mitochondrial *COI* genes were sequenced through one-way sequencing. Mitochondrial COI diversity parameters including polymorphic site (s), haplotypes (h), nucleotide diversity (Pi), haplotype diversity (hd), and the average number of nucleotide differences (k) were calculated using DNASP software [6]. The multiple sequence alignments, phylogenetic analysis, G/C contents, parsimonious sites, variable sites, and mean genetic distances were performed using EditSeq and MEGA software [7]. A median-joining network of haplotypes was constructed using NETWORK version 4.613 [8]. The fixation index (F_{ST}) value was carried out to assess the genetic differentiation among different geographical populations. The evolutionary history was inferred using the Neighbor-Joining method, and the evolutionary distances were computed using the Kimura two-parameter method and are in units of the number of base substitutions per site. The NJ phylogenetic tree of *COI* gene sequences of goldfish was conducted in MEGA11 [9].

Results. Genetic diversity analysis of COI gene sequences. The three goldfish population genetic diversity analysis results of *COI* gene are listed in Table 1. A total of 536 variation sites were detected in 225 *COI* sequences of goldfish. Compared with the proportion of haplotypes, the highest haplotype diversity was 0.931 ± 0.021 in TXR population, and the lowest was 0.866 ± 0.032 in OJR population. All sequences in the three populations were continuously encoded, and no base insertion or deletion was found. The Tajima's D values of the three populations were all negative with no significant difference, indicating that the sequence evolution pattern of the three populations may have experienced population expansion, and the target sequence followed the neutral model of evolution.

COI sequence haplotype analysis and population historical dynamics analysis. A total of 75 haplotypes were identified by genetic diversity analysis. Among them, only three haplotypes (Hap1, Hap2, and Hap3) were found in three populations, and two haplotypes (Hap8 and Hap9) were found in two populations. The median-joining network shows 75 haplotype relationships (Fig. 2).

Analysis of *C. auratus* population evolutionary relationship. *COI* gene sequences were used to construct evolutionary trees to infer population evo-

T a b l e 1

Statistical parameters of mtDNA *COI* gene in the three regions of *C. auratus*

Population	TXR	OJR	QTJ	Total
Number of samples	50	75	100	225
Sampling site	Huzhou	Qingtian	Hangzhou	
Sampling time	May 2020	May 2020	May 2020	
Sex ratio (F/M)	3.55	2.57	3.76	
Body weight (g)	18.2–368.0	16.4–398.5	17.2–302.7	
Body length (cm)	10.2–23.0	10.1–23.0	10.3–22.5	
Number of polymorphic sites (s)	509	469	535	536
Number of haplotypes (h)	25	20	38	75
Haplotype (gene) diversity (Hd)	0.931	0.866	0.923	0.959
SD of haplotype (gene) diversity	0.021	0.032	0.018	0.006
Nucleotide diversity (Pi)	0.388	0.477	0.549	0.552
Average number of nucleotide differences (k)	210.2	258.39	297.27	298.82
InDel sites analyzed	0	0	0	0
Tajima's D	0.1817	0.889	1.521	1.196
Fu's Fs	0.683	1.818**	1.933**	2.319
F_{ST} (TXR vs OJR)			0.1816	
F_{ST} (TXR vs QJR)			0.1409	
F_{ST} (QTJ vs OJR)			0.1775	

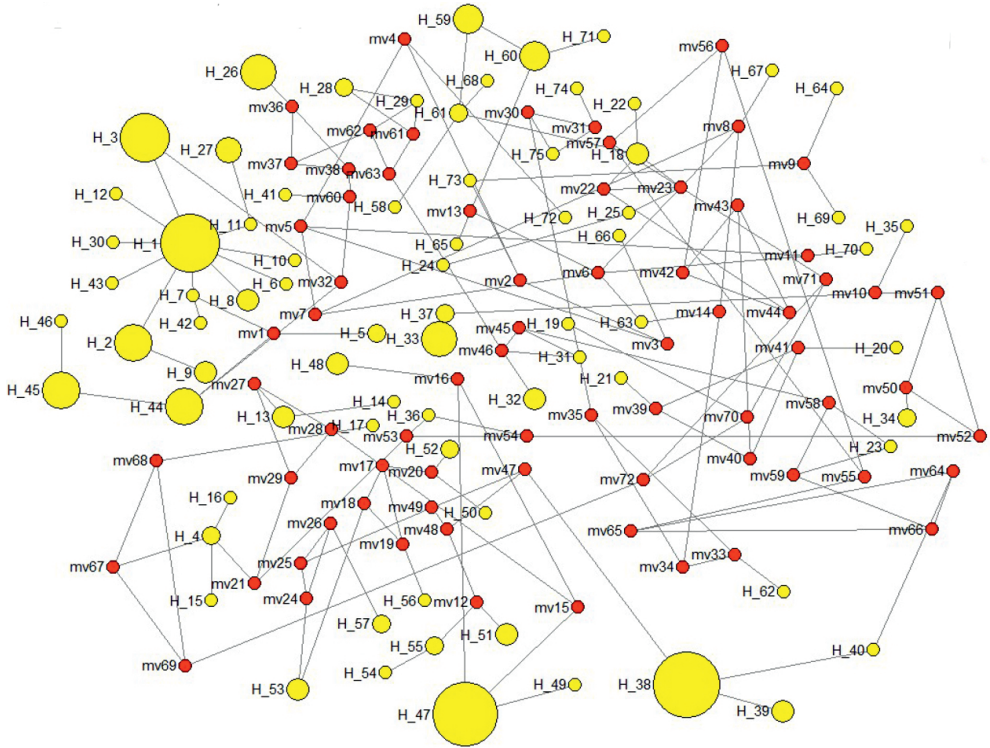


Fig. 2. The median-joining network of 75 haplotypes of mitochondrial *COI* gene sequences of *C. auratus*

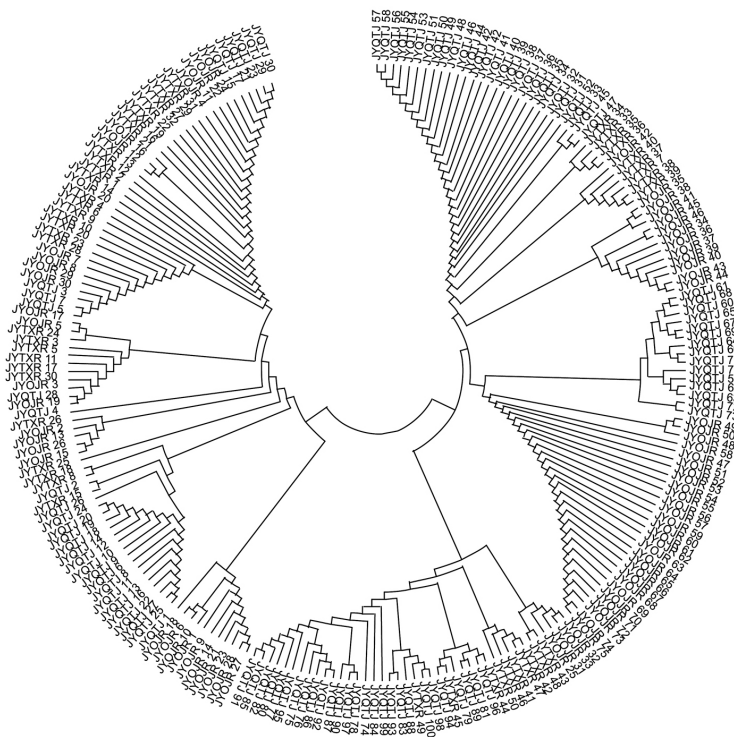


Fig. 3. The NJ phylogenetic tree of *COI* gene sequences of *C. auratus*.
MEGA 11 was used for the phylogenetic tree construction

lutionary history. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. This analysis involved 225 nucleotide sequences. The optimal tree is shown in Fig. 3. In TXR population, JYTXR8 and JYTXR18 individuals are clearly in the independent clade. Most individuals in QTJ population are in the same clade as JYOJR6, JYOJR10, and JYOJR18 individuals in OJR population. The number of individual clades in TXR population and OJR population is significantly higher than that in QTJ population.

Discussion. In this study, the lowest genetic diversity of OJR population was 0.866, which was lower than that of TXR and QTJ populations. This may be due to dam construction, overfishing, water pollution, and habitat destruction as factors contributing to the degradation of the genetic diversity of the *C. auratus* population [10]. However, the genetic diversity results of mitochondrial *COI* markers are inconsistent, which may be explained by the fact that mitochondrial markers belong to cytoplasmic traits [11]. GRANT and BOWEN [12] proposed the use of h and π to estimate the evolutionary history of populations. When h is more than 0.5, π is less than 0.5%, which indicates that the population expands rapidly after the bottleneck effect [12]. In this study, the three populations were

all of high haplotype diversity, possibly because the bottleneck of population size would also produce waves similar to those produced by sudden expansion.

Genetic differentiation is represented by F_{ST} coefficients, and paired F_{ST} values were calculated to distinguish the level of differentiation between populations. In this study, the F_{ST} values were 0.1816 (between TXR and OJR), 0.1409 (between TXR and QTJ), and 0.1775 (between QTJ and OJR), respectively, suggesting that TXR and OJR were in a state of lower genetic differentiation, while TXR and QTJ, QTJ and OJR were in a state of highest genetic differentiation. This is similar to the research on population differentiation of other species [13]. JIN [14] evaluated the genetic variation of six populations of mitten crabs in the Liaohe River, Yellow River, Yangtze River, Oujiang River, Pearl River, and Nanlijiang River by using RAPD technology. According to the UPGMA tree map of the populations in the six regions, the populations in southern China were significantly different from those in the north. WANG et al. [15] confirmed that the puffer fish was different from the Chinese puffer fish by using the *COI* gene, and the populations of the central region (Oujiang and MingJiang Rivers) had similar mixing with those of the Yangtze River.

The median connection network of 75 haplotypes suggests that the most widespread haplotype is Hap_38 (Fig. 2), while a clear answer to which haplotype is ancestral would require the use of a larger sample size. Goldfish has a high spreading ability during its growth and gene flow is a natural phenomenon. Gene flow can also be induced through aquaculture. Therefore, we speculate that goldfish are more likely to lose their environmental adaptability if they lose certain alleles, resulting in the death of the mutant or failure to reproduce successfully. Therefore, goldfish also have strong genetic protection.

TAJIMA's [16] and FU's [17] methods were performed for genetic diversity and population structure in this study. TXR, OJR, and QTJ populations are both negative values, indicating that there may have been population changes in the past, especially consistent with population expansion. Fu's FS test, which is devised specifically to detect population expansion and is more sensitive to the presence of singletons in the samples, showed highly significant values in the overall population. OJR, TXR and QTJ populations are all positive, QTJ and OJR population is significantly different from other populations ($P < 0.05$). In all, Tajima's and Fu's analyses indicated that OJR and QTJ *C. auratus* populations might have experienced expansion events in their history.

Conclusion. In conclusion, mitochondrial *COI* gene markers have explored the genetic diversity and population structure of the three river basin populations, which is an important prerequisite for the development and utilization of germplasm resources. Here, we reveal the genetic background of goldfish in TXR, OJR, and QTJ populations. We hope that this study will contribute to further study on the genetic selection of goldfish and make more contributions to the aquaculture industry. At the same time, this study will provide genetic infor-

mation for the further protection of goldfish germplasm resources along the river basin.

REFERENCES

- [1] SHENG J. Q., J. H. WANG, S. H. HE et al. (2014) Complete mitochondrial genome of a natural triploid crucian carp mutant, *Carassius auratus* var. *pingxiangnensis*, and phylogenetic analysis of different ploidies in crucian carp, *Genet. Mol. Res.*, **13**(3), 5849–5864.
- [2] ADDAI F. P., Y. LIU, J. WU et al. (2024) High-performance petal-like nickel organic framework for efficient dye sequestration, *J. Water Process Eng.*, **57**, 104622.
- [3] PUERTAS M. J., M. GONZÁLEZ-SÁNCHEZ (2020) Insertions of mitochondrial DNA into the nucleus-effects and role in cell evolution, *Genome*, **63**(8), 365–374.
- [4] CHEN W., Y. ZHAO, C. YANG (2016) Genetic variation and population history of three *Carassius auratus* populations in Huaihe River, China, *Mitochondrial DNA A DNA Mapp. Seq. Anal.*, **27**(6), 4252–4255.
- [5] WARD R. D., T. S. ZEMLAK, B. H. INNES et al. (2005) DNA barcoding Australia's fish species, *Philos. Trans. R. Soc. B: Biol. Sci.*, **360**(1462), 1847–1857.
- [6] LIBRADO P., J. ROZAS (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, *Bioinformatics*, **25**(11), 1451–1452.
- [7] TAMURA K., D. PETERSON, N. PETERSON et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.*, **28**(10), 2731–2739.
- [8] BANDELT H. J., P. FORSTER, A. RÖHL (1999) Median-joining networks for inferring intraspecific phylogenies, *Mol. Biol. Evol.*, **16**(1), 37–48.
- [9] TAMURA K., G. STECHER, S. KUMAR (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11, *Mol. Biol. Evol.*, **38**(7), 3022–3027.
- [10] TANG B., K. ZHOU, D. SONG et al. (2003) Molecular systematics of the Asian mitten crabs, genus *Eriocheir* (Crustacea: Brachyura), *Mol. Phylogenetics Evol.*, **29**(2), 309–316.
- [11] TAANMAN J. W. (1999) The mitochondrial genome: structure, transcription, translation and replication, *Biochim. Biophys. Acta*, **1410**(2), 103–123.
- [12] GRANT W. S., B. W. BOWEN (1998) Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation, *J. Heredity*, **89**(5), 415–426.
- [13] REN G., G. MIAO, C. MA et al. (2018) Genetic structure and historical demography of the blue swimming crab (*Portunus pelagicus*) from southeastern sea of China based on mitochondrial COI gene, *Mitochondrial DNA A DNA Mapp. Seq. Anal.*, **29**(2), 192–198.
- [14] JIN Z. (1999) Phylogenesis of populations of mitten crabs (*Eriocheir sinensis*, *E. japonicus*) in six rivers of mainland China: Biochemical genetic difference analysis, *Journal of Fisheries of China*, **23**, 325–330.
- [15] WANG C., C. LI, S. LI (2008) Mitochondrial DNA-inferred population structure and demographic history of the mitten crab (*Eriocheir sensu stricto*) found along the coast of mainland China, *Mol. Ecol.*, **17**(15), 3515–3527.

- [16] TAJIMA F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, *Genetics*, **123**(3), 585–595.
- [17] FU Y. X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection, *Genetics*, **147**(2), 915–925.

¹*State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Institute of Hydrobiology, Zhejiang Academy of Agricultural Sciences, Hangzhou 310000, China*

e-mails: liujd@zaas.ac.cn, xuxj@zaas.ac.cn, songhongjian2006@163.com, zlpshuxue2013@163.com, loubao6577@163.com

²*Wetland Protection Service Center, Reservoir Management Service Center, Hangzhou 310000, China*

e-mail: Hh11uo@163.com

³*Management Division of Provincial Natural Reserve for Asian Giant Soft-shelled Turtle, Lishui 323900, China*

e-mail: qtwlr@126.com

⁴*Zhejiang Institute of Freshwater Fisheries, Huzhou 313001, China*

e-mail: 499356345@qq.com