Reproduction and Breeding 4 (2024) 185–193

Contents lists available at ScienceDirect

Reproduction and Breeding

journal homepage: www.keaipublishing.com/en/journals/reproduction-and-breeding

Population genetic characteristics of two crucian carp varieties derived from distant hybridization

Wenjie Luo [a,1], Xuexue Huang [a,1], Xiaowei Xu [a], Chenghua Dai [a], Qiong Liu [a], Yating Zhu [a], Duansheng Wu [b], Shi Wang [a,c], Qingfeng Liu [a,c], Conghui Yang [a,c,*]

[a] State Key Laboratory of Developmental Biology of Freshwater Fish, College of Life Sciences, Hunan Normal University, Changsha, 410081, China
[b] Department of Laboratory Animal Science, University of South China, Hengyang, 421001, China
[c] Hunan Yuelu Mountain Scienece and Technology Co.Ltd. for Aquatic Breeding, Changsha, 410081, China

ARTICLE INFO

Keywords:
Crucian carp
Distant hybridization
transferrin
Genetic structure
Genetic diversity

ABSTRACT

The Hefang crucian carp (HFJ) derived from *Carassius cuvieri* (♀) × *C. auratus* red var. (♂) and the crucian carp-like homodiploid fish (NCRC) derived from *Cyprinus carpio* (♀) × *Megalobrama amblycephala* (♂) are important germplasm resources of crucian carp. To investigate the population characteristics and reveal the current genetic situations of these hybrid varieties, we sequenced one nuclear molecular marker *transferrin* (*Tf*) among six *Carassius* populations, which included two hybrid varieties (HFJ and NCRC), one laboratory variety (red crucian carp, RCC) and three wild populations. A total of 166 distinct *Tf* alleles (A1~A166) were identified, with sequence lengths ranging from 1145 bp to 1244 bp. Unexpectedly, the nucleotide diversity of *Tf* alleles in cultured populations (HFJ, NCRC, and RCC) exceeded that of wild populations. The phylogenetic reconstruction analysis results showed that HFJ and *C. cuvieri* were closely related, but they could not be separated from *C. auratus*. Besides, while most *Tf* alleles of NCRC and other *Carassius* populations were intermixed in five clades, two *Tf* alleles, together with *Cy. carpio*, formed a distinct monophyletic clade. Moreover, network and genetic structure analyses failed to distinguish between wild and cultured populations. Although AMOVA analysis based on *Tf* alleles indicated genetic differentiation among all six populations, the majority of the genetic variation (92.00 %) was observed within populations. Furthermore, the population dynamics analysis suggested that these populations have maintained relative stability in recent history. Therefore, the elevated nucleotide diversity of *Tf* alleles in cultured populations, the ambiguous population structure, and the limited genetic divergence among the six *Carassius* populations highlighted the hybrid genetic characteristics of HFJ and NCRC. These findings, based on *Tf* alleles, differ from the results obtained using mitochondrial gene markers, providing an alternative perspective for germplasm evaluation of hybrid varieties and crucian carp breeding.

1. Introduction

Crucian carp (*Carassius auratus*), commonly known as the goldfish, is a prevalent freshwater fish species in China, playing a pivotal role in aquaculture attributing to the appealing taste and high nutritional value of its meat. Cultivation of crucian carp relies heavily on superior varieties developed through artificial breeding, such as the Pengze crucian carp and Fangzheng crucian carp resulting from selective breeding of local breeds, and the Xiangyun crucian carp and Hybrids of Gold crucian carp produced through distant hybridization [1,2]. The aquaculture industry continually introduces new varieties, including the Hefang crucian carp (HFJ) obtained through the distant crossing of female Japanese white crucian carp (*C. cuvieri*) × male red crucian carp (*C. auratus* var. red), known for its rapid growth, strong resistance, tender meat, and high nutrition content [3–5]. Another recent addition, the crucian carp-like homodiploid fish (NCRC), generated from interspecies hybridization between female common carp (*Cyprinus carpio*) and male blunt snout bream (*Megalobrama amblycephala*), closely resembles crucian carp in appearance without the beard, featuring desirable traits like a small head and a high back [6,7]. While HFJ and NCRC

* Corresponding author. State Key Laboratory of Developmental Biology of Freshwater Fish, Engineering Research Center of Polyploid Fish Reproduction and Breeding of the State Education Ministry, College of Life Sciences, Hunan Normal University, Changsha, 410081, Hunan, China.
E-mail address: yangch@hunnu.edu.cn (C. Yang).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.repbre.2024.07.002
Received 25 February 2024; Received in revised form 4 June 2024; Accepted 18 July 2024
Available online 7 August 2024

serve as promising candidates for crucian carp breeding, existing research primarily focuses on improving their economic traits and genetic analysis [3–7], neglecting the crucial aspect of population genetics which is essential for the further development and enhancement of these new varieties.

Current studies have identified three *Carassius* species, *C. carassius* (distributed in Europe and the Irtysh River in China westwards), *C. cuvieri* (native to Japan), and the *C. auratus* complex (widespread across Eurasia and neighboring islands) [8,9]. Within the *C. auratus* complex, various subspecies exist due to its extensive geographical distribution and significant phenotypic diversity, including the gynogenetic triploid silver crucian carp (*C. auratus gibelio*), Japanese ginbuna (*C. auratus langsdorfii*), the domesticated goldfish (*C. auratus*), and numerous local varieties such as Pengze crucian carp (*C. auratus* var. Pengze), Puan crucian carp (*C. auratus* var. Puan), and Dianchi high-back crucian carp (*C. auratus*) [2]. Given the challenges in distinguishing species or subspecies within *Carassius* based on morphological characters, molecular markers are often used to resolve issues including species identification [10,11], phylogenetics [12–14], and population genetic differentiation [15,16]. While mitochondrial genes like *COI*, *COIII*, *Cytb* genes, and the D-loop sequence are frequently employed due to their maternal inheritance, rapid evolution and low recombination rate [17–19], recent studies focusing on genetic diversity and evolution have shown nuclear gene molecular markers to outperform mitochondrial ones [20].

Transferrin (Tf) is a protein that binds two Fe^{3+} ions in animals, crucial for iron transportation and elimination, directly influencing oxygen metabolism [21]. The *Tf* gene sequences are susceptible to mutation from various factors, exhibiting high genetic polymorphism, making it valuable for germplasm identification and population genetics in farmed and wild fish. Phylogenetic analyses of *Tf* sequences from thousands of *Carassius* species samples have shown higher genetic diversity in triploid populations compared with diploids, suggesting their origin from sympatric diploids through recurrent autoploidization events [22,23]. These researches underscore the *Tf* allele marker's significant role in elucidating the origin and evolutionary history of *Carassius* species.

In previous studies, we analyzed the population characteristics of HFJ and NCRC based on mitochondrial *COI* gene and D-loop data, revealing significant genetic differentiation among HFJ, NCRC, and other populations [24]. However, depending solely on mitochondrial molecular markers is inadequate for genetic diversity researches. In the present study, we utilized *Tf* alleles for genetic diversity analysis of HFJ, NCRC, and wild populations near breeding sites, reconstructing their phylogenetic relationships and investigating their evolutionary histories. The results obtained from our study will offer valuable insights for the utilization and conservation of germplasm resources in fish genetic breeding.

2. Materials and methods

2.1. Ethics statement

The guidelines established by the Administration of Affairs Concerning Animal Experimentation state that approval from the Science and Technology Bureau of China and Department of Wildlife Administration is not necessary when the fish in question are neither rare nor near extinction (first- or second-class state protection level). Therefore, approval was not required for the experiments conducted in this study.

2.2. Samples information

A total of 161 individuals in *Carassius*, sourced from two hybrid varieties, one laboratory variety, and three wild populations, were used in this study. The hybrid varieties, specifically HFJ (n = 30) and NCRC (n = 27), were obtained from the State Key Laboratory for Developmental Biology of Freshwater Fishes in Hunan Normal University, Changsha, China. The laboratory variety, red crucian carp C1HD (RCC, n = 16), was bred by the Department of Laboratory Animal Science in University of South China, Hengyang, China. Meanwhile, the wild populations were collected from Wangcheng District (Changsha, Hunan, China) (WC, n = 14), Changsha County (Changsha, Hunan, China) (CSX, n = 30), and Changde City (Hunan, China) (CD, n = 25). The caudal fin tissues from all samples were preserved in 100 % ethanol and stored at −20 °C.

2.3. DNA extraction, transferrin allele amplification, and sequencing

Total genomic DNA was extracted from the caudal fins of each sample using Tissue DNA Kits (OMEGA). A set of primers (*Tf*-F, 5′-CTCCTCAAAGAGCCTCGCCAT -3′; *Tf* –R, 5′- TACACCTGGCCACCATCAACTG -3′) was synthesized to amplify an approximate 1050 bp fragment spanning the 7th to 10th exons of the *Tf* gene, following previous studies [22,23]. The thermal cycling program consisted of an initial denaturation at 94 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 45 s, with a final extension at 72 °C for 7 min. PCR products were separated on a 1.8 % agarose gel using TAE buffer. The DNA fragments were purified using a gel extraction kit (Sangon, Shanghai, China), ligated into the pMD18-T vector (TaKaRa, Dalian, China), and transformed into *Escherichia coli DH5a* for further purification. At least 10 positive clones from each sample were selected and sequenced in both directions using an automated DNA sequencer (ABIPRISM 3730).

2.4. Data analyses

The sequences were checked manually and aligned by Clustal W [25]. We used DnaSP 6.12 [26] to identity the alleles of *Tf* gene and assess population genetic diversity, including allele diversity (h) and nucleotide diversity (π). The program MEGA11 [27] was used to calculate the genetic distance between and within populations. To better visualize the reticular relationships of six populations of *Carassius*, a median-joining allele network was constructed using PopART software [28]. The population genetic structure was determined by Structure 2.3.4 program [29,30], with the number of the most likelihood populations (K) ranging from 1 to 6 and 10 replicates for each K. The Arlequin 3.5.2.2 software [31] was applied to calculate the average values of population differentiation (F_{ST}) and conduct the analysis of molecular variance (AMOVA). Tajima's D [32], Fu's F_S [33] neutrality test statistics, and mismatch distribution analysis [34] were also performed using Arlequin to infer demographic history and identify expansion signals across the six *Carassius* populations.

We further reconstructed the phylogenetic relationships *Tf* alleles using Bayesian inference (BI), maximum likelihood (ML), and neighbor joining (NJ) methods. Additionally, four Tf alleles from C. cuvieri (amplified and identified in this study), four from Cy. carpio (GeneBank No. LN590718.1, JQ822196.1, JQ822197.1, and JQ822198.1), and two from *M. amblycephala* (identified from genome of GeneBank No. GCA_009869865.1) were included for a comprehensive analysis. The best-fit nucleotide substitution model for *Tf* alleles, the GTR + I + G model, was determined using the Akaike information criterion (AIC) in jModelTest 2.1.10 software [35]. Bayesian analysis was performed using MrBayes 3.2.7 [36], running four independent Markov chain Monte Carlo (MCMC) chains for 10,000,000 generations with a sample frequency of 1000 generations. The first 25 % of trees were discarded as burn-in, and the remaining samples were used to generate a consensus tree. The ML analysis was conducted using IQtree 1.6.12 [37], while the NJ analysis was performed in MEGA 11 applying the kimura 2-parameter (K2P) options. Nodal support values were assessed from 1000 nonparametric bootstrap replicates for both the ML and NJ trees. The final results were visualized and edited in FigTree 1.4.4 (http://tree.bio.ed.ac.uk/software/Figtree/).

3. Results

3.1. Genetic variation of Tf alleles in six carassius populations

About one thousand *Tf* sequences were obtained from 161 specimens of six *Carassius* populations, with sequence lengths ranging from 1145 bp to 1244 bp. A total of 166 distinct *Tf* alleles (A1∼A166) were defined based on 363 polymorphic sites, with an average nucleotide difference of 55.93 (Table 1). Specifically, 19 *Tf* alleles were detected in WC (n = 14), 35 in CSX (n = 30), 26 in CD (n = 26), 13 in RCC (n = 16), 64 in HFJ (n = 49), and 35 in NCRC (n = 26). The majority of *Tf* alleles (92 %) were unique to individual populations, with only 13 alleles being shared. Among the shared alleles, A5 was the most prevalent, present in five out of the six populations, followed by A1, A3, A4, and A9, each shared by four populations. Additionally, A79 was distributed in three populations. The overall allele diversity was $h = 0.983$, with no significant variation observed across populations (ranging from 0.922 to 0.972). The overall nucleotide diversity was $\pi = 0.05306$, ranging from 0.03379 to 0.05988. RCC had the highest π value, followed by HFJ and NCRC. Notably, the nucleotide diversity of *Tf* alleles in cultured populations (with an average of 0.05428) exceeded that in wild populations (with an average of 0.04383).

3.2. Phylogenetic relationship and network reconstruction

The allele network analysis was performed using the median-joining method (Fig. 1). The 166 *Tf* alleles were clustered into six discrete clades (Clade 1–6), with no single population exclusively belonging to a particular clade. Notably, Clade 1 consisted solely of two alleles (A121 and A122) from NCRC, exhibiting a relatively distant relationship with other clades, potentially attributed to the hybrid origin of NCRC. Clade 2 included 13 *Tf* alleles displaying a star-like structure centered around A1. Clade 3, with 30 *Tf* alleles, were roughly divided into two parts, one forming an obvious star-like structure network with A5 at its core, while the other predominantly consisted of CSX alleles. Clade 4 contained 26 *Tf* alleles, all from HFJ except for four from NCRC. In Clade 5, 16 *Tf* alleles were grouped, with A79 positioned centrally. The largest clade, Clade 6, encompassed representatives from all six populations and consisted of 63 *Tf* alleles. Remarkably, this clade harbored two conspicuous star-like structures centered around A3 and A4, respectively. Therefore, the network lacked distinct topological structure, as the *Tf* alleles from different populations exhibited an admixture distribution pattern with many shared alleles, indicating an absence of obvious lineage differentiation among these populations.

The phylogenetic trees for the six populations of *Carassius* were reconstructed by the BI, ML, and NJ methods. These methods yielded consistent topology structures, hence only the ML phylogenetic tree is shown (Fig. 2). The six clades identified in the network analysis were strongly supported in the phylogenetic tree. Significantly, A121 and A122 from NCRC clustered with four *Tf* alleles of *Cy. carpio*, forming a well-supported monophyly (Clade 1) located at the base of the phylogenetic tree, aligning with the known female parent of NCRC as *Cy. carpio*. However, the hybrid origin of HFJ was not clearly reflected in the phylogenetic tree, because the *Tf* alleles of *C. cuvieri* and *C. auratus* were intermixed and distributed across different clades, making it impossible to distinguish between them accurately. Similar to the median-joining network, the phylogenetic tree analysis revealed a complex intermixing pattern of *Tf* alleles among HFJ, NCRC, RCC, and the wild populations of *Carassius*. Not surprisingly, none of the populations possessed a distinct cluster, further emphasizing the relatively close relationship among all six populations based on *Tf* alleles.

3.3. Population structure and lineages divergence

Genetic distance and population differentiation among the six *Carassius* populations were calculated based on *Tf* alleles (Table 2). The average within-population genetic distance was 0.0514, with the smallest genetic distance observed in WC (0.0354) and the largest in RCC (0.0629). Notably, the within-population genetic distance of *Tf* alleles was higher in hybrid and cultured varieties compared to the wild populations. The average between-population genetic distance was 0.0555, with the smallest distance being between WC and CD (0.0432). Besides, genetic distances between RCC and other populations (0.0566–0.0627) were higher than the average, except for WC (0.0548). Pairwise F_{ST} values ranged from 0.05997 to 0.15814 among six populations of *Carassius*. It is worth noting that the genetic differentiation levels observed in the two hybrid varieties were quite different. The F_{ST} values for comparisons between NCRC and other populations were relatively low (0.02750–0.05997), with most being insignificant ($P > 0.05$). Conversely, the F_{ST} values were generally high (0.06789–0.15814) in comparisons between HFJ and other populations except for NCRC, with some achieving extremely remarkable level ($P < 0.001$).

The population structure of both cultured and wild *Carassius* populations was determined using *Tf* alleles data. Groupings (K) ranging from 1 to 6 were evaluated, with each grouping repeated 10 times. According to the ΔK criterion, the optimal number of clusters was determined to be 3 ($K = 3$), with results also presented for $K = 2$ and $K = 4$ for comparison (Fig. 3). Notably, no distinct visual pattern of population structure was evident across all cases. The *Tf* alleles from the six *Carassius* populations exhibited an apparent admixture of varying proportions, which is in line with results of the network and phylogenetic analyses. The samples from different populations, including both cultured and wild populations, cannot be clearly distinguished, suggesting a relatively limited genetic divergence among these populations.

AMOVA was performed based on *Tf* alleles from six populations of *Carassius* in order to further delineate genetic diversity (Table 3). The results showed that a mere 8 % of the genetic variance originated from variability among the populations, with a substantial 92.00 % of the variance attributed to within-population differences. This suggested that the overwhelming majority of genetic variation was accounted for by within-population factors. In addition, the statistical analysis revealed that the genetic differentiation coefficient (F_{ST}) among populations was 0.07999 and extremely significant ($P < 0.01$), indicating a moderate level of genetic differentiation in the six *Carassius* populations. Despite the existence of significant genetic variation among the six *Carassius* populations, the lack of distinct population groupings in the genetic structure analysis may be attributed to the presence of two hybrid

Table 1
Genetic diversity analysis based on *Tf* sequences.

Population code	Number of individuals	Number of alleles	Number of variable sites	Average number of nucleotide differences (k)	Allele diversity (h)	Nucleotide diversity (π)
WC	14	19	127	39.76923	0.932	0.03379
CSX	30	35	204	53.95286	0.972	0.05057
CD	26	26	206	50.18634	0.969	0.04712
RCC	16	13	184	63.83117	0.922	0.05988
HFJ	49	64	270	54.63114	0.969	0.05154
NCRC	26	35	241	54.56033	0.964	0.05142
Total	161	166	363	55.92915	0.983	0.05306

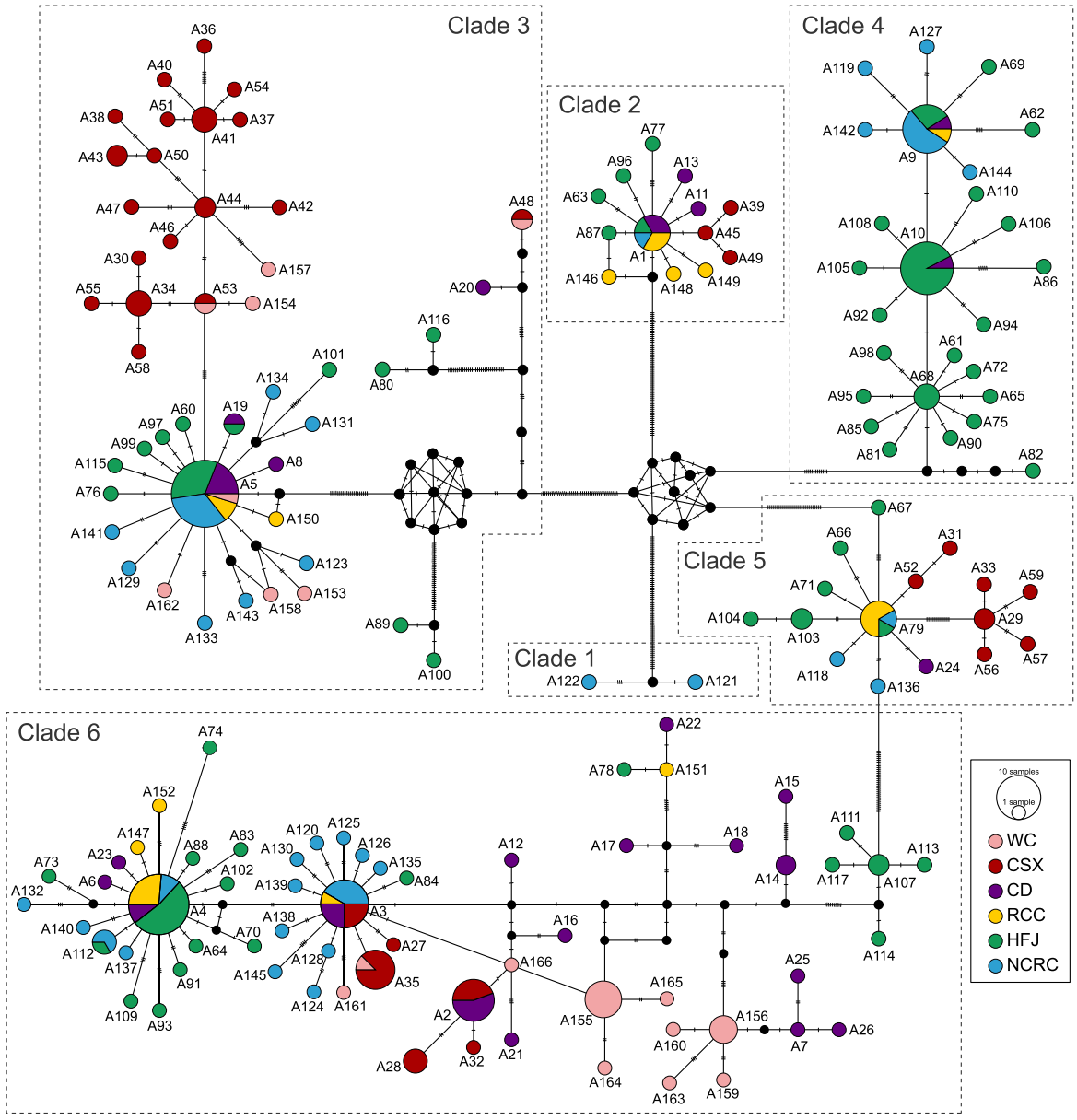

Fig. 1. Median-joining network based on *Tf* alleles. The six dashed boxes represent the six Clades. Circles represent different alleles and their corresponding occurrence frequency in all sampled populations.

varieties, HFJ and NCRC.

3.4. Demographic history of six carassius populations

The evolutionary history of the six *Carassius* populations was inferred using Tajima's *D* and Fu's F_S models (Table 4). Fu's F_S value for HFJ was non-significantly negative (−3.41294; $P = 0.243$), indicating a potential recent population expansion or selective pressure and bottleneck effect. Conversely, the results for all other populations exhibited small values (Tajima's *D*: 0.08293–1.11226; Fu's F_S: 1.1993–8.65078) with non-significant *P* values ($P > 0.05$), which was unlikely to be explained by population expansion. This finding was further supported by the analysis of mismatch distribution (Fig. 4), which revealed multiple peaks in the curves of all six *Carassius* populations, suggesting that these populations have likely remained relatively stable in recent history.

4. Discussion

The polymorphic gene, transferrin, plays an important role in physiological functions, including hypoxia tolerance, making it an extremely valuable candidate for studying local adaptation and population divergence. The polymorphism in fish serum transferrin was first reported in *Cy. carpio* [38], and has been subsequently confirmed in diverse fish species [39,40]. Cyprinid fish exhibited the most diverse transferrin polymorphisms [41]. As such, serum transferrin polymorphism emerges as a reliable biochemical genetic marker for germplasm evaluation in fish genetic breeding [42]. In this study, we amplified and sequenced an approximate 1150 bp fragment of the transferrin gene from six *Carassius* populations. We identified 166 *Tf* alleles and analyzed their genetic polymorphisms and population structure, and obtained several new insights that differed from those discovered using mitochondrial molecular markers.

Inbreeding is a common challenge often found in economically

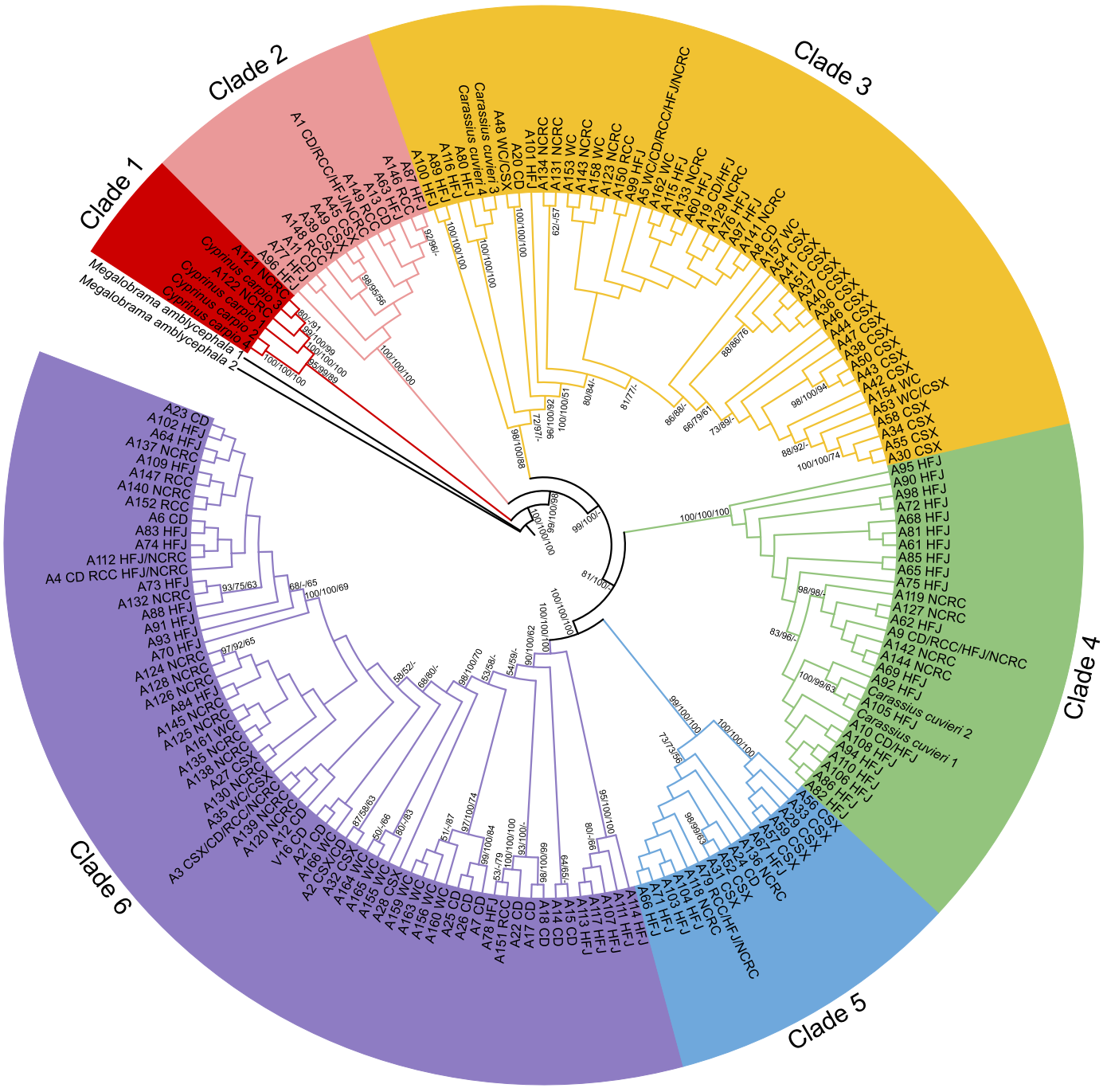

Fig. 2. Phylogenetic relationships obtained by *Tf* alleles. The six color shadows represent the six Clades. The numbers on nodes were posterior probability for BI, bootstrap support for ML, bootstrap support for NJ, respectively. The values less than 50 % were represented by "-".

Table 2
Genetic differentiation and genetic distances analysis of six *Carassius* populations based on *Tf* alleles.

	WC	CSX	CD	RCC	HFJ	NCRC
WC	**0.0354**	0.0484	0.0432	0.0548	0.0548	0.0482
CSX	0.09154*	**0.0522**	0.0557	0.0627	0.0617	0.0562
CD	0.01802	0.09175***	**0.0493**	0.0566	0.0573	0.0534
RCC	0.10981***	0.08785***	0.01553	**0.0629**	0.0626	0.0610
HFJ	0.15814***	0.13103***	0.09251***	0.06789*	**0.0541**	0.0561
NCRC	0.05997	0.05016*	0.02750*	0.04143	0.03088	**0.0545**

The data below diagonal were population genetic differentiation coefficient (F_{ST}), the data above diagonal were genetic distances among populations, and the data on diagonal were genetic distances within populations. *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, the same applies below.

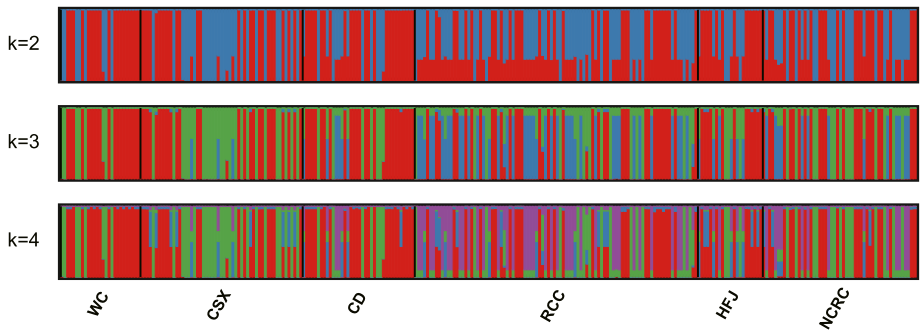

Fig. 3. Population structuring analysis based on *Tf* alleles.

Table 3
Analysis of molecular variance (AMOVA) based on *Tf* alleles of six *Carassius* populations.

Source of variation	*df*	Sum of squares	Variance components	Percentage of variation (%)	Fixation Indices
Among populations	5	652.336	2.27444 Va	8.00	
Within populations	284	7429.426	26.15995 Vb	92.00	F_{ST} = 0.07999**
Total	289	8081.762	28.43439		

Table 4
Neutrality test based on *Tf* alleles of six *Carassius* populations.

Index	WC	CSX	CD	RCC	HFJ	NCRC
Tajima's *D*	1.00261	0.78131	0.10698	1.11226	0.08668	0.08293
Tajima's *D* p-value	0.892	0.826	0.604	0.92	0.618	0.62
Fu's F_S	2.37474	1.88711	1.93277	8.65078	−3.41294	1.1993
Fu's F_S p-value	0.842	0.772	0.788	0.997	0.243	0.698

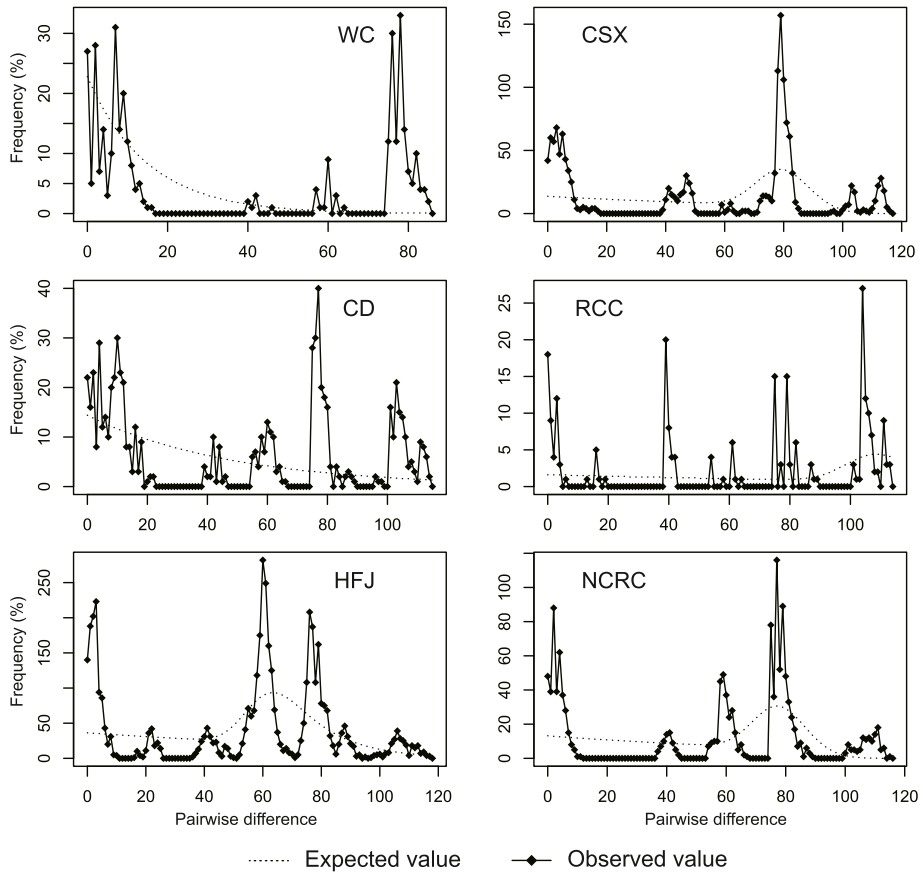

Fig. 4. Mismatch distribution analysis based on *Tf* alleles of six *Carassius* populations.

valuable fish during artificial propagation, leading to a decrease in genetic polymorphism and germplasm depression in cultured species. Therefore, genetic diversity evaluation is an important link in germplasm resources conservation and research. Our previous study accessed the genetic diversity of six *Carassius* populations using mitochondrial *COI* gene and D-loop data [24], and found that, when compared to wild populations, the cultured populations exhibited relatively low levels of haplotype diversity and nucleotide diversity, paralleling the trend seen in other farm-raised fish [43,44]. However, the present study, according to the criteria for assessing population diversity proposed by Grant and Bowen [45], indicated high levels of haplotype diversity ($h > 0.5$) and nucleotide diversity ($\pi > 0.005$) in both wild populations (WC, CSX, and CD) and cultured populations (RCC, HFJ, and NCRC). In addition, it is noteworthy that the nucleotide diversity of wild populations was lower than that of cultured populations despite the absence of a significant difference in haplotype diversity, which contradicts the previous conclusion based on mitochondrial genes.

The current study suggests a potential correlation between the high genetic diversity of *Tf* alleles observed in cultured populations of *Carassius* and their hybridization origins. Specifically, HFJ was derived from interspecies hybridization between *C. cuvieri* and *C. auratus* var. red, NCRC was derived from inter-subfamily hybridization between *Cy. carpio* and *M. amblycephala*, and RCC was the gynogenetic offspring of wild *C. auratus* var. red, whose eggs were activated by ultraviolet-irradiated *Cy. carpio* sperm [46]. A previous study focusing on serum transferrin polymorphisms in three *Clarias* catfish found that hybrids resulting from *Cl. gariepinus* and *Cl. fuscus* crossings had higher haplotype diversity than either of the parental species [47]. Likewise, a study on natural homoploid fish hybrids (Cyprinidae, Leuciscinae) revealed that all hybrids exhibited extensive rDNA polymorphism not observed in the parental species, which can be interpreted as transgressive phenotypes produced by hybridization [48]. In agreement with these studies, we hypothesized that the high genetic diversity of *Tf* alleles in the three cultured populations could be indicative of their heterosis. These populations effectively served as excellent germplasm of *Carassius* and demonstrated outstanding resistance to stress and infection [49–51]. It has been suggested that local adaptation and stress resistances in fish species might be partially related to transferrin polymorphism, given that serum transferrin plays a crucial role not only in iron metabolism but also in immune response. The sea bass (*Dicentrarchus labrax*), for instance, responded to bacterial infection and iron modulation by increasing transferrin expression in both the liver and brain [52]. In *Cy. carpio*, different transferrin genotypes were found to be associated with varying susceptibility to the blood parasite [53]. Accordingly, in this study, the elevated genetic diversity observed in cultured *Carassius* populations may also be attributed to the utilization of *Tf*, a molecular marker highly susceptible to selection pressure. We postulate that the high polymorphism of the nuclear gene in hybrid populations may be a manifestation of the genomic shock effect caused by hybridization.

The relationships among the six *Carassius* populations were inferred through phylogenetic analysis and network reconstruction. The results indicated that individuals from wild populations were not clustered according to the sampling sites, corroborating with our previous findings using mitochondrial data [24], thus suggesting insignificant divergences between wild *Carassius* populations from different regions. Similarity, a recent study that collected *Carassius* samples from 41 locations in Hunan Province, China, including the major water basins and Dongting Lake, found inadequate morphological differences and genetic differentiation among the populations to form stable local varieties [54]. Furthermore, although RCC (the red crucian carp C1HD) has been successfully distinguished based on molecular markers such as mitochondrial genes and microsatellite DNA [55], the *Tf* alleles failed to identify it in this study, because RCC samples were dispersed across all clades except Clade1, with nearly half of their *Tf* alleles found to share with other populations rather than forming an independent clade. In addition, most *Tf* alleles of NCRC and other *Carassius* populations were intermixed across five clades, but with two *Tf* alleles belonging to *Cy. carpio* detected. This not only confirmed the hybrid origin of NCRC but also highlighted its close relationship with *Cy. carpio* [14,56]. Regrettably, no *Tf* alleles belonging to *M. amblycephala*, the male parent of NCRC, were identified in the NCRC dataset, aligning with the discovery of the partial elimination of paternal 5S rDNA units and variations in the maternal 5S rDNA unit within the NCRC genome [6], potentially indicating genomic shock in distant hybridization offspring and leading to the NCRC genome exhibiting greater similarity to that of *C. auratus* than its parental species. In previous studies, analysis based on mitochondrial markers showed that HFJ could be grouped together with *C. cuvieri* to form monophyletic branches, indicating the species independence of *C. cuvieri* [10,12,13,18]. However, our present study showed that HFJ, together with *C. cuvieri*, presented an admixture distribution with other populations, which conflicts with the species classification within the genus *Carassius*. Generally, distinguishing different *Carassius* populations based on *Tf* alleles seems more challenging compared to mitochondrial molecular markers, as in this study, the populations that could not be distinguished by mitochondrial genes remained undifferentiated, while those that could be differentiated by mitochondrial genes became nondifferentiable. This could be attributed to the high polymorphism of *Tf* alleles in hybrid populations.

It has been reported that markers under selection may provide a significantly different perspective on population structure compared to neutral markers [20]. In a previous study, obvious genetic differentiation was detected between cultured and wild populations, with the three cultured populations of HFJ, NCRC, and RCC being distinctly separated from the other populations [24]. However, the results presented here suggest that there is no apparent population structure among the six *Carassius* populations. On the one hand, the genetic distance analysis indicated a relatively large within-population distance of greater than 3 %, with the cultured populations presenting higher values compared to the wild populations. Although the pairwise between-population distance was also large (4.32%–6.27 %), there was no distinct gap observed between within- and between-populations, indicating a challenge in distinguishing subpopulations based on genetic distance. On the other hand, according to Wright's criteria for measuring the extent of genetic differentiation among populations (F_{ST}) [57], it was found that, in addition to a significant genetic differentiation between HFJ and WC ($F_{ST} > 0.15$, $P < 0.001$), small ($F_{ST} < 0.15$) or undetectable ($P > 0.05$) genetic differentiation were observed among most of the other populations. Furthermore, the *Tf* alleles from the six populations of *Carassius* exhibited apparent admixture and failed to form a distinct pattern in STRUCTURE analysis, and the AMOVA analysis revealed that a significant proportion of the genetic variance was explained by differences within populations. Consequently, combined with the results of phylogenetic and network analysis, we determined that no distinct genetic divergence was observed among the six *Carassius* populations. This may result from the complex genetic background and genomic instability of HFJ and NCRC hybrid varieties.

In the population dynamic analysis, a negative value was only found in Fu's F_S test of HFJ, but it lacked any statistical significance. Combining these results with Tajima's *D* test and mismatch distribution, similar to the findings from mitochondrial data [24], it was suggested that the six *Carassius* populations have not experience population expansion or directional selection, remaining relatively stable in their recent history.

In summary, our research investigated the genetic diversity and population structure of six *Carassius* populations, relying on the nuclear marker *Tf*. The cultured populations, including RCC, HFJ, and NCRC, demonstrated greater genetic diversity compared to wild populations, contrasting with findings using mitochondrial gene markers. An unclear population structure between wild and cultured populations was observed, manifested by the apparent admixture distribution across the six *Carassius* populations. These results provide direct evidence, based on nuclear marker data, confirming the hybrid genetic characteristics of

two hybrid varieties, HFJ and NCRC, holding great significance for the conservation, improvement, and genetic breeding of crucian carp.

Funding

This work was supported by the National Key R&D Program of China (2022YFD2400102), the National Natural Science Foundation of China (Grant No. 32202912, 32293252, 32373119), the Training Program for Excellent Young Innovators of Changsha, China (Grant No. kq2200013, kq2107006), and the Hunan Provincial Natural Science Foundation (Grant No. 2022JJ30386, 2022JJ10035).

CRediT authorship contribution statement

Wenjie Luo: Writing – original draft, Visualization, Investigation, Data curation. **Xuexue Huang:** Visualization, Methodology, Investigation, Data curation. **Xiaowei Xu:** Investigation, Data curation. **Chenghua Dai:** Investigation, Data curation. **Qiong Liu:** Validation, Investigation. **Yating Zhu:** Validation, Investigation. **Duansheng Wu:** Resources. **Shi Wang:** Resources, Funding acquisition. **Qingfeng Liu:** Resources, Funding acquisition. **Conghui Yang:** Writing – original draft, Visualization, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

All authors declare that there are no competing interests.

References

[1] S. Wang, C. Tang, M. Tao, Q. Qin, C. Zhang, K. Luo, R. Zhao, J. Wang, L. Ren, J. Xiao, F.Z. Hu, R. Zhou, W. Duan, S.J. Liu, Establishment and application of distant hybridization technology in fish, Sci. China Life Sci. 62 (2019) 22–45, https://doi.org/10.1007/s11427-018-9408-x.

[2] C.J. Dong, X.J. Li, X.W. Sun, Research progress of the genetic diversity, origin and evolution of *Carassius auratus* in China, J. Fish. China 44 (2020) 1046–1062, https://doi.org/10.11964/jfc.20190811926.

[3] Q.F. Liu, Y.H. Qi, Q.L. Liang, X.J. Xu, F.Z. Hu, J. Wang, J. Xiao, S. Wang, W.H. Li, M. Tao, Q.B. Qin, R.R. Zhao, Z.Z. Yao, S.J. Liu, The chimeric genes in the hybrid lineage of *Carassius auratus cuvieri* (♀) × *Carassius auratus* red var. (♂), Sci. China Life Sci. 61 (2018) 1079–1089, https://doi.org/10.1007/s11427-017-9306-7.

[4] Q.F. Liu, J.M. Liu, Q.L. Liang, Y.H. Qi, M. Tao, C. Zhang, Q.B. Qin, R.R. Zhao, B. Chen, S.J. Liu, A hybrid lineage derived from hybridization of *Carassius cuvieri* and *Carassius auratus* red var. and a new type of improved fish obtained by backcrossing, Aquaculture 505 (2019) 173–182, https://doi.org/10.1016/j.aquaculture.2019.02.056.

[5] Q.F. Liu, Y.H. Qi, Q.L. Liang, J. Song, J.M. Liu, W.H. Li, Y.Q. Shu, M. Tao, C. Zhang, Q.B. Qin, J. Wang, S.J. Liu, Targeted disruption of tyrosinase causes melanin reduction in *Carassius auratus cuvieri* and its hybrid progeny, Sci. China Life Sci. 62 (2019) 1194–1202, https://doi.org/10.1007/s11427-018-9404-7.

[6] S. Wang, X.L. Ye, Y.D. Wang, Y.T. Chen, B.W. Lin, Z.F. Yi, Z.W. Mao, F.Z. Hu, R. R. Zhao, J. Wang, R. Zhou, L. Ren, Z.Z. Yao, M. Tao, C. Zhang, J. Xiao, Q.B. Qin, S. J. Liu, A new type of homodiploid fish derived from the interspecific hybridization of female common carp × male blunt snout bream, Sci. Rep. 7 (2017) 4189, https://doi.org/10.1038/s41598-017-04582-z.

[7] S. Wang, Q.L. Liu, X.X. Huang, C.H. Yang, L. Chen, M. Han, Y.Q. Shu, M. Wang, W. H. Li, F.Z. Hu, M. Wen, K.K. Luo, Y.D. Wang, R. Zhou, C. Zhang, M. Tao, R.R. Zhao, C.C. Tang, S.J. Liu, The rapid variation of *Hox* clusters reveals a clear evolutionary path in a crucian carp-like homodiploid fish lineage, Reprod. Breed. 1 (2021) 149–156, https://doi.org/10.1016/j.repbre.2021.09.002.

[8] Y.L. Luo, P.Q. Yue, Cyprinidae, in: P.Q. Yue (Ed.), Fauna Sinica Osteichthyes: Cypriniformes (III), Science Press, Beijing, 2000, pp. 399–433.

[9] A. Wheeler, Status of the crucian carp, *Carassius carassius* (L.), in the UK, Fish. Manag. Ecol. 7 (2000) 315–322, https://doi.org/10.1046/j.1365-2400.2000.007004315.x.

[10] L. Cheng, Y.M. Chang, C.Y. Lu, D.C. Cao, X.W. Sun, DNA barcoding and species and subspecies classification within genus *Carassius*, Zool. Res. 35 (2012) 463–472, https://doi.org/10.3724/SP.J.1141.2012.05463.

[11] S.Y. Wang, J. Luo, R.W. Murphy, S.F. Wu, C.L. Zhu, Y. Gao, Y.P. Zhang, Origin of Chinese goldfish and sequential loss of genetic diversity accompanies new breeds, PLoS One 8 (2013) e59571, https://doi.org/10.1371/journal.pone.0059571.

[12] M. Murakami, C. Matsuba, H. Fujitani, The maternal origins of the triploid ginbuna (*Carassius auratus langsdorfi*): phylogenetic relationships within the *C. auratus* taxa by partial mitochondrial D-loop sequencing, Genes Genet. Syst. 76 (2001) 25–32, https://doi.org/10.1266/ggs.76.25.

[13] M. Takada, K. Tachihara, T. Kon, G. Yamamoto, K. Iguchi, M. Miya, M. Nishida, Biogeography and evolution of the *Carassius auratus*-complex in east asia, BMC Evol. Biol. 10 (2010) 1–18, https://doi.org/10.1186/1471-2148-10-7.

[14] Q.H. Gu, S. Wang, H. Zhong, H. Yuan, J.L. Yang, C.H. Yang, X.X. Huang, X.W. Xu, Y.D. Wang, Z.H. Wei, J. Wang, S.J. Liu, Phylogeographic relationships and the evolutionary history of the *Carassius auratus* complex with a newly born homodiploid raw fish (2nNCRC), BMC Genom. 23 (2022) 242, https://doi.org/10.1186/s12864-022-08468-x.

[15] J. Luo, Y. Gao, W. Ma, X.Y. Bi, S.Y. Wang, J. Wang, Y.Q. Wang, J. Chai, R. Du, S. F. Wu, A. Meyer, R.G. Zan, H. Xiao, R.W. Murphy, Y.P. Zhang, Tempo and mode of recurrent polyploidization in the *Carassius auratus* species complex (Cypriniformes, Cyprinidae), Heredity 112 (2014) 415–427, https://doi.org/10.1038/hdy.2013.121.

[16] X.L. Liu, X.Y. Li, F.F. Jiang, Z.W. Wang, Z. Li, X.J. Zhang, L. Zhou, J.F. Gui, Numerous mitochondrial DNA haplotypes reveal multiple independent polyploidy origins of hexaploids in *Carassius* species complex, Ecol. Evol. 7 (2017) 10604–10615, https://doi.org/10.1002/ece3.3462.

[17] T. Komiyama, H. Kobayashi, Y. Tateno, H. Inoko, T. Gojobori, K. Ikeo, An evolutionary origin and selection process of goldfish, Gene 430 (2009) 5–11, https://doi.org/10.1016/j.gene.2008.10.019.

[18] Y. Gao, S.Y. Wang, J. Luo, R.W. Murphy, R. Du, S.F. Wu, C.L. Zhu, Y. Li, A. D. Poyarkov, S.N. Nguyen, P.T. Luan, Y.P. Zhang, Quaternary paleoenvironment oscillations drove the evolution of the eurasian *Carassius auratus* complex (cypriniformes, cyprinidae), J. Biogeogr. 39 (2012) 2264–2278, https://doi.org/10.1111/j.1365-2699.2012.02755.x.

[19] Y.D. Wang, D.K. Li, K.J. Chen, D.S. Wu, S.J. Liu, Characteristics analysis and application of mitochondrial DNA barcoding in laboratory red crucian carp (*Carassius Carassius*, red variety), J. Fish. China 45 (2021) 1–9, https://doi.org/10.11964/jfc.20191112046.

[20] Ø. Andersen, M.C. De Rosa, D. Pirolli, A. Tooming-Klunderud, P.E. Petersen, C. André, Polymorphism, selection and tandem duplication of transferrin genes in Atlantic cod (*Gadus morhua*)-Conserved synteny between fish monolobal and tetrapod bilobal transferrin loci, BMC Genet. 12 (2011) 51, https://doi.org/10.1186/1471-2156-12-51.

[21] J. Liu, Research advances on transfenins of vertebrates, Life Science Instruments 7 (2009) 11–14.

[22] F.F. Jiang, Z.W. Wang, L. Zhou, L. Jiang, X.J. Zhang, O.V. Apalikova, V.A. Brykov, J.F. Gui, High male incidence and evolutionary implications of triploid form in northeast Asia *Carassius auratus* complex, Mol. Phylogenet. Evol. 66 (2013) 350–359, https://doi.org/10.1016/j.ympev.2012.10.006.

[23] X.L. Liu, F.F. Jiang, Z.W. Wang, X.Y. Li, Z. Li, X.J. Zhang, F. Chen, J.F. Mao, L. Zhou, J.F. Gui, Wider geographic distribution and higher diversity of hexaploids than tetraploids in *Carassius* species complex reveal recurrent polyploidy effects on adaptive evolution, Sci. Rep. 7 (2017) 5395, https://doi.org/10.1038/s41598-017-05731-0.

[24] C.H. Yang, X.X. Huang, C.H. Dai, X.W. Xu, S. Wang, Q.F. Liu, C. Zhang, Y. Zhou, S. J. Liu, Population genetic characteristics of hybrid and wild crucian carps in China based on joint sequences of mitochondrial *COI* gene and D-loop region, Journal of Southern Agriculture 53 (2022) 3009–3019, https://doi.org/10.3969/j.issn.2095-1191.2022.11.001.

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

[26] J. Rozas, A. Ferrer-Mata, J.C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S. E. Ramos-Onsins, A. Sánchez-Gracia, DnaSP 6: DNA sequence polymorphism analysis of large data sets, Mol. Biol. Evol. 34 (2017) 3299–3302, https://doi.org/10.1093/molbev/msx248.

[27] K. Tamura, G. Stecher, S. Kumar, Mega 11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (2021) 3022–3027, https://doi.org/10.1093/molbev/msab120.

[28] J.W. Leigh, D. Bryant, POPART: full-feature software for haplotype network construction, Methods Ecol. Evol. 6 (2015) 1110–1116, https://doi.org/10.1111/2041-210X.12410.

[29] D. Falush, M. Stephens, J.K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies, Genetics 164 (2003) 1567–1587, https://doi.org/10.1093/genetics/164.4.1567.

[30] D. Falush, M. Stephens, J.K. Pritchard, Inference of population structure using multilocus genotype data: dominant markers and null alleles, Mol. Ecol. Notes 7 (2007) 574–578, https://doi.org/10.1111/j.1471-8286.2007.01758.x.

[31] L. Excoffier, H.E.L. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour 10 (2010) 564–567, https://doi.org/10.1111/j.1755-0998.2010.02847.x.

[32] F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, Genetics 123 (1989) 585–595, https://doi.org/10.1093/genetics/123.3.585.

[33] Y. Fu, Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection, Genetics 147 (1997) 915–925, https://doi.org/10.1093/genetics/147.2.915.

[34] S. Schneider, L. Excoffier, Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA, Genetics 152 (1999) 1079–1089, https://doi.org/10.1093/genetics/152.3.1079.

[35] D. Darriba, G.L. Taboada, R. Doallo, D. Posada, jModelTest 2: more models, new heuristics and parallel computing, Nat. Methods 9 (2012) 772, https://doi.org/10.1038/nmeth.2109.

[36] F. Ronquist, J.P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models, Bioinformatics 19 (2003) 1572–1574, https://doi.org/10.1093/bioinformatics/btg180.
[37] L.-T. Nguyen, H.A. Schmidt, A. Von Haeseler, B.Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (2015) 268–274, https://doi.org/10.1093/molbev/msu300.
[38] R. Creyssel, G.B. Richard, P. Silberzahn, Transferrin variants in carp serum, Nature 212 (1966) 1362, https://doi.org/10.1038/2121362a0.
[39] D. Møller, G. Naevdal, Serum transferrins of some gadoid fishes, Nature 210 (1966) 317–318, https://doi.org/10.1038/210317a0.
[40] E. Verspoor, Spatial correlations of transferrin allele frequencies in atlantic salmon (*Salmo salar*) populations from North America, Can, J. Fish. Aquat. Sci. 43 (1986) 1074–1078, https://doi.org/10.1139/f86-135.
[41] H. Long, F. Tang, Y. Zeng, C. Wang, Studies on genetic polymorphisms of freshwater fish serum transferrins, J. Fish. China 20 (1996) 168–174.
[42] L. Zhu, Y. Jiang, Intraspecific genetic markers of crucian carp (*Carassius auratus gibelio*) and their application to selective breeding, Acta Hydrobiol. Sin. 12 (1987) 105–111.
[43] W.W. Cheng, R.L. Xia, Q.Y. Wang, K.W. Zeng, W. Song, Q.W. Wei, G.Q. Deng, Y. H. Cheng, Genetic diversity of *Siniperca chuatsi* in wild and cultivated populations, Freshw. Fish. 50 (2020) 31–37, https://doi.org/10.13721/j.cnki.dsyy.2020.02.005.
[44] H.M. Yue, R. Ruan, H. Cao, L. Zhou, W. Jiang, S. Li, C.J. Li, Analysis on genetic diversity of cultured rock carp (*Procypris rabaudi*) from Chongqing City based on the mtDNA D-loop sequences, Acta Hydrobiol. Sin. 45 (2021) 846–850, https://doi.org/10.7541/2021.2020.082.
[45] W. Grant, B.W. Bowen, Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation, J. Hered. 89 (1998) 415–426, https://doi.org/10.1093/jhered/89.5.415.
[46] D.S. Wu, Z.G. Wang, X.J. Zhang, D.E. liu, F. Yao, X. Liu, S.P. Xu, Production of clones of homozygous diploid *Carrassius aurratus* red variety, Acta Lab. Anim. Sci. Sin. 9 (2001) 50–55, https://doi.org/10.3969/j.issn.1005-4847.2001.02.010.
[47] M. Xia, X. Mu, D. Luo, Y. Hu, H. Song, Y. Yang, J. Luo, Preliminary study on serum transferrin polymorphisms of three claris catfish, Guangdong Agricultural Sciences 39 (2012) 132–135, https://doi.org/10.3969/j.issn.1004-874X.2012.01.046.
[48] C.S.A. Pereira, M.A. Aboim, P. Ráb, M.J. Collares-Pereira, Introgressive hybridization as a promoter of genome reshuffling in natural homoploid fish hybrids (Cyprinidae, Leuciscinae), Heredity 112 (2014) 343–350, https://doi.org/10.1038/hdy.2013.110.
[49] Y. Tang, M. Feng, X. Zhu, J. Long, Z. Zhou, S. Liu, WR-GLP2, a glucagon-like peptide 2 from hybrid crucian carp that protects intestinal mucosal barrier and inhibits bacterial infection, Fish Shellfish Immunol. 122 (2022) 29–37, https://doi.org/10.1016/j.fsi.2022.01.035.
[50] N.X. Xiong, J. Ou, S.Y. Li, J.H. Zhao, J.F. Huang, K.X. Li, S.W. Luo, S.J. Liu, M. Wen, C. Wu, S. Wang, K.S. Luo, F.Z. Hu, Q.F. Liu, A novel ferritin L (FerL) in hybrid crucian carp could participate in host defense against *Aeromonas hydrophila* infection and diminish inflammatory signals, Fish Shellfish Immunol. 120 (2022) 620–632, https://doi.org/10.1016/j.fsi.2021.12.040.
[51] N.X. Xiong, W.S. Luo, X.Y. Kuang, F. Wang, Z.X. Fang, J. Ou, M.Z. Huang, L.F. Fan, S.W. Luo, S.J. Liu, Gut-liver immune and redox response in hybrid fish (*Carassius cuvieri* ♀ × *Carassius auratus red var.* ♂) after gut infection with *Aeromonas hydrophila*, Comp. Biochem. Physiol. C Toxicol. Pharmacol. 266 (2023) 109553, https://doi.org/10.1016/j.cbpc.2023.109553.
[52] J.V. Neves, J.M. Wilson, P.N.S. Rodrigues, Transferrin and ferritin response to bacterial infection: the role of the liver and brain in fish, Dev. Comp. Immunol. 33 (2009) 848–857, https://doi.org/10.1016/j.dci.2009.02.001.
[53] P. Jurecka, G.F. Wiegertjes, K. Rakus, A. Pilarczyk, I. Irnazarow, Genetic resistance of carp (*Cyprinus carpio* L.) to *Trypanoplasma borreli*: influence of transferrin polymorphisms, Vet. Immunol. Immunopathol. 127 (2009) 19–25, https://doi.org/10.1016/j.vetimm.2008.09.006.
[54] K.J. Liu, Study on the Distribution Patterns and Genetic Diversity of Crucian Carp Populations in Hunan Province, Hunan Agricultural University, Changsha, 2020.
[55] Y. Wang, K. Chen, D. Hong, D. Wu, Development and application of microsatellite DNA molecular markers for the laboratory red crucian carp (*Carassius auratus*) strain C1HD, Aquacult. Int. 30 (2022) 97, https://doi.org/10.1007/s10499-021-00786-3.
[56] S. Wang, N. Jiao, L. Zhao, M.W. Zhang, P. Zhou, X.X. Huang, F.Z. Hu, C.H. Yang, Y. Q. Shu, W.H. Li, C. Zhang, M. Tao, B. Chen, M. Ma, S.J. Liu, Evidence for the paternal mitochondrial DNA in the crucian carp-like fish lineage with hybrid origin, Sci. China Life Sci. 63 (2020) 102–115, https://doi.org/10.1007/s11427-019-9528-1.
[57] S. Wright, The interpretation of population structure by F-statistics with special regard to systems of mating, Evolution 19 (1965) 395–420.