Genetic Diversity Evaluation of Two Loach Fishes and their Artificial Hybrid Population Based on 19 Polymorphic Microsatellite Loci

Yanping Li, Fei Xu, Yongming Wang*, Yunyun Lv, Jinrong Shi, Biwen Xie, Wenyao Cai and Dian Liu

Key Laboratory of Sichuan Province for Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River, College of Life Sciences, Neijiang Normal University, Neijiang 641100, China.

ABSTRACT

Sinibotia superciliaris and *S. reevesae* are two economically important loaches in the middle and upper reaches of Yangtze River. Recently, wild populations of them have dropped sharply due to anthropogenic and environmental threats. There is an increasing need to implement captive propagation programs to protect their populations. Here, we firstly carried out artificial breeding of *S. superciliaris* and *S. reevesae*, and hybrid combinations were designed as *S. superciliaris* ♀ × *S. reevesae* ♂, and *S. superciliaris* ♂ × *S. reevesae* ♀. Secondly, 19 microsatellite loci were used to evaluate genetic diversity between hybrid offspring and their parents. The mean polymorphic information content (*PIC*) among groups ranged from 0.685 to 0.818, implying high polymorphism in all of the four groups (*PIC* > 0.5). The mean observed heterozygosity ranged from 0.779 to 0.887, and the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂ obtained the highest observed heterozygosity, while *S. reevesae* had the lowest. The high heterozygosity was helpful for protecting the genetic resources and preventing the genetic decline, which could reduce the risk of inbreeding due to loss of heterozygosity. Thus, we recommend the hybrid strategy of *S. superciliaris* ♀ × *S. reevesae* ♂ as the potential germplasm resources for further artificial reproduction. Population structure analysis and UPGMA tree indicated *S. superciliaris* and *S. superciliaris* ♀ × *S. reevesae* ♂ were clustered together that comprised of clade I, and *S. reevesae* and *S. superciliaris* ♂ × *S. reevesae* ♀ were clustered together that comprised of clade II, suggesting the close genetic relationship between the offspring and their female parents. This study provides reasonable comparison of genetic diversity and potential application of hybrid strategy of *S. superciliaris* and *S. reevesae*, and the population genetic information supplied here will provide valuable resources for breeding, germplasm resources preservation and conservation genetics of these two loaches.

INTRODUCTION

The *Sinibotia superciliaris* and *S. reevesae* are two loaches endemic to China, which are belong to the family Botiidae (Teleostei: Cypriniformes). *S. superciliaris* is primarily distributed in the main river and tributaries of the upper and middle Yangtze River. Differently, *S. reevesae* is exclusively distributed in the main streams and tributaries along the upper reaches of the Yangtze River, especially in the Tuojiang River (Yang, 1992). These two loaches usually inhabit in rocky, sandy and intermediate substrates with clear water and slow to moderate currents (Yang, 1992),

and have high economic and academic values due to their tender meat, delicious taste, and rich body surface mucus. Recently, their wild populations have declined sharply due to overfishing, water pollution, soil erosion, and loss of original spawning grounds and foraging grounds following completion of the Three Gorges Dam and several other dams along the upper Yangtze River. As a result, *S. reevesae* has been classified as vulnerable (VU) in the Red List by the Endangered Species Scientific Commission of China (Jiang *et al*., 2016). There is increasing need to develop captive propagation to conserve genetic diversity, and to recover the wild populations of the two loaches.

However, studies on these two species are limited and most studies have been restricted to common biology (Li *et al*., 2011; Pu *et al*., 2013; Yang and Ding, 2010), and karyotype analysis (Yue *et al*., 2013). Regarding genetic studies, a previous research related to molecular phylogeny revealed that Botiidae was divided into two monophyletic subfamilies of Leptobotiinae and the Botiinae (Šlechtová *et al*., 2006). The mitochondrial genomes of these two loaches were also released, which provided valuable

genetic markers for investigations on population genetics and conservation biology (Ye *et al*., 2013; Zou *et al*., 2017). Recently, two studies based on high-throughput sequencing developed microsatellite molecular markers in wild populations of *S. superciliaris* and *S. reevesae*, which provided important resources for further population and evolutionary studies (Liu *et al*., 2017; Wang *et al*., 2018).

Fortunately, our laboratory has successfully broken through the artificial reproduction of *S. superciliaris* and *S. reevesae* after conducting many years of researches involved in the gonadal development, reproduction habits, and spawning conditions (Huang *et al*., 2011; Wang *et al*., 2014a; Yue *et al*., 2011). Furthermore, in the early stage of artificial domestication of *S. superciliaris* and *S. reevesae*, we found that the mature males of *S. reevesae* had large testicles and thus could ejaculate a large amount of semen, and the egg fertilization rate was higher compared to the mature males of *S. superciliaris* (Wang *et al*., 2014b). In addition, we found *S. superciliaris* could hybridize with *S. reevesae* and produce fertile offspring (Wang *et al*., 2014b). Previous researches reported that distant hybridization was an important evolutionary mechanism of breeding, which could combine the excellent traits of the parents, increase genetic diversity, obtain excellent new hybrid varieties, increase the fitness and adaptive potential of populations (Lou and Li, 2006; Szűcs *et al*., 2012). Distant hybridization has been widely used in the genetic breeding of fish such as *Megalobrama amblycephala* (Qin *et al*., 2018), *Erythroculter ilishaeformis* (Li *et al*., 2019), *Epinephelus fuscoguttatus* (Chen *et al*., 2018), and *Siniperca scherzeri* (Shan *et al*., 2013), and a certain level of hybrid vigor was detected. Thus, we speculated that the hybridization of *S. superciliaris* and *S. reevesae* may help compensate for the small testis and low sperm of the mature male of *S. superciliaris*, thereby expanding the industrialization and the preservation of germplasm resources of these two loaches. In the present study, we firstly carried out artificial breeding of *S. superciliaris* and *S. reevesae*, and the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂, and *S. superciliaris* ♂ × *S. reevesae* ♀. However, during the process of artificial domestication, it is a common phenomenon that loss of genetic diversity in hatchery populations (Kohlmann *et al*., 2005; Skaala *et al*., 2004). This genetic diversity reduction may lead to negative impacts on the aquaculture industry like decreasing disease-resistance and reproductive capacity (Zhang and Gui, 2018). Therefore, one of the key scientific issues is how to maintain the genetic variation among populations, and it is necessary to exploit an appropriate breeding method based on genetic information to achieve the long-term success of domestication and to protect the genetic resources (Ma *et al*., 2020). Thus, we secondly

evaluated the genetic diversity and population genetic structure between the hybrid offspring and their parents.

Due to high level of polymorphism, codominance, abundance in the genome, relatively small size and rapid analysis protocol, microsatellites, or simple sequence repeats (SSRs), were usually used to investigate population genetic structure, conservation genetics, parentage analyses, and genetic variation of breeding stocks in farmed fish (Chistiakov *et al*., 2006; Ellegren, 2004; Jones *et al*., 2010; Mittal and Dubey, 2009). In addition, SSRs offer high allelic diversity and the relative ease of transfer between closely related species compared with other molecular markers (Guichoux *et al*., 2011).

In the present study, a total of 19 SSR loci were fluorescent-labeled multiple capillary electrophoresis to establish for allelic genotyping in four groups (*S. superciliaris*, *S. reevesae*, *S. superciliaris* ♀ × *S. reevesae* ♂, and *S. superciliaris* ♂ × *S. reevesae* ♀). The purpose of this study is threefold: (1) to evaluate the genetic diversity of *S. superciliaris*, *S. reevesae*, and their hybrid populations; (2) to analyze the genetic relationships and structure among these four groups and (3) to provide scientific and practical suggestions for the breeding program and ensure captive production.

MATERIALS AND METHODS

Sampling and DNA extraction

Firstly, live samples of *S. reevesae* and *S. superciliaris* were collected from the middle and upper reaches of Yangtze River. Subsequently, all the fishes were artificially domesticated at the key laboratory of Sichuan province for fish conservation and utilization in the upper reaches of the Yangtze River, Neijiang, China. The artificial reproduction of the two species was successful after the long-term domestication, and then we carried out artificial reproduction of *S. reevesae*, *S. superciliaris*, and hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂, and *S. superciliaris* ♂ × *S. reevesae* ♀. To estimate their genetic diversity of the offspring and their parents, a total of 96 samples were selected from the four populations comprising of pure F1 of *S. superciliaris* (group A), F1 of *S. reevesae* (group B), F1 of *S. superciliaris* ♀ × *S. reevesae* ♂ (group C), and F1 of *S. superciliaris* ♂ × *S. reevesae* ♀ (group D), the equal amount of 24 samples were randomly selected from the four groups respectively. The male and female parents of the four F1 populations were came from the artificial domestication of *S. superciliaris* and *S. reevesae*, which were cultured in two different independent ponds.

Tissue samples were taken from a small amount of caudal fin and immediately stored at 95% ethanol. After sampling the tissues, the fishes were returned to their

corresponding aquaculture pond. The caudal fins stored in 95% ethanol were used for genomic DNA extraction with DNeasy Blood Tissue Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions.

PCR amplification and genotyping

A total of 19 polymorphic microsatellite loci that can be amplified in both *S. superciliaris*, *S. reevesae* in our previous research (Wang *et al*., 2018) were chosen for genotyping and genetic assessment for this study (Table I). The 5′ ends of the forward primer of all the microsatellite loci were labelled with a FAM fluorochrome, and then used for PCR amplification and genotyping in four groups mentioned above.

Table I. Microsatellite loci of and their primers for genetic diversity evaluation in *S. superciliaris*, *S. reevesae*, and their hybrid populations.

Loci	Primer sequence (5'-3')	Primer 5′ label	Repeat motif	Size/bp	Tm/ ºC	GenBank accession No.
SS17	F: GGAGCAGACTTACAGACTTCACAC R: CACACAGTAATGCGGAGGAAGC	FAM	(AAAC)11	245	59	MF289948
SS19	F: CTCTCCATCATCTTCAGCACCAC R: GTATAGCAGCCCTAGAAGCACC	FAM	(GTTA)10	193	60	KY660315
SS20	F: CCTTCTCCTCTTTTAGGTGTGAC R: GAATCTTCAGAGCAGAGAGTGTG	FAM	(TCCA)10	168	58	KY660316
SS21	F: GGAATCTGGCTGTTGTTAAGCAC R: CGCTGTGGTGTTTCATTGGG	FAM	(AAGA)10	198	58	KY660317
SS24	F: CTTCTACACGAATGGGAAACGG R: GGCCTATGAAGTTATACAGTCCC	FAM	(CACT)10	316	60	KY660318
SS26	F: CCTCTCACAAGCCATGAACGTC R: GTATCAGTGCTGCCCTCTGGTG	FAM	(GATG)10	222	58	KY660320
SS27	F: CCAGTCTGTCAGGCTTATGCAG R: GCTTACAGACAGAATGCCAGCTC	FAM	(ATCT)10	290	58	KY660321
SS29	F: GTAAGTCCGTCAGTATCAGACAC R: CATGAAGAGATTGTATGACCGGC	FAM	(TGAA)10	217	60	MF289952
SS31	F: GAGAAGAGCTGTGCGGTTG R: GAGGAAGAGAGTTACGAGCAG	FAM	(TTCT)10	272	59	MF289954
SS33	F: CACTCATCAGATTTGCACTGGG R: GTCTGAAACAAGATACACTCACTCG	FAM	(TGAG)10	222	60	KY660322
SS34	F: CAACATCCAACATCAGCCCTG R: GCCATGTGTTGACTTTGTGTG	FAM	(GTTA)10	212	58	KY660323
SS37	F: GAACTAATTGCTGTGGAGCGG R: CCTTCTCTGCATAACCCACG	FAM	(ACAG)9	213	56	MF289955
SS38	F: CACCCAGAGACACAACTGTTCC R: GTTGTTGTTCCCAGGGTGTTCC	FAM	(TCTG)8	173	57	MF289956
SS39	F: CTAACCTGCATCTCATTGGCTCC R: GTCTTACGCTACCTACGGCTG	FAM	(TATC)9	236	57	MF289957
SS43	F: CAGAGAGGTGGTGGTAGAGAGG R: CCACAGATTGCAGTTATGTTCCC	FAM	(AATG)8	226	60	MF289959
SS50	F: GAAGAGATGGAGACAGGCGCTC R: CAGCATCATTGCGGAAGATCCAG	FAM	(GATA)8	274	59	MF289960
SS52	F: CGATCCTCTCCACCTCTACACAC R: GTCGCATCTGCTTCACAGCTTC	FAM	(GGAT)9	186	59	MF289961
SS56	F: CACACACGAAGATCCACACCTC R: CAAACGGTGCTCATGTCCTGC	FAM	(ACAA)8	202	59	MF289964
SS57	F: GAACTCTCTCCATCCCTTCCTG F: GAACTCTCTCCATCCCTTCCTG	FAM	(CTGT)9	266	56	MF289965

The final volume of polymerase chain reaction (PCR) was 25 μL, containing 12.5 μL 2×PCR Master Mix, 1.0 μL each primer (10 μM), 1.0 μL genomic DNA (100 ng/ μL) and distilled water up to 25 μL. The thermocycler programs for PCR procedure were conducted as follows: an initial denaturation stage at 94°C for 5 min, a total of 35 cycles of denaturation at 94°C for 30 s, an annealing stage at Tm for 30 s (the Tm was optimized according to different pairs of primers), an extension stage at 72°C for 30 s, and a final extension stage at 72°C for 7 min. PCR products were initially determined when an obvious PCR product was successfully amplified, which ensured the effectiveness of the corresponding primers using a 8% non-denaturing polyacrylamide gel (PAGE), and the pBR322 DNA/MspI molecular weight marker (TIANGEN; Beijing, China) was used as a standard ladder for the assessment to estimate the approximate interval product size. Then the successfully amplified PCR products were quantitatively diluted, and resolved on an ABI 3730 xl DNA analyzer (Applied Biosystems, USA) by multiple capillary electrophoresis method to analyze the number and size of the alleles of each microsatellite locus.

Genetic diversity and population differentiation analysis

GeneMarker software was used to analyze the allelic peaks and size of each microsatellite locus (Holland and Parson, 2011). In each group, the genetic diversity parameters, including the number of alleles (*Na*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), allelic richness (*Rs*), Hardy-Weinberg equilibrium test (*HWE*), and Nei's genetic distance among these four groups were assessed using POPGENE 1.3.1 (Yeh *et al.*, 1997). The polymorphic information content (*PIC*) was calculated using CERVUS 3.0.3 (Kalinowski *et al.*, 2007). The possible presence of null alleles was estimated at a 95% confidence interval using Micro-Checker 2.2.3 (Van Oosterhout *et al.*, 2004). Genetic variation in the two loaches and their hybrid populations were also calculated. Pairwise population fixation indices for F_{ST} values among four groups were performed using ARLEQUIN 3.5 (Excoffier and Lischer, 2010) with 1,000 random permutations. Moreover, analysis of molecular variance (AMOVA) was also computed in ARLEQUIN 3.5.

Genetic structure analysis

First, the UPGMA phylogenetic tree was constructed based on the Nei's genetic distance in MEGA 7 (Pritchard *et al.*, 2000). To further verify the reliability of the clustering relationship, the potential number of genetic clusters and the membership of each individual were further conducted using Bayesian clustering analyses in Structure 2.3.4 (Pritchard *et al.*, 2000). Admixture models were chosen

to assess possible clusters (K value). The lengths of the MCMC iterations were set to 200,000 repetitions with a burn in period of 20,000. The K value range was set to 1–4, and each K was replicated 15 times. The most likely K value was chosen according to peak value of the mean log likelihood [Ln P(X/K)] and the Delta K statistic for a given K using Structure Harvester 0.6.93 (Earl, 2012). The results from 15 replicates of the selected K values were then aligned and conducted by CLUMPP 1.1.2 (Jakobsson and Rosenberg, 2007). The results were finally visualized in DISTRUCT 1.1 (Rosenberg, 2004).

RESULTS

Genetic diversity parameters of the two loaches and their hybrids

All of the 19 microsatellite loci were amplified successfully and showed polymorphism in all groups of the present study. Genetic parameters of these 19 analyzed loci are shown in Table II. For group A of the population *S. superciliaris*, the number of alleles (*Na*) in per locus ranged from 6 to 22, the observed heterozygosity (*Ho*) and expected heterozygosity (*He*) ranged from 0.667 to 1.000 and 0.683 to 0.928, respectively, the allelic richness (*Rs*) ranged from 3.016 to 10.971, and the polymorphic information content (*PIC*) ranged from 0.645 to 0.903. For group B of the population *S. reevesae*, per locus of *Na* ranged from 2 to 19, *Ho* and *He* ranged from 0.042 to 1.000 and 0.042 to 0.922, respectively. *Rs* ranged from 1.043 to 10.286. *PIC* ranged from 0.040 to 0.895, and locus of SS21, and SS56 were lower than 0.5 (Table II). For the group C of the population *S. superciliaris* ♀ × *S. reevesae* ♂, per locus of *Na* fluctuated from 7 to 15, *Ho* and *He* fluctuated from 0.417 to 1.000 and 0.510 to 0.908, respectively, *Rs* fluctuated from 1.997 to 9.000. *PIC* fluctuated from 0.472 to 0.879, and the *PIC* value in all of these 19 was higher than 0.5 except for the locus of SS27. For the group D of the population *S. superciliaris* ♂ × *S. reevesae* ♀, per locus of *Na* fluctuated from 4 to 12, *Ho* and *He* fluctuated from 0.500 to 1.000 and 0.554 to 0.893, respectively, *Rs* fluctuated from 2.186 to 7.945. *PIC* fluctuated from 0.474 to 0.861, and the *PIC* value in all of these 19 was higher than 0.5 except for the locus of SS57 (Table II).

In general, the mean number of alleles per population varied from 8 (Group D of *S. superciliaris* ♂ × *S. reevesae* ♀) to 12 (Group A of *S. superciliaris*). Among the four groups, the lowest mean value of *PIC* (0.685) was observed in the Group B of *S. reevesae*, while the highest (0.818) was appeared in the group A of *S. superciliaris*. The lowest mean *Ho* (0.779) was observed in the Group B of *S. reevesae*, while the highest mean *Ho* (0.887) was observed in the Group C of *S. superciliaris* ♀ × *S. reevesae* ♂ (Table II).

Table II. Summary of genetic diversity based on 19 microsatellite loci in four groups including *S. superciliaris*, *S. reevesae*, and their hybrid populations.

Groups	Na	Ho	He	PIC	Rs	P_{HWE}
S. superciliaris						
SS17	7	0.917	0.796	0.752	4.535	0.966
SS19	17	0.875	0.927	0.901	10.688	0.299
SS20	14	0.958	0.914	0.886	9.521	0.460
SS21	17	1.000	0.915	0888	9.600	0.135
SS24	22	0.875	0.928	0.903	10.971	0.000*
SS26	7	0.792	0.780	0.750	4.608	0.433
SS27	11	0.792	0.683	0.645	3.016	0.857
SS29	6	0.667	0.778	0.726	4.189	0.065
SS31	14	0.958	0.903	0.873	8.597	0.031
SS33	14	0.958	0.905	0.876	8.794	0.816
SS34	11	0.917	0.883	0.850	7.385	0.811
SS37	12	0.750	0.881	0.849	7.291	0.088
SS38	9	0.958	0.850	0.813	5.969	0.509
SS39	13	0.875	0.880	0.848	7.245	0.953
SS43	10	0.750	0.895	0.863	8.056	0.031
SS50	15	0.958	0.919	0.892	10.017	0.715
SS52	10	0.958	0.832	0.791	5.383	0.943
SS56	9	0.708	0.700	0.645	3.182	0.011
SS57	10	0.917	0.840	0.799	5.620	0.960
Mean	12	0.873	0.854	0.818	7.097	-
S. reevesae						
SS17	6	0.792	0.768	0.712	4.028	0.003
SS19	14	0.917	0.907	0.878	8.930	0.096
SS20	8	0.917	0.840	0.800	5.647	0.039
SS21	5	0.833	0.571	0.462	2.268	0.348
SS24	9	0.917	0.871	0.835	6.777	0.028
SS26	3	0.833	0.646	0.555	2.723	0.035
SS27	10	0.708	0.704	0.657	3.218	0.123
SS29	5	0.542	0.560	0.500	2.215	0.009
SS31	14	0.875	0.891	0.861	7.937	0.028
SS33	16	1.000	0.916	0.888	9.681	0.001
SS34	12	0.917	0.898	0.867	8.288	0.529
SS37	3	0.750	0.649	0.561	2.743	0.042
SS38	9	0.708	0.599	0.569	2.420	1.000
SS39	16	0.917	0.922	0.895	10.286	0.028
SS43	6	0.667	0.684	0.614	3.032	0.837
SS50	19	0.750	0.920	0.895	10.105	0.000*
SS52	8	0.958	0.838	0.797	5.565	0.471
SS56	2	0.042	0.042	0.040	1.043	1.000
SS57	5	0.750	0.707	0.630	3.254	0.000*
Mean	9	0.779	0.733	0.685	5.266	-

Groups	Na	Ho	He	PIC	Rs	P_{HWE}
***S. superciliaris* ♀ × *S. reevesae* ♂**						
SS17	7	0.833	0.742	0.691	3.657	0.008
SS19	13	1.000	0.887	0.856	7.629	0.000*
SS20	10	0.870	0.895	0.862	8.015	0.009
SS21	14	0.913	0.838	0.800	5.539	0.994
SS24	12	0.917	0.831	0.793	5.358	0.163
SS26	8	0.958	0.826	0.784	5.236	0.015
SS27	7	0.417	0.510	0.472	1.997	0.693
SS29	8	0.875	0.820	0.775	5.075	0.656
SS31	11	0.958	0.895	0.863	8.056	0.238
SS33	10	1.000	0.845	0.807	5.789	0.010
SS34	10	1.000	0.875	0.841	6.982	0.081
SS37	11	0.958	0.770	0.737	4.056	0.991
SS38	8	0.958	0.805	0.757	4.721	0.515
SS39	14	0.792	0.886	0.854	7.529	0.000*
SS43	10	0.958	0.861	0.825	6.365	0.672
SS50	15	0.708	0.908	0.879	9.000	0.001
SS52	8	0.957	0.854	0.815	6.081	0.050
SS56	7	0.870	0.645	0.599	2.713	0.913
SS57	9	0.917	0.852	0.814	6.031	0.420
Mean	10	0.887	0.818	0.780	5.781	-
***S. superciliaris* ♂ × *S. reevesae* ♀**						
SS17	6	0.875	0.813	0.766	4.902	0.025
SS19	12	0.875	0.875	0.841	6.982	0.177
SS20	10	1.000	0.879	0.847	7.200	0.002
SS21	8	1.000	0.828	0.786	5.284	0.198
SS24	9	0.667	0.761	0.718	3.918	0.000*
SS26	5	0.875	0.808	0.757	4.780	0.062
SS27	8	0.875	0.796	0.748	4.535	0.229
SS29	5	0.958	0.679	0.627	2.985	0.036
SS31	11	1.000	0.893	0.861	7.945	0.035
SS33	10	0.917	0.875	0.841	6.982	0.007
SS34	10	0.958	0.854	0.817	6.095	0.060
SS37	8	0.792	0.760	0.716	3.905	0.000*
SS38	6	0.875	0.774	0.721	4.129	0.099
SS39	11	0.917	0.834	0.780	5.460	0.000*
SS43	6	0.875	0.718	0.663	3.368	0.578
SS50	10	0.792	0.881	0.848	7.291	0.000*
SS52	5	1.000	0.724	0.661	3.439	0.067
SS56	4	0.750	0.563	0.505	2.228	0.232
SS57	4	0.500	0.554	0.474	2.186	0.456
Mean	8	0.868	0.783	0.736	4.927	-

Note: *Na*, indicated the number of alleles; *Ho*, indicated observed heterozygosity; *He*, indicated expected heterozygosity; *PIC*, indicated the polymorphism information content; *Rs*, indicated allelic richness; and P_{HWE}, indicated significance of deviation from Hardy-Weinberg equilibrium at P-levels 0.001 (***).

Hardy-Weinberg balance analysis results showed that locus of SS24 in *S. superciliaris* and loci of SS50 and SS57 in *S. reevesae* exhibited departure from *HWE* after applying a Bonferroni correction (*P*<0.001). Loci of SS19 and SS39 in *S. superciliaris* ♀ × *S. reevesae* ♂, and loci of SS24, SS37, SS39, and SS50 in *S. superciliaris* ♂ × *S. reevesae* ♀ were also exhibited departure from *HWE*. For the estimation of null allele, SS50 in population of *S. reevesae* and *S. superciliaris* ♀ × *S. reevesae* ♂ showed evidence for a null allele. No locus showed evidence for a null allele in the population of *S. superciliaris* and *S. superciliaris* ♂ × *S. reevesae* ♀.

Population differentiation of the two loaches and their hybrids

Results of AMOVA analysis of *S. superciliaris*, *S. reevesae* and their hybrid populations showed that the variation among populations was 7.97%, variation among individuals within population was -6.68%, and variation within individuals was 98.71% (Table III). Most of the variance was due to variation within individuals.

Table III. Hierarchical analysis of molecular variance (AMOVA) for *S. superciliaris*, *S. reevesae*, and their hybrid populations based on 19 microsatellite loci.

| Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)
Among populations	3	114.807	0.65185 Va	7.97
Among individuals within population	92	642.208	-0.54620 Vb	-6.68
Within individuals	96	775.000	8.07292 Vc	98.71
Total	191	1532.015	8.17857	100

Nei's genetic distance (Ds) between groups varied from 0.3585 to 0.6575 (Table IV), suggesting the studied populations were in interspecific inheritance (0.1 < Ds < 2.0) (Thorpe, 1982). The genetic differentiation index (F_{st}) varied from 0.0545 to 0.1069 (Table IV), indicating all groups were moderately differentiated (0.05 < F_{st} < 0.15) (Wright, 1978).

Table IV. Matrix of pair-wise F_{st} values (above diagonal) and Nei's genetic distance (below diagonal) between groups of two loaches and their hybrid populations.

Pop ID	Group A	Group B	Group C	Group D
Group A	****	0.1069	0.0545	0.0709
Group B	0.6575	****	0.0845	0.0764
Group C	0.4461	0.4338	****	0.0763
Group D	0.4947	0.3585	0.4765	****

Population genetic structure of the two loaches and their hybrids

The results of UPGMA clustering tree based on Nei's genetic distance showed that group A (*Sinibotia superciliaris*) and group C (*S. superciliaris* ♀ × *S. reevesae* ♂) were firstly clustered together that comprised of clade I, group B (*S. reevesae*) and group D (*S. superciliaris* ♂ × *S. reevesae* ♀) were then gathered together that comprised of clade II (Fig. 1).
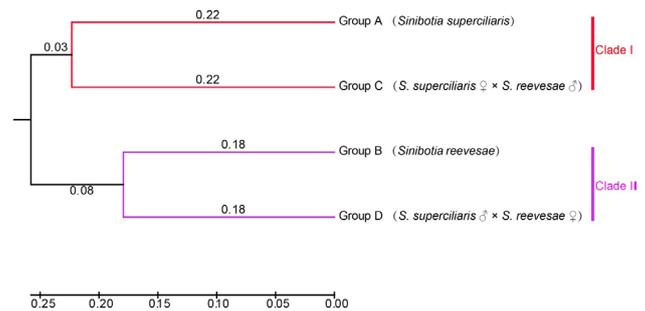
Fig.1. UPGMA clustering tree based on Nei's genetic distance among four groups of two loach fishes and their artificial hybrid population. The values above the branch represented branch length.

The difference among these four groups were also corroborated by Bayesian clustering, which reflected the presence of two clusters (K= 2; Fig. 2a) with a clear divide between group A, C, and group B, D. The genetic structure of Bayesian clustering analyses was consist with the UPGMA clustering tree. The mean Ln P(K) and ΔK change with the increase of the number of subgroups K. When K=2, mean Ln P(K) had the maximum inflection point and the maximum ΔK. Therefore, it was considered that K= 2 was the most probable model, and it was speculated that the 4 populations could be divided into 2 subgroups (Fig. 2b). However, isolation between the two clusters including group A, C, and group B, D was not complete, and some individuals from both clades showed mixed origin, especially for the hybrid populations.

DISCUSSION

Genetic diversity of the two loaches and their hybrids

In this study, a total of 19 microsatellite loci were successfully amplified and exhibited polymorphism in the four artificial breeding populations of *S. reevesae*, *S. superciliaris*, and their hybrid populations, implying that the microsatellite markers we developed in the previous research were effective (Wang *et al*., 2018). Among the 19 loci, only locus SS24 was deviated from *HWE* in Group A of *S. superciliaris*. Locus SS50 and SS57 in Group B

of *S. reevesae*, and locus SS19 and SS39 in Group C of *S. superciliaris* ♀ × *S. reevesae* ♂ were also deviated from *HWE*. Differently, Group D of *S. superciliaris* ♂ × *S. reevesae* ♀ have four loci (SS24, SS37, SS39, and SS50) that deviated from *HWE*. Among these loci, SS24, SS39 and SS50 deviated from *HWE* at least in two groups, and thus they should be used with caution in subsequent population genetic analyses. The potential presence of null alleles were observed at one locus (SS50) in two groups, including *S. reevesae* and *S. superciliaris* ♀ × *S. reevesae* ♂, and which may be caused by a mutation at the primer binding sites leading to amplification failure (Lehmann *et al.*, 1996). Thus, the deviation from *HWE* at SS50 can probably be explained by the presence of null alleles that mentioned in previous research (Li *et al.*, 2014). The reasons why other loci deviated from *HWE* need to be further explored.
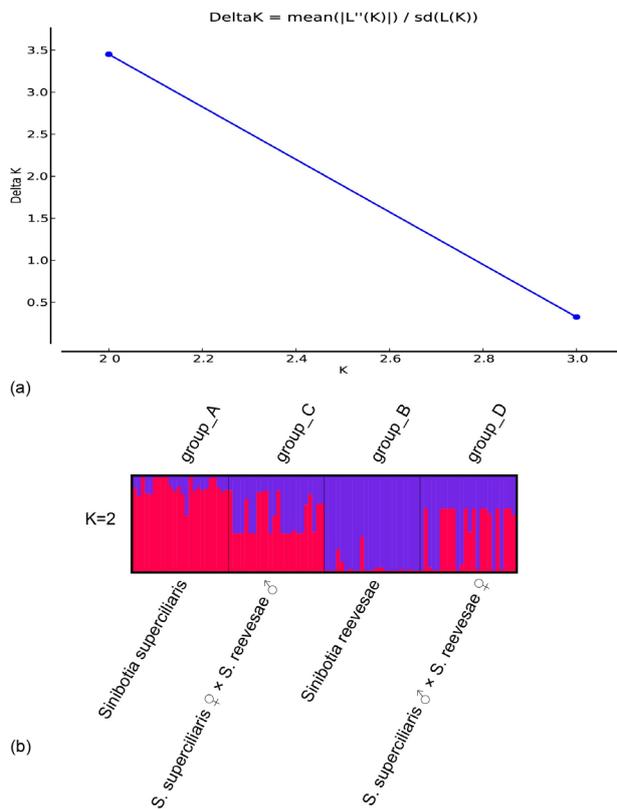

(a)

(b)

Fig.2. Results of Bayesian analysis (structure) of the two loaches and their hybrids offspring based on 19 microsatellite loci. (a) Delta K values from structure analysis; (b) Structure clustering analysis with K=2.

The number of alleles (*Na*), observed heterozygosity (*Ho*), and polymorphism information content (*PIC*) are three important parameters used to study population genetic diversity of organisms. Here, based on the analysis of 19 microsatellite loci, the mean values of *PIC* values were more than 0.5 in all groups. In detail, the mean *PIC* was 0.818, 0.685, 0.780, and 0.736 in *S. superciliaris*, *S. reevesae*, *S. superciliaris* ♀ × *S. reevesae* ♂, and *S. superciliaris* ♂ × *S. reevesae* ♀, respectively. Previous research indicated the locus was high polymorphic when *PIC* > 0.50 (Botstein *et al.*, 1980), the mean *PIC* in the four groups here was ranged from 0.685 to 0.818, suggesting high genetic diversity of the four groups in this study. The *PIC* in the wild populations of these two loaches also showed a high level of polymorphism, and the mean *PIC* in *S. superciliaris* and *S reevesae* was 0.54 and 0.5, respectively (Wang *et al.*, 2018). Compared with the wild populations, the culture F1 populations of these two loaches and their hybrid progeny of F1 here had also maintained high genetic diversity, indicating the domestication process of these two loaches were considered to be in the early stage, and the same pattern also detected in the red-tail catfish (Zhou *et al.*, 2021). The diversity of the both hybrid combinations were higher than the group B of *S. reevesa*, which may be the embodiment of the hybridization vigor among species. Furthermore, observed heterozygosity can represent the degree of individual genetic variation within a population, and the higher heterozygosity value signifies greater variation and higher genetic diversity. In the present research, the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂ obtained the highest value of *Ho* (0.887) compared to other three groups, which indicated its highest genetic diversity. In general, heterozygosity and genetic diversity of the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂ are higher than those of that in the combination of *S. superciliaris* ♂ × *S. reevesae* ♀, and the former has fewer loci deviating from *HWE*. Distant hybridization can effectively increase the genetic structure variation of the population and improve the inheritance of the population diversity. Thus, in the subsequent artificial breeding, it is recommended that the hybrids of *S. superciliaris* ♀ × *S. reevesae* ♂ can be utilized as the potential germplasm resources.

Genetic differentiation and genetic structure of populations

All of the four populations had moderate differentiation levels, and variation within individuals were the primary origin of variation. Genetic distance and genetic identity index were indicators that reflected the genetic relationship of the population. Comparing the Nei's genetic distance between *S. reevesae* and the hybrid combinations of *S. superciliaris* ♀ × *S. reevesae* ♂ and *S. superciliaris* ♂ × *S. reevesae* ♀, it was indicated that the Nei's genetic distance between *S. reevesae* and the hybrid combination of *S. superciliaris* ♂ × *S. reevesae* ♀

was the closest (0.3585), in comparison with the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂ (0.4338). Moreover, *S. superciliaris* also showed the similar pattern, with a closer genetic distance (0.4461) to *S. superciliaris* ♀ × *S. reevesae* ♂, in comparison with the combination of *S. superciliaris* ♂ × *S. reevesae* ♀ (0.4947). UPGMA clustering tree with Nei's genetic distance revealed two clusters of the four populations, which was consistent with the results from the Bayesian clustering analyses in structure. Meanwhile, gene flow was found among multiple individuals of the hybrid populations and their parents. In summary, although the hybrid offspring inherited the genetic characteristics of both parents, they were more closely related to the female parents and a closer genetic relationship was found between the offspring populations and their female parents, and a similar pattern was obtained in the interspecific hybridization research of *Quasipaa* (Zhang *et al.*, 2019).

Distant hybrization can integrate the genome of parents into the hybrid offspring to produce rich genetic variation. Thus, distant hybridization plays an important role in increasing species fitness and adaptive potential of populations. In this study, hybrid combinations of *S. superciliaris* ♀ × *S. reevesae* ♂ and *S. superciliaris* ♂ × *S. reevesae* ♀ obtained healthy crossbred offsprings. In our previous study, we observed the F1 embryonic development of *S. superciliaris* ♀ × *S. reevesae* ♂ showed obvious heterosis and could grow to be mature and reproduce normally to produce F2 hybrids (Wang *et al.*, 2014b). Here, after comparing the analysis of the genetic diversity indicators such as the number of alleles, polymorphic information content, and especially the observed heterozygosity of the two loaches and their hybrid populations, we observed the genetic diversity of hybrid offspring population was rich, especially for *S. superciliaris* ♀ × *S. reevesae* ♂ that presented the highest *Ho* than the other three populations, implying the highest genetic diversity in the population of *S. superciliaris* ♀ × *S. reevesae* ♂.

However, along with breeding environment and feeding ground destructions, and overfishing of *S. superciliaris* and *S. reevesae*, the wild resources of the two species would be severely declined. Thus, there is an increasing need to develop captive breeding to protect genetic diversity, support the strengthening of populations, and aid the two loaches recovery. In the artificial reproduction process, the shortage including the sparse of semen volume in *S. superciliaris* and low number of eggs amount in *S. reevesae* will block large amount of offspring reproduction. Artificial hybrids success of the two *S. superciliaris* and *S. reevesae* will overcome the shortage. Based on our results, we recommended the hybrid strategy of *S. superciliaris* ♀ × *S. reevesae* ♂ would be a suitable route, which displayed the highest genetic diversity compared with other strategies. In addition, the hybrid individuals of *S. superciliaris* ♀ × *S. reevesae* ♂ could be considered as a potentially alternative economic fish of *S. superciliaris* and *S. reevesae*, thus it would benefit the natural resource conversation of these two species.

CONCLUSIONS

In conclusion, four groups presented high polymorphic due to the average *PIC* values in all of them were greater than 0.5. Furthermore, we found that the hybrid combination of *S. superciliaris* ♀ × *S. reevesae* ♂ presented the highest *Ho* than the other three groups. The high heterozygosity was useful to conserve the genetic resources and prevent the genetic decline, which could reduce the risk of inbreeding due to loss of heterozygosity. The hybrid strategy of *S. superciliaris* ♀ × *S. reevesae* ♂ had great potentials for artificially propagated germplasm resources. Thus, we recommended the hybrid strategy of *S. superciliaris* ♀ × *S. reevesae* ♂ would be more suitable for further artificial reproduction compared with other strategies. This study provides reasonable comparison of genetic diversity and potential application of hybrid strategy of *S. superciliaris* and *S. reevesae*, and the captive propagation programs can aid in the recovery of *S. superciliaris* and *S. reevesae* by providing individuals that can be used to supplement wild populations, and to produce a hybrid strain for captive production to take the strain off the harvesting of wild populations. These data described in this study provided a foundation and guidelines for wild population protection and breeding population construction in future breed improvement of these two loaches.

ACKNOWLEDGMENTS

Statement of conflict of interest
The authors have declared no conflict of interest.

REFERENCES

Botstein, D., White, R.L., Skolnick, M., and Davis,

R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, **32**: 314.

Chen, X., Shao, T., Long, X., Esposito, A., 2018. Evaluation of the effects of different stocking densities on the sediment microbial community of juvenile hybrid grouper (♀*Epinephelus fuscoguttatus* × ♂ *Epinephelus lanceolatus*) in recirculating aquaculture systems. *PLoS One*, **13**: e0208544. https://doi.org/10.1371/journal.pone.0208544

Chistiakov, D.A., Hellemans, B., and Volckaert, F.A., 2006. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture*, **255**: 1-29. https://doi.org/10.1016/j.aquaculture.2005.11.031

Earl, D.A., 2012. Structure harvester: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.*, **4**: 359-361. https://doi.org/10.1007/s12686-011-9548-7

Ellegren, H., 2004. Microsatellites: Simple sequences with complex evolution. *Nat. Rev. Genet.*, **5**: 435-445. https://doi.org/10.1038/nrg1348

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

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., and Salin, F., 2011. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.*, **11**: 591-611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Holland, M.M., and Parson, W., 2011. GeneMarker® HID: A reliable software tool for the analysis of forensic STR data. *J. Forensic Sci.*, **56**: 29-35. https://doi.org/10.1111/j.1556-4029.2010.01565.x

Huang, Y., Yue, X., Wang, F., Gan, L., Xie, B., Wang, Y., Qi, Z., and Ge, Z., 2011. Studies on the individual fecundity of *Botia reevesae* in the Tuojiang River, a Tributary River of the Yangtze River. *Sichuan J. Zool.*, **30**: 916-920. (in Chinese).

Jakobsson, M., and Rosenberg, N.A., 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**: 1801-1806. https://doi.org/10.1093/bioinformatics/btm233

Jiang, Z., Jiang, J., Wang, Y., Zhang, E., Zhang, Y., Li, L., Xie, F., Cai, B., Cao, L., and Zheng, G., 2016. Red list of China's vertebrates. *Biodivers. Sci.*, **24**: 500. (in Chinese).

Jones, A.G., Small, C.M., Paczolt, K.A., and Ratterman, N.L., 2010. A practical guide to methods of parentage analysis. *Mol. Ecol. Resour.*, **10**: 6-30. https://doi.org/10.1111/j.1755-0998.2009.02778.x

Kalinowski, S.T., Taper, M.L., and Marshall, T.C., 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.*, **16**: 1099-1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Kohlmann, K., Kersten, P., Flajšhans, M., 2005. Microsatellite-based genetic variability and differentiation of domesticated, wild and feral common carp (*Cyprinus carpio* L.) populations. *Aquaculture*, **247**: 253-266. https://doi.org/10.1016/j.aquaculture.2005.02.024

Lehmann, T., Hawley, W.A., and Collins, F.H., 1996. An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics*, **144**: 1155-1163. https://doi.org/10.1093/genetics/144.3.1155

Li, Q., Yao, M., Zhou, B., Chen, X., Long, Y., Zhou, C., and Li, M., 2011. Technique of artificial reproduction of *Botia superciliaris* Günther. *Freshw. Fish.*, **41**: 92-95. https://doi.org/10.1007/s11056-010-9213-0

Li, S., Xie, L., Xiao, J., Yuan, L., and Liu, S., 2019. Diploid hybrid fish derived from the cross between female Bleeker's yellow tail and male topmouth culter, two cyprinid fishes belonging to different subfamilies. *BMC Genet.*, **20**: 1-13. https://doi.org/10.1186/s12863-019-0781-5

Li, Y., Wang, J., Zhao, K., Xie, B., and Peng, Z., 2014. Isolation and characterization of 16 polymorphic microsatellite loci for *Euchiloglanis kishinouyei*. *J. Fish Biol.*, **85**: 972-977. https://doi.org/10.1111/jfb.12472

Liu, S., Wang, Y., Xie, B., and Peng, Z., 2017. Twelve novel microsatellite markers for the Chinese golden loach *Sinibotia superciliaris* (Günther 1892). *J. appl. Ichthyol.*, **33**: 821-823. https://doi.org/10.1111/jai.13395

Lou, Y., and Li, X., 2006. Distant hybridization of fish and its application in aquaculture in China. *J. Fish. Sci. China*, **13**: 151-158.

Ma, Q., Wu, B., Jiang, J., and Song, Z., 2020. Genetic characterization of selected domestic populations of channel catfish (*Ictalurus punctatus*) using microsatellites. *Pakistan J. Zool.*, **52**: 1683-1689.

Mittal, N., and Dubey, A., 2009. Microsatellite

markers-A new practice of DNA based markers in molecular genetics. *Phcog. Rev.*, **3**: 235-246.

Pritchard, J.K., Stephens, M., and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945-959. https://doi.org/10.1093/genetics/155.2.945

Pu, J., Jia, L., Su, S., Wang, H., Sun, F., and Yao, W., 2013. Age and growth of *Botia superciliaris* in Hejiang section of the Yangtze River. *Freshw. Fish.*, **43**: 38-43.

Qin, Q., Liu, C., Wang, Y., Li, R., and Liu, S., 2018. Rapid genomic and genetic changes in the first generation of autotetraploid lineages derived from distant hybridization of *Carassius auratus* Red Var. (♀)× *Megalobrama amblycephala* (♂). *Mar. Biotechnol.*, **21**: 139-149. https://doi.org/10.1007/s10126-018-9859-8

Rosenberg, N.A., 2004. Distruct: A program for the graphical display of population structure. *Mol. Ecol. Notes*, **4**: 137-138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Shan, H., Liang, X.F., Sun, J., Li, L., Yu, Y., Huang, W., Qu, C., Cao, L., Bai, X., and Tao, Y., 2013. Insights into food preference in hybrid F1 of *Siniperca chuatsi* (♀)× *Siniperca scherzeri* (♂) mandarin fish through transcriptome analysis. *BMC Genom.*, **14**: 1-11. https://doi.org/10.1186/1471-2164-14-601

Skaala, Ø., Høyheim, B., Glover, K., and Dahle, G., 2004. Microsatellite analysis in domesticated and wild Atlantic salmon (*Salmo salar* L.): Allelic diversity and identification of individuals. *Aquaculture*, **240**: 131-143. https://doi.org/10.1016/j.aquaculture.2004.07.009

Šlechtová, V., Bohlen, J., Freyhof, J., and Ráb, P., 2006. Molecular phylogeny of the Southeast Asian freshwater fish family Botiidae (Teleostei: Cobitoidea) and the origin of polyploidy in their evolution. *Mol. Phylogenet. Evol.*, **39**: 529-541. https://doi.org/10.1016/j.ympev.2005.09.018

Szűcs, M., Eigenbrode, S.D., Schwarzländer, M., and Schaffner, U., 2012. Hybrid vigor in the biological control agent, *Longitarsus jacobaeae*. *Evol. Appl.*, **5**: 489-497. https://doi.org/10.1111/j.1752-4571.2012.00268.x

Thorpe, J.P., 1982. The molecular clock hypothesis: Biochemical evolution, genetic differentiation and systematics. *Annu. Rev. Ecol. Evol. Syst.*, **13**: 139-168. https://doi.org/10.1146/annurev.es.13.110182.001035

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P., and Shipley, P., 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes*, **4**: 535-538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Wang, Y., Liu, S., Peng, Z., and Xie, B., 2018. Characterization of 27 novel microsatellite loci in *Sinibotia superciliaris* (Günther 1892) and cross-species transferability in *Sinibotia reevesae* (Chang 1944). *J. appl. Ichthyol.*, **34**: 216-220. https://doi.org/10.1111/jai.13539

Wang, Y., Xie, B., Wang, F., Yue, X., Wang, Y., Qi, Z., Wu, Y., and Shen, S., 2014a. Reproductive characteristics of *Sinibotia reevesae* in the Tuojiang River. *Chin. J. Zool.*, **49**: 699-706. (In Chinese).

Wang, Y., Xie, B., Yue, X., Wang, F., Qi, Z., Wang, Y., and Wu, Y., 2014b. Embryonic development of hybrid of *Sinibotia superciliaris* (♀) and *Sinibotia reevesae* (♂). *Sichuan J. Zool.*, **33**: 90-93. (In Chinese).

Wright, S., 1978. *Variability within and among natural populations*. University of Chicago Press, Chicago, IL.

Yang, J., 1992. Revision of the subgenus *Botia* (*Sinibotia*) with description of a new species (Cypriniformes: Cobitidae). *Ichthyol. Explor. Freshw.*, **2**: 341-349.

Yang, M., and Ding, X., 2010. Study on the reproductive biology of *Botia supercilliaris*. *J. Hydroecol.*, **2**: 38-41.

Ye, H., Huang, Y., Wang, J., Peng, Z., and Zhang, Y., 2013. Determination and comparison of two complete mitochondrial genomes of golden Chinese loach, *Sinibotia superciliaris* (Teleostei, Cypriniformes). *Mitochondr. DNA*, **24**: 34-36. https://doi.org/10.3109/19401736.2012.716053

Yeh, F.C., Yang, R., Boyle, T.B., Ye, Z., and Mao, J., 1997. Popgene, the user-friendly shareware for population genetic analysis. *Corpus (Molecular biology and Biotechnology Centre, University of Alberta, Canada),* **10**: 295-301.

Yue, X., Wang, F., Xie, B., Qi, Z., Wang, Y., He, B., and Chen, X., 2011. Embryonic development of *Botia reevesae* in Tuojiang River. *Sichuan J. Zool.*, **30**: 390-395. (In Chinese).

Yue, X., Zou, Y., Liu, J., Wang, Y., Wang, F., Xie, B., and Qi, Z., 2013. Chromosome karyotype of *Sinibotia superciliaris*. *Sichuan J. Zool.*, **3**: 017. (In Chinese).

Zhang, D., Fei, X., Kong, S., Zhang, Q., and Zheng, R., 2019. Interspecific hybridization of Quasipaa and genetic characteristics of hybrid tadpole. *PeerJ.*, **7**: e27814. https://doi.org/10.7287/peerj.preprints.27814v1

Zhang, Q., and Gui, J., 2018. Diversity, evolutionary contribution and ecological roles of aquatic viruses.

Sci. China Life Sci., **61**: 1486-1502. https://doi.org/10.1007/s11427-018-9414-7

Zhou, Y., Wu, J., Wang, Z., Li, G., Zhou, L., and Gui, J., 2021. Microsatellite polymorphism and genetic differentiation of different populations screened from genome survey sequencing in red-tail catfish (*Hemibagrus wyckioides*). *Aquacult. Rep*., **19**: 100614. https://doi.org/10.1016/j.aqrep.2021.100614

Zou, Y., Xie, B., Qin, C., Wang, Y., Yuan, D., Li, R., and Wen, Z., 2017. The complete mitochondrial genome of a threatened loach (*Sinibotia reevesae*) and its phylogeny. *Genes Genom*., **39**: 767-778. https://doi.org/10.1007/s13258-017-0541-8

Online First Article