GMR

Analysis of genetic diversity of *Leuciscus leuciscus baicalensis* using novel microsatellite markers with cross-species transferability

D.J. Lei[1], G. Zhao[1], P. Xie[1], Y. Li[1], H. Yuan[1], M. Zou[1], J.G. Niu[2] and X.F. Ma[1]

[1]College of Fisheries, Huazhong Agricultural University, Wuhan, China
[2]Xinjiang Fisheries Research Institute, Urumqi, China

Corresponding author: X.F. Ma
E-mail: xufama@mail.hzau.edu.cn

Genet. Mol. Res. 16 (2): gmr16029376
Received September 23, 2016
Accepted March 8, 2017
Published May 4, 2017
DOI http://dx.doi.org/10.4238/gmr16029376

ABSTRACT. We used next-generation sequencing technology to characterize 19 genomic simple sequence repeat (SSR) markers and 11 expressed sequence tag (EST) SSR markers from *Leuciscus leuciscus baicalensis*, a small freshwater fish that is widely distributed in Xinjiang, China. Primers were used to test for polymorphisms in three *L. leuciscus baicalensis* populations in Xinjiang. There were 4-27 (average 11.3) alleles (N_A), the expected heterozygosity (H_E) was 0.36-0.94 (average 0.75 ± 0.14), the observed heterozygosity (H_O) was 0.37-1.00 (average 0.68 ± 0.18), and the polymorphism information content (PIC) was 0.31-0.93 (average 0.71). The averages of H_E and PIC for the EST-SSR markers were slightly lower than for the genomic SSR markers. Genetic analysis of the three populations showed similar results for PIC, H_E, and N_A. Amplifications were performed in nine other species; the top three transferability values were for *Rutilus lacustris* (80%), *Leuciscus idus* (76.7%), and *Phoxinus ujmonensis* (63.3%), with the following average values: PIC (0.56, 4.46, and 0.52); N_A (0.40, 3.00, and 0.32); and H_O (0.44, 2.74, and 0.22), respectively. *L. leuciscus*

baicalensis is one of the most important commercial fish in Xinjiang, but in recent years, fishery resources have decreased sharply owing to water conservation projects, unreasonable utilization, and invasion by alien species. These novel SSR markers are appropriate for studies involving fingerprinting, gene flow, genetic diversity, population structure, and molecular-assisted breeding, and could contribute to the conservation of *L. leuciscus baicalensis*.

Key words: *Leuciscus leuciscus baicalensis*; GSSR markers; EST-SSR markers; Genetic diversity; Cross-species transferability

INTRODUCTION

Leuciscus leuciscus baicalensis, a freshwater species of the Cyprinidae family (subfamily Leuciscinae), is widely distributed in Xinjiang, China (Huo et al., 2011). A few studies have been conducted on the artificial reproduction and development of *L. leuciscus baicalensis* and *Leuciscus idus* (Nowosad et al., 2014; Witeska et al., 2014; Siddique et al., 2016). Moreover, Chang et al. (2014) and Cui et al. (2015) have reported the differential gene expression of *Leuciscus waleckii.* Their studies indicate that transcriptome changes play a role in spawning migration and in acid-base homeostasis in fish under alkaline stress. *L. leuciscus baicalensis* is one of the most important commercial fish in Xinjiang, but in recent years, fishery resources have fallen sharply owing to water conservation projects, unreasonable utilization, and invasion by alien species. Therefore, researchers have begun to explore artificial domestication. Studies on the management of cultivation conditions and the impact of fishery drugs on fry seem to suggest that artificial culture is not difficult and could be popularized (Liu et al., 2015a; Lin and Tang, 2016).

Codominant simple sequence repeats (SSRs) are regarded as one of the most effective molecular markers for the examination of genetic diversity within and between populations, and they provide abundant genetic information. With the rapid development of next-generation sequencing (NGS) technology, both expressed sequence tag (EST)-SSRs and genomic SSRs (gSSRs) can be obtained cheaply and efficiently (Gao et al., 2012; Zheng et al., 2013; Liu et al., 2015b). In the present study, we developed novel SSR markers from both genomic DNA libraries and EST libraries, and focused on *L. leuciscus baicalensis* sampled from the Irtysh River, even though Dubut et al. (2009) have generated some polymorphic SSR markers based on European *L. leuciscus baicalensis*. According to the decryption of the evolutionary history and the genetic differentiation of the subfamily (Costedoat et al., 2006; Boron et al., 2009; Perea et al., 2010; Hu et al., 2015), Leuciscinae species, particularly *L. leuciscus baicalensis* and *L. idus*, are assumed to be closely related. Therefore, in the present study we tested the cross utility of polymorphic SSR primers mined from *L. leuciscus baicalensis* in other species to investigate transferability.

These novel SSR markers provide useful information for phylogenetic analysis and studies on population genetics. Moreover, they are appropriate for studies involving fingerprinting, gene flow, genetic diversity, population structure, germplasm characterization research, and molecular-assisted breeding in *L. leuciscus baicalensis* and related species.

MATERIAL AND METHODS

Fish materials and extraction of genomic DNA

Ninety-six specimens of *L. leuciscus baicalensis* comprising three populations were collected randomly using gill nets from the natural river systems of Xinjiang Province, including the tributary streams of the Habahe River (HBH: N48°04'546", E086°20'686") and the Buerjin River (BEJ: N47°42'875", E086°50'169"), and the Beiwan section of the main stream (BW: N48°01'486", E085°33'060"). All samples were examined and classified according to Ren et al. (2002). Based on previous studies on the effects of sample size on genetic diversity estimates in populations using SSR markers (Yan and Zhang, 2004; Pruett and Winker, 2008; Ou et al., 2009), we determined that approximately 30 individuals for each population was sufficient. The samples were preserved in 95% ethanol. Total genomic DNA was isolated from the fins by proteinase K digestion followed by the standard phenol/chloroform method, and visualized on a 1.5% agarose gel (Wang et al., 2011).

Primer design, SSR marker development, and detection

The assembled contigs and expressed sequences (from unpublished data) from NGS were used to detect SSR loci with a Perl script known as Microsatellite (MISA, http://pgrc.ipk-gatersleben.de/misa). The EST-SSR loci and gSSR loci were only considered if they contained at least six repeats for dinucleotides, five repeats for trinucleotides, and four repeats for tetranucleotides, pentanucleotides, and hexanucleotides. Differences in mononucleotide repeats were excluded for EST-SSRs, and at least 10 repeats were required for mononucleotides in gSSRs. The maximal number of bases interrupting two SSRs in a compound microsatellite was set at 100 bp. The primers flanking the SSR core sequences were designed using Primer Premier 5.0 software (Kamel and Abd-Elsalam, 2003).

Initially amplification was conducted to optimize the annealing temperature of the SSR markers. Polymerase chain reaction (PCR) amplifications were performed in a final volume of 10 μL [0.5 μL DNA, 0.5 μL each primer (Tsingke Biological Technology, Beijing, China), 5 μL 2X Es Taq Master Mix (CWBIO, Beijing, China), and 3.5 μL ddH_2O]. The optimized SSR primers were used to amplify DNA in 24 individuals from the Beiwan (BW) population. The PCR products were analyzed using 8% polyacrylamide gel electrophoresis (PAGE) and stained with silver to distinguish the polymorphisms (Creste et al., 2001). Each forward polymorphic primer was marked with 5-FAM, HEX, TET, or TAMRA fluorescent dye at the 5' end. Polymorphic loci were tested in the 96 individuals from the three populations (28 from HBH, 26 from BEJ, 42 from BW) with fluorescent primers. The accurate sizes of the target fragments were measured with GeneMarker version 1.51 software using an ABI 3730 capillary sequencer.

Cross-species transferability

To assess the transferability of polymorphic SSR loci developed above, cross-species amplifications were conducted in nine other Cyprinidae fish: *Leuciscus idus*, *Rutilus lacustris*, *Abramis brama orientalis*, *Phoxinus ujmonensis*, *Phoxinus brachyurus*, and *Tinca tinca* from Leuciscinae, *Gymnodiptychus dybowskii* and *Diptychus maculates* from Schizothoracinae,

and *Cyprinus carpio* from Cyprinidae. All of these species were also distributed in Xinjiang, and 12 individuals were sampled during our investigation. All the amplification systems and procedures were the same as above.

Evaluation of SSR polymorphism and genetic diversity analysis

We evaluated the following genetic parameters for both SSRs and populations using POPGENE version 1.31: the number of alleles (N_A), the number of effective alleles (N_E), the expected heterozygosity (H_E), the observed heterozygosity (H_O), and the genetic distance (Yeh et al., 1999). Polymorphism information content (PIC) was calculated by applying the PIC_CALC software package (version 0.6). A dendrogram was constructed by UPGMA clustering analysis on the basis of genetic distance (Pavlícek et al., 1999). The F-statistic (F_{ST}) was calculated using the Arlequin software package (version 3.11).

RESULTS AND DISCUSSION

Characterization of various SSRs in the genome

A total of 1,839,008 assembled genomic contigs were generated via Illumina for MiSeq™ 2000 sequencing, and 1373 of these contigs were longer than 1000 bp, as shown in Table 1. From 10,168 sequences, 1686 potential SSRs were identified through MISA. The largest groups of repeat motifs were mononucleotides and dinucleotides (both 38.7%), followed by trinucleotides (11.2%), tetranucleotides (6.4%), pentanucleotides (4.6%), and hexanucleotides (less than 1%) (Figure 1).

Table 1. Reads and assembled contig information for *Leuciscus leuciscus baicalensis*.

Total no. of contigs	Bases in all contigs	No. of large contigs (>1000 bp)	Bases of large contigs	Greatest length	Contig N50	Contig N90	GC percentage
1,839,008	5,099,360 bp	1373	2,085,150 bp	16,606 bp	1446 bp	1065 bp	40.89%

N50: scanned sequences were accumulated from large to small according to length. When the cumulative value was more than 50% of the entire sequence length, the length of the sequence was N50. N90: N90 was determined in the same way as N50. The average lengths of N50 and N90 expressed the stand or fall of splicing sequences more accurately.

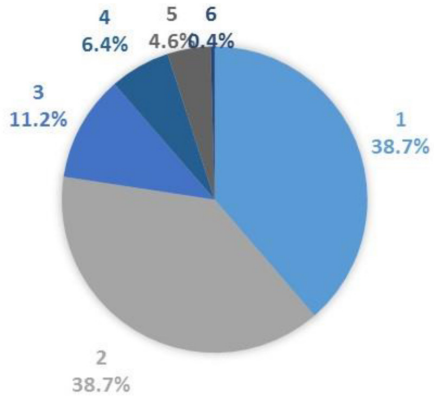

Figure 1. Distribution to different repeat type classes of gSSRs in Leuciscus leuciscus baicalensis. Mononucleotide (1), Dinucleotide (2), Trinucleotide (3), Tetranucleotide (4), Pentanucleotide (5), Hexanucleotide (6).

Evaluation of genetic diversity

Taking many factors into account, the use of mononucleotides was abandoned in the design of the primers. In total, 160 pairs of primers (64 from genomic libraries and 96 from EST libraries) were designed and amplified using the DNA from the BW samples, in which three pairs isolated from the genomic libraries and 25 pairs isolated from the EST libraries failed to provide PCR products. The high failure rate of the EST-SSR primers was possibly due to sequencing errors and the presence of introns. Thirty polymorphic SSR markers were identified among the successfully amplified primers, comprising 19 gSSRs and 11 EST-SSRs, which indicated high conservation in the transcribed regions of the various polymorphic regions (31.1 and 15.5%, respectively). The details of the 30 polymorphic SSRs are given in Table 2.

Table 2. Characteristics of 30 polymorphic simple sequence repeat (SSR) markers isolated from *Leuciscus leuciscus baicalensis*.

SSR	Primer sequences (5'-3')		GenBank accession No.	Fluorescent marker	SSR motif	Annealing temperature (°C)	Product size (bp)
BJZ6	R: GGGCAGCTGTTAGTCTGAGG	F: GGCCAAGTTATGTCTTTGAAATTGC	KX197890	FAM	$(AATTCA)_5$	60	165-200
BJZ71	R:GCGTCTCTGTCTGGTTTTGC	F: ATCTCTTCCCCTCGTCTGCT	KX197891	HEX	$(CA)_8(CG)_6$	56	193-209
BJZ78	R: AACTCTGTCCCTCCCGTCAT	F: AGTCCATGTGGTTGAGAGGC	KX197892	TET	$(TCT)_7$	56	121-145
BJZ66	R: CTTCCACCTTAACCAGCCCT	F: TCACCATCCAGGCTTAAACGT	KX197893	FAM	$(CAT)_{86}(AAT)_6$	56	219-246
BJZ62	R: GTGGAGGATTTGCATTGGGC	F: TGTCAGATGATGGGAGGCAAC	KX197894	FAM	$(AGTC)_5$	60	173-201
BJZ80	R: ATGTGAGGACATCTGCTGCC	F: GGAGCGAATCTGGACTGGAG	KX197895	HEX	$(GAA)_7$	52	226-244
BJZ33	R: CACGCCAAGACATGCTGAAC	F: ACTTCGCTCCCATTTGCTGT	KX197896	TET	$(GTCA)_6$	56	248-280
BJZ34	R: TCCTATGTGGTGATGCCCCT	F: AACACTGCGTGTAGGCTCTG	KX197897	FAM	$(GTTG)_6$	58	271-287
BJZ46	R: ACTGAAGGTGGCAAGCCTTA	F: TGGCACTGACAACCTCATCG	KX197898	FAM	$(CAA)_7$	56	131-146
BJZ88	R: ATCCTAGGTACCACACGGCT	F: TTTATGAAGTGCAGCGGGGT	KX197899	HEX	$(AC)_{10}$	60	176-194
BJZ89	R: CATCAGCCTGAAGGGGGTTT	F: TTGATCTCGCCGCTGAAACT	KX197900	TAMRA	$(GT)_{10}$	60	201-219
BJG21	R: CCTGATGCGTTACCTTCG	F: GCAATGCTCTGTTTGGGAT	KX197901	FAM	$(TC)_{10}g(CT)_{12}$	60	170-200
BJG25	R: CGCAGTGGCAGCATTTAT	F: CGGTTTAGGGTCAGGGTT	KX197902	HEX	$(TG)_{24}$	58	224-258
BJG13	R: CCACCCAATCCGCATCCT	F: CCCAGCCAAACAACCACC	KX197903	HEX	$(CACT)_{17}$	60	102-114
BJG20	R: CTCTGATGTGAGTGGGAAG	F: AATCGCCTGTAAGAATGAA	KX197904	FAM	$(TA)_{14}$	52	129-141
BJG23	R: GGTTGCTGATGGTTTAGAT	F: TCCTCACACAGATTTAGATAGA	KX197905	FAM	$(GT)_{16}$	56	120-150
BJG27	R: GACAAAAGCGTCTTCCAAAT	F: TGTAAAAGGTTAGGTGATAGCC	KX197906	HEX	$(AAT)_8(TAT)_6$	58	170-185
BJG31	R: CCTCACTCCAATGGTCTA	F: GTAATAAACAGGGGAATAAC	KX197907	TAMRA	$(GT)_{15}$	52	200-260
BJG60	R: GTAGGGTTTACCAGGACACA	F: GAGAGCACGGCAGCAT	KX197908	TAMRA	$(GT)_9$	60	214-244
BJG50	R: GCAAACAGAGCAACGATG	F: GTGAACTCAAACCAGGGG	KX197909	FAM	$(CA)_{12}$	56	128-158
BJG3	R: CCGTAGACAGAAAATCAACTT	F: GCACAAAACATTCAGCCA	KX197910	FAM	$(CTAT)_{29}$	60	205-257
BJG51	R: CTTCTGGTATTTCGGTAGC	F: AGTAATCAGGGGAGGAGG	KX197911	HEX	$(TTG)_8$	56	153-168
BJG41	R: TGCTGCGTCAAATGCGT	F: CACCCCTAAACTGGGATGT	KX197912	FAM	$(AT)_{14}$	54	231-275
BJG54	R: TGATTCCTTCAAATACACCG	F: CCCCTCTCTGCCAACTT	KX197913	HEX	$(TTA)_8$	60	155-185
BJG53	R: AAGAGAAGGAACAGAAAG	F: AAGACGAAAAAGAAGACT	KX197914	FAM	$(TC)_{12}$	56	268-302
BJG52	R: GTGGTGCGTCACGATTAT	F: TTGTGTGTCTGATTGGTCC	KX197915	HEX	$(CA)_{11}$	60	200-252
BJG62	R: GAACGAGCAGCAATCAAG	F: ATAGTAACGCCTGTGGTG	KX197916	FAM	$(AG)_{10}$	60	233-255
BJG57	R: CCTGATGGCGTCGTTACT	F: TCAAATGTTCCCCTGCTG	KX197917	HEX	$(AG)_{12}$	60	90-116
BJG26	R: CATTTTCAGGTTTTCCCC	F: CCGTTTTAGACACTTTGCTC	KX197918	FAM	$(TA)_{15}a(AT)_6$	58	252-282
BJG28	R: TAATCAAATAAAGGCAGGCT	F: GAACCGTTACATAATCCCAT	KX197919	HEX	$(TG)_{15}$	58	182-210

Genetic diversity was assessed using 30 polymorphic markers in 96 individuals from the three populations. Overall, the number of alleles (N_A) varied from 4 to 27 (with an average of 11.3), the expected heterozygosity was 0.36-0.94 (average 0.75 ± 0.14), and the observed heterozygosity was 0.37-1.00 (average 0.68 ± 0.18). The high values of mean H_O and H_E suggest that there was relatively high heterozygosity. The polymorphism information content (PIC) was 0.31-0.93 (average 0.71 ± 0.15), suggesting high genetic diversity, and these markers were of good quality. Furthermore, the average values of H_E and PIC for the EST-SSR set (0.67 and 0.62, respectively) were lower than those for the genomic SSR set (0.79 and 0.72, respectively) (Table 3), which confirms the hereditary conservation in the transcribed regions and the higher polymorphism of the gSSR marker, and corroborates the previous studies (Zhan et al., 2009; Molina-Luzón et al., 2012; Zhang et al., 2014).

The three populations displayed similar results for PIC and H_E: 0.71, 0.69, and 0.71, and 0.68, 0.69, and 0.68 for the HBH, BEJ, and BW populations, respectively. There were small differences in the N_A values: HBH (8.77), BEJ (8.57), and BW (9.47). Therefore, the three populations had almost the same level of genetic diversity. The details of the diversity parameters for the three populations are shown in Table 4.

Table 3. Averages of PIC, H_E, H_O, N_E, and N_A values of the expressed sequence tag-simple sequence repeat (EST-SSR) and genomic SSR (gSSR) markers of *Leuciscus leuciscus baicalensis*.

	PIC	H_E	H_O	N_E	N_A		PIC	H_E	H_O	N_E	N_A
BJZ6	0.73	0.77	0.65	4.28	7.00	BJG23	0.87	0.89	0.71	8.57	16.00
BJZ71	0.61	0.65	0.59	2.83	8.00	BJG27	0.71	0.76	1.00	4.06	8.00
BJZ78	0.65	0.68	0.65	3.07	9.00	BJG31	0.93	0.94	0.69	15.34	27.00
BJZ66	0.71	0.75	0.67	3.95	11.00	BJG60	0.50	0.57	0.55	2.29	11.00
BJZ62	0.44	0.47	0.43	1.87	7.00	BJG50	0.87	0.89	0.88	8.54	16.00
BJZ80	0.67	0.71	0.69	3.38	7.00	BJG3	0.88	0.90	0.45	9.22	13.00
BJZ33	0.70	0.73	0.53	3.59	9.00	BJG51	0.69	0.73	0.96	3.65	6.00
BJZ34	0.50	0.57	0.65	2.31	6.00	BJG41	0.90	0.91	0.74	10.61	22.00
BJZ46	0.57	0.64	0.54	2.77	6.00	BJG54	0.67	0.70	0.73	3.28	11.00
BJZ88	0.74	0.78	0.67	4.47	9.00	BJG53	0.83	0.85	0.52	6.45	17.00
BJZ89	0.53	0.58	0.37	2.38	6.00	BJG52	0.86	0.88	0.99	7.93	23.00
E-Mean ± SD	**0.62 ± 0.10**	**0.67 ± 0.10**	**0.59 ± 0.11**	**3.17 ± 0.84**	**7.73 ± 1.62**	BJG62	0.73	0.76	0.95	4.16	10.00
BJG21	0.70	0.74	0.57	3.86	12.00	BJG57	0.84	0.86	0.75	6.90	13.00
BJG25	0.69	0.73	0.62	3.70	12.00	BJG26	0.84	0.86	0.78	6.97	14.00
BJG13	0.31	0.36	0.39	1.56	4.00	BJG28	0.83	0.85	0.98	6.41	12.00
BJG20	0.81	0.84	0.69	5.87	7.00	**g-Mean ± SD**	**0.76 ± 0.15**	**0.79 ± 0.14**	**0.73 ± 0.19**	**6.28 ± 3.33**	**13.37 ± 5.87**
T-Mean ± SD	**0.71 ± 0.15**	**0.75 ± 0.14**	**0.68 ± 0.18**	**5.14 ± 3.07**	**11.30 ± 5.47**						

E-Mean ± SD: mean ± SD of EST-SSR markers; g-Mean ± SD: mean ± SD of genomic markers; T-Mean ± SD: mean ± SD of all markers.

Table 4. Details of diversity parameters for the three *Leuciscus leuciscus baicalensis* populations.

	HBH					BEJ					BW				
	PIC	H_E	H_O	N_E	N_A	PIC	H_E	H_O	N_E	N_A	PIC	H_E	H_O	N_E	N_A
BJZ6	0.73	0.79	0.71	4.37	6.00	0.74	0.79	0.50	4.36	6.00	0.73	0.76	0.69	3.97	6.00
BJZ71	0.61	0.62	0.68	2.57	5.00	0.63	0.67	0.65	2.93	7.00	0.57	0.66	0.50	2.90	7.00
BJZ78	0.65	0.76	0.77	4.00	8.00	0.63	0.67	0.58	2.88	8.00	0.72	0.62	0.63	2.55	7.00
BJZ66	0.71	0.81	0.71	4.96	9.00	0.60	0.65	0.58	2.77	7.00	0.77	0.76	0.69	4.02	7.00
BJZ62	0.44	0.39	0.30	1.61	5.00	0.44	0.48	0.50	1.88	5.00	0.36	0.51	0.48	2.03	7.00
BJZ80	0.67	0.76	0.78	3.93	7.00	0.58	0.64	0.50	2.67	6.00	0.72	0.72	0.76	3.41	6.00
BJZ33	0.70	0.68	0.59	3.00	7.00	0.66	0.70	0.38	3.22	7.00	0.64	0.76	0.59	4.08	9.00
BJZ34	0.50	0.56	0.52	2.24	4.00	0.42	0.51	0.62	2.00	3.00	0.48	0.62	0.76	2.59	6.00
BJZ46	0.57	0.69	0.57	3.10	6.00	0.58	0.66	0.62	2.81	5.00	0.62	0.61	0.48	2.53	5.00
BJZ88	0.74	0.77	0.63	4.07	8.00	0.65	0.70	0.58	3.14	7.00	0.72	0.80	0.76	4.78	8.00
BJZ89	0.53	0.58	0.19	2.31	5.00	0.54	0.61	0.56	2.48	5.00	0.52	0.58	0.38	2.35	5.00
E-Mean	**0.62**	**0.67**	**0.59**	**3.29**	**6.36**	**0.59**	**0.64**	**0.55**	**2.83**	**6.00**	**0.62**	**0.67**	**0.61**	**3.20**	**6.64**
±SD	**±0.10**	**±0.13**	**±0.19**	**±1.05**	**±1.57**	**±0.09**	**±0.09**	**±0.08**	**±0.66**	**±1.41**	**±0.13**	**±0.09**	**±0.13**	**±0.89**	**±1.21**
BJG21	0.70	0.75	0.61	3.73	10.00	0.70	0.75	0.63	3.79	9.00	0.69	0.74	0.51	3.69	10.00
BJG25	0.69	0.75	0.61	3.77	8.00	0.70	0.75	0.60	3.81	9.00	0.69	0.72	0.64	3.44	9.00
BJG13	0.31	0.31	0.36	1.44	3.00	0.40	0.49	0.62	1.91	4.00	0.28	0.30	0.26	1.42	3.00
BJG20	0.81	0.86	0.61	5.95	7.00	0.80	0.85	0.68	5.76	7.00	0.81	0.81	0.74	5.08	7.00
BJG23	0.87	0.87	0.68	7.00	10.00	0.84	0.88	0.73	7.08	13.00	0.84	0.89	0.71	8.58	13.00
BJG27	0.71	0.73	1.00	3.56	5.00	0.72	0.78	1.00	4.17	6.00	0.67	0.77	1.00	4.17	7.00
BJG31	0.93	0.94	0.58	12.26	18.00	0.90	0.92	0.72	10.68	16.00	0.91	0.95	0.73	16.93	23.00
BJG60	0.50	0.62	0.54	2.58	8.00	0.49	0.56	0.65	2.19	6.00	0.55	0.54	0.50	2.13	8.00
BJG50	0.87	0.90	0.89	8.48	14.00	0.85	0.88	0.88	7.19	13.00	0.87	0.88	0.86	7.93	12.00
BJG3	0.88	0.92	0.50	10.08	13.00	0.84	0.87	0.50	6.94	10.00	0.89	0.90	0.39	8.73	12.00
BJG51	0.66	0.75	0.96	3.73	6.00	0.70	0.75	0.92	3.84	6.00	0.69	0.71	0.98	3.30	6.00
BJG41	0.90	0.93	0.77	11.76	17.00	0.86	0.89	0.73	8.00	14.00	0.91	0.89	0.73	8.60	18.00
BJG54	0.67	0.69	0.67	3.08	9.00	0.64	0.68	0.77	3.03	9.00	0.65	0.72	0.74	3.46	10.00
BJG53	0.83	0.85	0.61	6.01	13.00	0.85	0.88	0.46	7.19	13.00	0.82	0.84	0.50	5.77	15.00
BJG52	0.86	0.86	0.96	6.56	13.00	0.88	0.91	1.00	9.01	16.00	0.83	0.87	1.00	7.03	18.00
BJG62	0.73	0.79	0.93	4.47	8.00	0.72	0.77	0.96	4.12	7.00	0.75	0.75	0.95	3.82	9.00
BJG57	0.84	0.84	0.75	5.72	10.00	0.84	0.87	0.88	7.01	11.00	0.80	0.86	0.67	6.53	10.00
BJG26	0.84	0.87	0.89	6.72	11.00	0.84	0.87	0.77	6.76	11.00	0.84	0.85	0.71	6.33	11.00
BJG28	0.83	0.86	1.00	6.43	10.00	0.82	0.85	1.00	6.06	11.00	0.83	0.85	0.95	6.18	10.00
g-Mean	**0.80**	**0.79**	**0.73**	**5.96**	**10.16**	**0.79**	**0.80**	**0.76**	**5.71**	**10.05**	**0.80**	**0.78**	**0.71**	**5.95**	**11.11**
±SD	**±0.12**	**±0.15**	**±0.19**	**±3.00**	**±3.87**	**±0.12**	**±0.12**	**±0.17**	**±2.37**	**±3.52**	**±0.11**	**±0.15**	**±0.21**	**±3.45**	**±4.75**
T-Mean	**0.71**	**0.75**	**0.68**	**4.98**	**8.77**	**0.69**	**0.74**	**0.69**	**4.66**	**8.57**	**0.71**	**0.74**	**0.68**	**4.95**	**9.47**
±SD	**±0.15**	**±0.15**	**±0.20**	**±2.77**	**±3.68**	**±0.15**	**±0.13**	**±0.17**	**±2.37**	**±3.51**	**±0.16**	**±0.14**	**±0.19**	**±3.09**	**±4.39**

E-Mean ± SD: mean ± SD of EST-SSR markers; g-Mean ± SD: mean ± SD of genomic markers; T-Mean ± SD: mean ± SD of all markers.

Analysis of cluster and genetic differentiation

Based on the Nei's genetic distances among the three populations, we constructed an unrooted dendrogram by UPGMA clustering analysis (Figure 2), showing an almost equal level of genetic differentiation among the populations. The three populations were located in continuous bodies of water, so the pelagic eggs were able to flow into the rivers, which led to the flow of genes. Comparatively speaking, the HBH and BW populations resembled each other most closely, possibly as a result of their geographical proximity to each other.

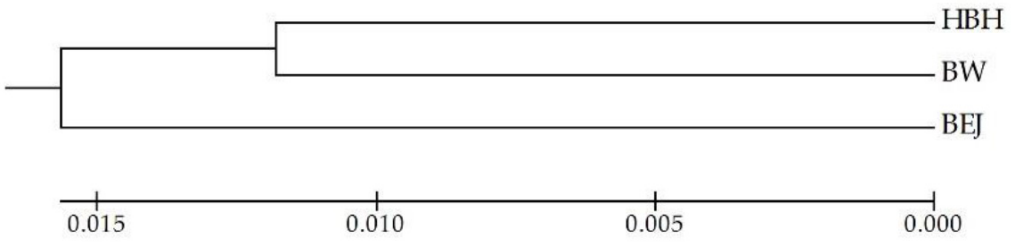

Figure 2. Dendrogram of the three populations of *Leuciscus leuciscus baicalensis* according to the genetic distance using UPGMA clustering analysis.

The F-statistic is an indicator of genetic differentiation among populations; the F_{ST} values for the three populations are given in Tables 5 and 6. According to the standard of Wright, the differentiation was defined as high when $F_{ST} > 0.25$, moderate when $0.15 < F_{ST} < 0.25$, low when $F_{ST} < 0.15$, and absent when $F_{ST} < 0.05$. In this study, the F_{ST} values were all below 0.05, which meant the HBH, BEJ, and BW populations were almost non-differentiated. The pairwise F_{ST} values between populations were also less than 0.05 with a P-value above 0.05, i.e., there were no significant differences between any two populations (Tables 5 and 6). Details of the analysis of molecular variance (AMOVA) are provided in Table 7; almost all the variation was between individuals (95.81%), and very little variation occurred among populations. This result was consistent with the clustering analysis, which indicated that a high level of gene exchange occurred between the three populations.

Table 5. F_{ST} values for the three populations.

Population	F_{ST}
HBH	0.00175
BEJ	0.00149
BW	0.00148

Table 6. Pairwise F_{ST} values between populations.

Population	HBH	BEJ	BW
HBH	0		
BEJ	0.00227	0	
BW	-0.00013	0.00289	0
P > 0.05			

Table 7. Analysis of molecular variance (AMOVA) analysis of *Leuciscus leuciscus baicalensis*.

Source of variation	Sum of squares	Variance component	Percentage of variation
Among populations	20.421	0.00833	0.09
Among individuals in populations	901.215	0.3817	4.1
Within individuals	857	8.92708	95.81
Total	1778.635	9.31711	100

Cross-species transferability

Among the 30 primer pairs developed from *L. leuciscus baicalensis*, only three EST-SSRs successfully amplified target fragments in all cross-species: BJZ6, BJZ34, and BJZ46. Furthermore, four gSSR markers failed to produce any bands or the expected size of DNA fragment in every species: BJG20, BJG41, BJG53, and BJG62. The transferability of the 30 markers is shown in Table 8. The top three species with respect to transferability values were *R. lacustris* (80%), *L. idus* (76.7%), and *P. ujmonensis* (63.3%); all three species are members of the Leuciscinae subfamily, as is *L. leuciscus baicalensis*. Therefore, they were subjected to further genetic diversity analysis. Among all the primer pairs used for further study, every gSSR marker was polymorphic, and the BJZ6 and BJZ62 primer pairs were monomorphic in all three species. Moreover, the BJZ13, BJZ80, and BJZ89 primer pairs also showed no polymorphism in *P. ujmonensis*. In conclusion, the EST-SSR marker was more transferable than the gSSR marker, which was consistent with earlier studies in aquatic species (Wang et al., 2007; Liu et al., 2011). According to the criterion previously described, the polymorphism was defined as high when PIC > 0.5, moderate when 0.25 < PIC < 0.5, and low when PIC < 0.25 (Botstein et al., 1980; Yadav et al., 2011). As shown in Figure 3, the polymorphic SSR markers demonstrated high polymorphism in *R. lacustris* (PIC = 0.56) and moderate polymorphisms in *L. idus* and *P. ujmonensis* (PIC = 0.40, 0.44, respectively). All the primers also exhibited higher H_E, H_O, N_E, and N_A values in *R. lacustris*. This result suggests that some of the microsatellite primers could be applied to other species, especially related species, and the selective use of the primers could be used effectively to discriminate species, particularly morphologically similar species (Chiang et al., 2012; Li et al., 2016).

Table 8. Transferability of 30 microsatellite markers for nine species of Cyprinidae fish.

	Leuciscus idus	*Gymnodiptychus dybowskii*	*Diptychus maculates*	*Rutilus lacustris*	*Cyprinus carpio*	*Abramisbrama orientalis*	*Phoxinus ujmonensis*	*Phoxinus brachyurus*	*Tinca tinca*
Transferability	76.7%	53.3%	20.0%	80.0%	56.7%	43.3%	63.3%	43.3%	43.3%

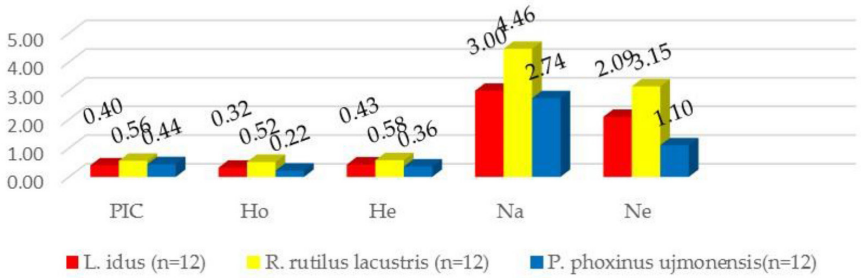

Figure 3. Averages of PIC, H_O, H_E, N_A, and N_E for cross-species.

CONCLUSION

In this study, 30 novel high-quality SSR markers were isolated to evaluate the genetic diversity of *L. leuciscus baicalensis*. The rate of successful amplification, the rate of polymorphism, and genetic diversity were lower in the EST-SSR markers. Among the three populations studied, the parameters N_A, N_E, H_O, H_E, and PIC showed the same level of genetic diversity, and the parameters of genetic distances and F_{ST} showed equal levels of genetic differentiation. With regard to cross-species transferability, the top three species were *R. lacustris*, *L. idus*, and *P. ujmonensis*.

Polymorphic SSR markers have been widely utilized in diverse areas (including in fish) for genetic research such as stock identification, parentage analysis, linkage map construction, evolutionary relationship analysis, and marker-assisted selection (Zhan et al., 2010; Liang et al., 2015; Wang et al., 2015; Zhang et al., 2016). Compared with gSSRs, EST-SSRs were more efficient in functional gene analysis, such as marker-related growth and immunity in aquatic animals (Zheng et al., 2014; Huang et al., 2015). The high levels of polymorphism and transferability of these novel markers for a number of important *Leuciscus* species, including *L. leuciscus baicalensis*, are very important attributes, which could be of vital significance with regard to genetic resource conservation and sustainable use in the future.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Special Project of the National Science and Technology Fundamental Work of China (grant #2012FY112700).

REFERENCES

Boron A, Porycka K, Ito D, Abe S, et al. (2009). Comparative molecular cytogenetic analysis of three *Leuciscus* species (Pisces, Cyprinidae) using chromosome banding and FISH with rDNA. *Genetica* 135: 199-207. http://dx.doi.org/10.1007/s10709-008-9269-3

Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Chang YM, Tang R, Dou XJ, Tao R, et al. (2014). Transcriptome and expression profiling analysis of *Leuciscus waleckii*: an exploration of the alkali-adapted mechanisms of a freshwater teleost. *Mol. Biosyst.* 10: 491-504. http://dx.doi.org/10.1039/c3mb70318e

Chiang YC, Huang BH, Shih HC, Hsu TW, et al. (2012). Characterization of 24 transferable microsatellite loci in four skullcaps (*Scutellaria*, Labiatae). *Am. J. Bot.* 99: e24-e27. http://dx.doi.org/10.3732/ajb.1100279

Costedoat C, Chappaz R, Barascud B, Guillard O, et al. (2006). Heterogeneous colonization pattern of European Cyprinids, as highlighted by the dace complex (Teleostei: Cyprinidae). *Mol. Phylogenet. Evol.* 41: 127-148. http://dx.doi.org/10.1016/j.ympev.2006.04.022

Creste S, Neto AT and Figueira A (2001). Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol. Biol. Report.* 19: 299. http://dx.doi.org/10.1007/BF02772828

Cui J, Xu J, Zhang S, Wang K, et al. (2015). Transcriptional profiling reveals differential gene expression of amur ide (*Leuciscus waleckii*) during spawning migration. *Int. J. Mol. Sci.* 16: 13959-13972. http://dx.doi.org/10.3390/ijms160613959

Dubut V, Martin JF, Gilles A, VAN Houdt J, et al. (2009). Isolation and characterization of polymorphic microsatellite loci

for the dace complex: *Leuciscus leuciscus* (Teleostei: Cyprinidae). *Mol. Ecol. Resour.* 9: 1179-1183. http://dx.doi.org/10.1111/j.1755-0998.2009.02594.x

Gao Z, Luo W, Liu H, Zeng C, et al. (2012). Transcriptome analysis and SSR/SNP markers information of the blunt snout bream (*Megalobrama amblycephala*). *PLoS One* 7: e42637. http://dx.doi.org/10.1371/journal.pone.0042637

Hu S, Niu J, Xie P, Liu C, et al. (2015). The complete mitochondrial genome of Leuciscus leuciscus baicalensis (Cypriniformes: Cyprinidae). *Mitochondrial DNA* 26: 751-752. http://dx.doi.org/10.3109/19401736.2013.848353

Huang L, Li G, Mo Z, Xiao P, et al. (2015). De novo assembly of the Japanese flounder (*Paralichthys olivaceus*) spleen transcriptome to identify putative genes involved in immunity. *PLoS One* 10: e0131146. http://dx.doi.org/10.1371/journal.pone.0131146

Huo TB, Yuan MY and Jiang ZF (2011). Length-weight relationships of 23 fish species from the Ergis River in Xingjiang, China. *J. Appl. Ichthyology* 27: 937-938. http://dx.doi.org/10.1111/j.1439-0426.2010.01528.x

Kamel A, Abd-Elsalam (2003). Bioinformatic tools and guideline for PCR primer design. *Afr. J. Biotechnol.* 2: 91-95. http://dx.doi.org/10.5897/AJB2003.000-1019

Li CY, Chiang TY, Chiang YC, Hsu HM, et al. (2016). Cross-species, amplifiable EST-SSR markers for *Amentotaxus* species obtained by next-generation sequencing. *Molecules* 21: 67. http://dx.doi.org/10.3390/molecules21010067

Liang Y, Han W, Sun P, Liang J, et al. (2015). Genetic diversity among germplasms of *Diospyros Kaki*, based on SSR markers. *Sci. Hortic. (Amsterdam)* 186: 180-189. http://dx.doi.org/10.1016/j.scienta.2015.02.015

Lin XY and Tang FQ (2016). Acute toxicity experiment of fishery drugs on the fry of *Leuciscus Leuciscus leuciscus baicalensis*. *Scient. Fish Farm.* 6: 56-58.

Liu X, Luo W, Zeng C, Wang W, et al. (2011). Isolation of new 40 microsatellite markers in mandarin fish (*Siniperca chuatsi*). *Int. J. Mol. Sci.* 12: 4180-4189. http://dx.doi.org/10.3390/ijms12074180

Liu XC, Zhuo RJ, Yin JG and Cai XQ (2015a). Experiment on pond culture for *Leuciscus Leuciscus leuciscus baicalensis* in Altay Prefecture, Xingjiang. *Scient. Fish Farm.* 31: 35-35.

Liu Y, Zhang P, Song M, Hou J, et al. (2015b). Transcriptome Analysis and Development of SSR Molecular Markers in *Glycyrrhiza uralensis* Fisch. *PLoS One* 10: e0143017. http://dx.doi.org/10.1371/journal.pone.0143017

Molina-Luzón MJ, López JR, Navajas-Pérez R, Robles F, et al. (2012). Validation and comparison of microsatellite markers derived from Senegalese sole (*Solea senegalensis*, Kaup) genomic and expressed sequence tags libraries. *Mol. Ecol. Resour.* 12: 956-966. http://dx.doi.org/10.1111/j.1755-0998.2012.03163.x

Nowosad J, Targońska K, Chwaluczyk R, Kaszubowski R, et al. (2014). Effect of temperature on the effectiveness of artificial reproduction of dace [Cyprinidae (*Leuciscus leuciscus* (L.))] under laboratory and field conditions. *J. Therm. Biol.* 45: 62-68. http://dx.doi.org/10.1016/j.jtherbio.2014.07.011

Ou LI, Zhao YY, Guo N, Lu CY, et al. (2009). Effects of sample size and loci number on genetic diversity in wild population of grass carp revealed by SSR. *Zoological Res.* 30: 121-130. http://dx.doi.org/10.3724/SP.J.1141.2009.02121

Pavlícek A, Hrdá S and Flegr J (1999). Free-Tree--freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus Frenkelia. *Folia Biol. (Praha)* 45: 97-99.

Perea S, Böhme M, Zupancic P, Freyhof J, et al. (2010). Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evol. Biol.* 10: 265. http://dx.doi.org/10.1186/1471-2148-10-265

Pruett CL and Winker K (2008). The effects of sample size on population genetic diversity estimates in song sparrows Melospiza melodia. *J. Avian Biol.* 39: 252-256. http://dx.doi.org/10.1111/j.0908-8857.2008.04094.x

Ren ML, Guo Y, Zhang RM, Zhang XS, et al. (2002). Fisheries resources and fishery of the Ertix River in China. Urumchi: Xinjiang Science and Technology, Hygiene Press.

Siddique MA, Linhart O, Krejszeff S, Żarski D, et al. (2016). Effects of preincubation of eggs and activation medium on the percentage of eyed embryos in ide (*Leuciscus idus*), an externally fertilizing fish. *Theriogenology* 85: 849-855. http://dx.doi.org/10.1016/j.theriogenology.2015.10.032

Wang D, Liao X, Cheng L, Yu X, et al. (2007). Development of novel EST-SSR markers in common carp by data mining from public EST sequences. *Aquaculture* 271: 558-574. http://dx.doi.org/10.1016/j.aquaculture.2007.06.001

Wang QB, Zhang L and Zheng PJ (2015). Genetic diversity and evolutionary relationship analyses within and among *Raphanus*, species using EST-SSR markers. *Mol. Breed.* 35: 62. http://dx.doi.org/10.1007/s11032-015-0261-1

Wang ZW, Zhu HP, Wang D, Jiang FF, et al. (2011). A novel nucleo-cytoplasmic hybrid clone formed via androgenesis in polyploid gibel carp. *BMC Res. Notes* 4: 82. http://dx.doi.org/10.1186/1756-0500-4-82

Witeska M, Sarnowski P, Ługowska K and Kowal E (2014). The effects of cadmium and copper on embryonic and larval development of ide *Leuciscus idus* L. *Fish Physiol. Biochem.* 40: 151-163. http://dx.doi.org/10.1007/s10695-013-9832-4

Yadav HK, Ranjan A, Asif MH, Mantri S, et al. (2011). EST-derived SSR markers in *Jatropha curcas L.*: development, characterization, polymorphism, and transferability across the species/genera. *Tree Genet. Genetics* 7: 207-219. http://dx.doi.org/10.1007/s11295-010-0326-6

Yan LN and Zhang DX (2004). Effects of sample size on various genetic diversity measures in population genetic study with microsatellite DNA markers. *Acta Zool. Sin.* 50: 279-290.

Yeh FC, Boyle T and Yang RC (1999). PopGene Version 131: Microsoft window-based freeware for population genetic analysis. *University of Alberta, Edmonton*.

Zhan A, Wang Y, Brown B and Wang HP (2009). Isolation and characterization of novel microsatellite markers for yellow perch (*Perca flavescens*). *Int. J. Mol. Sci.* 10: 18-27. http://dx.doi.org/10.3390/ijms10010018

Zhan QQ, Sui C, Wei JH, Fan SC, et al. (2010). Construction of genetic linkage map of Bupleurum chinense DC. using ISSR and SSR markers. *Yao Xue Xue Bao* 45: 517-523.

Zhang J, Ma WG, Wang WN, Gui JF, et al. (2016). Parentage determination of yellow catfish (Pelteobagrus Fulvidraco) based on microsatellite DNA markers. *Aquacult. Int.* 24: 567-576. http://dx.doi.org/10.1007/s10499-015-9947-y

Zhang M, Mao W, Zhang G and Wu F (2014). Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetan annual wild barley. *PLoS One* 9: e94881. http://dx.doi.org/10.1371/journal.pone.0094881

Zheng X, Pan C, Diao Y, You Y, et al. (2013). Development of microsatellite markers by transcriptome sequencing in two species of *Amorphophallus* (Araceae). *BMC Genomics* 14: 490. http://dx.doi.org/10.1186/1471-2164-14-490

Zheng X, Kuang Y, Lü W, Cao D, et al. (2014). Transcriptome-derived EST-SSR markers and their correlations with growth traits in crucian carp *Carassius auratus*. *Fish. Sci.* 80: 977-984. http://dx.doi.org/10.1007/s12562-014-0782-2