Taylor & Francis
Taylor & Francis Group

Characterization and pilot application of 14 microsatellite loci from the Korean shiner (*Coreoleuciscus splendidus*)

Sang Ki Kim[a], Ji Jung Kim[b], Shi Hyun Ryu[a], Ho Young Suk[b]* and Ui Wook Hwang[a]*

[a]Department of Biology, Teachers College and Institute for Phylogenomics and Evolution, Kyungpook National University, Daegu 702-701, Korea; [b]Department of Life Sciences, Yeungnam University, Gyeongsan Gyeongsangbuk-do 712-749, Korea

The Korean shiner (*Coreoleuciscus splendidus*, Gobioninae), a small freshwater fish native to Korea Peninsula, widely inhabits in most of the major drainages in South Korea. Here we describe the development of 14 novel and polymorphic microsatellites for this species and its effective utilization in estimation of genetic diversity using 72 individuals from three sampling sites in different drainages. Loci were isolated from a microsatellite enrichment procedure using probe-labeled magnetic beads. A total of 242 alleles were detected across all loci with an average of 17.3 alleles per locus ranging 4–32. The loci varied levels of polymorphism as evident from its expected heterozygosity ranging from 0.111 to 0.957. The average pairwise F_{ST} between two populations examined shows significant differentiation ($F_{ST} = 0.215$, $p < 0.05$). The 14 microsatellite loci developed here will also be useful to explain for the genetic structures among geographically isolated populations and gene flow dynamics within drainages in this species as well as the closely related species.

Keywords: microsatellites; Korean shiner; *Coreoleuciscus splendidus*; molecular markers; Gobioninae

Introduction

The Korean shiner (*Coreoleuciscus splendidus*, Gobioninae), a small freshwater fish native to Korea Peninsula, widely inhabits in most of the major drainages in South Korea (Choi et al. 1990). This species grows roughly 10–15 cm in length with slender body and exhibits iridescent colors on their body-side that make this species visually conspicuous. Although this species is not officially designated as an endangered species, it has dramatically declined throughout much of their range, at least partly because its habitats are generally confined to small fast-running shallow areas (Choi et al. 1990). The present distribution of the Korean shiner possibly reflects the natural population history, undisturbed by anthropogenic activities due to low commercial value of this species. Therefore, the Korean shiner provides a model system for investigating the relative contributions of historical events of dispersal, geographic isolation, water connectivity, and adaptive divergence to the population structure in Korea Peninsula.

A previous genetic study with AFLP revealed considerable genetic differentiation in the Korean shiner between the drainages flowing west (Han and Geum Rivers) and south (Seomjin and Nakdong Rivers) in Korea Peninsula (Song et al. 2010). Mitochondrial data also support the existence of two distinct genetic lineages (unpublished data). However, the comparison of genetic variability gauged by the use of multiple types of molecular markers (i.e. microsatellites) is necessary to obtain a precise view of the historical/contemporary dispersal and population genetic structure (e.g. Bos et al. 2008; Seifertová and Šimková 2010).

Microsatellites are tandemly repeating sequences of 1–6 base pairs of DNA found in all prokaryotic and eukaryotic genomes (Zane et al. 2002). Microsatellites are useful genetic markers because they can be easily amplified with the PCR and tend to be highly polymorphic (Schlötterer and Harr 2000; Bhatia and Arora 2007; An et al. 2010). By studying the variation of microsatellites among populations, inferences can be made about population structures, genetic drift, fragmentation-interaction dynamics of different populations, and even the date of a last common ancestor (e.g. Domínguez-Domínguez et al. 2008; Suk and Neff 2009). Microsatellites can also be used for the identification of new and incipient populations (Kashi et al. 1997; Li et al. 2002; Remya et al. 2010). To date, however, microsatellite marker has not been reported for the Korean shiner. We now describe the development of 14 novel and polymorphic microsatellites for this species, and examine genetic diversity of three populations collected from distinct drainages to compare our data with previous claims gained from different genetic markers.

Materials and methods

Sample collection

A total of 72 individuals of the Korean shiner were collected for the genetic analyses using seine and dip

EVOLUTION & SYSTEMATIC BIOLOGY

*Corresponding authors. Email: hsuk@ynu.ac.kr; uwhwang@knu.ac.kr

nets during 2010–2011, from three localities from Han River (37° 26′ 10.52″ N, 128° 10′ 36.42″ E; Anheung-myeon, Hoengseong-gun, Gangwon-do; $N = 30$), Geum River (36° 00′ 02.20″ N, 127° 41′ 18.20″ E; Muju-eup, Muju-gun, Jeollabuk-do; $N = 12$), and Nakdong River (35° 45′ 37.27″ N, 128° 59′ 06.85″ E; Sannae-myeon, Gyeongju-si, Gyeongsangbuk-do; $N = 30$) in Korea.

Isolation of microsatellites-containing DNA

We first constructed a microsatellite DNA-enriched library following the method of Hamilton et al. (1999). The total genomic DNA was extracted from the dorsal muscle tissue using an Exgene Tissue SV kit (GeneAll Co., Korea) following the manufacturer's protocol and was digested with *Hae* III, *Rsa* I, and *Nhe* I (Enzynomics Co., Korea). Fragments of 400–900 bp in length were ligated to a double strand SNX linkers (SNX-F, 5′-CTA AGG CCT TGC TAG CAG AAG C-3′; SNX-R, 5′-GCT TCT GCT AGC AAG GCC TTA GAA AA-3′) using T4 DNA ligase (Promega, USA). The fragments were recovered by polymerase chain reaction using the single-stranded forward and reverse linkers as primers. The PCR product was hybridized with five biotinylated $(GATA)_7$, $(GATC)_7$, $(GACA)_7$, $(GT)_{12}$, and $(CT)_{12}$ probes. After hybridizations, bound fragments were recovered using streptavidin-coated magnetic bead (Streptavidin Magnesphere Paramagnetic Particles, Promega, USA). Microsatellite-enriched DNA was amplified at 94°C for 2 min, followed by 34 cycles of 20s at 94°C, 40s at 58°C, 1 min at 72°C, and a final extension time of 72°C for 5 min using SNX forward primer.

PCR products were ligated into pGEM-T Easy vector (Promega, USA), transformed into competent *Escherichia coli* DH5a cells and spread onto LB agar plates. Approximately 400 positive clones were amplified with T7 and SP6 universal primers. PCR products were mixed with ExoSAP-IT endonuclease (USB, USA) and incubated for 15 min at 37°C to remove unused primers and nucleotides. Purified PCR fragments were sequenced in both directions with an ABI PRISM BigDye terminator system, and the reactants were analyzed on an ABI3700 model automatic sequencer (Genotech Co., Korea).

Microsatellite genotyping

Obtained sequences were used to design primers using Oligo-4.0-s software (http://www.oligo.net/). The PCR was performed in a volume of 20 µl containing 1 pg–1 µg DNA, 10 mM of each dNTP, 10 pM of forward and reverse primers, and 2.5 unit of DNA polymerase with the reaction buffer (SolGent Co., Korea). Forward primers were labeled with Applied Biosystems standard dye sets, TAMRA, FAM, HEX, or ROX (MACRO-GEN Co., Korea). The program setting includes a 2 min at 94°C for initial denaturation, 34 cycles of 20s for denaturation at 94°C, 40s for primer annealing at 58°C and 40s for extension at 72°C with a single further extension at 72°C for 5 min. The fragment analyses were conducted on an ABI3130xl DNA sequencer.

Statistical analysis

The utility of the polymorphic loci was examined using 72 individuals encompassing three sampling sites. For each site, Hardy–Weinberg equilibrium and the linkage disequilibrium were calculated using Genepop (ver. 4.0.10; Raymond and Rousset 1995). The test for the presence of null alleles was conducted using MICRO-CHECKER (ver. 2.2.3; Van Oosterhout et al. 2004). FSTAT (ver. 2.9.3.2; Goudet 1995) was used to calculate the gene diversity such as allelic frequencies, observed heterozygosities (H_O), and expected heterozygosities (H_E). Population differentiation among the three sampling sites based on pairwise F_{ST} estimates was examined using Genepop.

Cross-species amplification

The 14 polymorphic microsatellite loci were also tested for cross-species amplification with six Gobioninae species, including *Pseudorasbora parva*, *Pungtungia herzi*, *Sarcocheilichthys variegatus wakiyae*, *Squalidus gracilis majimae*, *Hemibarbus labeo*, and *Gobiobotia naktongensis*. PCR was carried out with two or three individuals for each species under the same conditions as described earlier.

Results and discussion

In 81 clones out of about 400, sufficient flanking sequences were available to design primers. Among them, 32 (40%) clones did not amplify consistently or showed smeared-band patterns. The rest of them, 14 (17%) clones were reliably amplified with polymorphism and were used as novel microsatellite loci in this study. The locus name, repeat motif, annealing temperature, and fluorescent label for each locus were shown in Table 1. From 72 individuals collected at three sampling sites, a total of 242 alleles were detected across all loci with an average of 17.3 alleles per locus, ranged from 4 (*Cos67*) to 32 (*Cos21*). The observed and expected heterozygosity ranged from 0.038 to 1.000 and from 0.111 to 0.957, respectively (Table 2). Three loci (*Cos9*, *Cos11*, and *Cos13*) in Anheung population, one locus (*Cos10*) in Muju population and one locus (*Cos13*) in Sannae population were found to be out of Hardy–Weinberg equilibrium after Bonferroni correction ($\alpha < 0.004$, $k = 14$). Null alleles may be present at

Table 1. Characterization of 14 polymorphic microsatellite loci for Korean shiner, *Coreoleuciscus splendidus*.

Locus	GenBank accession no.	Repeat motif	Primer sequences (5′-3′)	Ta (°C)	Fluorescent label
Cos9	JF298908	$(AC)_3GC(AC)_5$	F: TTC ACT TCT GGG CTT TGT TTG R: AAT CAT TTA AGC ACC TCT CAT	58	ROX
Cos10	JF298909	$(AC)_8$	F: ATC TCA TCA AAG TGG ATT GGC R: AAG CCC ACA CCT GAA TGA TAA	58	HEX
Cos11	JF298910	$(GT)_{13}$	F: AAT AAG CCC TCA CTA ACT GTA R: TAA AAA AAC AGC CCT CAA TGG	58	ROX
Cos13	JF298911	$(TG)_8$	F: CAC ACC TGA ATG ATA AAG TCG R: TCT CCA CAT GAA CAACAC GAA	58	ROX
Cos20	JF298912	$(AC)_6(AG)_{18}$	F: TAT TTG GTT GTG TCT TCT TGG R: TGT GTT TGG GTC TCT TTG TTT	58	HEX
Cos21	JF298913	$(ATAG)_{17} \sim (ACAG)_3 \sim (ATGG)_7$	F: TCA TCG TTG ACT TCC TCG TTC R: GGT CTT TAG ATT ATG TTA GGG	58	6-FAM
Cos26	JF298914	$(GT)_{14}$	F: GCT TCC CCA CTG TCT TCT TAA R: AGT CTC ACT CTT AGC CTT TGC	58	TAMRA
Cos35	JF298915	$(CA)_{14}$	F: AGT CTC ACT CTT AGC CTT TGC R: GTC TCC CCA CTG TCT TCT TTA	58	6-FAM
Cos45	JF298916	$(GACA)_7(CA)_{13}$	F: AAC ACA GGT GGA ATA TTG CTT R: CAA CAC GAC TCC ACA GAT ACT	58	TAMRA
Cos48	JF298917	$(TG)_6(AG)_4$	F: CCT TCG CAC AGA CTT CAC ATT R: CAC ACT CTC TCT CAC ACA CAC	58	6-FAM
Cos49	JF298918	$(GATA)_{16}$	F: ACA TTT ATT GCC TCA GCC TCA R: AAT TTG ATT GGT CCG AAC AGC	58	ROX
Cos56	JF298919	$(CA)_2CG(CA)_6$	F: CTT CAC TCT GGG CTT TGT TTG R: AAT GCG TCC CAA ATG TGT AAC	58	TAMRA
Cos67	JF298920	$(GT)_8GC(GT)_3(GCGT)_2(GTGC)_2$ $(GT)_4GC(GT)_9AT(GT)_4$	F: ATG GGT AGG TTT AGG GTT AGG R: CCC TTT AAT TTC ACC TCA GAT	58	6-FAM
Cos74	JF298921	$(GA)_{14}$	F: CCT GAT TTA TGG ACC TTC TGA R: AGA GAA CAG CAG GAA GAT GAA AGG	58	HEX

Note: Data comprise locus name, GenBank accession number, repeat motif, forward (F) and reverse (R) primer sequences, optimal annealing temperature (Ta), and fluorescent labels used for genotyping.

Table 2. Summary of genetic variability across 14 polymorphic microsatellite loci from three Korean shiner populations.

Locus	Alleles	Size range	Anheung (Han River)				Muju (Geum River)				Sannae (Nakdong River)				F_{IS}	F_{ST}
			n	H_O	H_E	p value	n	H_O	H_E	p value	n	H_O	H_E	p value		
Cos9	11	232–254	29	0.483	0.848	**0.000**	12	0.417	0.431	1.000	25	0.840	0.772	0.065	0.190	0.139
Cos10	29	188–278	29	0.724	0.928	0.004	12	0.583	0.949	**0.000**	25	0.640	0.644	0.127	0.195	0.155
Cos11	24	249–313	24	0.375	0.923	**0.000**	12	1.000	0.957	0.523	28	0.357	0.425	0.539	0.323	0.248
Cos13	27	95–205	26	0.385	0.922	**0.000**	12	0.917	0.913	0.625	20	0.050	0.573	0.000	0.532	0.177
Cos20	12	153–191	30	0.967	0.782	0.197	12	0.667	0.888	0.024	30	0.600	0.590	0.411	−0.064	0.255
Cos21	32	274–396	30	0.833	0.952	0.061	12	0.833	0.833	0.876	28	0.786	0.862	0.392	0.093	0.079
Cos26	11	242–266	30	0.633	0.614	0.957	12	0.417	0.359	1.000	30	0.500	0.470	0.645	−0.059	0.020
Cos35	7	223–235	29	0.655	0.596	0.851	12	0.583	0.540	0.524	26	0.038	0.112	0.019	−0.013	0.532
Cos45	12	184–206	30	0.867	0.864	0.289	12	0.500	0.511	1.000	30	0.900	0.788	0.006	−0.060	0.136
Cos48	5	262–272	30	0.500	0.488	1.000	12	0.583	0.641	0.665	25	–	–	–	0.012	0.633
Cos49	27	111–261	30	0.967	0.924	0.017	12	1.000	0.938	1.000	26	0.923	0.851	0.728	−0.065	0.041
Cos56	19	180–230	25	0.520	0.798	0.006	12	0.917	0.942	0.311	30	0.400	0.332	0.633	0.126	0.207
Cos67	4	103–111	26	0.115	0.111	1.000	12	–	–	–	27	0.333	0.297	1.000	−0.098	0.055
Cos74	22	181–227	30	0.633	0.811	0.022	12	0.833	0.946	0.671	28	0.393	0.378	1.000	0.137	0.167
		Total													0.104	0.215

Note: Data comprise the number of individuals examined (n), observed (H_O), expected (H_E) heterozygosity, and Hardy–Weinberg probability. p values showing significant deviations from Hardy–Weinberg equilibrium after Bonferroni correction ($\alpha < 0.004$, $k = 14$) are highlighted in bold.

loci *Cos13* and *Cos35* as suggested by the general excess of homozygotes for most allele size classes (MICRO-CHECKER). No linkage disequilibrium was detected between any pair of the 14 loci. The pairwise F_{ST} at each locus for three populations ranged 0.055 (*Cos67*) to 0.633 (*Cos48*), and the values across all loci show significant genetic differentiation ($F_{ST} = 0.215$, $p < 0.05$).

In the previous studies of the Korean shiner conducted using AFLP (Song et al. 2010) and two mtDNA markers (cytochrome *b*, and control region; unpublished data), a significantly high degree of genetic differentiation was exhibited between the populations from different two major drainages (i.e. Han River vs. Nakdong River). For instance, mean significant Φ_{ST} between Han River and Nakdong River was 0.960 (0.954–0.976, $p < 0.05$; cytochrome *b*) and 0.953 (0.938–0.967, $p < 0.05$; control region), respectively. The amount of divergence observed between those two populations is on the order of species-level

differentiations noted in other studies of cyprinids (e.g. Schmidt et al. 1998; McPhail and Taylor 1999; Zardoya and Doadrio 1999; Pfrender et al. 2004). On the other hand, the values within-Han and within-Geum were not significant and relatively low, suggesting weak genetic structuring within drainages (cytochrome *b*, and control region; unpublished data). The pairwise F_{ST} values obtained in the present study also confirmed that those three populations were significantly differentiated far beyond intra-specific level, although it is not yet clear if this pattern holds up across multiple drainages. In the test of the hierarchical AMOVA (analysis of molecular variance) significant genetic structure among populations was found across any grouping, although the grouping of Geum–Nakdong and Han showed negligible proportion between the two groups (negative value, Table 3). Relatively high proportion of genetic variance was found within populations in any grouping (Table 3), suggesting that 14 microsatellites developed here should particularly be

Table 3. Summary of the hierarchical AMOVA (analysis of molecular variance) results that partition genetic variation among and within three Korean shiner populations.

Source of variation	F_{ST}	F_{CT}	F_{SC}	%Among groups	%Within groups	p
One gene pool (populations)	**0.139**	–	–	13.93	86.07	<0.001
((Han, Geum), Nakdong)	**0.148**	0.046	**0.107**	4.56	10.24	<0.001
(Han, (Geum, Nakdong))	**0.113**	−0.136	**0.219**	−13.61	24.94	<0.001
((Han, Nakdong), Geum)	**0.167**	0.058	**0.116**	5.83	10.89	<0.001

Note: F_{ST}, the proportion of genetic variation between all sample sites; F_{SC}, the proportion of genetic variation between site within regions; F_{CT}, the proportion of genetic variation between regions within the total sample; NS, not significant; $p < 0.05$ is highlighted in bold.

Table 4. Cross-species amplification tests of polymorphic microsatellite loci obtained from the Korean shiner in six other species of Gobioninae.

Locus	Ta (°C)	Ppa	Phe	Sva	Sgr	Hla	Gna
Cos9	54	P (2)	P (2)	P (2)	P (2)	P (2)	P (2)
Cos10	54	NA	NA	NA	NA	NA	NA
Cos11	54	P (2)	P (2)	M	NA	M	P (2)
Cos13	54	NA	NA	NA	NA	NA	NA
Cos20	54	NA	M	M	NA	P (4)	NA
Cos21	54	NA	P (2)	M	NA	P (2)	P (3)
Cos26	54	P (2)	P (2)	P (4)	P (2)	P (3)	P (3)
Cos35	54	P (2)	P (4)	P (2)	P (2)	P (2)	P (4)
Cos45	54	P (2)	P (2)	M	P (2)	P (2)	P (3)
Cos48	54	P (2)	P (4)	P (2)	P (2)	P (2)	P (3)
Cos49	54	P (2)	P (2)	NA	P (2)	P (2)	P (2)
Cos56	54	P (2)	P (2)	P (2)	P (2)	P (2)	P (2)
Cos67	54	NA	NA	NA	NA	NA	NA
Cos74	54	P (2)	P (4)	P (2)	P (2)	P (2)	M

Note: Ta, annealing temperature; Ppa, Pseudorasbora parva; Phe, Pungtungia herzi; Sva, Sarcocheilichthys variegatus wakiyae; Sgr, Squalidus gracilis majimae; Hla, Hemibarbus labeo; Gna, Gobiobotia naktongensis; P, polymorphic; M, monomorphic; NA, No amplification product; number in parentheses indicates number of the allele.

useful to explain population structures and gene flow dynamics within drainages.

Fourteen polymorphic microsatellite loci were further tested for cross-species amplifications with six other Gobioninae species including *P. parva*, *P. herzi*, *S. variegatus wakiyae*, *S. gracilis majimae*, *H. labeo*, and *G. naktongensis* under the same amplification condition as used for the Korean shiner (Table 4). Although three loci (*Cos10, Cos13*, and *Cos67*) were not successful to amplify in all 6 Gobioninae species, the remaining 11 microsatellite loci were successfully working with polymorphism in at least one of the 6 species, indicating the possibility that those loci may also be used in closely related other species.

Acknowledgements

References

An HS, Hong SW, Kim EM, Lee J, Noh JK, Kim HC, Park CJ, Min BH, Myeong J. 2010. Comparative genetic diversity of wild and released populations of Pacific abalone *Haliotis discus discus* in Jeju, Korea, based on cross-species microsatellite markers including two novel loci. Anim Cells Syst. 14:205–313.

Bhatia S, Arora R. 2007. Genetic characterization and differentiation of Indian sheep breeds using microsatellite marker information. Korean J Genet. 29:297–306.

Bos DH, Gopurenko D, Williams RN, DeWoody JA. 2008. Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. Evolution. 62:1458–1468.

Choi KC, Choi SS, Hong YP. 1990. On the microdistribution of fresh-water fish, *Coreoleuciscus splendidus* (Gobioninae) from Korea. Korean J Ichthyol. 2:63–76 (in Korean).

Domínguez-Domínguez O, Alda F, Pérez-Ponce de León G, García-Garitagoitia JL, Doadrio I. 2008. Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* (Bean, 1898) (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography based on mitochondrial and nuclear DNA data. BMC Evol Biol. 8:161.

Goudet J. 1995. FSTAT: a computer program to calculate F-statistics. J Hered. 86:485–486.

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. Biotechniques. 27:500–507.

Kashi Y, King D, Soller M. 1997. Simple sequence repeats as a source of quantitative genetic variation. Trends Gen. 13:74–78.

Li YC, Korol AB, Fahima T, Beiles A, Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol. 11:2453–2465.

McPhail JD, Taylor EB. 1999. Morphological and genetic variation in northwestern longnose suckers, *Catostomus catostomus*: the Salish sucker problem. Copeia. 4:884–893.

Pfrender ME, Hicks J, Lynch M. 2004. Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). Mol Phylogenet Evol. 30:490–502.

Raymond M, Rousset F. 1995. GENEPOP (ver. 1.2): population genetics software for extract tests and ecumenicism. J Hered. 86:248–249.

Remya KS, Joseph S, Lakshmi PK, Akhila S. 2010. Microsatellites in varied arenas of research. J Pharm Bioall Sci. 2:141–143.

Schlötterer C, Harr B. 2000. *Drosophila virilis* has long and highly polymorphic microsatellites. Mol Biol Evol. 17:1641–1646.

Schmidt TR, Bielawski JP, Gold JR. 1998. Molecular phylogenetics and evolution of the cytochrome *b* gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). Copeia. 1998:14–22.

Seifertová M, Šimková A. 2010. Structure, diversity and evolutionary patterns of expressed MHC class II*B* genes in chub (*Squalius cephalus*), a cyprinid fish species from Europe. Immunogenetics. 63:167–181.

Song H, Kim K, Yoon M, Nam YK, Kim DS, Bang I. 2010. Genetic variation of *Coreoleuciscus splendidus* populations (Teleostei; Cyrpriniformes) from four major

river drainage systems in South Korea as assessed by AFLP markers. Genes Genom. 32:199–205.

Suk HY, Neff BD. 2009. Microsatellite genetic differentiation among populations of the Trinidadian guppy. Heredity. 102:425–434.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 4:535–538.

Zane L, Bargelloni L, Patarnello T. 2002. Strategies for microsatellite isolation: a review. Mol Ecol. 11:1–16.

Zardoya R, Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J Mol Evol. 49:227–237.