© Indian Academy of Sciences

ONLINE RESOURCES

Development of twenty-nine polymorphic microsatellite loci from largemouth bronze gudgeon (*Coreius guichenoti*)

MEIHUA XIONG, SHUXIANG YAN, KE SHAO, WEITAO LI, FANG SHI, BIN ZHU and NIAN XU*

Key Laboratory of Ecological Impacts of Hydraulic-Projects and Restoration of Aquatic Ecosystem of Ministry of Water Resources, Institute of Hydroecology, Ministry of Water Resources and Chinese Academy of Sciences, 578# Xiongchu Avenue, Wuhan 430079, People's Republic of China

[Xiong M., Yan S., Shao K., Li W., Shi F., Zhu B. and Xu N. 2014 Development of twenty-nine polymorphic microsatellite loci from largemouth bronze gudgeon (*Coreius guichenoti*). *J. Genet.* **93**, e100–e103. Online only: http://www.ias.ac.in/jgenet/OnlineResources/93/e100.pdf]

Introduction

Largemouth bronze gudgeon (*Coreius guichenoti*) is an endemic freshwater fish in the upper reaches of the Yangtze river in China (Ding 1994). As a typical migratory river fish, broodstocks of *C. guichenoti* spawn in the middle and lower reaches of the Jinsha river, which is part of the upper reaches of the Yangtze river. The eggs hatch in the process of drifting downstream, then the juvenile fish is traced towards upstream of the Yangtze river to spawn (Cao 2008). The impoundment of Xiangjia dam in 2012 and Xiluodu dam in 2013 on the lower reaches of the Jinsha river have blocked the migration routes of *C. guichenoti* supplementary stocks because of which the number of broodstocks declined significantly. Artificial propagation of *C. guichenoti* has been developed to conserve the resources. Besides increasing the number of *C. guichenoti* populations, maintaining the genetic diversity of populations should also be concerned in the conservation strategy.

To date, 29 polymorphic microsatellites have been published for *C. guichenoti*, and 25 of their motifs are dinucleotide repeats (Liao *et al.* 2007; Xu *et al.* 2007). Genetic diversity in seven populations of *C. guichenoti* has been examined using 11 loci isolated by Liao *et al.* (2006, 2007). Dramatic population differentiation among populations which were seperated by dams was observed (Zhang and Tan 2010). No significant variation occurred among populations in continuous habitat (Zhang and Tan 2010), which was consistent with the study by Yuan *et al.* (2008). Therefore it is necessary to analyse the genetic diversity of populations seperated by Xiangjia dam and Xiluodu dam.

Dinucleotide microsatellites have been demonstrated to be less polymorphic and stable than trinucleotide and tetranucleotide repeated microsatellites (Edwards *et al.* 1991). Polynucleotide-repeated microsatellites are less stuttering and it is easier to distinguish different alleles (Shan *et al.* 2014). Moreover, the number of loci has also played an important role in classifying individuals by relatedness. One can exclude nonparent–offspring pairs with almost complete certainty by sampling more loci (Blouin *et al.* 1996). Therefore, more loci from *C. guichenoti* would be needed to guide genetic management decisions for artificial propagation populations. In this study 29 novel microsatellites including 27 polynucleotide repeats were isolated from *C. guichenoti*. These loci could be useful for the studies of genetic diversity, population genetic structure, estimating relatedness between individuals in artificial propagation and release.

Materials and methods

Genomic DNA for genomic library construction was extracted using the standard proteinase K / phenol extraction protocol (Sambrook and Russell 2001) from fin clips of two wild *C. guichenoti* individuals previously preserved in 100% ethanol. Microsatellite loci were isolated from two-enriched partial genomic libraries for the repeat motifs $(AC)_n$ and $(GATA)_n$ following essentially the FIASCO protocol (Zane *et al.* 2002) with a few modifications developed by Xiong *et al.* (2011). DNA samples were digested with restriction enzyme *Mse*I (New England Biolabs, Ipswich, UK). Fragments 100–800 bp in size were purified and ligated to *Mse*I adaptors (OligoA: 5′-TAC TCA GGA CTC AT-3′, OligoB: 5′-GAC GAT GAG TCC TGA G-3′) using T4 DNA ligase (Takara, Dalian, China). DNA concentration was elevated through PCR amplification using *Mse*I -N (5′-GAT GAG TCC TGA GTA AN-3′) as the primer. The PCR products were then hybridized to 5′-biotinylated probes $(AC)_n$ or $(GATA)_n$. Single-stranded DNA fragments contained

*For correspondence. E-mail: xunianbubu@gmail.com.

Keywords. microsatellite; polymorphic; *Coreius guichenoti.*

Table 1. Characterization of 29 polymorphic microsatellite loci in *Coreius guichenoti*.

Locus	Accession no.	Primer sequence (5′–3′)	Repeat sequence	Size range (bp)	T_a	N_A	H_O	H_E	PIC	P (HWE)	Null present	P (null alleles)
CGU001	KF586769	AACCAAAACTCCTCCCTTCC (tailed) AACAGACAACGATGCAAACG	$(TTTC)_4(TTCC)(TTTC)_{16}$	161–265	54	24	0.871	0.959	0.941	0.027	No	>0.05
CGU003	KF586771	GTAACAGGAACGGACAAC (tailed) GACGGAAGGACAGATGG	$(ATCT)_5$	204–232	52	7	0.567	0.772	0.721	0.115	Yes	<0.025
CGU006	KF586774	AACCAGAAGATCAGCAAA (tailed) AAGGAATGTTCACAAGG	$(GACA)_3(GATA)_{23}$	227–323	52	20	0.833	0.94	0.92	0.031	No	>0.05
CGU041	KF586809	TGGAAATGACTCTGGTA (tailed) ATGGCTGACTTCACTAA	$(TCTA)_4(TCCA)(TCTA)_{13}$ $(TCCA)(TCTA)_2$	136–220	46	18	0.781	0.946	0.927	0.040	No	>0.05
CGU046	KF586814	ACAGATAGACGGACGAA (tailed) CACTCTGCCCTACTACA	$(GATA)_{10}$	232–280	50	21	0.875	0.948	0.929	0.224	No	>0.05
CGU060	KF586828	CATTTTCTCACTGGCTG (tailed) ATTGCTGTATTCTGATGAT	$(TCTG)_4(TCTA)(TCTG)_2$ $(TCTA)(TCTG)_5$	112–160	48	12	0.679	0.818	0.779	0.062	No	>0.05
CGU068	KF586836	ATGGTGTTCCTAATAATCCTTT (tailed) CAATGTAATGATGAAATCCTGT	$(GATA)_9$	148–180	52	9	0.719	0.851	0.819	0.192	No	>0.05
CGU069	KF586837	GTCAACAGCGTAACAACAGAAT (tailed) AATACCCTAAACCGCAAGAA	$(CTTT)_6$	195–211	56	5	0.156	0.682	0.623	0.000*	Yes	<0.001
CGU070	KF586838	AGGATGATGCTGCTGCTTT (tailed) AGAGGCTGTTAGTGATGATGTT	$(TTTC)_{10}$	152–240	56	17	0.906	0.938	0.919	0.616	No	>0.05
CGU072	KF586840	CGCTGAACTCCGTGACCTGAT (tailed) GCTCTTTCCTCCCTGTAACTCG	$(GT)_{15}$	230–266	58	10	0.31	0.863	0.829	0.000*	Yes	<0.001
CGU073	KF828901	ACCGTAACAACTGCTGAC (tailed) GAACCGTGGGTGGAGAA	$(TCTA)_3(TCCA)_8$	217–337	50	20	0.806	0.94	0.92	0.038	No	>0.05
CGU074	KF828902	GGCGTAGAGTTGATGTA (tailed) CATGGGATTCAAGTGTT	$(TCTA)_8$	122–278	53	25	0.781	0.952	0.934	0.000*	No	>0.05
CGU075	KF828903	CACAAACAACAATCCAC (tailed) TCCATATCTATCCATCG	$(GATA)_{11}$	106–142	50	9	0.633	0.878	0.848	0.007	Yes	<0.01
CGU076	KF828904	TGAGGAGTATTATTTTGGACC (tailed) TGCAACAAGCTACAGGATC	$(GACA)_7$	100–120	53	6	0.25	0.711	0.656	0.000*	Yes	<0.001
CGU077	KF828905	AGAACCTTGCCAGAACG (tailed) GCCGCAGCACTGCTGAC	$(TTGTGTG)_3(TG)_3$	240–256	53	5	0.313	0.729	0.67	.000*	Yes	<0.001
CGU078	KF828906	GAAGAGGCAATTACAAG (tailed) TAGTGCGATAGATAGGT	$(TCTA)_{16}$	103–231	54	24	0.733	0.946	0.926	0.000*	No	>0.05
CGU079	KF828907	ATGGGAAACAATAGAGTG (tailed) GAGGAGGATTATGTGGAA	$(TAGA)_5(CAGA)_7$	165–197	50	8	0.500	0.818	0.778	0.000*	Yes	<0.001
CGU103	KF828931	GCCTTAGGAGGGTGGAT (tailed) ATTGGAATGACAAGAGTGGA	$(GACA)_7(GACG)$ $(GATG)_2(GATA)_{20}$	281–433	50	24	0.75	0.952	0.934	0.000*	No	<0.05
CGU104	KF828932	ACCATTACACCACCAGCAGC (tailed) TAGGCAGTCAGATAGATAGAACGA	$(TCTA)_9$	104–162	62	11	0.813	0.846	0.813	0.354	No	>0.05
CGU114	KF828942	ACTGAATGAACATCCGACT (tailed) GTCTTGTGGAGTGAAACG	$(GATA)_{14}$	165–241	54	18	0.903	0.943	0.924	0.555	No	>0.05
CGU115	KF828943	CCATCCTATTATCACAAGCGTC (tailed) AAGGGAACTGGCAGGTAAAG	$(CATA)_7$	187–223	62	9	0.313	0.82	0.784	0.000*	Yes	<0.001
CGU124	KF828952	GTAGAGGGTCACGCTATGTTT (tailed) GTTTAGGCACTGGGTAGGATT	$(GATA)_7$	328–492	62	26	0.813	0.963	0.945	0.047	No	>0.05

Table 1 (*contd*)

Locus	Accession no.	Primer sequence (5′–3′)	Repeat sequence	Size range (bp)	T_a	N_A	H_O	H_E	PIC	P (HWE)	Null present	P (null alleles)
CGU127	KF828955	CGCTACCCAGAAGAGGC (tailed) GCTTGAAGATGTTACGA	$(TCTA)_{37}$	204–324	50	24	0.806	0.962	0.944	0.042	No	>0.05
CGU135	KF828962	AGAATGAATTGATCGGATGT (tailed) GGTGTACCTAATAAAGTGGC	$(TATC)_5$	212–240	57	8	0.594	0.776	0.731	0.001*	Yes	<0.01
CGU137	KF828964	TTTTCTTCCTGAGGTGGC (tailed) ACTGAATGAACATCCGACT	$(TCTA)_8$	211–287	54	17	0.800	0.941	0.92	0.085	No	>0.05
CGU144	KF828971	ATGGCACTTTTCTTGGC (tailed) GGTATCGGCAGTTCACA	$(TGTC)_2(TCTG)_3(TCTA)_4$	98–170	46	10	0.531	0.859	0.828	0.003	Yes	<0.001
CGU150	KF828977	TGAGGGAAAGACAGGAAAAAAT (tailed) TATTTGTAAACTTGCGGCTTG	$(TATC)_{11}$	103–183	54	19	0.844	0.928	0.907	0.166	No	>0.05
CGU155	KF828982	TCTGGGCTGATAACAAA (tailed) AAGTGAACTGCGAATGG	$(CA)_{24}$	254–354	48	20	0.968	0.935	0.915	0.953	No	>0.05
CGU162	KF828989	CTCTAAGGAGCGAAACT (tailed) TAATCTGAAAGCGAAAG	$(GATA)_{16}$	196–404	48	30	0.833	0.972	0.954	0.000*	No	>0.05
Mean						15.7	0.679	0.882	0.853			

T_a, optimal annealing temperature; N_A, number of alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; P (HWE), P values for HWE were calculated by exact test ($N_A \leq 4$) or Markov chain method ($N_A > 4$); *significant deviation from HWE after Bonferroni correction (adjusted $P = 0.0016$); P (null alleles), combined P values for all classes of alleles.

microsatellites were captured by Streptavidin Magnetic beads (Roche, Basel, Switzerland). Fragments containing microsatellites were ligated to a PMD-18T (Takara) and transformed into DH5-α competent cells. Then the recombinants were screened by PCR using a probe primer and the M13 primers. Positive clones were selected and sequenced using ABI 3730 XL sequencer. A total of 998 positive clones were identified, 210 of which were found with microsatellite repeats. One hundred and seventy-three primer pairs were designed using Primer Premier 5 software (Zhang and Gao 2004) and were commercially synthesized (Tianyibiotech, Beijing, China).

PCR conditions for each primer pair were optimized in an initial test on four *C. guichenoti* individuals, and 36 primers produced discriminable products. A 5′-M13 tail (5′-CAC GAC GTT GTA AAA CGA C) was then added to the forward primer of 36 primer pairs to allow fluorescent labelling during the amplification reaction. The 36 primer pairs were further tested on 32 individuals, of which, 10 were captured from Yibin section of Yangtze river, 12 from the Chongqing section of Yangtze river and 10 from Wu river, the largest tributary on the southern bank of Yangtze river. PCR amplification was performed in a volume of 10 μL containing 1 μL 10× PCR buffer (Takara), 2 mM dNTP, 1 μM the reverse primer and the fluorescent-labelled M13 primer (LI-COR, Lincoln, USA), 0.5 U r*Taq* DNA polymerase (Takara), 20 ng DNA template. PCR amplification was conducted with an initial denaturation at 94°C for 5 min, 35 cycles of 94°C for 40 s, primer-specific annealing temperature (table 1) for 40 s, 72°C for 40 s, and followed by a final 10 min extension at 72°C. The PCR products were loaded on a LI-COR 4300 automated DNA Sequencer together with a size standard (50–500 bp) and genotyping was subsequently carried out using SAGAGT software (LI-COR).

The variability at each locus was measured by numbers of alleles, observed heterozygosity and expected heterozygosity. The polymorphism of each locus was measured by polymorphism information content (PIC). Numbers of alleles, observed heterozygosity, expected heterozygosity and PIC were calculated using Cervus 3.0 (Marshall *et al.* 1998). Deviations from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were tested using GenePop 3.4 (Raymond and Rousset 1995). Null allele tests in each locus were performed using Micro-Checker 2.2.3 (Oosterhout *et al.* 2004).

Results and discussion

Thirty-six loci were successfully genotyped, 29 of which showed clearly polymorphic banding patterns in target regions with a maximum of two alleles for each locus per individual. GenBank accession, primer sequences, number of alleles, allele size range, observed and expected heterozygosities, PIC and optimal annealing temperatures are listed in table 1. The 29 polymorphic loci produced 456 alleles

ranging from 5 to 30 with an average of 15.7 alleles per locus (table 1). The observed heterozygosity values ranged from 0.156 to 0.968 with an average of 0.679, and the expected heterozygosity ranged from 0.682 to 0.972 with an average of 0.882. The PIC ranged from 0.623 to 0.954 at the average of 0.853. Analyses using GenePop revealed that significant LD was observed between three locus pairs, CGU068 and CGU069; CGU060 and CGU073; CGU006 and CGU124. Eleven loci deviated from HWE (table 1), with the estimation of exact P values by the Markov chain method ($P < 0.0016$) after sequential Bonferroni correction (Rice 1989). Null allele tests using Micro-Checker showed 10 loci had null alleles (details in table 1). The results of 11 loci deviated from HWE may be related to the 32 individuals tested.

Acknowledgements

Authors are grateful to Huiyuan Tang, Zhi Yang, and Yanfu Que for sample collection, Dongmei Xu for her assistance in the experiment. This work was supported by NSF of China (51409176) the research of population dynamics and genetic structure of *Coreius guichenoti* (0701 1032), NSF of China (51109145 51309166) and projects.

References

Blouin M., Parsons M., Lacaille V. and Lotz S. 1996 Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.* **5**, 393–401.

Cao W. 2008 Several issues related to conservation of fish resources in the Yangtze River. *Resour. Environ. Yangtze Bas.* **17**, 163–164.

Ding R. 1994 *Fishes of Sichuan*. Sichuan Sciences Press, Chengdu, China.

Edwards A., Civitello A., Hammond H. and Caskey C. 1991 DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**, 746–756.

Liao X., Zhu B., Yu X., Tan D., Chang J. and Tong J. 2006 Isolation and characterization of polymorphic microsatellites in a Yangtze River fish, brass gudgeon (*Coreius heterodon Bleeker*). *Mol. Ecol. Notes* **6**, 393–395.

Liao X., Yu X., Chang J. and Tong J. 2007 Polymorphic microsatellites in largemouth bronze gudgeon (*Coreius guichenoti*) developed from repeat-enriched libraries and cross-species amplifications. *Mol. Ecol. Notes* **7**, 1104–1107.

Marshall T. C., Slate J., Kruuk L. E. B. and Pemberton J. M. 1998 Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7** 639655.

Oosterhout C. V., Hutchinson W. F., Wills D. P. M. and Shipley P. 2004 Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538.

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

Rice R. W. 1989 Analyzing tables of statistical tests. *Evolution* **43**, 223–225.

Sambrook J. and Russell D. W. 2001 *Molecular cloning: a laboratory manual*, 3rd edition. Cold Spring Harbor Laboratory Press, New York, USA.

Shan Y., Lu C., Li C., Cheng L. and Sun X. 2014 Isolation and characterization of 57 novel polynucleotide microsatellites from yellow catfish (*Pelteobagrus fulvidraco*) genome for genetic analysis. *Conser. Genet. Resour.* **6**, 73–77.

Xiong M., Shi F., Xu N., Que Y., Yang Z., Hu J. and Zhu B. 2011 Isolation and characterization of microsatellite loci in *Spinibarbus sinensis* (Bleeker). *J. Hydroecol.* **32**, 40–44.

Xu S., Zhang Y., Wang D., Li Z. and Chen D. 2007 Genetic diversity in largemouth bronze gudgeon (*Coreius guichenoti* Sauvage et Dabry) from Yibin section of Yangtze River based on sequence of microsatellite DNA. *Freshwater Fisher.* **37**, 76–79.

Yuan X., Yan L., Xu S., Wang D., Zhang Y. and Chen D. 2008 Genetic diversity of bronze gudgeon (*Coreius heterodon*) and largemouth bronze gudgeon (*C. guichenoti*) in Yangtze River Basin. *J. Fish. Sci. Chin.* **15**, 377–385.

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

Zhang F. and Tan D. 2010 Genetic diversity in population of largemouth bronze gudgeon (*Coreius guichenoti* Sauvage et Dabry) from Yangtze River determined by microsatellite DNA analysis. *Genes Genet. Syst.* **85**, 351–357.

Zhang X. and Gao Y. 2004 To design PCR primers with Oligo 6 and Primer Premier 5. *China J. Bioinf.* **2**, 15–18.

Received 26 November 2013, in revised form 27 March 2014; accepted 3 June 2014
Unedited version published online: 25 June 2014
Final version published online: 6 November 2014