GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Development and characterization of microsatellite markers for the lizardfish known as the Bombay duck, *Harpadon nehereus* (Synodontidae)

T.J. Xu, D.Q. Sun, H.Y. Li and R.X. Wang

Key Laboratory for Marine Living Resources and Molecular Engineering, College of Marine Science, Zhejiang Ocean University, Zhoushan, P.R. China

Corresponding author. R.X. Wang
E-mail: wangrixin1123@126.com

Genet. Mol. Res. 10 (3): 1701-1706 (2011)
Received August 18, 2010
Accepted July 21, 2011
Published August 16, 2011
DOI http://dx.doi.org/10.4238/vol10-3gmr1045

ABSTRACT. The Bombay duck, or bummalo (*Harpadon nehereus*), is a lizardfish native to the Arabian sea, but also common in the China sea. It is normally dried and salted before consumption and export. To provide molecular information on this economically important fish species, we developed and characterized microsatellite markers. Ninety positive clones from the $(CA)_{15}$-enriched genomic library were sequenced; 62 sequences contained sufficient repeat motifs (di-, tri- and tetra-nucleotide). Twenty-eight primer pairs were designed and 21 were successfully amplified; five loci were polymorphic, but with a low number of alleles (three or four). The observed and expected heterozygosities ranged from 0.3500 to 0.8421 and from 0.5244 to 0.6244, respectively. All of the five polymorphic loci were at Hardy-Weinberg equilibrium (adjusted), while linkage disequilibrium between Hane-97 and Hane-175 ($P < 0.05$, adjusted P value = 0.01) was significant. The low degree of polymorphism of microsatellite markers may be due to the large size of the repeat motifs. These markers will be useful for genetic diversity analysis of *H. nehereus*.

Key words: *Harpadon nehereus*; Microsatellite; Molecular marker

INTRODUCTION

Harpadon nehereus (Synodontidae), which is usually known as the Bombay duck, despite being a lizardfish, is an inshore shallow water fish widely distributed along the coastal of China. It has been an important commercial fish for domestic use and also a valuable export item, in dried or laminated form due to its highly perishable body composition. However, no information on population genetics using molecular markers has been reported for understanding population differentiation and clarifying species identity.

Microsatellites, also known as simple sequence repeats (SSRs), are regions of DNA that exhibit short repetitive sequence motifs (Degnan and Arévalo, 2004). Microsatellites were widely employed in population genetic studies of numerous species, and this application is continuously expanding (Kohlmann et al., 2005; Mia et al., 2005). Those motifs are often composed of 1-6 bp repeats and there is often a high degree of polymorphism for different numbers of repeats. As microsatellites have some unique characteristics, such as co-dominance of alleles, high allelic diversity, and relatively simple polymerase chain reaction (PCR)-based screening methods that are reproducible, they have been widely used in recent years for population genetic studies of numerous species (Sekino and Hara, 2001; Kohlmann et al., 2005; Mia et al., 2005; Selkoe and Toonen, 2006; Liu et al., 2009). At the same time, microsatellites have been successfully used in genetic mapping and genome analysis (Chen et al., 1997; Li et al., 2000), genotype identification, variety protection seed purity evaluation (Senior et al., 1998), germplasm conservation (Brown et al., 1996), and marker-assisted breeding (Weising et al., 1998). Here, we report on the development and genetic characterization of microsatellite loci for *H. nehereus*.

MATERIAL AND METHODS

DNA extraction and enrichment for microsatellites

DNA was extracted from 36 samples collected from nine different geographic sampling sites. The nine sites are located along the East Sea and the South Sea of China (Figure 1). Total genomic DNA was isolated from the muscle tissue using the standard phenol-chloroform method with some modifications; this was subsequently dissolved in 100 µL TE buffer and then its quality was checked on a 1% agarose gel, stained with ethidium bromide, and compared with known molecular weight standards (Tiangen). DNA was stored at -20°C before use. An enriched partial genomic library for the repeat motif $(CA)_{15}$ was constructed using a DNA pool composed of four individuals from different sites, and following the modified FIASCO protocol (Zane et al., 2002).

The total genomic DNA pool was simultaneously digested with the *Mse*I restriction enzyme (NEB). DNA fragments ranging from 250 to 1000 bp were isolated from a 1.5% agarose gel and purified using a TIANgel Midi Purification Kit (Tiangen). These fragments were ligated to *Mse*I amplified fragment length polymorphism (AFLP) adapters OligoA (5'-TAC TCA GGA CTC AT-3') and OligoB (5'-GAC GAT GAG TCC TGA G-3') and then were amplified in a total volume of 50 µL by PCR using adaptor-specific primers (*Mse*I-N: 5'-GATGAGTCCTGAGTAAN-3'). PCR conditions were: 1X *Taq* DNA polymerase buffer (Tiangen), 300 ng *Mse*I-N, 200 µM of each dNTPs, 0.8 µL *Taq* DNA

Figure 1. Sample sites for the *Harpadon nehereus*.

polymerase (Tiangen, 5 U/μL) and 20 μL of a 1/10 dilution of digested-ligation DNA. The cycling conditions were 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 53°C for 30 s, and 72°C for 1 min, followed by 1 cycle of 72°C for 10 min and then holding at 10°C. The DNA fragments between 300 and 1000 bp were isolated.

Genomic DNA fragments containing SSR were captured by hybridization to $(CA)_{15}$ biotin-labeled probes. For enrichment, the adaptor-ligated DNA fragments was denatured at 95°C for 8 min, and then hybridized to biotin-labeled probes in 70 μL hybridization solution at 65°C for 30 min. The DNA hybridized to the probe was separated and captured by streptavidin magnetic beads at room temperature for 30 min, followed by nonspecific and specific washing. The final captured fragments were ligated to pMD19-T vectors (Takara), according to manufacturer instructions, then cloned using the DH5α cells, following the standard protocol.

Primer design and PCR amplification

Ninety positive clones, which were screened via PCR with M13-F/M13-R and $(CA)_{15}$ primers, were sequenced using M13-F primer on an ABI 3730 automated sequencer; 62 sequences contained sufficient repeat motifs (di-, tri- and tetra-nucleotide). Considering that the repeat size will affect the polymorphism, 28 primers were designed for loci containing repeat motifs (12 bp ≤ repeating size ≤ 52 bp) using the PRIMER PREMIER 5.0 software. The optimal conditions of PCR amplifications for each primer were adjusted by the annealing temperature and Mg^{2+} concentration. Then, polymorphism at each locus was determined us-

ing 36 individuals, four individuals from each of the nine sites. PCR amplifications were carried out in 25 μL volumes containing 1X PCR buffer (Tiangen), 2 μL dNTPs, 0.6 μL forward and reverse primers, and 0.4 μL *Taq* polymerase (Tiangen, 5 U/μL). The cycling conditions were 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, annealing temperature (Table 1) for 30 s, and 72°C for 30 s, followed by 1 cycle of 72°C for 5 min and then holding at 4°C. PCR amplification was performed on an ABI 9700 thermal cycler.

The products of PCR amplifications were determined on 1.5% agarose gels, and then were denatured at 96°C for 8 min using denaturant (98% formamide, 10 mM EDTA, 0.25% bromophenol blue, 0.25% FF). The denatured amplified products were separated on 6% denaturing polyacrylamide gels and visualized using silver staining. A denatured pBR322 DNA/*Msp*I molecular weight marker (Tiangen) was used as a size standard to identify alleles (Xu et al., 2009).

Data analysis

The number of alleles, and observed (H_O) and expected (H_E) heterozygosities were estimated using POPGENE32 (Yeh and Boyle, 1997), and the ARLEQUIN 3.11 software (Schneider et al., 2000) was used to calculate Hardy-Weinberg equilibrium expectations and genotypic linkage disequilibrium. All results for multiple tests were corrected using the Bonferroni's correction (Rice, 1989).

RESULTS AND DISCUSSION

Details of the newly developed microsatellite loci are listed in Table 1. Twenty-one of the 28 primer pairs were amplified successfully and five loci were shown to be polymorphic in *H. nehereus*. The variability measures across 36 individuals of Bombay duck are also summarized in Table 1. The number of alleles per locus was three or four, with a mean of 3.4, and H_O and H_E heterozygosities ranged from 0.3500 to 0.8421 and from 0.5244 to 0.6244, respectively. The polymorphism information content per locus ranged from 0.4403 to 0.5538. Hardy-Weinberg equilibrium probability tests showed that all of the five polymorphic loci were at Hardy-Weinberg equilibrium (adjusted P value >0.01). According to pairwise tests for linkage disequilibrium, significant pairwise test values were found between Hane-97 and Hane-175 ($P < 0.05$, adjusted P value <0.01).

Our results suggest that *H. nehereus* genome has abundant microsatellites (69%), but the polymorphism of the loci was low; only five of the 21 loci showed polymorphism and the mean number of alleles was 3.4. The low polymorphism maybe due to the motif repeat size, which is thought to be related to the polymorphism per locus (Weber, 1990; Brandström and Ellegren, 2008). Some microsatellites found in *H. nehereus* repeated either more than 20 times or only six times, while the size of polymorphic loci in *H. nehereus* all ranged from 8 to 13 times, and there were many compound and imperfect types of microsatellites, which may limit the mutation rates. These polymorphic microsatellite DNA markers that we developed could be useful to evaluate genetic variation in *H. nehereus*.

ACKNOWLEDGMENTS

Research supported by the Open Foundation of Ocean Fishery Science and Technology in the Most Important Subjects of Zhejiang (#20100209).

Table 1. Characterization of 21 microsatellite loci of *Harpadon nehereus* developed from an enriched genomic library.

Locus	Motif	Type	Primer sequence	Size (bp)	*Tm* (°C)	Mg^{2+}(mM)	N_A	H_O/H_E	PIC	Accession No.
Hane-22	$(AGAC)_{18}$	P	TAATACTGAATAGAATGGAC GTCTTGGTAACTTGACTTT	184	54	1.5	1	- -	-	HQ185582
Hane-23	$(AG)_8$	P	GTGTTTCCTTTGGTAATG ATACAACTCGTCACTCGC	169-173	56	1.0	3	0.8000 0.5244	0.4403	HQ185583
Hane-38	$(TC)_8TT(TC)_4$	I	ATGAGGAAACCTGAAATAA TTCTCAGCACATCCGTAG	128	50	1.5	1	- -	-	HQ185584
Hane-50	$(GA)_{12}$	P	GCAGCCGTACAGTCAA GGCAGTTCTAATGGGATA	164-168	50	1.5	4	0.4000 0.6244	0.5450	HQ185585
Hane-59	$(CCT)_3CT(CCT)_4CT$ $(CCT)_2GT(CCT)_3$	I	GACGGAACCGAGCAAACT GACCGCAGCAAAGCATCT	133	50	1.6	2	- -	-	HQ185586
Hane-68	$(GA)_7(CA)_6$	C	ATACCGATTTTACTTATTTCC ATCAGCCTCATTAGCCAA	169	48	1.5	1	- -	-	HQ185587
Hane-69	$(GA)_5A(GA)_3$	I	ATAGTAAGTCGCAAGGTA CAGGCAGTATAGGTAAAT	132	50	0.7	1	- -	-	HQ185588
Hane-74	$(CT)_{26}\text{^}(CT)_8$	P	TCATTTTCTGTCTTTCACCCTC GCCCCAACCTGATCTTCT	156	56	1.5	1	- -	-	HQ185589
Hane-78	$(AG)_9$	P	CTTCCCATCATAGCATAA CATTCCTACCCATACACC	120	50	1.5	1	- -	-	HQ185590
Hane-87	$(GA)_{11}$	P	CCAGAACTGAGGGTGT CTGCCAAACTAATGAAA	164	52	1.5	1	- -	-	HQ185591
Hane-97	$(TC)_{13}$	P	ATGAGGGCACAGGAGGAA AGCAAGGAGAAGGCAAGA	163-169	60	1.5	4	0.4000 0.6038	0.5430	HQ18559
Hane-107	$(TG)_6$	P	CTTATAGTTGGAGCCGTGGTT GGCAGCCGTGCTTTAGTG	103	52	1.0	2	- -	-	HQ185593
Hane-119	$(CT)_8$	P	TGACGCCGCTATGACCTG TGACCTCCTTGCCCTGCT	206	56	1.5	2	- -	-	HQ185594
Hane-129	$(AC)_6$	P	ACACTCTGGATTTCACAT GTTTACTCACAGCACTCA	182	50	1.5	1	- -	-	HQ185595
Hane-130	$(CT)_6$	P	ATTCCTACCCATACACCC GCTTCCCATCATAGCATA	120	50	1.5	1	- -	-	HQ185596
Hane-131	$(AG)_{10}$	P	TCCGTCTGGTCCCATTAG ATTCTCCCGTTTCCTTCC	214	50	1.5	1	- -	-	HQ185597
Hane-154	$(GT)_{14}$	P	GACCCACCCAACAGCAAC CTGGCAAGTCATAAGGATTAGA	238-242	50	0.9	3	0.8421 0.5914	0.5236	HQ185598
Hane-175	$(GA)_{11}$	P	AGCCGTACAGTCAAACAC GGCAGTTCTAATGGGATA	160-166	50	1.5	4	0.3500 0.6175	0.5538	HQ185599
Hane-181	$(AGG)_7$	P	CCCTTCAAGTCAAGCAGA ATGGGTGGTTTGGATGTA	227	51	1.5	1	- -	-	HQ185600
Hane-188	$(TC)_4TT(TC)_3$	I	GCAGGCAGAGCCAGAAAA ACGAATGAGGCGGGGAGG	181	50	1.5	1	- -	-	HQ185601
Hane-192	$(TC)_{11}$	P	CTCGCTCTGGCTCTAATC AAGCAAGTGTCCCAACAA	123	50	1.5	1	- -	-	

P, I and C = perfect, imperfect and compound, respectively; *Tm* = annealing temperature; N_A = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content.

REFERENCES

Brandström M and Ellegren H (2008). Genome-wide analysis of microsatellite polymorphism in chicken circumventing the ascertainment bias. *Genome Res.* 18: 881-887.

Brown SM, Hopkins MS, Mitchell SE, Senior ML, et al. (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* 93: 190-198.

Chen X, Temnykh S, Xu Y, Cho YG, et al. (1997). Development of a microsatellite framework map providing genome wide coverage in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 95: 553-567.

Degnan PH and Arévalo E (2004). Isolation of microsatellite loci in *Sceloporus grammicus* (Squamata, Phrynosomatidae). *Am. J. Undergrad. Res.* 2: 113-120.

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

Li CD, Rossnagel BG and Scoles GJ (2000). The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. *Theor. Appl. Genet.* 101: 1259-1268.

Liu JY, Lun ZR, Zhang JB and Yang TB (2009). Population genetic structure of striped mullet, *Mugil cephalus*, along the coast of China, inferred by AFLP fingerprinting. *Bioch. Syst. Ecol.* 37: 266-274.

Mia MY, Taggart JB, Gilmour AE, Gheyas AA, et al. (2005). Detection of hybridization between Chinese carp species (*Hypophthalmichthys molitrix* and *Aristichthys nobilis*) in hatchery broodstock in Bangladesh, using DNA microsatellite loci. *Aquaculture* 247: 267-273.

Rice WE (1989). Analyzing tables of statistical tests. *Evolution* 43: 223-225.

Schneider S, Roessli D and Excoffier L (2000). ARLEQUIN: A Software for Population Genetics Data Analysis, Version 2.000. Genetics and Biometry Laboratory, Department of Anthropology. University of Geneva, Switzerland.

Sekino M and Hara M (2001). Application of microsatellite markers to population genetics studies of Japanese flounder *Paralichthys olivaceus*. *Mar. Biotechnol.* 3: 572-589.

Selkoe KA and Toonen RJ (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9: 615-629.

Senior ML, Murphy JP, Goodman MM and Stuber CW (1998). Utility of SSRs for determining genetic similarities and relationships in maize using an agarose system. *Crop Sci.* 38: 1088-1098.

Weber JL (1990). Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. *Genomics* 7: 524-530.

Weising K, Winter P, Huttel B and Kahl G (1998). Microsatellite markers for molecular breeding. *J. Crop Prod.* 1: 113-143.

Xu TJ, Shao CW, Liao XL, Ji XS, et al. (2009). Isolation and characterization of polymorphic microsatellite DNA markers in the rock bream (*Oplegnathus fasciatus*). *Conserv. Genet.* 10: 527-529.

Yeh FC and Boyle TJB (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J. Bot.* 129: 157.

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