Conservation Genet Resour
DOI 10.1007/s12686-010-9365-4

TECHNICAL NOTE

Rapid development of microsatellites for the endangered Neotropical catfish *Conorhynchus conirostris* using a modest amount of 454 shot-gun pyrosequencing

Daniel C. Carvalho · Luciano B. Beheregaray

Received: 24 November 2010 / Accepted: 25 November 2010

Abstract A set of 13 polymorphic microsatellites were developed using a next generation sequencing approach for the endangered Neotropical catfish *Conorhynchus conirostris*. Using only 1/16 of a shot-gun pyrosequencing reaction, we were able to obtain 3,796 sequences containing putative microsatellites motifs. Out of 20 selected loci for further optimization, 13 loci were successfully genotyped in 25 specimens from a locality in the São Francisco River, Southeast Brazil. All microsatellite loci were polymorphic, with 4–18 alleles per locus (mean = 10.5) and generally high values of heterozygosity (mean = 0.796). These polymorphic markers will be a valuable tool for captive breeding and stoking programs and for analyses of population connectivity and genetic structure in this endangered Neotropical catfish.

Keywords Neotropical · Catfish · Population genetics · Conservation · Endangered species

The rich Neotropical fish fauna has been greatly threatened by habitat fragmentation, pollution, over-exploitation (Agostinho et al. 2005) and introduction of invasive species (Carvalho et al. 2009). In this scenario, molecular ecological studies are of great value since they can be used to estimate population connectivity in the wild and to assist conservation breeding programs aimed at restocking depleted populations. *Cornohynchus conirostris* is an endemic migratory catfish from the São Francisco River (SFR) listed as an endangered species (Lins et al. 1997). This species was once an important commercial species, reaching up to 100 cm in total length and 13 kg in body mass (Sato 1999). However, most of its populations are now considered as commercially extinct and the species is essentially restricted to the middle section of the SFR. Therefore, breeding and supportive stoking programs are required to supplement *C. conirostris* populations. Such programs would greatly benefit by the development of species-specific microsatellite markers.

For this study, a total of 10 μg of genomic DNA extracted from fin clips of one specimen of *C. conirostris* was sent to the Australian Genome Research Facility (www.agrf.com.au). This samples was subjected to high throughput DNA sequencing on 1/16 of a 70 9 75 PicoTiterPlate using the Roche GS FLX (454) system as described elsewhere (Margulies et al. 2005). Using the command "Design Primers" of MSATCOMMANDER 0.8.2 (Faircloth 2008), we screened contigs for microsatellites using the default settings of a total of 35,724 reads, resulting in a total of 3,796 sequences containing putative microsatellites motifs. MICROFAMILY (Meglecz 2007) was used to group contigs' flanking regions into families, according to its similarity in all-against-all BLAST (Basic Local Alignment Search Tool) search. Contigs classified as unique (i.e. no similarities to any other sequences of the same dataset) or UnBLASTable (i.e. no similarity to any sequences in GenBank database) were screened by eye for loci with at least 8 repeats long and with enough flanking region for primer development. In some cases, primers were manually designed to a specific size range using PRIMER 3 (Rozen and Skaletsky 1997) because of multiplexing genotyping reasons. A M13 universal sequence (5′-TGT AAA ACG ACG GCC AGT) was appended to the 5′ end of each forward primer to facilitate subsequent fluorescent labeling.

D. C. Carvalho (✉) · L. B. Beheregaray
Molecular Ecology Laboratory, School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia
e-mail: daniel.carvalho@flinders.edu.au;
carvalhodcc@yahoo.com.br

Published online: 14 December 2010

In total, MICROFAMILY recovered 673 unique loci, from which the best 20 loci (17 dinucleotides, 1 tri and 2 tetras) were chosen for polymerase chain reaction (PCR) trials. Amplification followed the method described by Schuelke (2000). Reactions were performed in a final volume of 10 µl containing 1× Mango *Taq* Reaction Buffer (Bioline), 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.1 U Mango *Taq* DNA polymerase (Bioline), BSA (0.1%), 0.05 µm forward primer, 0.2 µm reverse primer and 0.2 µm fluorescent M13 primer. The PCR program followed Beheregaray and Sunnucks (2000). After the PCR optimization procedure with the initial panel of 20 loci, 13 microsatellite sets amplified consistently and were selected for further analyzes. A total of 25 individuals of *C. conirostris* sampled from one site at the São Francisco River (S18°8′9.76″ W45°14′41.78″) were used for characterization of the microsatellites.

Amplification products were detected on an ABI 3130 Sequencer (Applied Biosystems) at the SouthPath and Flinders Sequencing Facility. PCR products were combined in two separated runs, namely Plex1 and Plex2. Plex 1 consisted of PCR products from loci Con3, Con5, Con7, Con11, Con14, Con23 and Con33, whereas Plex2 consisted of loci Con13, Con20, Con26, Con29, Con39 and Con41. The resulting microsatellite profiles were examined using GENEMAPPER 4.0 (Applied Biosystems) and peaks were scored manually. GENEPOP v3.3 (Raymond and Rousset 1995) was used to estimate expected (H_E) and observed (H_O) heterozygosity, number of alleles (N_A), linkage disequilibrium and Hardy–Weinberg proportions. Bonferroni corrections were applied when conducting multiple statistical tests (Rice 1989). The program MICROCHECKER (Van Oosterhout et al. 2004) was used to check for null alleles and scoring errors.

Table 1 Primers sequences and characteristics of 13 microsatellite loci isolated from *C. conirostris*

Locus	*Primer sequences (5′–3′)	Repeat structure	N_A	Size range (bp)‡	H_O/H_E	H–W	GenBank Accesion no.
Con3	F CGCCGAGGAGATCGTAGAG	GT(11)	6	224–236	0.800/0.715	0.9445	HQ625029
	R GAGGCTGGGGAGGGTATTG						
Con5	F GGTGGAGCAGCATCTTTTG	AC(14)	14	183–215	0.880/0.892	0.6682	HQ625030
	R CGTCCGTCTTAACCTCTCCC						
Con7	F GCCTGTAGAGCTGCTGGG	AC(22)	16	76–106	0.833/0.910	0.0893	HQ625031
	R TGCAGCGTGATCTGATTGG						
Con11	F ATGCTCCTAACACCCCCTCT	AC(9)	4	114–120	0.600/0.614	0.4142	HQ625032
	R CGTAAAGACGCACAGACGAA						
Con13	F TGTAGGCAACAAAGAAAGGC	CTT(9)	8	208–229	0.760/0.804	0.4459	HQ625033
	R ATGTGTGAGGAGGGCTGTG						
Con14	F AGACAACAGGTGCTCCCTC	ATGG(8)	7	199–223	0.760/0.794	0.2746	HQ625034
	R AGCGAGCATGCATAAGCTAC						
Con20	F CAAAGTCGGAGGTTTTGGGATG	GA(16)	15	154–190	0.880/0.894	0.2234	HQ625035
	R AGTGAAAACCGACAAGGTTGC						
Con23	F AGAAGGACAACAGGTGAAAGG	GA(14)	12	156–186	1.000/0.850	0.4692	HQ625036
	R TGCTGGGTCACAGAACTCC						
Con26	F TGGGCCTTTCACGAGTAGG	AC(14)	8	255–274	0.720/0.727	0.5709	HQ625037
	R TGCGACCTGAAAGCATCTC						
Con29	F ACAGCCTATGAGGTGAAAGC	CA(13)	18	179–235	0.920/0.896	0.9122	HQ625038
	R TACGAGCACTCAGACAGCC						
Con33	F TGACATGTATTTAGTCAGGGGC	CA(16)	15	238–278	0.920/0.894	0.4308	HQ625039
	R TGTGTTTGCGCTTTTGTGTTC						
Con39	F GTTGAGCTGCTCCTCCAGAC	AG(12)	8	106–120	1.000/0.827	0.3607	HQ625040
	R AGAGTGGAATCCCAGCAGTG						
Con41	F TGCACAACACAGCAGTAAAACA	TC(8)	6	108–118	0.280/0.256	1.0000	HQ625041
	R GCAGAATGCGTCAGGTTACA						

Number of alleles (N_A), range of allelic size, observed (H_O) and expected (H_E) heterozygosity and Hardy–Weinberg *P* values (*H–W*) are based on 25 individuals

* Forward primers were tagged with a 5′M13 universal sequence (5′TGTAAAACGACGGCCAGT-3′)

‡ Size range not excluding 5′M13 universal sequence

All loci were polymorphic with an average of 10.5 alleles per locus (between 4 and 18 alleles per locus) and no significant deviation from Hardy–Weinberg equilibrium or linkage disequilibrium locus-pair/population was observed. Moreover, MICROCHECKER did not detect evidence of null alleles or scoring errors. In order to determine reliability of genotyping results, PCRs were repeated for 35% of samples. No scoring errors were detected. Although *C. conirostris* is considered an endangered species, no heterozygosity deficiency was observed in our sample (Table 1). Analyses of genetic connectivity and population structure using the markers described here are strongly recommended for this species to guide conservation breeding and restocking programs of this endangered Neotropical catfish.

Acknowledgments We would like to thank José Vanderval (IEF) and Francisco Andrade (CEMIG) for helping with the fieldwork, Mike Gardner for assistance with data analysis and Minami Sasaki and Mariana Bertelli for laboratory assistance. DC is grateful to CAPES, an institution from the Brazilian Ministry of Education, for a postdoctoral fellowship (# 4095-09-0). This work was supported by funds from Flinders University to LBB.

References

Agostinho AA, Thomaz SM, Gomes LC (2005) Conservation of the biodiversity of brazil's inland waters. Conserv Biol 19(3): 646–652. doi:10.1111/j.1523-1739.2005.00701.x

Beheregaray LB, Sunnucks P (2000) Microsatellite loci isolated from odontesthes argentinensis and the o-perugiae species group and their use in other south american silverside fish. Mol Ecol 9(5):629–631. doi:10.1046/j.1365-294x.2000.00882.x

Carvalho DC, Oliveira DAA, Santos JE, Teske P, Beheregaray LB, Schneider H, Sampaio I (2009) Genetic characterization of native and introduced populations of the neotropical cichlid genus cichla in Brazil. Genet Mol Biol 32(3):601–607. doi: 10.1590/S1415-47572009005000060

Faircloth BC (2008) Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8:92–94

Lins LV, Machado ABM, Costa CMR, Herrmann G (1997) A guidebook for elaboration for endangered species lists: with the official list for the fauna of minas gerais. Belo Horizonte, fundação biodiversitas (in Portuguese)

Margulies M, Egholm M, Altman WE, Attiya S et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437(7057):376–380. doi:10.1038/nature03959

Meglecz E (2007) Microfamily (version 1): a computer program for detecting flanking-region similarities among different microsatellite loci. Mol Ecol Notes 7(1):18–20. doi:10.1111/j.1471-8286.2006.01537.x

Raymond M, Rousset F (1995) Genepop (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86(3): 248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43(1): 223–225

Rozen S, Skaletsky H (2000) Primer3 on the www for general users and for biologist programmers. Methods Mol Biol 132:365–386

Sato Y (1999) Reproduction of São Francisco river basin fishes: induction and characterization of patterns. PhD thesis, Federal University of São Carlos, SP, Brazil (in Portuguese)

Schuelke M (2000) An economic method for the fluorescent labeling of pcr fragments. Nat Biotechnol 18(2):233–234

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(3):535–538. doi:10.1111/J.1471-8286.2004.00684.X