Molecular Ecology Notes (2005) **5**, 84–86 doi: 10.1111/j.1471-8286.2004.00840.x

PRIMER NOTE

Microsatellite DNA markers for population genetic studies and parentage assignment in cobia, *Rachycentron canadum*

CHRISTIN L. PRUETT,* ERIC SAILLANT,* MARK A. RENSHAW,* JOHN C. PATTON,* CAIRD E. REXROAD III† and JOHN R. GOLD*

**Center for Biosystematics and Biodiversity, Department of Wildlife and Fisheries Sciences, Texas A and M University, College Station, TX 77843–2258, USA, †USDA/ARS National Center for Cold and Cool Water Aquaculture, 11861 Leetown Road, Kearneysville, West Virginia 25430, USA*

Abstract

Twenty nuclear-encoded microsatellites from a genomic DNA library of cobia, *Rachycentron canadum*, were isolated and characterized. The microsatellites include two tetranucleotide, one trinucleotide, three combination tetranucleotide/dinucleotide, nine dinucleotide, and five imperfect (dinucleotide) repeat motifs. Gene diversity ranged between zero to 0.910; the number of alleles among a sample of 24 fish ranged from one to 15. Cobia support an important recreational fishery in the southeastern United States and recently have become of interest to aquaculture. The microsatellites developed will be useful tools for studying both population genetics (e.g. stock structure, effective population size) and inheritance of traits important to aquaculture.

Keywords: cobia, genomic library, microsatellites, polymerase chain reaction primers, *Rachycentron canadum*

Received 01 September 2004; revision accepted 06 October 2004

Cobia, *Rachycentron canadum*, is an economically important, pelagic fish distributed in tropical warm waters worldwide (Shaffer & Nakamura 1989). It is a highly prized food and recreational trophy fish, and is considered a prime candidate for aquaculture (Benetti *et al*. 2003). Because of the popularity of cobia as a 'game' fish, methods to identify or distinguish products harvested in cobia aquaculture from 'wild' stocks will be needed in order to ensure legal sale and alleviate potential conflicts. Nuclear-encoded microsatellites (Weber & May 1989) are especially well suited for this purpose because of their codominant, Mendelian inheritance and their high levels of polymorphism. Microsatellites also have many applications in breeding programs (Garcia de Leon *et al*. 1998) and for assessing population structure of 'wild' populations as a means to improve assessment and allocation of resources. In this note, we report the development from a genomic library of cobia DNA of polymerase chain-reaction (PCR) primers for 20, nuclear-encoded microsatellites.

Correspondence: J. R. Gold, Fax: + 1979 8454096; E-mail: goldfish@tamu.edu

Whole genomic DNA was extracted from cobia muscle tissue using a standard phenol-chloroform method and digested with *Dpn*II (New England BioLabs). Fragments ranging in size from 500 to 2000 bp were size selected by extraction from a 1% agarose gel and purified using a gel extraction kit (Qiagen). Fragments were ligated into a *Bam*HI (New England BioLabs) digested and dephosphorylated (Calf Intestinal Alkaline Phosphatase, New England BioLabs) pBluescript vector using T4 DNA ligase (New England BioLabs) and transformed into XL10-Gold ultracompetent cells (Stratagene). Transformed cells were plated on X-Gal – IPTG Luria-Bertani (LB) agar with 50 μg/mL of ampicillin and grown overnight at 37 °C. Recombinant colonies were picked using a GENETIX QBOT, inoculated into 384 well plates that contained 50 μL of LB freezing media [36 mm K_2HPO_4, 13.2 mm KH_2PO_4, 1.7 mm sodium citrate, 0.4 mm $MgSO_4$, 6.8 mm $(NH_4)_2SO_4$, 4.4% (v/v) glycerol, 50 μg/mL ampicillin, LB], and incubated overnight at 37 °C before freezing at –80 °C.

A total of 19 200 clones (50 × 384-well plates) were spotted in a 4 × 4 array onto 22.5 cm × 22.5 cm Hybond nylon membranes (Amersham), with each clone being spotted twice to eliminate false positives. Membranes were placed

on LB agar plates with 50 μg/mL of ampicillin and incubated at 37 °C until colonies were approximately 1–2 mm (18–24 h). Membranes were placed on chromatography paper (3M) and soaked as follows to fix colonies: 10% SDS for 3 min, denaturing solution (1.5 mm NaCl, 1.5 Tris) for 5 min, and 2 × SSC for 5 min. Filters were incubated for 5 h at 65 °C.

Resulting colonies were probed with two cocktails of [gamma] ^{32}P-labelled oligonucleotides: (i) tetranucleotides $[GATA]_9$, $[CATA]_8$, $[GACA]_8$ and trinucleotide $[CAA]_8$, and (ii) trinucleotides $[GAA]_8$, $[TAA]_{13}$, and dinucleotides $[GA]_{13}$ and $[CA]_{13}$. A total of 164 positive clones were screened as follows. Frozen glycerol stocks arrayed in 96-well plates were used to inoculate 1 mL cultures of Luria Broth selective media (ampicillin) and incubated overnight at 37 °C. Plasmid DNA was isolated (alkaline lysis) with a BioRobot 8000 (QIAGEN, CA). Miniprep DNA was quantified, normalized, and both strands sequenced, using M13 forward and reverse sequencing primers and ABI BigDye Terminator v3.1. Products were purified and electrophoresed on an ABI 3100 DNA Analyser (Applied Biosystems, CA). Phred (CodonCode, MA) was used for DNA sequence base calling and vector trimming. A total of 54 complete sequences containing microsatellite arrays were obtained from the positive clones. Three of the clones contained two repeat sequences from which primer pairs were designed. A total of 45 primer pairs were developed using AMPLIFY 1.2 (Engels 1993) and Netprimer® (http://www.premierbiosoft.com/netprimer).

Unlabelled PCR primers were purchased from Invitrogen (Carlsbad, CA) and tested for amplification by screening DNA isolated from eight individuals obtained from offshore of Ocean Springs, Mississippi. PCR amplifications were performed in 10 μL reaction volumes containing 1 μL (100 ng) DNA, 1 μL 10× reaction buffer (500 mm KCl, 100 mm Tris, 10% Triton-X 100), 0.1 U of *Taq* DNA polymerase (Gibco-BRL), 0.5 μm of each primer, 200 μm of each dNTP, and 1 mm $MgCl_2$. PCR conditions consisted of an initial denaturation of 94 °C for 3 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 45 °C–65 °C for 45 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min. PCR products were electrophoresed in a 2% agarose gel, stained with ethidium bromide. Once appropriate annealing temperatures for each primer were determined, microsatellite arrays were tested for polymorphisms by end-labelling (using T4 polynucleotide kinase) one primer from each pair with [γ^{32}P]-dATP and employing the same conditions as above. PCR products were electrophoresed in 6% polyacrylamide gels and visualized by autoradiography; 20 microsatellite repeats were chosen for further screening (Table 1). These microsatellites included one trinucleotide, two tetranucleotide, three combination tetranucleotide/dinucleotide, nine dinucleotide, and five imperfect dinucleotide repeats. Two of the microsatellites were on the same clone, *Rca*1B-E08A and *Rca*1B-E08B. Lengths of cloned alleles ranged in size from 122 to 308 base pairs. Optimal annealing temperatures ranged from 48 °C–60 °C.

Screening involved a total of 24 cobia obtained from offshore of Ocean Springs, Mississippi. One primer from each pair was fluorescently labelled with one fluorescent label of set D (Applied Biosystems) for run on an automated sequencer ABI-377. Alleles were sized using the GENESCAN® 400 HD Size Standard (Applied Biosystems); allele sizing and calling were performed using GENESCAN® 3.1.2 and GENOTYPER®, version 2.5 software. Genetic variability of the markers was measured by the number of alleles, gene diversity (expected heterozygosity), and the observed heterozygosity. Wright's F_{IS}, estimated as Weir and Cockerham's f in Genetic Data Analysis (GDA) (Lewis & Zaykin 2001), was used to measure departure of genotype proportions from Hardy–Weinberg expectations at each microsatellite. Fisher exact tests, as performed in GDA, were used to test significance of departures from Hardy–Weinberg equilibrium (genotype) expectations at each microsatellite and for departure from genotypic equilibrium at pairs of microsatellites. The effect of Hardy–Weinberg departures (within locus disequilibrium) on significance of between locus linkage disequilibrium tests was removed by preserving genotypes in GDA (Lewis & Zaykin 2001).

Summary data are presented in Table 1. The number of alleles detected per microsatellite ranged from one to 15. Expected heterozygosity ranged from zero to 0.910, while observed heterozygosity ranged from zero to 0.957. Genotypes at four microsatellites deviated significantly from Hardy–Weinberg expectations following sequential Bonferroni correction (Rice 1989). Three of the microsatellites exhibited heterozygote deficiency (*Rca* 1B-F06, $F_{IS} = 0.324$; *Rca* 1-B12, $F_{IS} = 0.578$; and *Rca* 1-E06, $F_{IS} = 0.468$), while one (*Rca* A-10) exhibited heterozygote excess ($F_{IS} = -0.079$). All pairwise comparisons of microsatellites did not deviate significantly from genotypic equilibrium following Bonferroni corrections (Rice 1989). The 20 microsatellites developed in this work will prove useful in future studies of population genetics and quantitative genetics of 'wild' and domesticated cobia, respectively.

Acknowledgements

We thank J. Franks of the Gulf Coast Research Laboratory and J. Graves of the Virginia Institute of Marine Science for providing the cobia tissues, and C. Bradfield for technical assistance in the laboratory. Funding was provided by the Saltonstall-Kennedy Program of U.S. Department of Commerce (Grant NA17FD2371) and by the Texas Agricultural Experiment Station (Project H-6703). This paper is number 42 in the series 'Genetic Studies in Marine Fishes' and Contribution no. 127 of the Center for Biosystematics and Biodiversity at Texas A & M University.

Table 1 Summary data for microsatellites developed from cobia, *Rachycentron canadum*

Microsatellite	Primer sequence (5′–3′)*	Repeat sequence†	T_A	N‡/N_A§	Size Range¶	H_E/H_O**	P_{HW}††
Rca 1B-A10	**GCAGCCCAATGCTAACAAGCC**¶¶ CATGTAGTCAAGCGAGCCACG	$(GTT)_6$	60	24/4	174–186	0.735/0.792	**0.003**
Rca 1B-D09	**CAGCCTGCTTAGCCTATCA** GAAGGATGGACCACTTGTGAC	$(GT)_9(CTGT)_2(CT)_2(GT)_2$	60	23/1	169	0.000/0.000	1.000
Rca 1B-E02	**GTGTTGCAGCCAAATGCTA** CTCCCTAGTGCCACTACAGCTC	$(CT)_{18}$	60	24/7	298–314	0.598/0.667	0.483
Rca 1B-E08A§§	**CATATCAAGTCAATATCACAGACC** CCACGGAATAGCAGACTTTCTC	$(CA)_3GA(CA)_5A(CA)_{16}$	55	24/5	181–225	0.582/0.458	0.030
Rca 1B-E08B	**GCAGTTGATTCTGATTGCTACAC** CTAATGCCAGCTCATTATGTCC	$CA)_8GA(CA)_3$	60	24/2	120–122	0.496/0.417	0.672
Rca 1B-F06	**CAAGCAAATGCGTGGCCGA** CGTTAGCAACCACACGAGCTTG	$(CTAT)_{15}$	55	24/11	260–300	0.796/0.542	**0.000**
Rca 1B-F07	**GGAATCTGGTGGTGAGTCAT** CTGTGGCTGAAGCGTGTGTT	$(GACA)_6(CA)_{12}$	55	24/3	132–140	0.082/0.083	1.000
Rca 1B-G10	**GGAAACTCTATAACAGCATGTC** GTAGACAGAGCAACACATGAG	$(CT)_5TT(CT)_4$	55	23/2	154–156	0.043/0.043	1.000
Rca 1B-H09	**CATGTTATTCTCCAACTCATGG** GTGTATCCGCATACTTTCAG	$(GATA)_{31}$	48	23/12	176–224	0.910/0.957	0.351
Rca 1-A04	**CACGCACATGCACTACTTTAACC** GCTGTTGATGTGGCGAAGCAAC	$(CA)_9(CACT)_4$	60	24/6	196–206	0.722/0.625	0.095
Rca 1-A11	**CTACAGTGGTGTTCCCTGTTAG** CAGTACATAGAGAAACAGGAGG	$(GT)_{24}$	55	24/15	167–201	0.889/0.792	0.271
Rca 1-B12	**GCTTCAGGCAAGTGAGACC** GGGAGGTAATTATGTCCTGT	$(AC)_9$	55	24/7	177–193	0.780/0.333	**0.000**
Rca 1-C04	**GACATCAAGTGGCACTTTG** CACTAAACTTGTTCCTCCTG	$(GT)_{17}$	48	24/10	223–253	0.641/0.667	0.188
Rca 1-D04	**GCTGAACTTGTCGCCGCT** GGACTGAACCTCCCTATCCTC	$(TG)_9AC(TG)_5$	60	24/3	125–129	0.551/0.667	0.723
Rca 1-D11	**CGTAACACCTTTGGAAGACATC** CTCCATTGAGGCTGACTAGTG	$(GT)_8$	55	24/4	204–212	0.295/0.333	1.000
Rca 1-E04	**CCAAGAACAGGCGGGCAAC** GCCACCATTGTGTGTGGGTGA	$(CA)_8$8 bp$(CA)_5$	55	23/4	216–238	0.336/0.391	1.000
Rca 1-E06	**GGCACCAATCACTCACTACTG** TGTTGAGGTCTATCAGTGCC	$(CA)_{39}$	48	24/10	144–188	0.853/0.458	**0.000**
Rca 1-E11	**GTCCCAGCTCCAGCCCAAAC** GACACTGGCTGCGTGAGCA	$(CA)_{12}$	55	23/7	167–181	0.757/0.783	0.245
Rca 1-F01	**GCTCATTTCACTAAGTGTGTTGTAGC** CCATGAATCTACATTCACCTGCCA	$(TG)_{12}$	60	24/2	202–206	0.120/0.125	1.000
Rca 1-G05	**GGGCTGTCTGCTGGCTGTAA** GCATCTGTGTCCTGGTGAGAGTCCC	$(GT)_{17}$	60	24/7	274–282	0.697/0.667	0.148

*Primer sequences are forward (top) and reverse (bottom); †Repeat sequence indicates the repeat motif; T_A is annealing temperature in °C; ‡N is the number of individuals assayed; §NA is number of alleles detected; ¶Size range refers to alleles thus far uncovered; **H_E and H_O are expected and observed heterozygosity, respectively; ††PHW represents the probability of deviation from Hardy–Weinberg expectations (significant values after Bonferroni correction are in bold) and sequences of clones are listed in GenBank (Accession numbers AY721664–AY721682); §§Primers Rca 1B-E08A and Rca 1B-E08B were developed from the same clone; ¶¶The fluorescently labelled primer is in bold.

References

Benetti DD, Alarcon JF, Stevens OM *et al.* (2003) Advances in hatchery and growout technology of marine finfish candidate species for offshore aquaculture in the Caribbean. *Proceedings of the Gulf and Caribbean Fisheries Institute*, **54**, 473–487.

Engels WR (1993) Contributing software to the Internet: the AMPLIFY program. *Trends in Biochemical Sciences*, **18**, 448–450.

Garcia de Leon FJ, Canonne M, Quillet E, Bonhomme F, Chatain B (1998) The application of microsatellite markers to breeding programmes in sea bass, *Dicentrarchus labrax*. *Aquaculture*, **159**, 303–306.

Lewis PO, Zaykin D (2001) GENETIC DATA ANALYSIS: computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.htm.

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

Shaffer RV, Nakamura EL (1989) Synopsis of the biological data on the cobia, *Rachycentron canadum* (Pisces: Rachycentridae). US Department of Commerce, NOAA Technical Report NMFS 82.

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388–396.