

PRIMER NOTE

New tetranucleotide microsatellites for fine-scale discrimination among endangered chinook salmon (*Oncorhynchus tshawytscha*)

CAROLYN GREIG,* DAVID P. JACOBSON† and MICHAEL A. BANKS†

*Wellcome Trust Center for Cell Matrix Research, University of Manchester, Oxford Road, Manchester M13 2PT, UK, †Coastal Oregon Maine Experiment Station, Oregon State University Hatfield Marine Science Center, 2030 SE Marine Science Drive, Newport, OR 97365, USA

Abstract

The unambiguous identification of Central Valley spring-run chinook salmon has become imperative since their proposed listing in 1998. The accuracy of methods used to assign individuals to their stock of origin is critical for understanding juvenile migration patterns and determining the success of protection measures. Existing microsatellites discriminate between the endangered winter-run and other chinook but are insufficient to characterize phylogenetically less distinct runs. Here, we isolated and developed highly variable tetranucleotide microsatellites for the specific goal of increasing discriminatory power among closely related populations, providing a new power towards the reliable differentiation of nonwinter runs

Keywords: chinook, microsatellite, *Oncorhynchus tshawytscha*, polymorphic marker, population discrimination, tetranucleotide

Received 11 February 2003; revision accepted 24 March 2003

Previously developed microsatellites enable the identification of winter-run chinook salmon (Banks *et al.* 1999) but cannot completely classify fish from other runs. Tetranucleotide microsatellites, first used in our laboratory, had a higher discriminatory power than most dinucleotides. We created a library enriched for four tetranucleotide repeats (Table 1).

The method was modified from that of Kandpal *et al.* (1994): 10 µg of good quality DNA from four individuals were digested with MboI, ligated to Sau L linkers, size selected on an agarose gel for 350–800-bp fragments, hybridized to 3.5 µM of each biotinylated tetranucleotide probe, captured with 50 mg of a biotin-binding matrix (Vectrex-avidin D; Vector Laboratories) and washed twice at 50 °C in 150 mM NaCl/100 mM Tris, pH 7.5. Fragments were eluted and concentrated before polymerase chain reaction with the SauLA primer to generate ds-DNA and linker removal by MboI digestion. The DNA was ligated into BamHI-pUC18 vector (Pharmacia), transformed into

XL-2-Blue-ultracompetent cells (Stratagene) and plated on 20 × 20-cm agar plates. Colony lifts were hybridized with the biotinylated probes and detected on X-ray film using a streptavidin–alkaline phosphatase conjugate and luminescent CSPD® (BRL). Positive colonies were rechecked, occurring at 4.7%.

DNA from 100 chinook clones was prepared using a Qiaprep spin-miniprep kit (Qiagen) and size-screened by polymerase chain reaction with M13 primers (Table 1). Primers were designed using Oligo_5.0 (MBI), standardizing the T_m and G/C range for subsequent multiplexing. The polymerase chain reactions were optimized for each primer (Table 2). An additive was developed which enhanced the relative yield of larger alleles, preventing their occasional dropout.

Utility for run discrimination was assessed using a panel of 64 fish, 16 from each of the four runs. Products were labelled with TAMRA-dNTPs and loci were typed on 8% denaturing polyacrylamide gels visualized using a FMBIO II scanner (Hitachi).

The power to detect between-run variability was initially assessed using population pairwise values for Nei's genetic

Correspondence: Carolyn Greig. Fax: + 44 161 275 5082; E-mail: carolyn.greig@man.ac.uk

Repeat unit	(GATA)*	(GAAA)*	(CATA)†	(CATC)†	Mixed	Total
Clones sequenced	12	8	27	27	8	82
Average no. of repeats	28.5	21	28.2	—	—	25.9
Range of repeat nos	6–64	7–38	12–46	—	—	6–64
Complex interrupted	0	2	0	25	8	35
Short flanking sequence	5	3	5	13	2	28
Primers developed	5	4	18	2	1	30
Rejected: PCR/scoring	2	0	6	0	0	8
Monomorphic or low diversity	2	2	6	2	1	13
Diagnostic‡	1	2	3	0	0	6

Table 1 Library development

*Frequently used vertebrate microsatellites (EMBL database).

†More rarely used vertebrate microsatellites (EMBL database).

‡Three polymorphic markers not developed for run diagnosis EMBL nos AJ534364, AJ534365 and AJ534366.

PCR, Polymerase chain reaction.

distance (Nei 1972), since heterozygosity-based distances (e.g. F_{ST}) may be less accurate with highly polymorphic loci. (Hedrick 1999). These were combined for nonwinter populations using

$$D_{\text{combined}} = -\ln(\sum(e^{-D})/n)$$

where n = no. distances and ranked alongside loci currently used (Banks *et al.* 1999). Loci were also ranked using the critical population routine described in the program WHICHLOCI (Banks *et al.* 2003). Repeated iterations for assignment (Banks & Eichert 2000) were performed for each locus, scoring by correct assignment to source populations. Equivalent results were subsequently obtained with frequencies from over 100 fish per population.

Six loci ranked above all but our two most diagnostic loci and were developed in three colour-multiplexed sets (Table 2). They all have alleles of moderate frequency with strong run biases so, despite the presence of low frequency alleles in some individuals, when combined they have strong run discriminatory powers.

G-tests (Sokal & Rohlf 1995) were performed on transmission results in three or more control families described previously (Banks *et al.* 1999). Normal Mendelian inheritance was affirmed for all six loci. Cross-species testing was carried out on five other *Oncorhynchus* species (Table 2).

Homology searches (<http://www.ncbi.nlm.nih.gov/BLAST>) revealed that *Ots-208* had been isolated simultaneously as G68 (Accession no. AF393187). Also, *Ots-204*, amplifying exclusively in chinook, is situated immediately proximal to a presumed inactive Tc1-like transposable element (Accession no. L41171). These have been previously associated with microsatellites (Goodier & Davidson 1998).

Tetranucleotide microsatellites have been reported to have lower mutation rates than dinucleotide repeats in several species (Kruglyak *et al.* 1998; Schug *et al.* 1998)

although hypermutable tetranucleotides have been found (e.g. Primmer & Ellegren 1998). Polymerase mutations have been shown to have differential effects on the mutation rate of di- and tetranucleotide repeats (Eckert & Yan 2000; Bayliss *et al.* 2002), providing a mechanism for species differences in relative mutation rates. Motif sequence bias may also affect mutation rate since tetranucleotides with a lower G/C content (as here) can be more variable (Eckert *et al.* 2002). Larger, more variable, tetranucleotides may also be more prevalent in fish, which have longer microsatellites (Brooker *et al.* 1994).

In this study, highly polymorphic tetranucleotides have been isolated and comprehensively characterized, proving to be valuable tools for fine-scale population discrimination.

Acknowledgements

We are grateful to V. K. Rashbrook for technical assistance. This study was supported by funding from CALFED and the California Department of Water Resources.

References

- Banks MA, Blouin MS, Baldwin BA *et al.* (1999) Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tshawytscha*). *Journal of Heredity*, **90**, 281–288.
- Banks MA, Eichert W (2000) WHICHRUN (version 3.2): A computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity*, **91**, 87–89.
- Banks MA, Eichert W, Olsen JB (2003) Which genetic loci have greater population assignment power? *Bioinformatics*, **in press**.
- Bayliss CD, van de Ven T, Moxon ER (2002) Mutations in poll but not mutSLH destabilize *Haemophilus influenzae* tetranucleotide repeats. *EMBO (European Molecular Biology Organization) Journal*, **21**, 1465–1476.
- Brooker AL, Cook D, Bentzen P, Wright JM, Doyle RW (1994) Organization of microsatellites differs between mammals and

Table 2 Locus details for multiplexed primer sets

	Set 1	Set 2	Set 3
5'-3' Primer sequences (F-5'-labelled)	204F: CTGGCTAAGGTGTATGTAACCTTC 204R: CAGTATGTCAAGAAAGGGTCGAG 209F: CCAAGTGACCTGCTGTGTAGTTAC 209R: TCTCAGTTGGACAGCTGTAACAGC	211F: TAGTTACTGCTTCCGTCGAATG 211R: GAGAGGTGGTAGGATTTGCAG 212F: TCTTCCCTGTTCTCGCTTC 212R: CCGATGAAGACGAGAAGAGAC	208F: GGATGAACCTGCAGCTTTGTTATG 208R: GGCAATCACATACTTCAACTCC 213F: CCCCTACTCATGTCTCTATTTGGTG 213R: AGCCAAGGCAATTTCTAAGTGAC
10 µL polymerase chain reaction mixes	Ots 209 and Ots 204 primers (0.5 mM each), 2 mM MgCl ₂ , 0.125 mM dNTPs (Promega), 0.01 U Taq (Promega), Taq buffer (supplied). Additive: 0.5% formamide, 0.05 mg/mL BSA (Sigma A7906), 0.1% Tween 100 (94 °C 3 min, 32 cycles of 94 °C 30 s, 63 °C 20 s, 72 °C 30 s then 72 °C 2 min	Ots 211 and Ots 212 primers (0.5 mM each), 2 mM MgCl ₂ , 0.125 mM dNTPs (Promega), 0.01 U Taq (Promega), Taq buffer (supplied). Additive: 0.5% formamide, 0.05 mg/mL BSA (Sigma A7906), 0.1% Tween 100 (94 °C 3 min, 32 cycles of 94 °C 30 s, 63 °C 20 s, 72 °C 30 s then 72 °C 2 min	Ots 208 and Ots 213 primers (0.5 mM each), 1.625 mM MgCl ₂ , 0.125 mM dNTPs (Promega), 0.01 U Taq (Promega), Taq buffer (supplied). Additive: 0.5% formamide, 0.05 mg/mL BSA (Sigma A7906), 0.1% Tween 100 (94 °C 3 min, 32 cycles of 94 °C 30 s, 57 °C 40 s, 72 °C 45 s then 72 °C 2 min
Cycling conditions (MJ Research 'Tetra' Thermocycler)			
Locus name	OTS-204	OTS-211	OTS-208
EMBL Accession no.	AJ534360	AJ534361	AJ534362
Repeat structure	GTAAT ₍₉₇₎ GTAAG-GTAT ₍₆₎	CATA ₍₁₂₎	CATA ₍₂₈₎
Size of cloned allele	190 bp	178 bp	202 bp
Allele size range	174-302 bp	196-292 bp	150-426 bp
Inferred repeat no.	33-65	17-41	7-76
Allele no. (5 pops, <i>n</i> = 675-842)	28	25	52
Average per population allele no. (<i>n</i> = 100-189)	21	19	36
Butte spring, <i>n</i> =	149	104	148
<i>H</i> exp., <i>H</i> obs.	0.908, 0.812	0.906, 0.885	0.948, 0.892
*Coho (<i>O. kisutch</i>), <i>n</i> =	5	5	14
Allele no.	0	1	11
Size range			174-242
<i>H</i> exp., <i>H</i> obs.			0.860, 0.714
*Steelhead (<i>O. mykiss</i>), <i>n</i> =	32	32	20
Allele no.	0	1	10
Size range			142-190
<i>H</i> exp., <i>H</i> obs.			0.804, 0.750
*Cut-throat (<i>O. clarkii</i>), <i>n</i> =	36	36	17
Allele no.	0	1	6
Size range			114-134
<i>H</i> exp., <i>H</i> obs.			0.792, 0.412†
*Chum (<i>O. keta</i>), <i>n</i> =	36	36	36
Allele no.	0	0	0
Size range			
<i>H</i> exp., <i>H</i> obs.			
*Pink (<i>O. gorbuscha</i>), <i>n</i> =	12	9	11
Allele no.	0	1	12
Size range			126-266
<i>H</i> exp., <i>H</i> obs.			0.901, 1.000

*Cross-species tests; amplification attempted at T_m for chinook then $T_m - 5$ °C.

†Possible null allele(s); > 50% samples produced no amplified product.

BSA, Bovine serum albumin.

- cold-water teleost fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 1959–1966.
- Eckert KA, Yan G (2000) Mutational analyses of dinucleotide and tetranucleotide microsatellites in *Escherichia coli*: Influence of sequence on expansion mutagenesis. *Nucleic Acids Research*, **28**, 2831–2838.
- Eckert KA, Yan G, Hile SE (2002) Mutation rate and specificity analysis of tetranucleotide microsatellite DNA alleles in somatic human cells. *Molecular Carcinogenesis*, **34**, 140–150.
- Goodier JL, Davidson WS (1998) Characterization of novel mini-satellite repeat loci in Atlantic salmon (*Salmo salar*) and their phylogenetic distribution. *Journal of Molecular Evolution*, **46**, 245–255.
- Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.
- Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 88–92.
- Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 10 774–10 778.
- Nei M (1972) Genetic distance between populations. *The American Naturalist*, **106**, 283–292.
- Primmer CR, Ellegren H (1998) Patterns of molecular evolution in avian microsatellites. *Molecular Biology and Evolution*, **15**, 997–1008.
- Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TFC, Aquadro CF (1998) The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **15**, 1751–1760.
- Sokal RR, Rohlf FJ (1995) *Biometry*. Freeman, San Francisco.