Molecular Ecology Notes (2005) **5**, 220–222 doi: 10.1111/j.1471-8286.2005.00882.x

PRIMER NOTE

Atlantic capelin (*Mallotus villosus*) tetranucleotide microsatellites

K. GORDOS,*† E. L. KENCHINGTON,*‡ L. C. HAMILTON,‡ B. S. NAKASHIMA§ and C. T. TAGGART*¶

Marine Gene Probe Laboratory, Dalhousie University, Halifax, NS, Canada B3H 4J1, †Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, BC, Canada, V9T 6N7, ‡Fisheries and Oceans Canada, Box 1006 Dartmouth, NS, Canada B2Y 4A2, §Fisheries and Oceans Canada, Box 5667, St John's, NF, Canada A1C 5X1, ¶Department of Oceanography, Dalhousie University, Halifax, NS, Canada B3H 4J1

Abstract

Twelve microsatellite loci developed for Atlantic capelin (*Mallotus villosus*) using magnetic bead hybridization enrichment for tetranucleotide microsatellites revealed five loci composed of single repeat elements and six composed of complex repeats. Forty-four beach-spawning females from three different northwestern Atlantic Newfoundland beach-spawning populations were screened at each locus. Loci were polymorphic (two to 59 alleles per locus) and all but two exhibited high heterozygosity (0.86–1). The loci are considered suitable for addressing questions related to fine-scale population structure, spawning fidelity and survivorship/kinship issues.

Keywords: Atlantic capelin, microsatellite, population genetics, tetranucleotide

Received 30 September 2004; revision received 26 October 2004; accepted 18 November 2004

Capelin (*Mallotus villosus* Müller) is a small, circumpolar, pelagic fish species ranging across the North Atlantic and Pacific oceans. Their ecological role as a primary forage species and their economic role in commercial fisheries led Carscadden & Vilhjalmsson (2002) to recommend genetic methods for clarifying stock structure. Such methods can also address fine-scale structure, spawning-site fidelity and kinship-related survivorship. Allozyme, isozyme, mitochondrial DNA (mtDNA) and dinucleotide microsatellite methods illustrate varying differentiation among capelin populations (Mork & Friis-Soerensen 1983; Dodson *et al*. 1991; Dushchenko 1993; Roed *et al*. 2003). Tetranucleotide markers accrue advantages over dinucleotide markers, particularly when less prone to polymerase chain reaction (PCR) induced stutter (O'Reilly *et al*. 2002). We report on 12 such markers.

Muscle tissues collected in 2002 from three capelin populations spawning in Trinity and Notre Dame bays, Newfoundland, were preserved in 95% ethanol. Genomic DNA was isolated using Qiagen DNeasy Tissue kit spin columns. PCR product purification used QIAquick PCR purification (QIAGEN). Genomic DNA from five beach-spawning females (age 3: years) was pooled for microsatellite isolation. Loci were subsequently tested on additional samples of 15, 15, and 14 beach-spawning females (age 3: years) from each of three populations, respectively.

Loci were isolated using a modified Hamilton *et al*. (1999) enrichment. Genomic DNA was initially digested with *Hinc*II restriction endonuclease (New England Biolabs) with mung bean nuclease treatment omitted. The post SNX-adaptor ligation PCR reaction (Hamilton *et al*. 1999) was limited to 20 cycles based on parallel tests (not shown) to avoid over-amplification (Zane *et al*. 2002). Streptavidin-coated beads (Dynal) were prepared by washing twice with 1× PBS (phosphate buffered saline) –0.1% bovine serum albumin and twice with 5× SSC (standard saline citrate); 5′ biotinylated oligonucleotide probes [$(GACA)_4$ $(GATA)_4$ and $(GTAT)_7$] were added, incubated at room temperature for 15 min, washed four times with 5× SSC and resuspended in 10× SSC. Denatured hybridization mixture was added to the bead preparation and incubated for 20 min at the probe annealing temperature {50 °C for [$(GACA)_4$ and $(GTAT)_7$], 40 °C for [$GATA]_4$}. Washes contained 0.5 ng/μL SNX forward linker and were performed at the probe annealing temperature as follows (4 each):

Correspondence: E. L. Kenchington, Fax: (902) 426 1862; E-mail: kenchingtone@mar.dfo-mpo.gc.ca

2× SSC, 0.1% SDS; 1× SSC; and 0.5× SSC. Microsatellite-enriched DNA was released in 60 µL TE (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0) incubating for 15 min at 96 °C. Recovered DNA was amplified as in Hamilton *et al*. (1999) using 30 cycles for $(GATA)_4$ and 25 cycles for $(GACA)_4$ and $(GTAT)_7$ enriched DNA. Optimum cycle number was determined by parallel PCR cycling (not shown). PCR products were purified, ligated into pDrive (QIAGEN) and transformed into Invitrogen's ONE Shot TOP10 competent cells.

Plasmid DNA from 192 random colonies was isolated using QIAprep 96 Turbo (QIAGEN). Cycle sequencing used universal M13 Forward and Reverse primers, DTCS Quick Start at 1/4 reaction volume and ethanol precipitation (Beckman Coulter). Sequencing (Beckman Coulter CEQ 8000) showed ~60% contained di- or tetranucleotide microsatellite repeats. Fifteen useable and unique repeat sequences were identified. The enrichment protocol was repeated using DNA from one individual and the screening of 96 plasmids accrued 10 additional unique sequences. Eighteen of the 25 sequences were suitable for development. Seven of these shared either the 3′ or the 5′ flanks.

PCR primers were designed using PRIMER 3 (Rozen & Skaletsky 2000) and GENE RUNNER (version 3, Hastings Software). PCR reactions, electrophoresis and visualizations were performed as in O'Reilly *et al*. (2002) using locus-annealing temperatures (T_a °C) in Table 1. Because of product sizes (Table 1), the *Mvi*1 and *Mvi*31 products were visualized using a BASESTATION (MJ Research) and an internal ladder (GeneFlo 625, Chimerx). Analyses used CARTOGRAPHER (MJ Research).

Thirteen loci showed nonspecific amplification or monomorphism. Two loci shared a common flanking area (below). Observed and expected heterozygosities, conformity to Hardy–Weinberg expectation and genotypic disequilibrium among the three populations were assessed using GENEPOP version 3.3 (Raymond & Rousset 1995), ARLEQUIN version 2.0 (Schneider *et al*. 2000) and Guo & Thompson (1992) exact test. When locus-specific *n*-alleles in a collection was $< n = 5$, complete enumeration was used (Louis & Dempster 1987). When $> n = 4$ alleles, a Markov chain random walk algorithm with 800 batch-runs (1000 iterations batch^{-1}) estimated the exact *P* (Guo & Thompson 1992).

All loci produced distinguishable alleles at optimum annealing temperatures with ≤ 2 alleles per individual. Many loci had alleles differing by < 4 bp (Table 2). Use of internal ladder (*Mvi*1, *Mvi*31) and control samples (all loci) allowed alleles to be reproducibly sized. Loci ranged from two to 59 alleles per locus (mean = 28; median = 33). All

Table 1 Summary of 12 capelin tetranucleotide microsatellite loci listing locus identifiers, the associated repeat array and primer sequences (*F*orward and *R*everse), National Centre for Biotechnology Information, GenBank Accession no. and annealing temperature (T_a)

Locus	Repeat array sequence	Primer sequence (5′–3′)	GenBank Accession no.	T_a (°C)
*Mvi*1	$(GTCTCT)_{25}$	F: CTg AgA CTg TCT AgT ATg CTg R*: CAC TgC TAT ATC AAA AAg gCT CA	AY686620	56
*Mvi*2	$(GATA)_4(GACA)_1(GATA)_{15}$	F: ATT CCT gAC AAg AgT CTg TAT CC R*: ATg TTg gAg gAg CTg TgA gg	AY686621	54
*Mvi*3	$(GATA)_2(GATG)_1(GATA)_5$	F: CTT CCT ATC AAg gCg ATT AAg A R*: CCC CCA AAA CTA CTC TCT TCA	AY686622	54
*Mvi*5	$(GATA)_{17}$	F*: gTT TCA gAA TgT TCC TCA AgA T R: TTT gTC CTg AAT TTC CCT ACA	AY686623	56
*Mvi*9	$(GACA)_{15}$	F: gAC AgT CCT gCA TTC gTC Tg R*: gTC gTg TTT CTg TTT gCC TgT	AY686624	65
*Mvi*10	$(GACA)_{19}$	F: gAA gAg ACA gCA CTg Agg R*: CTC ggC ACA AgC ATT TAg TC	AY686625	65
*Mvi*12	$(GACA)_2\ (CACA)_1\ (GACA)_{19}$	F: gCT TCC CTg TCA CTg CAC R*: CAT TCA CCT gAT CAg Cag TgT T	AY686626	65
*Mvi*14	$(GACA)_5\ (GACC)_1\ (GACA)_{11}$	F: TTT gAA ggT AAg AgA gAg ATC C R*: Tgg ACA TAC TTT gTg AAA ACg AC	AY686627	62
*Mvi*16	$(GTAT)_{20}$	F*: CAA AgT Agg ggT gTA TAA CTg AAT R: CTg ATT gAA AgC CAC Agg g	AY686628	65
*Mvi*22	$(GTAT)_{29}$	F: CAT ACC CTA ACT ATT AAg TgT gAA CA R*: gCT ggA gCA ACT TCA TTC Ag	AY686629	56
*Mvi*31	$(GACA)_{11}(GGCA)_6$	F: ggg ATT AgT CTA AgA gCT gCT g R*: TCT ACT TgT CTT ATC TAC CAg CCT AC	AY686630	56
*Mvi*33	$(GGCA)_4(GACA)_{28}$	F: ggg ATT AgT CTA AgA gCT gCT g R*: gTT TCT gCC TgC CAg TCg	AY686631	56

*HEX labelled primer.

Table 2 Summary of 12 capelin tetranucleotide microsatellite loci listing base-pair (bp) size range, number of alleles, number of base-pair variations, observed (H_O) and expected (H_E) heterozygosity, and *P*-value for deviation from Hardy–Weinberg expectation (standard error, SE of *P*-value in parentheses). Estimates are based on a sample of 44 beach-spawning capelin collected from three spawning populations along eastern Newfoundland in 2002

Locus	Size range (bp)	Number of alleles	Base pair variation	H_O	H_E	HWE (SE)
Mvi1	90–372	59	2	0.932	0.946	0.458 (0.018)
Mvi2	112–230	33	1, 2, 4	0.977	0.959	0.552 (0.016)
Mvi3	160–216	14	4	0.864	0.916	0.314 (0.009)
Mvi5	100–136	10	4	0.864	0.852	0.713 (0.006)
Mvi9	94–232	38	1, 2, 4	0.977	0.973	0.765 (0.014)
*Mvi*10	116–200	21	4	1.000	0.945	0.995 (0.001)
*Mvi*12	100–304	38	1, 2, 4	0.886	0.965	0.009 (0.003)
*Mvi*14	246–400	36	1, 2, 4	0.977	0.967	0.701 (0.015)
*Mvi*16	158–386	47	1, 2, 4	0.932	0.981	0.056 (0.008)
*Mvi*22	152–308	32	1, 4	0.932	0.966	0.513 (0.016)
*Mvi*31	372–483	6	1, 2, 4	0.568	0.533	0.611 (0.006)
*Mvi*33	164–168	2	4	0.023	0.023	—

loci, except *Mvi*31 (0.568) and *Mvi*33 (0.023), exhibited high H_O (0.864–0.977). *Mvi*12 showed heterozygote deficiency though no other departed from Hardy–Weinberg expectation (Table 2). Genotypic disequilibrium tests between locus pairs were largely insignificant. Evidence of linkage disequilibrium between *Mvi*5 and *Mvi*16 ($P = 0.0024$) proved insignificant after multiple-test correction.

Acknowledgements

We thank B. Slaney, M. Clark and DFO staff in Newfoundland for field efforts; S. Wu for lab assistance; P. O'Reilly and P. Bentzen for suggestions; and CBS for funding to CTT, ELK and BSN.

References

Carscadden JE, Vilhjalmsson H (2002) Capelin — What are they good for? Introduction. *ICES Journal of Marine Science*, **59**, 863–869.

Dodson JJ, Carscadden JE, Bernatchez L, Colombani F (1991) Relationship between spawning mode and phylogeographic structure in mitochondrial DNA of North Atlantic capelin *Mallotus villosus*. *Marine Ecology Progress Series*, **76**, 103–113.

Dushchenko VV (1993) On the causes of spatial and temporal differentiation of the spawning concentrations of capelin *Mallotus villosus* in the Barents Sea. *Voprosy Ikhtiologii*, **33**, 248–257.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.

Hamilton MB, Pincus EL, Fiore DI, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, **27**, 500–507.

Louis EJ, Dempster ER (1987) An exact test for Hardy–Weinberg and multiple alleles. *Biometrics*, **43**, 805–811.

Mork J, Friis-Soerensen E (1983) Genetic variation in capelin *Mallotus villosus* from Norwegian waters. *Marine Ecology Progress Series*, **12**, 199–205.

O'Reilly PT, McPherson AA, Kenchington E, Taggart CT, Jones MW, Bentzen P (2002) Isolation and characterization of tetranucleotide microsatellites from Atlantic haddock (*Melanogrammus aeglefinus*). *Marine Biotechnology*, **4**, 418–422.

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

Roed KH, Fevolden SE, Christiansen JS, Midthjell L (2003) Isolation and characterization of DNA microsatellites for capelin *Mallotus villosus* population genetic studies. *Marine Ecology Progress Series*, **262**, 299–303.

Rozen S, Skaletsky HJ (2000) PRIMER 3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, NJ.

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

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.