Conservation Genet Resour (2014) 6:349–351
DOI 10.1007/s12686-013-0091-6

MICROSATELLITE LETTERS

Development and characterization of 16 polymorphic microsatellite loci for the Alaska blackfish (Esociformes: *Dallia pectoralis*)

Matthew A. Campbell · George K. Sage · Rachel L. DeWilde · J. Andrés López · Sandra L. Talbot

Received: 22 August 2013 / Accepted: 23 November 2013 / Published online: 5 December 2013
© Springer Science+Business Media Dordrecht (outside the USA) 2013

Abstract Blackfishes (Esociformes: Esocidae: *Dallia*), small fishes with relictual distributions, are unique in being the only primary freshwater fish genus endemic to Beringia. Although the number of species of *Dallia* is debated, disjunct populations and distinct mitochondrial divisions that predate the end of the last glacial maximum are apparent. We developed sixteen polymorphic microsatellites from the Alaska blackfish (*Dallia pectoralis*) to study genetic diversity in *Dallia*. Genotypes from two populations, Denali (n = 31) and Bethel (n = 35), demonstrated the usefulness of the loci for population-level investigation. Observed and expected heterozygosity averaged 18.6 and 19.8 % in Denali and 61.1 and 63.7 % in Bethel. Number of alleles at each locus averaged 3.50 in Denali and 9.63 in Bethel. The observed signature of variability and structuring between populations is consistent with mitochondrial data.

Keywords Esociformes · *Dallia* · 454 sequencing · Microsatellite loci

We isolated novel microsatellite loci from Alaska blackfish (*Dallia pectoralis*), by: (1) standard construction of repeat enriched genomic libraries, and (2) shotgun sequencing with 454 technology and subsequent bioinformatics analysis (Abdekrim et al. 2009).

Genomic DNA from muscle tissue of one Alaska blackfish obtained from a pond in Bethel, Alaska was extracted and used to create GC and AT-enriched libraries. For 454 sequencing, genomic DNA was extracted from fin tissue from an individual collected at Ballaine Pond, Fairbanks, Alaska (64.8692, −147.8252, WGS 84) and sequenced on one-eighth of a picotiter plate. Quality control and assembly was conducted with GS De Novo Assembler v 2.6. 14,770 reads were assembled into 1,175 contigs of at least 100 base pairs; 82,275 singleton reads were greater or equal to 200 bases. Assembled and singleton reads were mined for microsatellites using MSATCOMMANDER v 1.0.8 (Faircloth 2008). Microsatellite loci were further screened manually to identify duplicates based on repeat motif and flanking sequences. Among 433 contigs greater than 500 base pairs, twelve dinucleotide microsatellites with nine or more repeats were identified. Singleton reads were filtered for an average quality score of 35, then searched for di-, tri-, or tetranuclotide microsatellites containing thirteen or more repeats. Thirty-one candidate reads were identified.

Three polymorphic microsatellite loci from standard cloning and thirteen polymorphic microsatellite loci from the 454 sequencing effort were genotyped in Alaska blackfish from Bethel, AK (Hangar Lake Drainage; n = 35) and Denali National Park and Preserve (Minchumina Lake Basin; n = 31). GENEPOP v 3.3 (Raymond and Rousset

Electronic supplementary material The online version of this article (doi:10.1007/s12686-013-0091-6) contains supplementary material, which is available to authorized users.

M. A. Campbell (✉) · R. L. DeWilde
Department of Biology and Wildlife, University of Alaska Fairbanks, 101 Murie Building, 982 North Koyukuk Drive, Fairbanks, AK 99775, USA
e-mail: macampbell2@alaska.edu

G. K. Sage · S. L. Talbot
Alaska Science Center, US Geological Survey, 4210 University Drive, Anchorage, AK 99508, USA

J. A. López
University of Alaska Museum, 907 Yukon Drive, Fairbanks, AK 99775, USA

J. A. López
School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 207A O'Neill Building, Fairbanks, AK 99775, USA

Table 1 Characterization of 16 microsatellite loci developed for Alaska blackfish (*Dallia pectoralis*) from Bethel ($n = 35$) and Denali ($n = 31$) Alaska, including repeat motif, primer sequence (universal tail in parentheses), allele size range in base pairs (bp), number of alleles (A), observed (H_O) and expected (H_E) heterozygosity, and GenBank Accession numbers

Locus	Repeat motif	Primer sequence (5′–3′)[a]	*Bethel*			*Denali*			GenBank accession no.
			Size	A	H_O/H_E	Size	A	H_O/H_E	
Dpe001	$GT_{(11)}$	F: (SP6) AGAACAGAACGGTGAGTC R: TGCCAAGATTAGGGTAGC	121–125	3	0.49/0.42	123	1	–	KF437968
Dpe002	$CA_{(23)}$	F: (M13F) CAGGACATTGCAGATGAC R: GTGAGTTTGTGTACATGC	111–149	12	0.83/0.86	123–131	5	0.23/0.24	KF437969
Dpe003	$CA^{b}_{(10)}$	F: (M13R) AGATGCAGGCCATTATGC R: CAGTCAGACCCATAGGGC	184–194	3	0.14/0.14	192	1	–	KF437970
Dpe004	$GT_{(14)}$	F: (SP6) ACTGACTGACTGATTGAG R: CCACACACACTGGGATAC	268–276	5	0.46/0.47	272–274	2	0.10/0.09	KF437971
Dpe005	$GT^{bc}_{(13)}$	F: (M13R) GAGTGGTGAAAGAGTGGC R: ATTTGGCCAGACTCCTGC	152–162	7	0.54/0.53	156	1	–	KF437972
Dpe006	$GT_{(14)}$	F: (SP6) CTCTTGTGTTTGTCATGC R: ATAGAACCACTCCGGAGC	119–141	10	0.63/0.73	133–137	2	**0.00/0.06**	KF437973
Dpe007	$GT^{b}_{(15)}$	F: (M13F) AGAGAAAGCTGACTGTGC R: TGACTGGTGGGAAGTCAG	115–123	4	0.34/0.35	117–127	4	0.26/0.23	KF437974
Dpe008	$GT_{(13)}$	F: (M13R) AGATCCGCTCTGGACAGC R: ACCTGTCCAGTTAGTTGC	144–158	5	0.34/0.41	144–150	2	0.03/0.03	KF437975
Dpe009	$GT_{(13)}$	F: (SP6) GTGAGTTGATCTTTCTGC R: TCATACCTAAGAGTTCGC	100–110	6	0.46/0.53	102–112	4	0.13/0.13	KF437976
Dpe010	$GT^{bc}_{(13)}$	F: (M13F) CGATCTGATCGTCTGTCC R: ACACACCGAGACAGATGC	129–199	23	**0.80/0.93**	127–145	8	0.48/0.46	KF437977
Dpe011	$GT_{(13)}$	F: (SP6) CACGGTACGACCCTGAGC R: GTACGCGTACATCCCAAC	177–191	8	0.82/0.86	175–183	4	0.19/0.18	KF437978
Dpe012	$GT^{b}_{(13)}$	F: (SP6) CCATTCGAGAACTATAGC R: GTGAGTTGACAAACCCGG	134–158	11	0.80/0.82	136–152	5	0.48/0.62	KF437979
Dpe013	$GT_{(14)}$	F: (M13F) ACTATCAGCTGGCAATGC R: TACCATGTCAGCAACGGC	130–150	9	0.63/0.54	134–142	3	0.13/0.12	KF437980
Dpe014	$GT^{b}_{(20)}$	F: (SP6) TGAAGAGGGCCCATCTGG R: ACGCTCACTGTCATCCGC	154–178	12	0.80/0.88	154–174	5	0.23/0.24	KF437981
Dpe015	$GT^{b}_{(14)}$	F: (SP6) TTATTAGGCTGCCACAGC R: TTCTTTGAGTGCCAACTG	143–159	9	**0.71/0.78**	151–157	3	0.48/0.52	KF437982
Dpe016	$GT_{(20)}$	F: (M13F) TGTTGAGGTGGCAAGTGG R: TCGTTCATATGAAGTCCC	178–248	27	1.00/0.96	174–186	6	0.26/0.24	KF437983

Values in bold text denote comparisons that deviated from Hardy–Weinberg equilibrium expectation

[a] Universal tail primer sequences; M13F (CACGACGTTGTAAAACGAC), M13R (GGATAACAATTTCACACAGG), and SP6 (GATTTAGGTGACACTATAG)

[b] Signifies an imperfect repeat. The longest perfect repeat motif is listed

[c] Signifies the presence of 1 base pair repeat

1995) was used to calculate observed (H_O) and expected (H_E) heterozygosity and test for linkage disequilibrium (LD) and deviations from Hardy–Weinberg equilibrium (HWE) within populations. Potential null alleles and scoring errors were identified with MICRO-CHECKER (Van Oosterhout et al. 2004). We used GENEPOP and ARLEQUIN v 3.1 (Excoffier et al. 2005) to assess levels of population differentiation (χ^2 distribution of genotypes and θ, respectively). Bonferroni corrections for multiple tests ($\alpha = 0.05$) were applied.

Genetic diversity metrics are listed in Table 1. Two loci deviated from HWE (Dpe006: $\chi^2 = 13.82$, df = 4, $P = 0.008$; Dpe010: $\chi^2 =$ infinity, df = 4, $P < 0.001$) and three locus pairs were found to be significantly linked:

Dpe011 × Dpe012 ($\chi^2 = 12.42$, df = 4, $P < 0.015$); Dpe005 × Dpe007 ($\chi^2 = 10.38$, df = 2, $P = 0.006$); Dpe005 × Dpe008 ($\chi^2 = 7.82$, df = 2, $P = 0.020$).

In Denali, thirteen loci were polymorphic, the number of alleles per locus averaged 3.50; and H_O and H_E averaged 18.6 and 19.8 %, respectively. Denali was in overall HWE ($\chi^2 = 23.68$, df = 24, $P = 0.480$); one locus (Dpe006; $P = 0.016$) deviated from HWE expectations (Table 1), perhaps due to the presence of a null allele ($0.01 < P < 0.05$). Genotypes from Denali produced more significant (n = 5; $P < 0.05$) locus-by-locus pairwise comparisons than would be expected among seventy-eight possible comparisons, involving nine of the thirteen polymorphic loci (data not shown). In Bethel, sixteen loci were polymorphic, the number of alleles at each locus averaged 9.63; and H_O and H_E averaged 61.1 and 63.7 %, respectively. Alaska blackfish from Bethel deviated from overall HWE (χ^2 = infinity, df = 32, $P < 0.001$), due to heterozygote deficit in two loci (Dpe010, $P < 0.001$; Dpe015, $P = 0.020$; Table 1); MICRO-CHECKER identified a possible null allele in Dpe010 ($0.01 < P < 0.05$). Bethel was in linkage equilibrium overall, with fewer significant (n = 4; $P < 0.05$) locus-by-locus comparisons than expected randomly among 120 possible comparisons. The four significant pairwise comparisons observed in Bethel involved six loci (data not shown).

None of the significant locus pairs in LD in Denali were among pairs significantly linked among Bethel samples. We found that population differentiation between Denali and Bethel is high and significant (χ^2_{genotype} = infinity; df = 32, $P < 0.001$; $\theta = 0.334$, $P < 0.001$). We are currently applying these loci to additional populations to investigate population structure.

Acknowledgments The Alaska Experimental Program to Stimulate Competitive Research (AK EPSCoR) funded work by RLD and MAC, and 454 sequencing. MAC was also supported by a Discover Denali graduate fellowship, and an Institute of Arctic Biology summer fellowship. Additional funding was provided by the United States Geological Survey Alaska Science Center, the Department of Biology and Wildlife at the University of Alaska Fairbanks and a National Science Foundation Research Initiation Grant to JAL (for enriched library cloning and sequencing). We wish to thank Dr. Naoki Takebayashi for allowing MAC to use his computational resources. The University of Alaska Museum and the United States Fish and Wildlife Service Yukon-Kuskokwim National Wildlife Refuge provided specimens. Lee Tibbitts and Ann Riddle provided reviews of this manuscript prior to submission. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Abdekrim J, Robertson B, Stanton JA, Gemmell N (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. Bio Techniques 46:185–192

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

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

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

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 4:535–538