Molecular Ecology Notes (2005) **5**, 259–262 doi: 10.1111/j.1471-8286.2005.00903.x

PRIMER NOTE

Characterization of 13 single nucleotide polymorphism markers for chum salmon

C. T. SMITH,* J. BAKER,† L. PARK,† L. W. SEEB,* C. ELFSTROM,* S. ABE‡ and J. E. SEEB*
**Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518, USA, †National Marine Fisheries Service, 2725 Montlake Blvd., Seattle, WA 98122, USA, ‡Laboratory of Breeding Science, Hokkaido University, Hakodate 041-8611, Japan*

Abstract

We report the characterization of 13 single nucleotide polymorphism (SNP) genotyping assays for chum salmon (*Oncorhynchus keta*). These assays are based on the 5′-nuclease reaction and thus facilitate high-throughput genotyping with minimal optimization time. Because data generated using these markers may be transported and combined across laboratories, SNPs offer the potential to reduce the amount of redundant work being done in mixture and migratory studies of chum salmon.

Keywords: 5′-nuclease assay, chum salmon, *Oncorhynchus keta*, PCR, SNP, TaqMan

Received 15 September 2004; revision received 11 October 2004; accepted 2 December 2004

Chum salmon (*Oncorhynchus keta*) has the widest natural distribution of all the Pacific salmon species, ranging northward from Korea in the western Pacific Ocean, through the Chukchi Sea in the Arctic, and south to the northwest United States in the eastern Pacific Ocean. Chum salmon provides a significant component of many commercial and subsistence fisheries in North America and Asia. Some wild populations on both sides of the Pacific Ocean have become increasingly threatened as a result of hatchery production, habitat loss, decadal climate change, and fishing (Seeb *et al*. 2004). Consequently, management agencies have increasingly relied upon gene markers to clarify population structure and study oceanic distribution and migration to attempt to conserve depressed populations of this fish species (cf. Seeb & Crane 1999).

Mixture and migratory studies of chum salmon have been accomplished in the past using allozymes, mtDNA PCR-RFLPs, and microsatellite markers (e.g. Park *et al*. 1993; Scribner *et al*. 1998; Seeb *et al*. 2004). Each of these markers suffers from attributes that negatively impact either throughput rates or interlaboratory standardization (reviewed in Morin *et al*. 2004) and have led to the development of redundant baselines. Standardization difficulties are particularly well illustrated by microsatellite markers. At least four baselines describing allele frequencies in chum salmon have been produced in different laboratories in the United States and Canada, yet managers are unable to compare or combine data across laboratories.

Correspondence: Christian Smith, Fax: + 9072672442; E-mail: christian_smith@fishgame.state.ak.us

Single nucleotide polymorphisms (SNPs) are a class of genetic marker for which data may be compared to external DNA sequences, and thus data are automatically standardized across chemistries, hardware platforms, and laboratories. For example in chinook salmon (*Oncorhynchus tshawytscha*), we obtained sequence variants from four different laboratories using three different genotyping techniques and developed 5′-nuclease assays that yielded SNP genotypes common to all (Smith *et al*. 2005). The 5′-nuclease reaction is a SNP genotyping method in which the 5′-exonuclease activity of DNA polymerase is used to digest allele-specific probes in the course of the polymerase chain reaction (PCR) (Holland *et al*. 1991). The use of fluorogenic probes further allows amplification and genotyping to take place simultaneously (Lee *et al*. 1993) negating the need for electrophoresis. Here, we present conditions for using the 5′-nuclease reaction to genotype 13 SNPs in chum salmon.

Two mitochondrial and five nuclear loci were amplified in ≥ 32 individuals using PCR conditions cited in Table 1, sequenced using BigDye 3.1 Chemistry (Applied Biosystems) and analysed on an Applied Biosystems 3100 DNA Sequencer. Novel sequences were deposited in GenBank (Accession nos AB091514–AB091531, AY771785–AY771786, AY773960–AY773963, AY779330–AY779335, AY780634–AY780641, AY791847–AY791849).

DNeasy 96 tissue kits (QIAGEN) were used to extract DNA from 695 chum salmon representing eight collections

Table 1 Thirteen SNP markers in chum salmon

Locus (locus name, assay name, reference GenBank Accession no., source of PCR primers)	Oligonucleotide sequences (5′–3′) (forward primer, reverse primer, allele-specific probe 1*, allele-specific probe 2*)	Heterozygosity		
		H_E	H_O	F_{ST}
mtDNA control region	ACTACTCTCTGGCGGCTACA	0.00	0.00	0.20
Oke_Cr30	AACCTTGGATTAGTGCTGATGTATGAG			
AB039894, position 30	VIC-CAAGTTATGGCATTTACA			
Sato *et al*. 2001	FAM-TTACAAGTTATAGCATTTACA			
mtDNA control region	CCCTCATACATCAGCACTAATCCAA	0.00	0.00	0.00
Oke_Cr231	GCTGAAAGTTGGTGGGTAAAGAC			
AB039894, position 231	VIC-CTGCAACCGATTAAT			
Sato *et al*. 2001	FAM-CTGCAACTGATTAAT			
mtDNA control region	CTTAATGTAGTAAGAACCGACCAACGA	0.00	0.00	0.71
Oke_Cr386	ACTTAGGAACCAAATGCCAGGAAT			
AB039894, position 386	VIC-ATCGTATTAGGTCGCATCT			
Sato *et al*. 2001	FAM-AAATCGTATTAGTCGCATCT			
mtDNA NADH-3	TGGTATTGAATTTGTCGTAGAAGGCAAA	0.00	0.00	0.73
Oke_ND3-69	CCACGGCCTACACGTAATCATC			
AY771785, position 69	VIC-CAGCCAGAAAGGTAGAG			
Domanico & Philips 1995	FAM-CAGCCAGAAAAGTAGAG			
Cytokeratin S	GGGCCATTCTCTGAGTTCAGT	0.40	0.38	0.08
Oke_CKS-389	GAGCACCGGTTGTCATGGA			
AY773960, position 389	VIC-AAATGAATGATAATGTGTTCTG			
Moran 2002	FAM-AAATGAATGATAATATGTTCTG			
Gonadotropin releasing hormone	CCTGAGGAGACAAGTGCACATG	0.49	0.35	0.24
Oke_GnRH-373	ATTGGCCATTAGGAATACAATGAATACAATATCT			
AY779333, position 373	VIC-ATGTAGTAAGTAGTCATATTCATTTA			
Baker *et al*. 2002	FAM-TAGTAAGTAGTCATATTAATTTA			
Gonadotropin releasing hormone	GGTAAAATATGTTGTAAACAATATACCAAAGTATCTGAAATAA	0.10	0.10	0.08
Oke_GnRH-527	GCAACATTGTCTCCCTACATTTATAGGTA			
AY779333, position 527	VIC-AAGATTTCCATTTGTCAGCCT			
Baker *et al*. 2002	FAM-AAGATTTCCATTAGTCAGCCT			
unknown	AGGTTTGTATGCGGCTGCTT	0.34	0.28	0.14
Oke_u1–519	CAACTCAGCACAAGAACTGTTCAC			
AY791847, position 519	VIC-TCATGAAATGGGTTTCTAT			
Moran 2002	FAM-ATGAAATGGGATTCTAT			
Interleukin 8 receptor	CACACAAGCACTCTCGCTTTG	0.43	0.36	0.11
Oke_IL8r-35	TGGTCGCAAGTGAAAGGTTGT			
AY780640, position 35	VIC-AAGGCACGCTCATAGT			
Moran 2002	FAM-AAGGCACGGTCATAGT			
Interleukin 8 receptor	TGAGGTCTGTACAAACTTGTTTGCA	0.42	0.37	0.10
Oke_IL8r-125	GGCAACCGGGAGGGAAAA			
AY780640, position 125	VIC-CTCTGCATCATGGAATA			
Moran 2002	FAM-TCTGCATCGTGGAATA			
Interleukin 8 receptor	AGGCTGAGGCAGTGACATG	0.28	0.24	0.11
Oke_IL8r-272	ACCGACGCTTCTAGCAATGTAC			
AY780640, position 272	VIC-CAAGAGGTCTTTGTCGACAT			
Moran 2002	FAM-AAGAGGTCTTTGCCGACAT			
Interleukin 8 receptor	GGATGGACATTCACAGTCTGGTT	0.43	0.37	0.09
Oke_IL8r2–406	TTTCCAATCCCTGGCATCGT			
AY780634, position 406	VIC-AAACACAAAACCCC			
Moran 2002	FAM-AAACACAAACCCCC			
Myelin proteolipid protein	CACTCCCCTCGCTTACTGATATCTA	0.48	0.47	0.08
Oke_DM20-548	ACGGACAGCTCATTTACATACACAA			
AY779331, position 548	VIC-CTAGATCATGTTCACTATCTATA			
Moran 2002	FAM-TAGATCATGTTCACTATATATA			

*Each allele-specific probe was labelled with either VIC or 6FAM on its 5′ end and bore a minor groove binder and a nonfluorescent quencher on its 3′ end.

Table 2 Observed allele frequencies for the rarer allele at 13 SNP loci. Collections were taken from Sasauchi River (Honshu Island, Japan), Tokachi River (Hokkaido Island, Japan), Bolshaya River (Kamchatka Peninsula, Russia), Sheenjek River (Upper Yukon drainage, United States), Anvik River (Lower Yukon drainage, United States), Volcano River (southern Alaska Peninsula, United States), Wally H. Noerenberg Hatchery (WHN, northern Gulf of Alaska, United States), and Puget Sound (Washington State, United States)

	Population							
Locus	Sasauchi	Tokachi	Bolshaya	Sheenjek	Anvik	Volcano	WHN	Puget Sound
n	78	80	95	95	95	64	95	91
Oke_Cr30	0.347	0.182	0.032	0.000	0.011	0.079	0.000	0.000
Oke_Cr231	0.026	0.025	0.032	0.000	0.000	0.016	0.000	0.011
Oke_Cr386	0.883	0.722	0.063	0.000	0.011	0.078	0.000	0.000
Oke_ND3-69	0.885	0.734	0.032	0.000	0.011	0.078	0.000	0.000
Oke_CKS-389	0.135	0.222	0.382	0.274	0.274	0.375	0.483	0.099
Oke_GnRH-373	0.867	0.763	0.500	0.145	0.315	0.422	0.389	0.216
Oke_GnRH-527	0.000	0.000	0.043	0.126	0.168	0.023	0.006	0.016
Oke_u1–519	0.026	0.081	0.274	0.069	0.150	0.234	0.391	0.450
Oke_IL8r-35	0.571	0.559	0.321	0.263	0.213	0.359	0.183	0.148
Oke_IL8r-125	0.563	0.526	0.289	0.263	0.216	0.359	0.174	0.144
Oke_IL8r-272	0.387	0.365	0.205	0.116	0.074	0.141	0.067	0.071
Oke_IL82r-406	0.545	0.519	0.299	0.263	0.216	0.359	0.176	0.159
Oke_DM20-548	0.149	0.237	0.479	0.606	0.426	0.391	0.472	0.330

taken from throughout the species' range (Table 2). Genotyping assays were performed in 384-well reaction plates, with four wells in each plate as negative controls (no-template). Each reaction was conducted in a 5-μL volume consisting of 0.15 μL of template DNA in 1× TaqMan PCR cocktail (proprietary PCR buffer, thermal-stable DNA polymerase exhibiting 5′-nuclease activity, and a passive reference dye) (Applied Biosystems), 900 nM of each PCR primer, and 200 nM of each probe (Table 1). Thermal cycling was conducted at ramp rates of 1 °C per second on either a 7900 real-time PCR instrument (Applied Biosystems) or a DNA Engine Tetrad (MJ Research) as follows: an initial denaturation of 10 min at 95 °C was followed by 45 cycles of 92 °C for 15 s and 60 °C for 1 min. Following amplification, all plates were read and scored on the 7900 real-time PCR instrument using SEQUENCE DETECTION software 2.1 (Applied Biosystems).

GENEPOP (Raymond & Rousset 1995) was used to calculate expected and observed heterozygosities (H_E and H_O) and to test for genotypic linkage disequilibrium (LD) and departures from Hardy–Weinberg equilibrium (HWE). Significance levels for all tests were corrected for multiple simultaneous tests (Rice 1989). No significant ($\alpha = 0.05$) departures from HWE were detected. Significant LD was observed for SNPs within loci (i.e. within the mtDNA, within the interleukin 8 receptor and within gonadotropin releasing hormone), but LD was not detected among SNPs from different sequences. Examining tightly linked loci will result in diminishing returns for population discrimination studies. For example, over 98% of the compound mtDNA haplotypes observed in this study could be identified by genotyping two SNPs (*Oke_Cr30* + either *Oke_Cr386* or *Oke_ND3-69*). While this argues that it would be most efficient to collect data from two mtDNA SNPs and five nuclear SNPs rather than the entire set presented here, rare compound haplotypes could become more informative as additional populations are examined.

Data collected using the present assays directly reflect underlying DNA sequences and are thus readily combined with data collected from different hardware, chemistry platforms, and laboratories. This transportability of data as well as the rapid rate at which SNP data are generated renders these markers well suited to mixture and migratory studies of chum salmon.

Acknowledgements

This work was funded by NOAA Cooperative Agreement NA03NMF4380185 and by North Pacific Research Board grant R0303. We are grateful for comments made on earlier drafts by A.J. Gharrett and A. Surridge, and for technical assistance provided by J. Rhydderch.

References

Baker J, Bentzen P, Moran P (2002) Molecular markers distinguish coastal cutthroat trout from coastal rainbow trout/steelhead and their hybrids. *Transactions of the American Fisheries Society*, **131**, 404–417.

Domanico MJ, Phillips RB (1995) Phylogenetic analysis of Pacific salmon (genus *Oncorhynchus*) based on mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution*, **4**, 366–371.

Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of *Thermus aquaticus* DNA polymerase.

Proceedings of the National Academy of Sciences of the United States of America, **88**, 7276–7280.

Lee LG, Connell CR, Bloch B (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Research*, **21**, 3761–3766.

Moran P (2002) Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. *Ecology of Freshwater Fish*, **11**, 30–55.

Morin PA, Luikart G, Wayne RK, SNP Workshop Group (2004) Application of SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**, 208–216.

Park LK, Brainard MA, Dightman DA, Winans GA (1993) Low levels of intraspecific variation in the mitochondrial DNA of chum salmon (*Oncorhynchus keta*). *Molecular Marine Biology and Biotechnology*, **2**, 362–370.

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

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

Sato SJ, Ando H, Ando S, Urawa S, Urano A, Abe S (2001) Genetic variation among Japanese populations of chum salmon inferred from nucleotide sequences of the mitochondrial DNA control region. *Zoological Science*, **18**, 99–106.

Scribner KT, Crane PA, Spearman WJ, Seeb LW (1998) DNA and allozyme markers provide concordant estimates of population differentiation: analyses of U.S. and Canadian populations of Yukon River fall-run chum salmon (*Oncorhynchus keta*). *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1748–1758.

Seeb LW, Crane PA (1999) Allozymes and mitochondrial DNA discriminate Asian and North American populations of chum salmon in mixed-stock fisheries along the south coast of the Alaska Peninsula. *Transactions of the American Fisheries Society*, **128**, 88–103.

Seeb LW, Crane PA, Kondzela CM *et al.* (2004) Migration of Pacific Rim chum salmon on the high seas: insights from genetic data. *Environmental Biology of Fishes*, **69**, 21–36.

Smith CT, Seeb JE, Schwenke P, Seeb LW (2005) Use of the 5′-nuclease reaction for SNP genotyping in Chinook salmon. *Transactions of the American Fisheries Society*, in press.