Chinook Baseline Expansion with SNP Markers

Final Report to the Pacific Salmon Commission
November 9, 2009

by

Shawn Narum, Andrew Matala, and Jeff Stephenson

Columbia River Inter-Tribal Fish Commission
Hagerman Fish Culture Experiment Station
3059-F National Fish Hatchery Road
Hagerman ID 83332

Summary:

The objective of this project was to genotype single nucleotide polymorphisms (SNPs) in multiple collections of Chinook salmon from the Columbia River to add to coastwide and regional genetic baselines for GSI purposes. In this multi-year project, a total of 52 collections were genotyped with a panel of 96 SNP markers. Collections were selected from the existing microsatellite baseline to allow both types of genetic markers to be combined as appropriate for GSI applications. Each of the three major genetic lineages were represented with 5 populations from the lower Columbia River, 15 from the interior ocean-type lineage (i.e., fall-run), and 32 from the interior stream-type lineage (i.e, spring-run). This genetic data will be added to the shared GAPS database for use by various agencies.

Introduction

Both microsatellite and single nucleotide polymorphism (SNP) markers have proven to be effective in differentiating populations of salmon (Narum et al. 2008) and determining the proportion of stock origin in mixed stock fisheries (Beacham et al. 2006; Smith and Seeb 2008). SNP discovery and assay development has provided the resources (e.g., Campbell and Narum 2008a) to include additional genetic markers to the existing GAPS microsatellite baseline (Seeb et al. 2007). The availability of these new markers, combined with high throughput genotyping instruments, enables the capability of producing large amounts of SNP genotypes in a short time-frame. Using this technology, we genotyped SNP markers in 52 collections of Chinook salmon from the Columbia River Basin to include in the GAPS baseline. These additional markers are expected to greatly increase power for Genetic Stock Identification (GSI) of mixed stock fisheries. This study was intended to improve upon the existing genetic baseline, and follows from Recommendation 12 in the PSC Expert Panel Report (Hankin et al. 2005), that there is a need for support of an "immediate evaluation of a coordinated transition for all salmon species from GSI based on the use of microsatellite markers to GSI based on SNP markers."

Current Project Objectives

The specific objective of this project was to genotype SNPs in multiple collections of Chinook salmon from the Columbia River to add to coastwide and regional genetic baselines. Current microsatellite markers do not adequately distinguish populations of summer/fall run Chinook salmon in the Columbia River, but inclusion of SNP markers

used in tandem with existing microsatellite data will not only improve the resolution of stocks for GSI, but will also improve accuracy of stock assignments.

SNP Genotyping Methods

Tissue samples from each individual were processed with Qiagen DNeasy® kits to extract DNA from fin clips stored in 100% ethanol. Isolated DNA from each sample was genotyped for 96 SNP markers (Table 1) with Taqman chemistry (Applied Biosystems) and Fluidigm 96.96 dynamic array chips (reaction volumes of ~7nL) for SNP genotyping. Since genotyping in nL reaction volumes reduces the average starting copy number to a range where genotyping accuracy becomes less reliable (Campbell and Narum 2008b), a pre-amplification protocol was used to increase the number of starting copies. Pre-amplification occurred in 7µl reactions with 2µl of genomic DNA and 5µl of PCR cocktail (3.5µl of Qiagen Multiplex Mastermix, 0.875µl of 96 pooled primer sets at 0.36uM, and 0.625µl water) under the following thermal cycling program: initial denature at 95°C for 15 minutes, 14 cycles of 95°C for 15 seconds and 60°C for 4 minutes, hold at 4°C. Immediately after cycling, 133uL of nuclease free H_2O or TE buffer was added to each PCR reaction and stored at 4°C.

Pre-amplified template DNA was then genotyped with Fluidigm 96.96 dynamic array chips that included a three step process: 1) SNP assays (Taqman primers/probes) and DNA samples were mixed according to manufacturers protocols and loaded onto the chip with a Fluidigm IFC Controller instrument, 2) target SNPs were amplified for 50 cycles on a Eppendorf thermal cycler specially formatted for the Fluidigm 96.96 chip,

and 3) chips were scanned with a Fluidigm EP-1 instrument to detect fluorescently labeled allele-specific probes. Genotypes for each assay were auto-scored with Fluidgim SNP Analysis v.2.1.1 software and verified by eye with scoring guides provided by an assay database and a heterozygous indicator sample for each SNP.

Table 1. List of 96 SNPs genotyped in Chinook salmon collections.

	Chinook SNP assays:	Alleles V/F		Chinook SNP assays:	Alleles V/F
1	Ots_113242-216	C/T	49	Ots_P53	G/A
2	Ots_113457-40R	C/T	50	Ots_PGK-54	T/A
3	Ots_123048-521	A/C	51	Ots_Prl2	A/G
4	Ots_128757-61R	A/-	52	Ots_RAG3	C/T
5	Ots_94857-232R	T/C	53	Ots_RFC2-558	A/-
6	Ots_94903-99R	G/T	54	Ots_S7-1	T/C
7	Ots_96222-525	C/T	55	Ots_SClkF2R2-135	A/T
8	Ots_96500-180	G/T	56	Ots_SERPC1-209	A/T
9	Ots_96899-357R	T/A	57	Ots_SL	A/G
10	Ots_97077-179R	G/T	58	Ots_SWS1op-182	T/A
11	Ots_AldB1-122	C/T	59	Ots_TAPBP	C/T
12	Ots_aldb-177M	TTG/ATA	60	Ots_TGFB	C/T
13	Ots_AsnRS-60	T/C	61	Ots_TLR3	C/T
14	Ots_aspat-196	G/C	62	Ots_Tnsf	A/G
15	Ots_C3N3	T/G	63	Ots_u07-07.161	C/T
16	Ots_CD59-2	G/A	64	Ots_u07-18.378	A/T
17	Ots_CD63	A/C	65	Ots_u07-25.325	T/C
18	Ots_cox1-241	C/T	66	Ots_u07-49.290	G/A
19	Ots_EndoRB1-486	G/A	67	Ots_u07-53.133	C/T
20	Ots_EP-529	A/G	68	Ots_u07-57.120	A/T
21	Ots_ETIF1A	A/C	69	Ots_u202-161	T/A
22	Ots_FARSLA-220	G/A	70	Ots_u211-85	C/T
23	Ots_FGF6A	G/T	71	Ots_u4-92	T/C
24	Ots_FGF6B_1	A/C	72	Ots_u6-75	C/T
25	Ots_GDH-81x	C/-	73	Ots_unk526	A/G
26	Ots_GH2	A/T	74	Ots_zP3b-1	G/T
27	Ots_GnRH-271	C/T	75	Ots_ZR-575	G/A
28	Ots_GPDH-338	G/A	76	Ots_102414-395	A/G
29	Ots_GPH-318	C/T	77	Ots_ARNT	G/T
30	Ots_GST-375	C/T	78	Ots_arp-436	AA/TT
31	Ots_GTH2B-550	C/G	79	Ots_Cath_D141	T/C
32	Ots_HSP90B-100	C/T	80	Ots_CCR7	C/T
33	Ots_IGF-I.1-76	A/T	81	Ots_CRB211	A/C
34	Ots_Ikaros-250	G/A	82	Ots_E2-275	A/G
35	Ots_IL11	T/C	83	Ots_GST-207	G/A
36	Ots_IL8R_C8	C/T	84	Ots_hsc71-3'-488	C/T
37	Ots_MHC1	G/A	85	Ots_hsc71-5'-453	C/T
38	Ots_MHC2	T/G	86	Ots_hsp27b-150	G/A
39	Ots_mybp-85	C/T	87	Ots_105105-613	C/G
40	Ots_Myc-366	T/C	88	Ots_106747-239	C/A
41	Ots_myo1a-384	A/C	89	Ots_mapK-3'-309	T/G
42	Ots_myoD-364	T/G	90	Ots_mapKpr-151	A/T
43	Ots_nkef-192	C/T	91	Ots_RAS1	C/T
44	Ots_NOD1	C/G	92	Ots_TNF	C/T
45	Ots_nramp-321	G/A	93	Ots_u07-17.135	A/G
46	Ots_OPLW173_1	T/C	94	Ots_u07-20.332	A/C
47	Ots_Ots311-101x	AA/--	95	Ots_u07-64.221	G/C
48	Ots_P450	T/A	96	Ots_110064-383	C/T

Results & Discussion

Nearly 5,000 individual Chinook salmon from 52 collections in the Columbia River were genotyped with 96 SNP markers. The SNP markers were variable throughout the region with an unbiased heterozygosity average of 0.238 and range of 0.190 to 0.311 in Camas Cr. and Kalama R., respectively (Table 2).

Markers were tested for deviation from Hardy-Weinberg in each population with GENEPOP (Raymond and Rousset 1995). Deviations from Hardy-Weinberg expectations were not consistent for any locus or population tested, indicating that the data reasonably represented these collections in the baseline.

The collections in this project represent the majority of populations from the Columbia River, with geographic coverage extending throughout the basin. However, future efforts will be needed to add some collections that were not available (i.e., more lower Columbia River populations). This project provides SNP data to add to the microsatellite baseline that should be useful for GSI of Chinook salmon.

Quality Control

Genetic data was tested under standard quality control procedures in CRITFC's genetic laboratory. This includes confirmation of assay genotypes, repetitive genotyping, positive and negative controls, and automated allele conversion.

Project Benefits / Monitoring and Evaluation

The objective of this study was intended to improve both regional and coastwide GSI applications. The additional SNPs that were genotyped will benefit all agencies as these markers are available for the entire genetics community.

Table 2. Columbia River collections of Chinook salmon genotyped with SNP markers.

Population (Collection ID)	Sample Size (n)	Region	Lineage	Run	Origin	Year	life stage	H_o	H_e	F_{is}
Cowlitz River	86	L. Columbia R.	Lower	Fall	HAT	2004	Adult	0.2779	0.2762	-0.0072
Lewis River	93	L. Columbia R.	Lower	Fall	NOR	2003	Adult	0.2814	0.2858	0.0033
Kalama River	81	L. Columbia R.	Lower	Spring	HAT	2004	Adult	0.3003	0.3110	0.0363
McKenzie River	84	L. Columbia R.	Lower	Spring	HAT	2004	Adult	0.2577	0.2547	-0.0022
North Santiam River	87	L. Columbia R.	Lower	Spring	HAT	2004	Adult	0.2519	0.2495	-0.0014
Clearwater River (CherryLane)	213	Clearwater R.	Ocean	Fall	NOR	2008	Adult	0.2688	0.2599	-0.0249
Clearwater River	118	Clearwater R.	Ocean	Fall	NOR	2008	juvenile	0.2458	0.2583	0.0288
Clearwater River	73	Clearwater R.	Ocean	Fall	NOR	2000	Adult	0.2553	0.2585	0.0015
Nez Perce Tribal Hatchery	86	Clearwater R.	Ocean	Fall	HAT	2003	Adult	0.2514	0.2531	0.0146
Deschutes River (lower)	90	M. Columbia R.	Ocean	Fall	NOR	1999	Adult	0.2519	0.2589	0.0149
Hanford Reach	93	M. Columbia R.	Ocean	Fall	NOR	2000	Adult	0.2494	0.2496	-0.0115
Klickitat River	92	M. Columbia R.	Ocean	Fall	NOR	2004	Adult	0.2490	0.2549	0.0113
Little White Salmon River	91	M. Columbia R.	Ocean	Fall	HAT	2006	Juvenile	0.2508	0.2549	0.0187
Lyons Ferry Hatchery	90	M. Columbia R.	Ocean	Fall	HAT	2000	Adult	0.2495	0.2536	0.0182
Spring Creek (Tule)	88	M. Columbia R.	Ocean	Fall	HAT	2006	Juvenile	0.2221	0.2185	-0.0055
Umatilla River	82	M. Columbia R.	Ocean	Fall	HAT	2006	Adult	0.2495	0.2475	0.0021
Priest Rapids	85	U. Columbia R.	Ocean	Fall	HAT	2001	Juvenile	0.2420	0.2458	0.0063
Wells Dam	89	U. Columbia R.	Ocean	Fall	HAT	1993	Adult	0.2456	0.2514	0.0413
Deschutes River (Upper)	90	M. Columbia R.	Ocean	summer	HAT	1998	juvenile	0.2539	0.2636	0.0491
Methow River	88	U. Columbia R.	Ocean	summer	NOR	1993	juvenile	0.2450	0.2494	0.0103
Lolo Creek	89	Clearwater R.	Stream	Spring	NOR	2001	Juvenile	0.2190	0.2177	0.0100
Dworshak-NFH	88	Clearwater R.	Stream	Spring	HAT	2005	Adult	0.2238	0.2225	0.0061
Lochsa River	77	Clearwater R.	Stream	Spring	HAT	2005	Adult	0.2180	0.2195	0.0164
East Fork Salmon River	94	E. Fork Salmon R.	Stream	Spring	NOR	-?-	Adult	0.2088	0.2044	-0.0099
Looking-Glass Creek	89	Grand Rhonde	Stream	Spring	HAT	1994	Juvenile	0.2202	0.2197	0.0020
Lostine River	82	Grand Rhonde	Stream	Spring	-?-	2003	-?-	0.2153	0.2105	0.0007
Minam River	82	Grand Rhonde	Stream	Spring	NOR	2002	Juvenile	0.2321	0.2371	0.0292
Wenaha River	44	Grand Rhonde	Stream	Spring	NOR	2002	juvenile	0.2370	0.2331	-0.0119
Big Creek	92	Grand Rhonde	Stream	Spring	NOR	2001	Adult	0.1995	0.1991	0.0188
Catherine Creek	85	Grand Rhonde	Stream	Spring	NOR	2003	Adult	0.2241	0.2265	0.0111
Imnaha River	92	Grand Rhonde	Stream	Spring	NOR	1998	Adult	0.2206	0.2177	0.0035

Rapid River	93	L. Salmon R.	Stream	Spring	HAT	1999	Adult?	0.2139	0.2152	0.0270
Tucannon River	87	L. Snake R.	Stream	Spring	NOR	2003	Adult	0.2521	0.2431	-0.0154
Little White Salmon River	92	M. Columbia R.	Stream	Spring	HAT	-?-	Adult	0.2254	0.2289	0.0213
Carson-NFH	91	M. Columbia R.	Stream	Spring	HAT	2007	Juvenile	0.2235	0.2260	0.0124
John Day River	84	M. Columbia R.	Stream	Spring	NOR	2005	Adult	0.2349	0.2375	0.0212
Klickitat River	157	M. Columbia R.	Stream	Spring	HAT	2002	Adult	0.2847	0.2886	0.0180
Klickitat River (broodstock)	129	M. Columbia R.	Stream	Spring	HAT	2006	Adult	0.2905	0.3018	0.0405
Klickitat River	187	M. Columbia R.	Stream	Spring	NOR	2006	Adult	0.2964	0.3093	0.0409
Camas Creek	47	M. Fork Salmon R.	Stream	Spring	NOR	2006	juvenile	0.1957	0.1896	-0.0116
CapeHorn Creek	88	Snake R.	Stream	Spring	NOR	2005	Juvenile	0.1881	0.1915	-0.0017
Newsome Creek	90	S. Fork Clearwater R.	Stream	Spring	NOR	2001	Adult	0.2147	0.2179	0.0208
Johnson Creek (McCall Stock)	88	S. Fork Salmon R.	Stream	Spring	HAT	2002	Adult	0.2024	0.2013	-0.0059
Secesh River	81	S. Fork Salmon R.	Stream	Spring	NOR	2001	Juvenile	0.1970	0.1984	0.0058
Johnson Creek (Weir)	92	S. Fork Salmon R.	Stream	Spring	NOR	2002	Adult	0.2007	0.2012	0.0118
Methow River	93	U. Columbia R.	Stream	Spring	HAT	1998	Juvenile	0.2121	0.2216	0.0693
Wenatchee River	85	U. Columbia R.	Stream	Spring	-?-	1993	-?-	0.2134	0.2159	0.0068
Winthrop-NFH (Carson stock)	84	U. Columbia R.	Stream	Spring	HAT	2001	Adult	0.2222	0.2232	0.0150
Pahsimeroi River	93	U. Salmon R.	Stream	Spring	-?-	2004	-?-	0.2099	0.2068	-0.0174
Sawtooth Hatchery	91	U. Salmon R.	Stream	Spring	HAT	2003	Adult	0.2040	0.2025	0.0047
West Fork Yankee Fork	75	U. Salmon R.	Stream	Spring	NOR	2005	juvenile	0.1998	0.2028	0.0168
Cle Elum Hatchery	88	Yakima R.	Stream	Spring	HAT	2006-2007	-?-	0.2624	0.2676	0.0273

H_o = observed heterozygosity
H_e = unbiased expected heterozygosity
F_{is} = Inbreeding coefficient

References

Beacham, T. D., J. R. Candy, K. L. Jonsen, J. Supernault, M. Wetklo, L. Deng, K. M. Miller, and R. E. Withler. 2006b. Estimation of stock composition and individual identification of Chinook salmon across the Pacific Rim using microsatellite variation. Transactions of the American Fisheries Society 135:861-888.

Campbell, N. R., and S. R. Narum. 2008a. Identification of novel SNPs in Chinook salmon and variation among life history types. Transactions of the American Fisheries Society 137:96-106.

Campbell, N. R., and S. R. Narum. 2008b. Quantitative PCR assessment of microsatellite and SNP genotyping with variable quality DNA extracts. Conservation Genetics DOI 10.1007/s10592-008-9661-7.

Hankin, D. G., J. H. Clark, R. B. Deriso, J. C. Garza, G. S. Morishima, B. E. Riddell, C. Schwarz, and J. B. Scott. 2005. Report of the Expert Panel on the future of the coded wire tag recovery program for Pacific salmon. Pacific Salmon Comm.

Narum, S. R., M. Banks, T.D. Beacham, M.R. Bellinger, M.R. Campbell, J. DeKoning, A. Elz, C.M. Guthrie III, C. Kozfkay, K.M. Miller, P. Moran, R. Phillips, L.W. Seeb, C.T. Smith, K. Warheit, S.F. Young, J.C. Garza. 2008. Differentiating salmon populations at broad and fine geographic scales with microsatellites and SNPs. Molecular Ecology 17:3464-3477.

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

Seeb, L. W, A. Antonovich, M.A. Banks, T.D. Beacham, M.R. Bellinger, S. M. Blankenship, M. Campbell, N.A. Decovich, J.C. Garza, C.M. Guthrie III, T. A. Lundrigan, P. Moran, S.R. Narum, J.J. Stephenson, K.J. Supernault, D.J. Teel, W.D. Templin, J.K.Wenburg, S.F. Young, C.T. Smith. 2007. Development of a Standardized DNA Database for Chinook Salmon. Fisheries 32:540-552.

Smith CT, Seeb LW (2008) Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single-nucleotide-polymorphism (SNP) baselines for chum salmon. Transactions of the American Fisheries Society, 137, 751–762.