Chinook Baseline Expansion with Additional Genetic Markers

Final Report to the Pacific Salmon Commission
June 30, 2008

Prepared by

Shawn R. Narum & Jeff Stephenson
Columbia River Inter-Tribal Fish Commission
3059-F National Fish Hatchery Road
Hagerman ID 83332

Summary:

Genetic stock identification (GSI) is a tool with great potential to provide stock composition data for managing mixed fisheries. The existing genetic baseline for Chinook salmon includes 13 microsatellite markers for 165 populations, but lacks adequate power to distinguish some closely related stocks. The objective of this project was to include additional genetic loci (SNP markers) to the current GAPS (Genetic Analysis of Pacific Salmonids) baseline to provide an increase in power to differentiate stocks. From a pool of 82 possible markers, we selected the 50 most informative SNPs to expand the baseline. Of these, 5 had technical problems so 45 SNP markers were genotyped in 20 populations of Chinook salmon (n = 1,893) from the Columbia River. These genotypes will be combined with existing microsatellite data to provide a more powerful baseline for genetic stock ID purposes.

Introduction

Current GAPS Baseline

The current Chinook salmon microsatellite baseline constructed by the GAPS (Genetic Analysis of Pacific Salmonids) consortium for the CTC provides a powerful tool for evaluating mixed stock fisheries (Seeb et al. 2007). The current GAPS baseline (version 2.1) with contains a total of 165 coastwide populations (~22,000 samples) that have been genotyped with 13 highly polymorphic microsatellite loci (~500 alleles). Initial power analyses indicate that 41 of 43 reporting groups in the baseline have at least 90% accuracy in 100% mixture simulations. Further baseline expansion and maintenance planned in 2008 by GAPS laboratories will lead to a microsatellite baseline (v. 3.1) with even more dense geographic coverage (greater than 200 populations) with at least 50 populations from the Columbia River. Thus, a highly effective tool for mixed stock analysis of Chinook salmon is now in place and is available to co-managers through an online database.

Despite large, representative sample sizes from many populations and very high microsatellite allelic diversity, the resolution of specific stocks and populations in the baseline is limited in some cases. For example, fall Chinook salmon in the Columbia River are closely related and remain difficult to distinguish even with this powerful set of 13 microsatellite markers. Several other closely related populations in the baseline are similarly difficult to distinguish and thus have been pooled into a single reporting unit for GSI applications. In some cases (i.e., Washington coast troll fishery), a finer level of stock discrimination is necessary for management of fisheries. Additional genetic markers should increase stock assignment reliability when greater resolution is required.

Current Project Objectives

As a contributing GAPS laboratory, CRITFC recognized the need to increase the number of genetic markers in baseline populations. The objective of this project was to genotype 50 additional SNP markers in 20 Chinook salmon populations for inclusion with the microsatellite baseline (Narum et al. 2008).

Methods

The goal of this project was to improve stock differentiation for populations in the GAPS (Genetic Analysis of Pacific Salmonids) baseline. To achieve this goal, our objective was to genotype 50 SNPs in 20 baseline populations. The process included DNA extraction from tissue samples, polymerase chain reaction (PCR) to amplify 50 SNP markers (Table 1), and plate scans to detect fluorescently labeled PCR products. Since SNPs only have two alleles, no standardization was necessary. The SNP data will be incorporated to the GAPS baseline for improved mixed fishery genetic stock identification.

Table 1. Ranking of 50 most informative SNPs for Chinook salmon.

	Avg Rank	Locus	He	Global Fst	Rank (Fst)	Rank (WL)	Rank (ln)
1	1.0	Ots_C3N3	0.230	0.549	1	1	1
2	2.3	NRAMP	0.229	0.548	2	2	3
3	2.7	GTH2B-550	0.236	0.522	3	3	2
4	4.0	Ots_u211-85	0.239	0.517	4	4	4
5	5.3	Ots_FARSLA-220	0.252	0.462	5	6	5
6	5.7	Ots_P450	0.255	0.461	6	5	6
7	8.7	Ots_SL	0.290	0.403	12	7	7
8	9.0	Ots_u07-57.120	0.256	0.419	9	10	8
9	9.7	Ots_HSP90B-100	0.265	0.412	10	9	10
10	12.0	Ots_TAPBP	0.221	0.454	7	18	11
11	12.3	CCR7	0.195	0.451	8	20	9
12	12.3	IL8R-C8	0.242	0.404	11	14	12
13	12.3	COX1	0.325	0.367	13	11	13
14	13.0	Ots_u202-161	0.334	0.340	17	8	14
15	14.3	Ots_Tnsf	0.305	0.348	16	12	15
16	16.0	IL11	0.327	0.335	18	13	17
17	18.0	Ots_MHC2	0.305	0.326	20	16	18
18	18.0	Ots_hnRNPL-533	0.309	0.332	19	15	20
19	18.7	NKEF	0.202	0.365	14	26	16
20	20.0	RAG3	0.359	0.293	21	17	22
21	22.0	ASPAT	0.196	0.364	15	30	21
22	22.7	Ots_RFC2-558	0.340	0.272	24	21	23
23	22.7	Ots_ETIF1A	0.352	0.269	25	19	24
24	24.0	Ots_u07-18.378	0.276	0.278	23	24	25
25	25.0	Ots_FGF6A	0.310	0.265	26	23	26
26	26.3	PGK-54	0.333	0.237	27	25	27
27	26.3	Ots_MHC1	0.372	0.227	29	22	28
28	30.3	NOD1	0.278	0.234	28	34	29
29	32.3	Ots_u07-53.133	0.414	0.176	38	27	32
30	33.0	MYOD	0.398	0.186	36	28	35
31	33.3	Ots_u07-17.373	0.144	0.284	22	48	30
32	33.7	TLR3	0.362	0.187	35	33	33
33	34.0	Ots_u07-07.161	0.397	0.177	37	29	36
34	35.0	Otsu07-25.325	0.196	0.213	30	41	34
35	35.7	ALDB	0.252	0.196	32	38	37
36	36.7	Ots_SClkF2R2-135	0.387	0.166	40	32	38
37	37.3	CD63	0.412	0.165	41	31	40
38	38.7	Ots_IGF-I1-76	0.136	0.192	33	52	31
39	39.0	Ots311	0.193	0.192	34	44	39
40	40.0	Ots_u4-92	0.234	0.171	39	39	42
41	41.3	Ots_SWS1op-182	0.416	0.135	45	36	43
42	44.0	Ots_GnRH-271	0.094	0.205	31	57	44
43	44.3	Ots_Prl2	0.440	0.118	48	37	48
44	46.3	Ots_HSP90B-385	0.281	0.126	47	46	46
45	48.7	MYO1	0.192	0.134	46	53	47
46	49.0	TGFB	0.386	0.088	51	45	51
47	49.0	S7-1	0.459	0.081	53	40	54
48	49.3	Otsu07-49.290	0.436	0.080	54	42	52
49	49.3	Ots_P53	0.424	0.081	52	43	53
50	50.7	Ots_GPH-318	0.186	0.107	49	54	49

Results

A total of 1,893 Chinook salmon samples were genotyped from the Columbia River basin with additional SNP markers. Five SNPs with technical difficulties were removed from the data set, leaving a total of 45 SNPs for 1,893 samples that were analyzed. Across all populations, each SNP had 2 alleles for a total of 90 alleles in the data set. Of 900 tests for Hardy-Weinberg equilibrium, six were statistically significant after Bonferroni corrections. Populations and loci that deviated from equilibrium were: three populations with heterozygote deficits at hnRNPL (McKenzie, North Santiam, and Priest Rapids Hatchery), upper Deschutes with a heterozygote deficit at Ots_NOD1, Wells Hatchery with heterozygote excess at Ots_NRAMP, and Priest Rapids Hatchery with heterozygote deficit at Ots_S7-1.

Genetic diversity as measured by expected heterozygosity (Table 2) was highest in lower Columbia River populations (average of 0.325), moderate in stream-type populations (average of 0.276), and lowest in ocean-type populations (0.246). This was a particularly interesting trend given that this pattern of diversity is not consistent with microsatellite markers that suggest ocean-type populations have the highest levels of diversity.

Tests of genetic differentiation with pairwise F_{ST} indicated that all but three pairwise collections were significantly different from one another ($P < 0.0068$), lower Deschutes R. vs. upper Deschutes R., North Santiam vs. McKenzie, and Methow R. (summer) vs. Priest Rapids H. Overall, population relationships were consistent with expectations and suggest the data can be merged with the existing GAPS baseline.

Table 2. Sample size (n) and expected heterozygosity (He) for 20 populations of Chinook salmon.

	Population	n	He
1	Carson stock (WNFH) spring	94	0.266
2	Kalama Hatchery spring	95	0.358
3	Lewis River Hat. fall	95	0.396
4	Cowlitz Hatchery fall	95	0.305
5	lower Deschutes River (fall)	95	0.245
6	upper Deschutes River (sum/fall)	95	0.252
7	Methow River summer	95	0.235
8	Methow River spring	95	0.264
9	Wells Dam summer	96	0.259
10	John Day spring	95	0.280
11	Catherine Creek spring	96	0.270
12	Warm Springs H. spring	92	0.271
13	McKenzie River	96	0.282
14	North Santiam	96	0.281
15	Priest Rapids Hat. (fall)	96	0.241
16	Lostine River (spring)	96	0.265
17	Lewis River-wild spring	95	0.331
18	Chiwawa River spring	88	0.250
19	CleElum Hat. spring	96	0.335
20	Tucannon River spring	92	0.287
		1893	

Discussion

The objective of this project was to expand the number of genetic markers for existing populations of Chinook salmon in the GAPS genetic baseline. We successfully completed this objective and closely followed the anticipated time schedule from the proposal. This additional data will now be incorporated into the genetic baseline with plans to post all baseline genotypes to an online database (to be included in next version).

Quality Control

Genetic data was tested under standard quality control procedures in CRITFC's genetic laboratory. This includes confirmation of raw genotypes through repetitive genotyping, positive and negative controls, and use of indicator alleles.

Project Benefits

This project was intended to enhance PSC genetic stock identification of mixed stock fisheries since this is essential information related to the Pacific Salmon Treaty. The additional SNP markers genotyped in this project provide further power to distinguish populations in the baseline. We expect this baseline to be a long standing tool for genetic stock identification of mixed fisheries under the Pacific Salmon Treaty. All agencies and organizations that utilize the GAPS baseline for genetic stock identification will benefit from a more powerful baseline.

Future directions for the Genetic Baseline

In order to increase the power of the genetic baseline across the coastwide range, the GAPS consortium has received funding from the Chinook Technical Committee to select 75 of the most informative SNPs available and purchase assays. Members of the GAPS group met in May, 2008 and selected 75 SNPs for baseline genotyping. As part of the recently completed PSC project to add markers, CRITFC has completed genotyping for 45 of the 75 coastwide SNPs. Thus 30 more SNPs will need to be genotyped in

Columbia River populations to maintain a standard set of SNPs for coastwide baseline and GSI applications.

References

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.

Seeb LW, Antonovich A, Banks MA, Beacham TD, Bellinger MR, Blankenship SM, Campbell M, Decovich NA, Garza JC, Guthrie III CM, Lundrigan TA, Moran P, Narum SR, Stephenson JJ, Supernault KJ, Teel DJ, Templin WD, Wenburg JK, Young SF, Smith CT (2007) Development of a Standardized DNA Database for chinook Salmon. *Fisheries*, **32**, 540-552.