

# Expanded Bilateral Chum Salmon SNP Genetic Baseline for Genetic Stock Identification

<sup>1</sup>Maureen P. Small WDFW and <sup>2</sup>John Candy DFO

*Final Report to the Pacific Salmon Commission Southern Boundary Restoration and Enhancement Fund*

SF-2018-SP-22

WDFW #18-10957

Contract dates: 07/01/2018 to 06/30/2020, extended to 12/31/2020

November 24, 2020

<sup>1</sup>Washington Department of Fish and Wildlife, 1111 Washington St. SE., Olympia, WA 98501

<sup>2</sup>Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, B.C., Canada, V9T 6N7

## Table of Contents

Acknowledgements.....	2
Abstract.....	2
Introduction.....	2
Materials and Methods.....	3
Amplicon Sequencing Assays and Genotyping .....	3
Pairwise genetic comparisons .....	3
Neighbor-joining dendrogram .....	3
Assignment tests and mixed fishery analyses .....	3
Results and Discussion .....	4
Genotyping and collections.....	4
Pairwise genetic comparisons .....	4
Neighbor-joining dendrogram .....	5
Assignment tests and mixed fishery analyses .....	5
Literature Cited:.....	7
Figures. ....	9
Tables.....	14

## Acknowledgements

This project was funded by the Southern Boundary Restoration and Enhancement Fund and Washington State General Funds. Archived and newly acquired samples added to the Chum salmon baseline were collected by tribal and state personnel. We thank the Chum Technical Committee for ideas and support. We thank the WDFW lab personnel for genotyping work. We thank Ben Sutherland (DFO) for his contributions to the project. We thank Todd Kassler (WDFW) for administering the contract. We thank the Southern Boundary Restoration and Enhancement Fund for supporting this project.

## Abstract

In this project we expanded the initial SNP baseline generated by WDFW in the Southern Boundary region to include better representation of Canadian (CA) Chum salmon populations in the Fraser River and along Strait of Georgia, and Puget Sound populations. The United States (US) had developed a comprehensive panel of 500 SNPs for a Chum salmon baseline funded in part by the Southern Fund and NOAA Saltonstall Kennedy funding (Seeb et al. 2014, Rawson et al. 2014, Small et al. 2017a). The US incorporated 350 of these SNPs in a SNP panel using the Genotyping-in-Thousands by sequencing (GT-seq) technology (Campbell et al. 2015). and genotyped an initial 31 Chum salmon populations with the Chum GT-seq panel. Under the current grant, the US filled in gaps with six new populations and increased temporal representation and collection sizes for 13 existing populations. Also under this grant, the DFO lab in Nanaimo conducted a parallel project where they tested the 500 Chum SNPs and added SNPs from other species to design a Chum SNP panel specific to their Ion Torrent genotyping technology. The DFO lab genotyped additional CA Chum salmon populations to develop a comprehensive Chum SNP baseline in CA. There are 300 SNPs in common between the US and CA SNP panels: the data are compatible allowing cooperative research and data sharing after minimal standardization.

## Introduction

Chum salmon escapements have trended downward for the past 10 years in most stocks along the Southern Boundary creating a moderate level of concern for these stocks. In response to concerns over stock abundance levels commercial exploitation was reduced or suspended in some areas and in some years. Our goal was to expand the SNP baseline to improve the ability of the parties to manage populations of low abundance and to distinguish populations in the lower Fraser from population in North Puget Sound that have been difficult to resolve. With the expanded baseline both countries will be able to evaluate catch composition in domestic fisheries and in mixed fisheries in the Southern Boundary Region. Because the data are compatible, fisheries analyses can be conducted in either or both countries and data can be shared directly. We set priority areas in the Southern Boundary Region (Figure 1) to focus baseline additions including the Strait of Juan de Fuca, Puget Sound, and the WA coast. The US has improved baseline coverage by adding in populations genotyped for other projects under other funding to further fill in gaps in these areas (Figure 2). In this project we describe the baseline improvements in the Southern Boundary Region on the US side and the performance of the current baseline in simulated mixed stock fisheries.

## Materials and Methods

### Amplicon Sequencing Assays and Genotyping

An amplicon is composed of thousands of copies of the same single piece of a DNA sequence that has been amplified (copied) by the polymerase chain reaction (PCR). The amplicon sequence contains a region of interest that requires the high number of copies in order to assess the region of interest. For our purposes, the region of interest is a single nucleotide within the amplicon sequence that is variable (SNP) among individuals and among populations. The actual sequence of the amplicon containing the variable site is obtained by next-generation sequencing methods. In next-generation sequencing hundreds of amplicons are sequenced concurrently in hundreds of individuals, simultaneously creating genotypes for hundreds of individuals consisting of hundreds of SNPs. The WDFW Molecular Genetics Laboratory uses an Illumina next-generation sequencer to genotype SNPs with GT-seq technology.

The GT-seq protocol (Campbell et al. 2015) genotypes hundreds of samples at hundreds of loci in a single effort, greatly improving genotyping efficiency and decreasing costs. Genomic DNA is extracted from tissue samples using silica membrane kits (Macherey-Nagel). The SNP genotypes are generated from the extracted DNA as follows: briefly, samples are subjected to an initial tailed, multiplex PCR reaction that adds sequencing primer sites to target DNA, in a second PCR reaction unique barcode DNA sequences (barcodes identify individual fish) are added to the target DNA (amplicon), barcoded amplicons are pooled and sequenced on the next-generation sequencer (WDFW-MGL Illumina MiSeq). Using the barcodes, perl scripts split pooled sample sequences into individual files and assemble the SNP genotypes for each individual. We perform QA/QC on the genotypes, removing samples missing more than 80% of their data. We used GenAIEx (Peakall and Smouse 2012) to format genotypic data for analyses.

### Pairwise genetic comparisons

To explore spatial and temporal genetic relationships between populations we calculated pairwise  $F_{ST}$  values using ARLEQUIN (Excoffier and Lischer 2010). Pairwise  $F_{ST}$  is an estimate of genetic variation among collections (higher genetic variation indicates higher genetic distinction and lower gene flow or longer time since sharing common ancestors). The pairwise  $F_{ST}$  values were first calculated with all temporal collections from a tributary separate and then with temporal collections combined within tributaries. Pairwise  $F_{ST}$  values were tested for whether they were significantly different from zero with a permutation test (1000 permutations).

### Neighbor-joining dendrogram

We plotted Nei's genetic distances among collections in a neighbor-joining tree using programs within the PHYLIP software package (Felsenstein 1993). We assessed the repeatability of the groupings on the tree with 1000 bootstrap replications.

### Assignment tests and mixed fishery analyses

We tested the baseline for its power to identify fish back to its home collection using ONCOR from Kalinowski (<http://www.montana.edu/kalinowski/>). For the assignments, we conducted 100% mixture simulations which simulated 200 fish from a single population 100 times and estimated the mixture proportions in comparison to the full baseline. Assignments were conducted with all temporal collections from a tributary separate and then with temporal collections combined within tributaries. While the 100%

simulation is an unrealistic scenario (100% of the fish in a fishery sample of 200 fish are all from the same population), it gives an indication of whether individuals from a population can be assigned reliably back to their population.

We also used ONCOR to simulate international and domestic mixed fisheries in North Puget Sound (Skagit versus Fraser, Skagit versus Nooksack, and Nooksack versus Fraser), South Puget Sound (Green River Hatchery versus Diru Hatchery, Diru Hatchery versus Nisqually River, and Green River Hatchery versus Hoodspout Hatchery). Two hundred fish were simulated 100 times for each test and results were averaged over all 100 simulations to show average assignments to individual collections and then averaged over regional run groups. Each test consisted of fisheries composed of 50% each of two populations that appeared closely related according to their pairwise  $F_{ST}$  values and that were near each other.

## Results and Discussion

### Genotyping and collections

Genotyping went well for most collections (see Table 2), with 92% of the samples genotyping at 80% or more of the loci. The genotyping added six new collections to the baseline and increased representation in 13 collections that had been too small for reliable estimations. The genotyping also added new temporal collections to the baseline which are important for monitoring changes in populations over time.

### Pairwise genetic comparisons

WDFW has been building the Chum SNP baseline as WDFW conducted research in tributaries absent in the initial baseline (see Table 1). We examined the full Chum SNP baseline for genetic, run group, and geographic structure by calculating the amount of genetic variation using pairwise  $F_{ST}$  tests. Because of the size and complexity of the data set, we present only the pairwise  $F_{ST}$  values in the Southern Boundary Region with temporal collections combined per tributary (Table 3). The values are colored with a heat map to make it easier to see relationships: green for low values, red for high values and yellow for intermediate values. The values were tested for whether they were significantly different from zero with a permutation test (1000 permutations). Following sequential Bonferroni correction for multiple simultaneous tests, corrected alpha was 0.0008. Pairwise  $F_{ST}$  values that were not significantly different from zero are in boxes.

The heat map illustrates the relationships among Chum salmon collections where collections are similar to each other within geographic region and run group and different from other regions and run groups. The triangles of green boxes along the diagonal consist of collections from the same geographic region and run group that are genetically similar. For instance, the Strait of Juan de Fuca (SJF) fall Chum salmon are the first six collections in the matrix and they are similar to each other (green) and to the North Puget Sound (NPS) fall Chum salmon (next seven collections) but both the SJF and NPS fall Chum salmon are different from the South Puget Sound (SPS) fall Chum salmon, the next collection of samples (yellow). One comparison to note is between Chum salmon collected as summer and fall-run in 1994 from Sherwood Creek in SPS: these were not significantly different from each other when they were tested separately (pairwise  $F_{ST} = 0.0004$ ,  $p = 0.2471$ ) and were combined in Table 3. I discussed the collections with the WDFW Biologist in the area (Gabe Madel, pers comm.) and he said the distribution of spawner returns is unimodal in the creek, suggesting they are all fall Chum. There may have been some bimodality to the run in the 1990's, but they were not

genetically distinct, suggesting that there was sufficient gene flow to prevent genetic divergence, even in the face of divergent return times. The winter Chum salmon from SPS are similar to each other (significantly different), and distinct from the SPS fall Chum salmon. The Hood Canal (HC) fall Chum salmon are mostly indistinguishable from each other (green color and box indicates insignificant value). The heat map also shows some effects of small collection size. For instance, the collection from Green River Hatchery ( $N = 47$ ) is half the size we recommend for characterizing a population and it is weakly distinguished from the HC fall Chum salmon. Although that relationship could also result if Green River Hatchery chum salmon were planted in HC or the same broodstock was used in Green River and HC hatcheries. Similarly, the North Fork Skokomish fall ( $N = 30$ ) from HC is not significantly different from Salmon Creek falls from Strait of Juan de Fuca (SJF) and Salmon Creek falls are weakly distinguished from the other HC fall Chum salmon. There is little significant variation (green values in boxes) within each of the region/run groups. The HC summer Chum salmon (including Salmon Creek from SJF) are highly divergent from all fall and winter Chum salmon (red).

### Neighbor-joining dendrogram

We plotted Nei's genetic distances among all collections in the full WDFW Chum SNP baseline in a Neighbor-Joining tree (Figure 3 and Figure 4). Because the size of the data set makes it difficult to see collection names, we present the tree in a radial format (Figure 3) that better shows the geographic and run group clusters in the data set, and also in a square format (Figure 4) where names are more visible. These trees are identical, it is just the architectural structure that differs. The trees illustrate the genetic relationships we saw in the pairwise  $F_{ST}$  values: collections from each of the geographic and run groups cluster onto supported branches in the tree. There is little substructure indicated within the clusters, with the exception of the three Fraser collections and the three temporal collections from the Nooksack that each occupied supported subbranches on the NPS branch. The SPS winters were on their own supported branch in between the SPS fall branch (which included Sherwood Creek summers) and the HC fall branch. The West Coast Vancouver Island branch inserted on the main branch next to the SJF collections, which inserted individually along the main branch. The placement of the SJF collections could reflect their location along the SJF Chum salmon corridor. The CA collections in the WDFW baseline are uniformly small (average  $N = 41$ ) and thus more generally characterize the populations. This was seen in the lack of differentiation in the pairwise  $F_{ST}$  tests and in their lack of substructure in the SOG cluster in the trees.

### Assignment tests and mixed fishery analyses

The self-assignments were lower than the regional assignments and varied widely. The assignments back to home collection (self) averaged 60% ( $SD = 0.315$ ) and ranged from a low of 1% in the North Fork Skokomish fall collection to a high of 100% in the Salmon summer collection (Figure 5 and Table 4). The collection size was positively associated with self-assignment. If temporal collections were left separate before assignment, collection size explained ~45% of the self-assignment value. If temporal collections were grouped before assignment, collection size explained ~32% of the self-assignment value. The HC fall regional group had the lowest self-assignments on average (Table 5). The pairwise  $F_{ST}$  tests showed that most of these collections were not differentiated and the tree showed that they clustered together and they assigned strongly to their regional group (Table 5). There was nearly universally high assignment to region and run group. The assignments to regional and run groups averaged 97% ( $SD = 0.075$ ) and ranged from a high of 100% for 31 collections to a low of 53% for Sooke. The Sooke collection was in the WCVI regional group and is located along the north side of the SJF (Area 20 in Figure 1). In the assignment test, the fish that did not assign back to Sooke assigned to SJF\_F (30%) and to the SOG (16%), supporting a genetic relationship to Chum salmon

on the south side of the SJF and in SOG, either by including fish from these regions in the collection or from gene flow to nearby populations.

The simulated mixed fishery analyses showed that the baseline will be powerful for identifying unknown components of real mixed fisheries (Table 6). In the first test, the fishery was composed of equal numbers of fish from the lower Skagit and from Squawkum in the Fraser River. We chose these collections because the rivers are near each other and we suspected that these fish might return through the same pathway at the same time and be caught in a fishery together. The analysis indicated that roughly half of the simulated fish from each collection could be assigned back to their collection of origin and that the rest of the fish would assign to other collections in their regional group. For instance, 23% assigned back to Squawkum and 16% assigned to Peach and 7% assigned to Hopedale, for an average of 45% (4% SD) assigning back to the Fraser regional group. The Fraser group was slightly under-assigned and the NPS group was slightly over-assigned (53%, 4% SD). We suspect that the under-assignment to the Fraser could be a function of the small collection sizes in our baseline data set (avg N = 38 for Fraser collections). When testing the mix of Nooksack and Skagit, and mix of Nooksack and Fraser, the Nooksack is so distinct that all the simulated Nooksack fish assigned back to Nooksack and less than half of the simulated Skagit and Fraser fish assigned back to their home collection but they assigned well to their regional group. In the other domestic fisheries, the winter run Chum salmon from SPS all assigned correctly, but only half of the Green River and Hoodspport hatcheries assigned correctly, although regional assignments were high for both hatcheries.

The Chum SNP genetic baseline is already in use in recovery and management projects. For instance, we have used the baseline to identify origins of juvenile Chum salmon using nursery areas in Hood Canal and Strait of Juan de Fuca and co-managers have been able to estimate production per run group in streams supporting both summer and fall Chum salmon. We are also using the baseline to identify components of fisheries in Puget Sound, providing co-managers with data to better manage their fisheries to protect sensitive stocks (Litz et al. 2020). We re-analyzed components of a mixed fishery in South Puget Sound that had been analyzed previously with an earlier version of the baseline. An important change in the fishery components was an increase in the winter-run assigned Chum salmon from 92/1447 to 103/1447. Even more important was that with the improved baseline, the distribution of the Chum salmon that had assigned as winter-run changed from 60% Diru Hatchery and 40% Nisqually River, to 82% assigned to Diru Hatchery and 18% assigned to Nisqually River. The Nisqually winter run Chum salmon are a sensitive stock and Puget Sound fisheries are timed to avoid catching them. The updated baseline improved the resolution of this important stock. Another current use of the SNP panel is a parent-based tagging (PBT) project in the Lower Columbia River to assess an experimental supplementation and reintroductions program to recover the Columbia River Chum salmon.

#### SNP baseline future applications

The SNP baseline will provide higher resolution data for the Chum TC to input into the ChumGEM fisheries model. Because the model incorporates genetic and environmental information to manage fisheries and some SNPs are located in selected regions of the genome, SNP data will help us understand population responses to environment. The growing SNP baseline will be used in traditional genetic stock identification studies (cf., Larson et al. 2014, Beacham et al. 2017a, Beacham et al. 2017b, Litz et al. 2020), parentage based tagging (PBT) (Steele et al. 2012, Small et al. 2017b), genetic mark recapture (GMR) (Pearse et al. 2001, Seamons et al. 2012, Small et al. 2016), and ecological modeling and investigations, especially those involving genomic association to environmental factors, such as climate change (Hecht et al 2012, Johnston et al. 2014, Hand et

al. 2016). The most immediate use for the Southern Boundary Region will be mixed stock fisheries analysis (e.g Candy et al. 2016, Litz et al. 2020). Understanding stock proportions moving through the area and how these proportions vary over time are key components to managing fisheries in the Southern Boundary Region that involve Chum salmon originating in both countries. Initial analyses conducted by the Chum TC in the Strait of Juan de Fuca showed that stock proportions varied over time and location and also varied between years (Van Will et al. 2016). Treaty allocations require fisheries to shut down based on assumptions of the origins of fish in the harvest. More precise estimates of fishery compositions will allow better management of fisheries in the border region. The timing and composition of fish moving through Southern Boundary waters are crucial parameters for the ChumGEM model under development by the Chum Tech Committee, which will also inform fish management in the border region.

## Literature Cited:

Beacham TD, Wallace C, MacConnachie C, Jonsen K, McIntosh B, Candy JR, Devlin RH, Withler RE. 2017a. Population and individual identification of coho salmon in British Columbia through parentage-based tagging and genetic stock identification: an alternative to coded-wire tags. *Canadian Journal of Fisheries and Aquatic Sciences* 74:1-20.

Beacham TD, Wallace C, MacConnachie C, Jonsen K, McIntosh B, Candy JR, Withler RE. 2017b. Population and individual identification of Chinook Salmon in British Columbia through parentage-based tagging and genetic stock identification with single nucleotide polymorphisms. *Canadian Journal of Fisheries and Aquatic Sciences* 75(7): 1096-1105.

Candy, J.R. P. Van Will, A. H. Araujo, B. Patton (2016) Joint US and CA mixed-stock Chum fisheries sampling design and analysis. Report to the Southern Endowment Fund Project 57536.

Campbell, N. R., S. A Harmon, S. R. Narum. 2015. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Mol. Ecol. Resour.* 15(4):855-867.

Hand BK, et al. (11 co-authors) 2016. Climate variables explain neutral and adaptive variation within Salmonid metapopulations: the importance of replication in landscape genetics. *Molecular Ecology* 25(3):689-705.

Johnston SE, Orell P, Pritchard VL, Kent MP, Lien S, Niemela E, Erkinaro J, Primmer CR. 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (*Salmo salar*). *Molecular Ecology* 23(14):3452-3468.

Larson, W. A., L. W. Seeb, M. V. Everett, R. K. Waples, W. D. Templin, and J. E. Seeb. 2014. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications* 7(3):365-369.

Litz, M., M. Small, K Addae, B Patton, and A. Spidle. 2020. Report to Southern Boundary Restoration and Enhancement Fund: Puget Sound Chum Salmon GSI. 18 pp.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics* 28: 2537-2539.

Pearse, D. E., C. M. Eckerman, F. J. Janzen, and J. C. Avise. 2001. A genetic analogue of ‘mark–recapture’ methods for estimating population size: an approach based on molecular parentage assessments. *Molecular Ecology* 10:2711-2718.

Rawson K, Candy J, Beacham T, Seeb L, Small M, Warheit K, Winans G. 2014. Final Report to Southern Fund Panel, Chum SNP baseline Part I.

Seamons, T. R., D. J. Rawding, P. Topping, D. Wildermuth, S. Peterson, and M. Zimmerman. 2012. Progress Report, Spawner abundance estimates for Green River Chinook salmon submitted to Sentinel Stocks Committee of the Pacific Salmon Commission, July 24, 2012, Washington Department of Fish and Wildlife, Olympia.

Seeb, L. W., J. Seeb, and K. Warheit. 2014. Development of high resolution DNA markers to manage fishery interactions of chum salmon in Western Washington. Final Report Saltonstall-Kennedy Grant

Small, M. P., C. Scofield, J. Griffith, A. Spidle, P. Verhey, J. Whitney, and A. Whitcomb. 2016. 2014Broodyear Report: Abundance estimates for Stillaguamish River Chinook salmon using trans-generational genetic mark recapture, year 1. Report to the Southern Boundary Restoration and Enhancement Program.

Small, M. P., Warheit, K. W, Pascal, C., Seeb, L., Ruff, C., Winans, G, and J. Seeb. 2017a. Chum salmon Southern Boundary Area RAD sequencing project. Report to the Southern Boundary Restoration and Enhancement Fund, 29 pp.

Small, M. P. T. Hillson and G. Gee. 2017b. Determining hatchery or natural origin of Chum salmon spawners in the Cascade and Lower Gorge regions of the Columbia River with parent-based genetic tagging. WDFW Molecular Genetics Laboratory Report, 9 pp.

Steele, C. A., M. W. Ackerman, J. McCane, and M. R. Campbell. 2012. Parentage Based Tagging of Snake River hatchery steelhead and Chinook salmon. Bonneville Power Administration. Annual Progress Report July1, 2011 - June 30, 2012. IDFG report Number 12-09

Van Will, P, J.R. Candy, C. McConnell, L. Kearey, and B. Patton. (2016) Joint US and CA Juan de Fuca Chum Sampling Program Report to the Southern Endowment Fund.

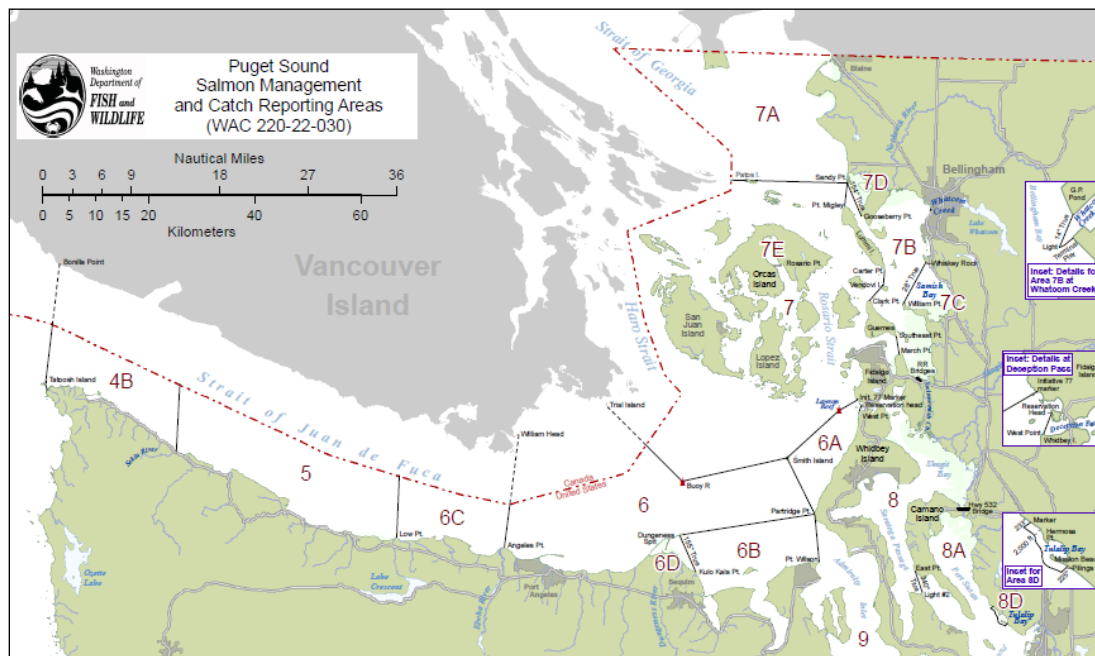
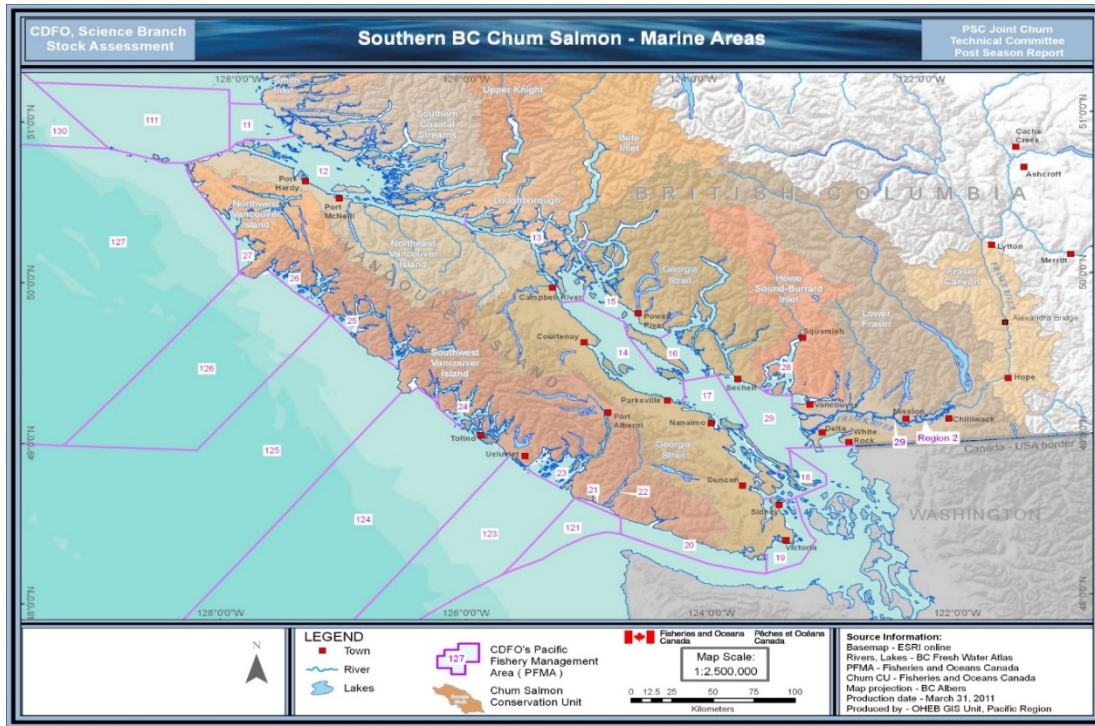


Figure 1. Maps showing Canadian Fishery Management areas (top) and US catch areas (bottom) in the Southern Boundary Region (CTC Annual Report, 2011).

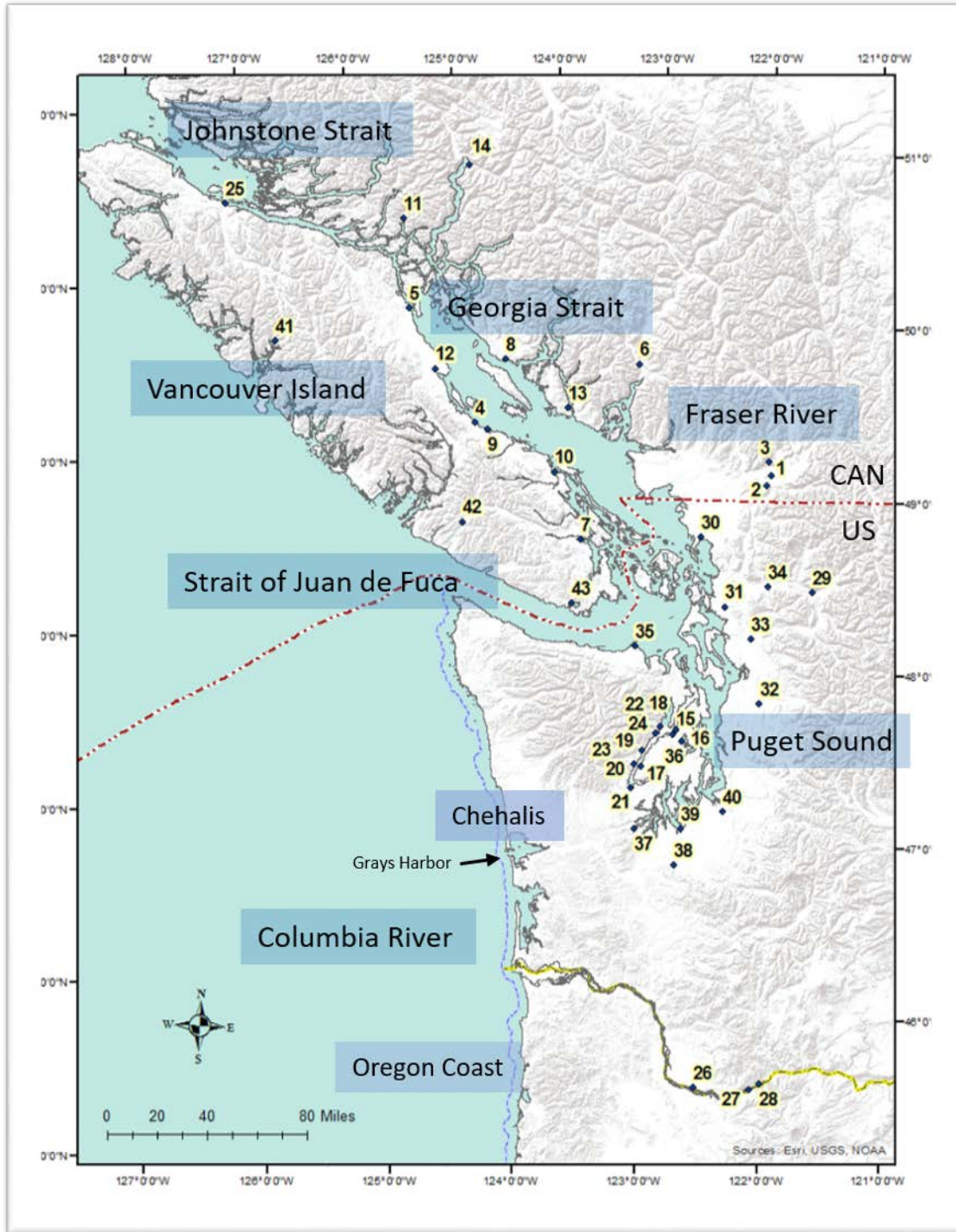


Figure 2. Map showing Chum salmon regions represented in the full WDFW Chum SNP baseline. Numbers on the map are locations of some of the individual sample locations but are not coordinated with collection names.





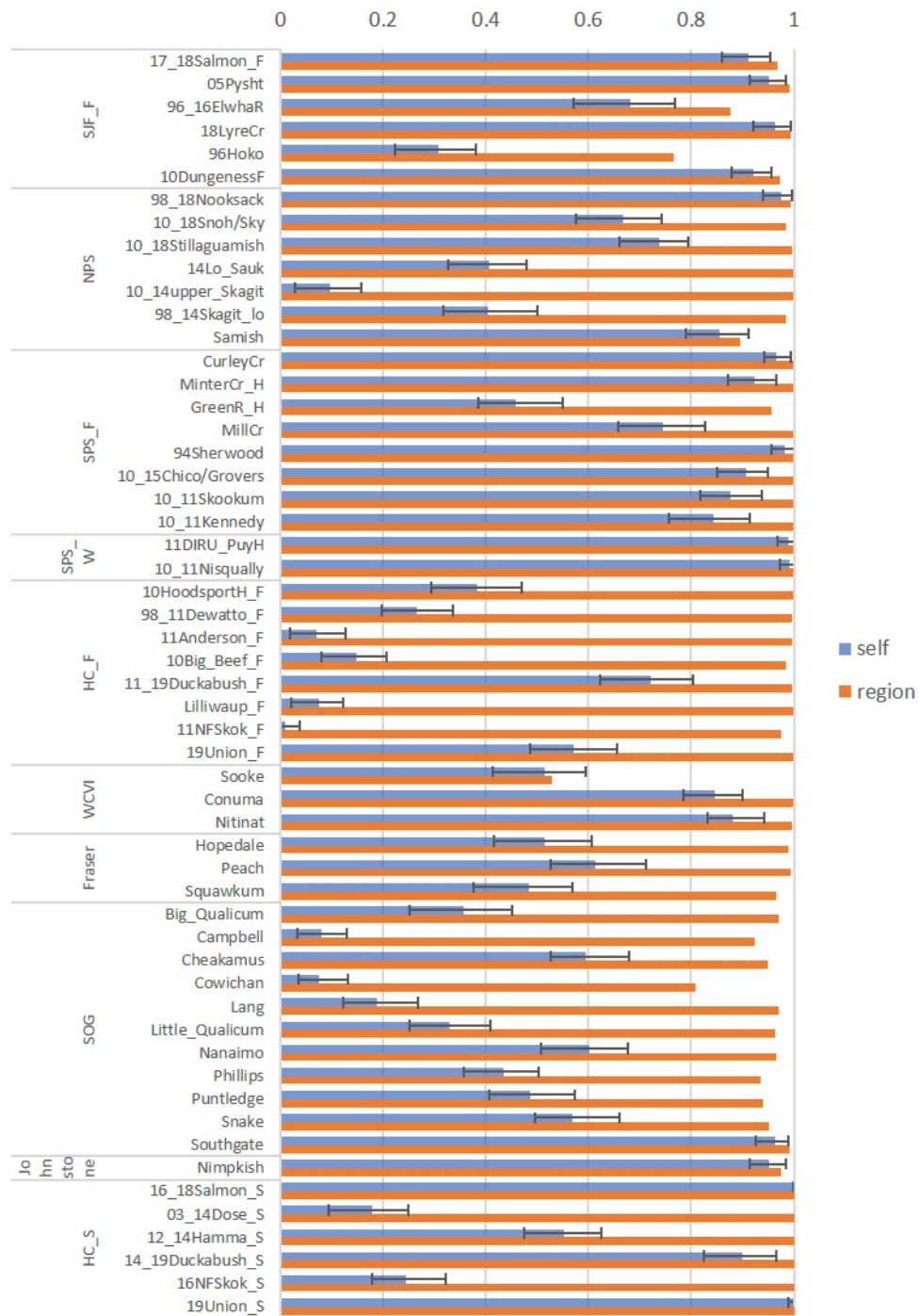


Figure 4. Plot of the percentage of assignments back to collection of origin (self) and the 95% confidence interval, and assignments to regional group. Numerical data is in Table 4.

WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

**Table 1. List of Chum salmon collections from the US and CA in the Chum SNP genetic baseline (WDFW baseline). Collections that were funded under this current grant by the Southern Fund are indicated by “SF” in the funding column. Assignments were calculated to individual collections and to regional groups in ONCOR. The abbreviations are as follows: NPS = North Puget Sound, CenPS = Central Puget Sound, SPS = South Puget Sound, SPS\_W = South Puget Sound winter run, HC = Hood Canal, SJF = Strait of Juan de Fuca, L\_Col = Lower Columbia River, SOG = Strait of Georgia, VCI\_W = West Coast Vancouver Island, \_S = summer run, and \_F = fall run.**

Funding	US/region	collection	N	CA/region	collection	N
SF	NPS_F	98Nooksack	60	Fraser	Hopedale	39
	NPS_F	10Nooksack	32	Fraser	Peach	43
SF	NPS_F	18Nooksack	84	Fraser	Squawkum	32
SF	NPS_F	18Snoh/Skykomish	53	Johnstone	Nimpkish	41
	NPS_F	10Snohomish	34	SOG	Big Qualicum	48
SF	NPS_F	12Snohomish	55	SOG	Campbell	32
SF	NPS_F	12Stillaguamish	64	SOG	Cheakamus	42
	NPS_F	10Stillaguamish	39	SOG	Cowichan	25
	NPS_F	18Stillaguamish	52	SOG	Lang	33
SF	NPS_F	14Lo_Sauk	67	SOG	Little Qualicum	46
SF	NPS_F	98Skagit_lo	48	SOG	Nanaimo	50
	NPS_F	14Skagit_lo_main	29	SOG	Phillips	44
	NPS_F	10_14upper_Skagit	45	SOG	Puntledge	48
SF	NPS_F	98Samish	73	SOG	Snake	44
	SPS_F	02CurleyCr	95	SOG	Southgate R	48
	SPS_F	07GreenR_H	47	VCI_W	Sooke	32
	SPS_F	03MinterCr_H	91	VCI_W	Conuma	44
	SPS_F	11MillCr (byShelton)	66	VCI_W	Nitinat	48
	SPS_S	Sherwood_S	76			
	SPS_F	Sherwood_F	94			
	SPS_F	10_15Chico/Grovers	93			
	SPS_F	10_11Skookum	94			
	SPS_F	10_11Kennedy	95			
SF	SPS_W	11DIRU_PuyH_W	94			
	SPS_W	10_11Nisqually_W	95			
	HC_F	19Union_F	107			
SF	HC_F	10HoodsportHat_F	88			
	HC_F	11NFSkokomish_F	30			
SF	HC_F	98Dewatto_F	47			
	HC_F	11Dewatto_F	47			
	HC_F	11Anderson_F	48			
	HC_F	10Big Beef_F	47			
	HC_F	19Duckabush_F	76			
SF	HC_F	11Duckabush_F	46			
	HC_F	11Lilliwaup_F	47			
	SJF_F	17_18Salmon_F	106			
SF	SJF_F	05Pysht	93			

WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

Funding	US/region	collection	N	CA/region	collection	N
SF	SJF_F	96Elwha	45			
	SJF_F	16Elwha	42			
SF	SJF_F	18LyreCr	94			
SF	SJF_F	96Hoko	43			
	SJF_F	10DungenessF	101			
	SJF_S	16_18Salmon_S	164			
	HC_S	14Duckabush_S	47			
	HC_S	19Duckabush_S	221			
	HC_S	19Union_S	186			
	SF	HC_S	03Dosewallips_S	46		
SF	HC_S	12_14Dosewallips_S	36			
	HC_S	12_14Hamma_S	58			
	HC_S	16NFSkokomish_S	41			
SF	Chehalis	00Willapa	70			
	Chehalis	17Satsop	44			
	Chehalis	17Wishkah	48			
	Chehalis	18Cloquallum	29			
	Chehalis	18Hoquiam	45			
	Chehalis	18Humptulips	32			
	Chehalis	18Wynoochee	37			
	L_Col	16_I205	91			
	L_Col	St Cloud	39			
	L_Col	Multnomah	46			
SF	L_Col	17Ives	100			
	L_Col	17Ham/Hardy	82			
	L_Col	17Duncan H	79			
	L_Col	17LewisH	76			
	WA_Coast	17GraysH	93			
	WA_Coast	01QuinaltH	66			

WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

**Table 2. List of Chum salmon samples genotyped under this Southern Fund grant for this project in 2019 and 2020. The total number of samples genotyped is under “N to run” column and the total number of samples that had at least 80% genotypic data that were added to the baseline are under the “>80%” column. WDFW genotyped 95 Chum salmon samples from DFO (internal funds) to allow a direct comparison of genotypes from GT-seq and Ion Torrent platforms.**

		region	Collection	N to run	Code	>80%
run in 2019	under SF-2018-SP-22	HoodCanal_F	Dewatto_F	48	98JR	47
	under SF-2018-SP-22	HoodCanal_S	NF Skokomish_S	42	16LA	41
	under SF-2018-SP-22	HoodCanal_F	Hoodsport Hat_F	95	10LH	88
	under SF-2018-SP-22	HoodCanal_S	Dosewallips_S	48	03FF	46
	under SF-2018-SP-22	NorthPS	Nooksack	65	98LC	60
	under SF-2018-SP-22	NorthPS	Lower_Sauk	50	14DC	43
	under SF-2018-SP-22	NorthPS	Skagit_lower_mainstem	48	98HF	48
	under SF-2018-SP-22	NorthPS	Snohomish/Skykomish	62	18MH	53
	under SF-2018-SP-22	NorthPS	Stillaguamish	57	18MI	52
	under SF-2018-SP-22	SJF	Dungeness (Beebe Cr)	39	10LK	39
	under SF-2018-SP-22	SJF	Elwha River	53	96EH	45
	under SF-2018-SP-22	SJF	Lyre Cr	51	18JR	51
	under SF-2018-SP-22	SJF	Pysht	95	05MT	93
	under SF-2018-SP-22	SJF	Hoko	43	96EJ	41
	under SF-2018-SP-22	SJF	Hoko	2	95GM	2
	under SF-2018-SP-22	SouthPS_W	DIRU_Puyallup H	52	11KR	51
run in 2020	under SF-2018-SP-22	NorthPS	Nooksack Hatchery	94	18MM	84
	under SF-2018-SP-22	NorthPS	Stillaguamish	73	12QE	64
	under SF-2018-SP-22	NorthPS	Snohomish	73	12QD	55
		For comparison	DFO Chum	95	20GS	60



WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

Table 4. Percentage of simulated samples in the 100% simulation that assigned back to its home collection (self) and to its region and run group (region/run), followed by the 95% confidence interval for the estimation. Temporal collections were grouped before assignment test. Data are plotted in Figure 5.

Region/run	self	95% CI		region	95% CI		
SJF_F	17_18Salmon_F	0.9109	0.8602	0.9546	0.9671	0.9381	0.9915
	05Pysht	0.9519	0.9140	0.9843	0.9905	0.9725	1.0000
	96_16ElwhaR	0.6816	0.5723	0.7678	0.8771	0.7856	0.9420
	18LyreCr	0.9637	0.9201	0.9944	0.9936	0.9792	1.0000
	96Hoko	0.3086	0.2230	0.3808	0.7666	0.6794	0.8483
NPS_F	10DungenessF	0.9201	0.8785	0.9570	0.9724	0.9403	0.9953
	98_18Nooksack	0.9751	0.9410	0.9964	0.9936	0.9776	1.0000
	10_18Snoh/Sky	0.6671	0.5763	0.7415	0.9856	0.9621	0.9997
	10_18Stillaguamish	0.7376	0.6596	0.7945	0.9968	0.9869	1.0000
	14Lo_Sauk	0.4079	0.3270	0.4792	0.9978	0.9880	1.0000
SPS_F	10_14upper_Skagit	0.0972	0.0277	0.1577	0.9976	0.9887	1.0000
	98_14Skagit_lo	0.4042	0.3170	0.5017	0.9842	0.9583	0.9989
	Samish	0.8563	0.7887	0.9125	0.8959	0.8276	0.9467
	CurleyCr	0.9663	0.9416	0.9933	0.9984	0.9893	1.0000
	MinterCr_H	0.9225	0.8730	0.9668	0.9996	0.9954	1.0000
SPS_W	GreenR_H	0.4583	0.3856	0.5493	0.9574	0.9152	0.9875
	MillCr	0.7448	0.6577	0.8262	0.9992	0.9937	1.0000
	94Sherwood	0.9811	0.9565	0.9999	0.9996	0.9955	1.0000
	10_15Chico/Grovers	0.9063	0.8496	0.9494	0.9991	0.9918	1.0000
	10_11Skookum	0.8760	0.8188	0.9368	0.9996	0.9948	1.0000
HC_F	10_11Kennedy	0.8437	0.7556	0.9134	0.9996	0.9950	1.0000
	11DIRU_PuyH	0.9882	0.9686	1.0000	0.9978	0.9900	1.0000
	10_11Nisqually	0.9917	0.9729	1.0000	0.9993	0.9950	1.0000
	10Hoodsporth_F	0.3833	0.2949	0.4695	0.9983	0.9921	1.0000
	98_11Dewatto_F	0.2661	0.1971	0.3352	0.9961	0.9831	1.0000
WCVI	11Anderson_F	0.0708	0.0191	0.1271	0.9973	0.9897	1.0000
	10Big_Beef_F	0.1474	0.0793	0.2067	0.9841	0.9540	1.0000
	11_19Duckabush_F	0.7224	0.6235	0.8041	0.9963	0.9843	1.0000
	Lilliwaup_F	0.0758	0.0226	0.1229	0.9988	0.9919	1.0000
	11NFSkok_F	0.0100	0.0000	0.0374	0.9752	0.9512	0.9956
Fraser	19Union_F	0.5705	0.4866	0.6555	0.9985	0.9887	1.0000
	Sooke	0.5153	0.4143	0.5950	0.5282	0.4422	0.5958
	Conuma	0.8461	0.7853	0.8992	0.9988	0.9940	1.0000
	Nitinat	0.8810	0.8326	0.9432	0.9973	0.9867	1.0000
	Hopedale	0.5154	0.4150	0.6069	0.9886	0.9729	0.9996
SOG	Peach	0.6131	0.5254	0.7116	0.9932	0.9766	1.0000
	Squawkum	0.4846	0.3766	0.5680	0.9649	0.9282	0.9890
	Big_Qualicum	0.3584	0.2525	0.4516	0.9700	0.9217	0.9964
	Campbell	0.0799	0.0335	0.1306	0.9244	0.8702	0.9681
	Cheakamus	0.5947	0.5260	0.6783	0.9496	0.9068	0.9811
Johnstone	Cowichan	0.0762	0.0364	0.1324	0.8079	0.7358	0.8664
	Lang	0.1888	0.1215	0.2674	0.9707	0.9337	0.9950
	Little_Qualicum	0.3296	0.2522	0.4089	0.9633	0.9174	0.9941
	Nanaimo	0.6028	0.5070	0.6758	0.9667	0.9288	0.9900
	Phillips	0.4352	0.3578	0.5035	0.9358	0.8905	0.9733
HC_S	Puntledge	0.4872	0.4059	0.5740	0.9407	0.8936	0.9770
	Snake	0.5681	0.4963	0.6601	0.9508	0.9141	0.9868
	Southgate	0.9635	0.9268	0.9901	0.9922	0.9771	1.0000
	Nimpkish	0.9518	0.9150	0.9849	0.9753	0.9428	0.9953
	16_18Salmon_S	0.9999	0.9997	1.0000	1	1.0000	1.0000
Johnstone	03_14Dose_S	0.1779	0.0944	0.2484	1	1.0000	1.0000
	12_14Hamma_S	0.5524	0.4738	0.6258	1	1.0000	1.0000
	14_19Duckabush_S	0.9008	0.8253	0.9652	1	1.0000	1.0000
	16NFSkok_S	0.2449	0.1787	0.3224	1	1.0000	1.0000
	19Union_S	0.9980	0.9892	1.0000	1	1.0000	1.0000

Table 5. Averages of assignment back to home collection (self) and to region and run group (region) for each region and run group.

	avg self	avg region
SJF_F	0.7895	0.9279
NPS	0.5922	0.9788
SPS_F	0.8374	0.9941
SPS_W	0.9900	0.9986
HC_F	0.2808	0.9931
WCVI	0.7475	0.8414
Fraser	0.5377	0.9822
SOG	0.4259	0.9429
Johnstone	0.9518	0.9753
HC_S	0.6457	1.0000

WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

Table 6. Simulated mixed stock fishery analyses from ONCOR. Under “actual value” is the proportion of the fishery from selected stocks to be simulated and under “avg” is the average proportions of stocks assigned to the fishery and the standard deviation (st dev) for the value. The assignment proportions to regional run groups for each simulated mixed fishery is at the bottom of the table. The Hood Canal summers were excluded from the table because no fish assigned to them in any analysis.

Region/run	collection	ACTUAL			ACTUAL			ACTUAL			ACTUAL			ACTUAL					
		VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV			
SJF_F	17_18Salmon_F	0	0.0003	0.0013	0	0.0005	0.0015	0	0.0002	0.0011	0	0	0.0003	0	0.0004	0.0017	0	0.0003	0.0014
SJF_F	05Pysht	0	0.0005	0.0017	0	0.0006	0.0019	0	0.0003	0.0014	0	0	0	0	0	0	0	0	0.0004
SJF_F	96_16ElwhaR	0	0.0025	0.0048	0	0.0018	0.0042	0	0.0014	0.0030	0	0	0.0001	0	0	0.0003	0	0.0002	0.0008
SJF_F	18LyreCr	0	0.0003	0.0011	0	0.0003	0.0011	0	0.0002	0.0010	0	0	0	0	0	0	0	0	0
SJF_F	96Hoko	0	0.0071	0.0084	0	0.0019	0.004	0	0.0030	0.0044	0	0	0	0	0	0.0002	0	0	0
SJF_F	10DungenessF	0	0.0013	0.0029	0	0.0008	0.0021	0	0.0010	0.0023	0	0	0	0	0.0001	0.0007	0	0	0
NPS	98_18Nooksack	0	0.0113	0.0118	0.5	0.5076	0.0400	0.5	0.4989	0.0370	0	0	0	0	0	0	0	0	0
NPS	10_18Snoh/Sky	0	0.1245	0.0350	0	0.1137	0.0356	0	0.0060	0.0074	0	0	0.0003	0	0	0	0	0.0001	0.0004
NPS	10_18Stillaguamish	0	0.1221	0.0346	0	0.1322	0.0333	0	0.0068	0.0087	0	0	0.0004	0	0.0001	0.0006	0	0.0001	0.0004
NPS	14Lo_Sauk	0	0.0230	0.0180	0	0.0213	0.0155	0	0.0022	0.0044	0	0	0	0	0	0	0	0.0001	0.0005
NPS	10_14upper_Skagit	0	0.0192	0.0158	0	0.0161	0.0139	0	0.0004	0.0013	0	0	0	0	0	0	0	0	0
NPS	98_14Skagit_lo	0.5	0.2262	0.0416	0.5	0.1960	0.0372	0	0.0080	0.0109	0	0	0	0	0	0	0	0	0.0002
NPS	Samish	0	0.0021	0.0038	0	0.0021	0.0040	0	0.0005	0.0015	0	0	0	0	0.0001	0.0006	0	0	0.0002
SPS_F	CurleyCr	0	0	0	0	0	0	0	0	0	0	0.0002	0.001	0	0.0155	0.0123	0	0.0136	0.0114
SPS_F	MinterCr_H	0	0	0	0	0	0	0	0	0	0	0.0001	0.0005	0	0.0026	0.0044	0	0.0025	0.0045
SPS_F	GreenR_H	0	0	0	0	0	0	0	0	0	0	0.0002	0.0009	0.5	0.2308	0.0337	0.5	0.2139	0.0301
SPS_F	MillCr	0	0	0	0	0	0	0	0	0	0	0	0.0002	0	0.0028	0.0041	0	0.0023	0.0039
SPS_F	94Sherwood	0	0	0	0	0	0	0	0	0	0	0.0001	0.0006	0	0.0144	0.0126	0	0.0137	0.0124
SPS_F	10_15Chico/Grovers	0	0	0	0	0	0	0	0	0	0	0.0002	0.0008	0	0.1974	0.036	0	0.1800	0.0323
SPS_F	10_11Skookum	0	0	0	0	0	0	0	0	0	0	0.0001	0.0006	0	0.0077	0.0089	0	0.0070	0.0076
SPS_F	10_11Kennedy	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0046	0.0071	0	0.0050	0.0066
SPS_W	11DIRU_PuyH	0	0	0	0	0	0	0	0	0	0.5	0.4917	0.0376	0.5	0.4949	0.0326	0	0.0005	0.0016
SPS_W	10_11Nisqually	0	0	0	0	0	0	0	0	0	0.5	0.5069	0.0375	0	0.0055	0.0067	0	0.0006	0.0018
HC_F	10HoodsportH_F	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0052	0.0062	0.5	0.2066	0.0376
HC_F	98_11Dewatto_F	0	0	0	0	0	0	0	0	0	0	0.0001	0.0005	0	0.0037	0.0051	0	0.1555	0.0376
HC_F	11Anderson_F	0	0	0.0001	0	0	0	0	0	0	0	0	0	0	0.0013	0.0037	0	0.0140	0.0141
HC_F	10Big_Beef_F	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0019	0.0037	0	0.0153	0.0140
HC_F	11_19Duckabush_F	0	0	0	0	0	0	0	0	0	0	0.0001	0.0004	0	0.0053	0.0077	0	0.0471	0.0245
HC_F	Lilliwaup_F	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0007	0.0019	0	0.0339	0.0217
HC_F	11NFSkok_F	0	0	0.0005	0	0.0001	0.0004	0	0	0	0	0.0002	0.0008	0	0.0018	0.0036	0	0.0023	0.0044
HC_F	19Union_F	0	0	0	0	0	0	0	0	0	0	0.0002	0.0008	0	0.0033	0.0056	0	0.0854	0.0304

WDFW Final Report to Southern Fund – Bilateral Chum salmon SNP baseline, 2020

Region/run	collection	ACTUAL			ACTUAL			ACTUAL			ACTUAL			ACTUAL					
		VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV	VALUE	AVG	ST DEV			
WCVI	Sooke	0	0.0002	0.0008	0	0.0002	0.0008	0	0.0001	0.0005	0	0	0	0	0	0	0	0	
WCVI	Conuma	0	0	0.0004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
WCVI	Nitinat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Fraser	Hopedale	0	0.0653	0.0249	0	0.0001	0.0004	0	0.0677	0.0249	0	0	0	0	0	0	0	0	
Fraser	Peach	0	0.1569	0.029	0	0.0002	0.0014	0	0.1638	0.0326	0	0	0	0	0	0	0	0	
Fraser	Squawkum	0.5	0.2310	0.0352	0	0.0002	0.0009	0.5	0.2330	0.0335	0	0	0	0	0	0	0	0	
SOG	Big_Qualicum	0	0.0005	0.0015	0	0.0004	0.0013	0	0.0003	0.0014	0	0	0	0	0	0	0	0	
SOG	Campbell	0	0.0004	0.0012	0	0.0005	0.0017	0	0.0010	0.0023	0	0	0	0	0	0	0	0	
SOG	Cheakamus	0	0.0003	0.0012	0	0.0006	0.0019	0	0.0002	0.0008	0	0	0	0	0	0	0	0	
SOG	Cowichan	0	0.0006	0.0019	0	0.0002	0.0008	0	0.0002	0.0009	0	0	0	0	0	0	0	0	
SOG	Lang	0	0.0005	0.0016	0	0.0001	0.0006	0	0.0003	0.0011	0	0	0	0	0	0	0	0	
SOG	Little_Qualicum	0	0.0010	0.0023	0	0.0009	0.0025	0	0.0009	0.0023	0	0	0	0	0	0	0	0	
SOG	Nanaimo	0	0.0006	0.0021	0	0.0004	0.0012	0	0.0011	0.0033	0	0	0	0	0	0	0	0	
SOG	Phillips	0	0.0005	0.0017	0	0.0003	0.0014	0	0.0005	0.0016	0	0	0	0	0	0	0	0	
SOG	Puntledge	0	0.0011	0.0026	0	0.0005	0.0016	0	0.0016	0.0032	0	0	0	0	0	0	0	0	
SOG	Snake	0	0.0003	0.0015	0	0.0002	0.0009	0	0.0004	0.0013	0	0	0	0	0	0	0	0	
SOG	Southgate	0	0.0001	0.0008	0	0.0001	0.0008	0	0	0	0	0	0	0	0	0	0	0	
Johnstone	Nimpkish	0	0.0001	0.0008	0	0	0	0	0	0.0002	0	0	0	0	0	0	0	0	
GROUPS																			
	SJF_F	0	0.0121	0.0099	0	0.0059	0.0072	0	0.0062	0.0055	0	0	0.0003	0	0.0006	0.0019	0	0.0005	0.0019
	NPS	0.5	0.5282	0.0407	1.0	0.9891	0.0085	0.5	0.5228	0.0374	0	0.0001	0.0005	0	0.0002	0.0008	0	0.0003	0.0008
	SPS_F	0	0	0	0	0	0	0	0	0	0	0.0008	0.0020	0.5	0.4757	0.0319	0.5	0.4380	0.0343
	SPS_W	0	0	0	0	0	0	0	0	0	1.0	0.9986	0.0026	0.5	0.5003	0.0333	0	0.0011	0.0025
	HC_F	0	0.0001	0.0005	0	0.0001	0.0004	0	0	0	0	0.0004	0.0013	0	0.0232	0.0133	0.5	0.5601	0.0344
	WCVI	0	0.0002	0.0009	0	0.0002	0.0008	0	0.0001	0.0005	0	0	0	0	0	0	0	0	0
	Fraser	0.5	0.4533	0.0392	0	0.0005	0.0016	0.5	0.4645	0.0371	0	0	0	0	0	0	0	0	0
	SOG	0	0.0061	0.0055	0	0.0042	0.0051	0	0.0065	0.0064	0	0	0	0	0	0	0	0	0
	Johnstone	0	0.0001	0.0008	0	0	0	0	0	0.0002	0	0	0	0	0	0	0	0	0