Columbia River Inter-Tribal Fish Commission
729 NE Oregon, Suite 200
Portland, OR 97232
503.238.0667
www.critfc.org

CRITFC

TECHNICAL REPORT 10-12

2009 Annual Report Genetic Assessment of Columbia River Stocks

Shawn Narum, Nate Campbell, Andrew Matala, and Jon Hess

February 16, 2010

2009 Annual Report

GENETIC ASSESSMENT OF COLUMBIA RIVER STOCKS

Prepared by:

Shawn Narum
Nate Campbell
Andrew Matala
Jon Hess

Columbia River Inter-Tribal Fish Commission
Hagerman Fish Culture Experiment Station

Prepared for:
U.S. Department of Energy
Bonneville Power Administration
Division of Fish and Wildlife
P.O. Box 3621
Portland, OR 97283-3621
Project Number: 2008-907-00
Contract Number: 41224
Reporting Period: 3/1/09-3/31/10
Submitted: February 16, 2010

ABSTRACT
This project combines four inter-related studies from the Fish & Wildlife Program Accords that address these current and future objectives: 1) discover and evaluate SNP markers in salmon and steelhead; 2) expand and create genetic baselines for multiple species (Chinook, steelhead, sockeye, and coho); 3) implement Genetic Stock Identification (GSI) programs for mainstem Chinook fisheries and 4) GSI of fish passing Bonneville Dam (steelhead and Chinook). These four projects are highly related since SNP markers are needed to complete species specific baselines, and these baselines are requisite to complete GSI. In the first year of this project, SNP discovery goals (Objective 1) were achieved with successful development of 10 new assays for steelhead, 30 new assays for coho salmon, and 21 new assays for sockeye salmon. These newly discovered SNP markers were combined with existing SNP markers to generate genetic baselines and for two applications of genetic stock identification (GSI). For genetic baseline expansion (Objective 2), we successfully genotyped 96 SNP markers in 52 Chinook salmon and 54 steelhead populations from the Columbia River Basin. Results from population genetics analyses suggest SNPs are a class of markers that perform well for distinguishing populations, and these baselines will be useful for estimating stock composition in GSI applications. Results also indicated that some loci may be candidate markers and valuable for landscape genetic analyses (based on selective divergence). The first year of the project included two broad applications of GSI, stock composition of Chinook salmon fisheries (Objective 3), and stock composition of Chinook salmon and steelhead passing Bonneville Dam (Objective 4). Results of Objective 3 indicate that spring-run Chinook salmon harvested in commercial, sport, and test fisheries were primarily composed of three adipose-clipped stocks (in descending order of stock composition): Rapid River Hatchery/Clearwater R., Upper Columbia R. (i.e., Carson stock), and Mid-Columbia R and these Chinook salmon stocks were also the most strongly represented at Bonneville Dam. During the spring Chinook test and sport fisheries, a fourth stock, Williamette R., was found because these fisheries include harvests spanning an earlier part of the season and locations closer to the mouth of the Columbia R. For fall Chinook fisheries, the sport fishery at Buoy10 had predominantly Lower Columbia fall stocks (>60% composition), and less than 20% composition of the following stocks (in descending order): Snake R. fall, upper Columbia R. summer/fall, and Deschutes R. fall. The entire Zone 6 tribal Chinook fishery was heavily comprised of Upper Columbia R. summer/fall stock (60-80% depending on region), but Region 1 (closest region to Bonneville Dam) of Zone 6 contained more Lower Columbia R. fall stock (~30%) than Region 2 (< 5%), whereas Snake R. fall stock was similar in both regions (12-15%). Weekly composition of Chinook salmon stocks were highly variable and differed among the two regions in Zone 6. Results of Objective 4 indicate that four "major" stocks of steelhead were sampled as they migrated past Bonneville Dam: upper Columbia R. (0.135 ± 0.013), middle Columbia R./lower Snake R. (0.254 ± 0.050), upper Clearwater R (0.320 ± 0.069), and upper Salmon R. (0.154 ± 0.037). These four steelhead stocks varied considerably in peak run timing (weeks 24-25, 28-29, 38-39, and 32-33, respectively), and clear transitions occurred when each stock of steelhead was most abundant in the mainstem Columbia R.

Table of Contents

List of Figures

List of Tables

Acknowledgements

The authors of this report would like to thank the following individuals and organizations for providing time and expertise toward this project: The Bonneville Power Administration for providing funding for this research project and to its staff for their assistance. Tissue samples were collected and provided by William Ardren, Bill Arnsberg, Bobby Begay, Steve Boe, Chris Brun, Peter Cleary, Roger Dick Jr., Roger Dick Sr., Dani Evenson, Chris Kern, James Kiona, Paul Kuneki, Jacinda Mainord, Alfred McConnville, Jim Nagler, Winfred Perez, Terra Schultz, Jason Vogel, Devin Volenec, John Whiteaker, Marc Whitman, Bill Young, and Joe Zendt. Collection of samples from the commercial, sport, and test fishery were coordinated with Joe Hymer. We are grateful for extensive contributions in the genetics laboratory from Stephanie Harmon, Vanessa Jacobson, Amanda Matala, Lori Maxwell, Megan Moore, and Jeff Stephenson.

Introduction

Microsatellites have been the molecular marker of choice in ecological and conservation genetics studies in the last decade due to their high variability and power to resolve population structure. However, complicated mutation models (i.e., two-phase mutation model), high incidence of homoplasy, high potential error rate and low genotyping throughput have led researchers to consider alternative marker types. With increasing genomic information available for non-model organisms like salmon, single nucleotide polymorphisms (SNPs) have begun to see increased use as genetic markers for population genetic studies. These polymorphisms represent the most abundant variation in the genome of most organisms, and are spread throughout the entire genome at high density. Further, mutation rates, mutation models and error rates for SNPs are generally well understood, providing a foundation for estimating parameters that are important in ecological and conservation genetics. Both microsatellite and SNP markers have been shown to be useful for differentiating Chinook salmon populations (e.g., Narum et al. 2008). However, SNP markers are particularly well suited for applications of Genetic Stock Identification (GSI) due benefits of high-throughput and low error rates. Thus, this project is focused on implementing SNP markers for differentiating salmon and steelhead populations and GSI of unknown mixture samples from the mainstem Columbia River.

This project combines four inter-related studies from the Fish & Wildlife Program Accords that address these current and future objectives: 1) discover and evaluate SNP markers in salmon and steelhead; 2) expand and create genetic baselines for multiple species (Chinook, steelhead, sockeye, and coho); 3) implement Genetic Stock Identification (GSI) programs for mainstem Chinook fisheries and 4) GSI of fish passing Bonneville Dam (steelhead and Chinook). These four projects are highly related since SNP markers are needed to complete species specific baselines, and these baselines are requisite to complete GSI. The results of these four objectives address needs for distinguishing specific stocks, determining genetic diversity, stock specific run timing, and estimating stock composition to provide information for fisheries management and harvest.

Objective 1) SNP Discovery

One of the highest priorities in the full-scale implementation of SNPs for salmon genetics is the discovery and development of a sufficient number of these markers to characterize population variability. These polymorphisms represent the most abundant variation in the genome of most organisms, and are spread throughout the entire genome at high density (Morin et al. 2004). Thus SNPs can be discovered through sequencing known regions of DNA and converted to high throughput assays (e.g., Campbell and Narum 2008a). Further, mutation rates, mutation models and error rates for SNPs are generally well understood, providing a foundation for estimating genetic divergence between populations. SNP markers also offer the potential of a more cost-effective and less error-prone alternative to existing genetic tools that may be used independently or in tandem with existing microsatellite markers to improve accuracy and precision of stock assignments. The combined power of these two marker types is expected to improve

stock composition accuracy (Narum et al. 2008a) and allow researchers to meet rigorous stock composition and assessment needs for timely management of fisheries.

Objective 2) Baseline Expansion

Currently, genetic baselines of microsatellite markers are in place for Chinook salmon across the coastwide range (Seeb et al. 2007) and steelhead in the interior Columbia River Basin. Despite large, representative sample sizes from many populations and very high microsatellite allelic diversity, the resolution of specific stocks and populations in these baselines is limited in some cases. For example, fall Chinook salmon in the Columbia River are closely related and remain difficult to distinguish even with a powerful set of 13 microsatellite markers. Several other closely related populations in the Chinook salmon baseline are similarly difficult to distinguish and thus have been pooled into a single reporting unit for GSI applications. In some cases (i.e., mainstem Columbia R. Chinook fisheries), a finer level of stock discrimination is necessary for management of fisheries. Additional SNP loci will increase stock assignment reliability where greater resolution is required. Given the difficulty and expense of inter-laboratory standardization, additional microsatellite markers may not the most efficient choice. In this regard, SNP markers are the preferred option for additional loci since they offer many beneficial characteristics that make them amenable to adding loci to existing baselines.

Objectives 3 &4) Genetic Stock Identification

Genetic Stock Identification (GSI) methods have proven to be effective in determining the proportion of stock origin in several mixed stock applications (Shaklee et al. 1999, Beacham et al. 2006, Narum et al. 2008b). This proposal includes two GSI projects that will utilize genetic baselines: 1) GSI to Evaluate Catch; and 2) GSI of fish passing Bonneville Dam.

This study will include GSI analysis of Chinook salmon collected from commercial, recreational, and tribal fisheries in the Columbia River. (Subsequent years of the study will include steelhead and coho fisheries as possible.) Implementation of GSI technology could make monitoring individual production units in mixed stock areas possible. Tissues will be sampled annually from fisheries with existing programs in place with Washington Department of Fish and Wildlife (WDFW) and Oregon Department of Fish and Wildlife (ODFW). We plan to genotype representative samples from fisheries of primary interest. The GSI estimates may help refine CWT based estimates of stock composition used in fishery management.

The second application of GSI analysis in this proposal includes sampling unknown origin salmon and steelhead at Bonneville Dam for genetic analysis. Samples will be collected over the entire length of the run on a weekly basis, and genetic baselines will be utilized to determine the stock composition of these runs. Few studies have been able to determine the extent of overlap among life history types of salmon and steelhead, but GSI of each life history type will allow us to determine the stock composition of the different runs through Bonneville Dam with greater accuracy than current methods. Population genetic methods and statistical assignment models have advanced dramatically in recent years, and estimating stock composition is now possible using either Bayesian or

Maximum Likelihood methods (Anderson et al. 2008). Therefore, we plan to estimate stock composition of multiple species passing Bonneville Dam and provide this information on a timely basis to fisheries managers.

Report Structure

This report is divided into four sections, one for each of the objectives of the study. The first section reports on SNP discovery efforts and the second section on genotyping SNP markers in Chinook salmon and steelhead to create genetic baselines. The third section contains GSI results from Chinook salmon sampled in mainstem fisheries and adults migrating over Bonneville Dam in 2009. The fourth section includes GSI results of adult steelhead migrating over Bonneville Dam over multiple years.

Section 1: SNP Discovery

Introduction

Conservation genetics projects such as genetic stock identification (GSI), reproductive success, and selection studies require the use of an informative suite of genetic markers. Although many types of markers have been used historically, single nucleotide polymorphisms (SNPs) are the most common type of variation in the genome, genotyping assays can be designed for a number of high throughput platforms, and the resulting genetic information can be easily exchanged between collaborating laboratories (Melton, 2003). This section describes the sequencing and identification of SNPs in three species of Pacific salmonids (*O. mykiss*, *O. nerka*, and *O. kisutch*), and the development of genotyping assays. Target goals for this objective were to produce 10 informative assays for *O. mykiss* and 15 each for *O. kisutch* and *O. nerka*.

Methods

DNA sequences from *O. mykiss* and *O. tshawytscha* were collected from online database entries in TIGR and NCBI for primer design. Primers were designed using either the embedded primer design tool on NCBI's database or the primer3 program (http://frodo.wi.mit.edu/primer3/) to amplify a PCR product of 400 to 800 base pairs in length (Table 1). All primer pairs (N = 107) were tested for amplification of a PCR product suitable for sequencing using 2uL of extract DNA from 2 individuals per species. PCR conditions for testing were 1x Qiagen multiplex master mix, 250nM primers, and 1-2uL of extract DNA in a 10uL reaction. PCR were done in 96-well PCR plates and thermal cycling conditions were [95ºC for 15 min; (95ºC for 30 sec, 60ºC for 30 sec, 72ºC for 45 sec, repeated 40x); 4ºC hold]. Amplified products separated on a 1% agarose gel and stained with Sybr-green. Products were visualized using a UV light source and images were captured on a CCD camera. Primer pairs that produced suitable products were used to amplify products from an ascertainment panel of 32 individuals per species. Individuals chosen for ascertainment panel were from across each species' range but with the greatest representation within the Columbia River drainage (Table 2).

For amplification of the ascertainment samples, primers were multiplexed such that each reaction produced four products. PCR and cycling conditions remained the same as in the test. PCR products were cleaned by treatment with Exonuclease 1 and Shrimp alkaline phosphatase (New England Biolabs) to remove unincorporated dNTPs and primers. Sequencing reactions were performed in one direction using the Big Dye v.1.1 sequencing kit (Applied Biosystems) and MJ research and Applied Biosystems thermal cyclers using 2uL of each product. The reactions were cleaned up by ethanol-EDTA-NaAcetate precipitation prior to separation by capillary electrophoresis using a 3730 instrument (Applied Biosystems). Chromatogram data was analyzed by using the Sequencing Analysis program v.5.4 (Applied Biosystems) to assign base calls followed by Sequencher v.4.7 (Gene Codes) to align and edit the data. Observed SNPs and small indels were recorded and minor allele frequencies calculated for each.

Suitable SNP sites were used to design Taqman™ assays by submitting to Applied Biosystems via their FILE BUILDER program. Once received, the assays were used for genotyping using the Fluidigm EP-1 96.96 system using the manufacturers suggested protocol, but modified slightly by including a sample pre-amplification step and increasing the PCR cycles to 50. The assays were validated by comparison of genotyping data to sequencing data in the ascertainment samples. Assays were further evaluated by genotyping one or more populations within the Columbia River drainage to examine allele frequencies, deviations from Hardy-Weinberg expectations, and linkage disequilibrium using the program GENEPOP (Raymond and Rousset 1995).

Results

For *O. mykiss*, 22 loci were amplified and sequenced, producing 17 sequence alignments totaling 8,037 unique consensus bases with an average contig length of approximately 470 base pairs. 127 total SNP sites were identified and the number of SNPs per locus ranged from 1 to 17. Based on minor allele frequency and distribution in the ascertainment panel, 15 SNP sites were chosen for assay design. The 15 assays were used to genotype the ascertainment samples as well as 188 additional samples representing 7 regions (lower, middle, and upper Columbia R., S.E. Alaska, Snake R., McCloud hatchery strain, and a mixture of steelhead taken from Bonneville Dam 2004). Thirteen of the assays were validated by comparison of genotyping data to sequencing data and 10 of the validated assays showed a high occurrence of the minor allele in Columbia River populations. The 10 validated and putatively informative assays are listed in Table 3. The genotypes showed no significant linkage with one another and no significant deviations from Hardy-Weinberg expectations were observed.

In *O. kisutch*, 83 loci were amplified and sequenced producing 53 sequence alignments totaling 20,784 unique consensus bases with an average contig length of approximately 392 base pairs. 149 total SNP sites were identified with several loci revealing no variations within the contig. The number of observed SNPs per base pair was less than half that of the *O. mykiss* panel. SNP sites were chosen for assays design based on minor allele frequency and distribution within the panel. Thirty-five SNP sites were selected and submitted to Applied Biosystems for assay design. Once received, the assays were used to genotype the 32 ascertainment samples as well as 31 adult samples from Cowlitz hatchery (1998) and 31 samples from the Wenatchee River (2008) using the Fluidigm EP-1 96.96 system. Wenatchee River samples were poor quality and required that they be excluded in order to evaluate the assays. The ascertainment panel genotypes were compared to sequencing data to validate the assays. Thirty-two of the 35 designed assays passed validation and their utility was further evaluated. Twenty-three of the validated assays had easily scorable genotyping plots and minor allele frequencies of over 10% between the Cowlitz samples and the ascertainment panel. Two assays showed heterozygote excess within the Cowlitz collection; assays Oki_afp4-10 ($p = 0.0118$) and Oki_spf30-119 ($p = 0.000$). Significant linkage was detected between 2 sets of assays [Oki_gdh-189; Oki_hsc713-56 ($p = 0.0002$] and [Oki_hsc713-56; Oki_hsc71p-31 ($p = 0.00778$)]. The latter two assays are from opposite ends of the same gene but alleles are not in complete phase. Thirty validated assays with easily interpreted genotyping plots

are listed in Table 3. Further testing of these assays on other Columbia River collections is ongoing.

For the *O. nerka* ascertainment panel, 80 loci were amplified and sequenced producing 57 sequence alignments totaling 21,647 unique bases with an average contig length of 380 bases. 93 total SNP sites were identified with numerous loci containing no variation across the ascertainment panel. The number of observed SNP sites per consensus base screened was 3.7 times lower than that observed in the *O. mykiss* panel. Twenty-nine SNP sites were chosen for assay design based on minor allele frequency and distribution within the ascertainment panel. The new assays were used to genotype the ascertainment samples as well as 100 samples from Suttle Lake, WA (Kokanee), 31 sockeye samples collected at Tumwater Dam (Wenatchee R.), and 31 sockeye samples collected at Wells dam (Columbia R. below confluence of the Okanogan and Methow rivers). Twenty-three of the assays were validated by comparison of the ascertainment sample genotypes with sequencing data. Two of the assays were validated against the sequencing data but were difficult to score. The remaining 21 assays are listed in Table 3. Two of the assays, One_tshB-92 and One_taf12-248, had minor allele frequencies below 10% in all 3 of the collections tested. A single validated assay showed a possible heterozygote excess in the Suttle lake collection (assay One_cin-177; $p = 0.0163$). All other assays showed no deviations from Hardy-Weinberg expectations in any of the 3 collections. No significant linkage was observed between any of the assays tested.

Discussion

SNPs were detected at a much lower rate in the *O. kisutch* and *O. nerka* panels than in the *O. mykiss* panel (2.2 and 3.7 times lower respectively), and much lower than previous SNP discovery efforts in other salmonid species (Campbell and Narum 2008; Campbell et al. 2009). Since similar strategies were used in choosing samples for each panel and care was taken to include samples from distant populations, the disparity in the occurrence of SNPs suggests that the species-wide genetic diversity is lower within *O. kisutch* and *O. nerka* versus *O. mykiss*. If this is the case, then screening for informative SNP sites will be more difficult in *O. nerka* and *O. kisutch* regardless of the ascertainment samples chosen. Future genetic screening in these species therefore might be more suited for next generation pyrosequencing (i.e., Wiedmann et al. 2008) than the targeted gene approach employed here. However, that approach has drawbacks also, such as a higher rate of false discovery, lower conversion rate from SNP to assay, and the majority of SNPs are from unknown regions of the genome (Sanchez et al. 2009).

Initial analysis of the validated assays showed that for each species the target number of informative assays was achieved. However, further testing will be required to determine which assays are the most effective for differentiation of Columbia River stocks and allow us to compare them to assays developed by others. Ultimately, these assays will be used to generate genotypes from a larger set of collections and rank the most informative markers with descriptive statistics such as F*st* and locus ranking programs (e.g., BELS; Bromaghin 2008). Some of the assays that showed very low minor allele frequencies in the populations tested to this point may show much higher frequencies in other

populations and prove to be the most informative. Similarly, some of the least informative assays for the Columbia River may work well in other parts of the species' range. All functional assays will be made available by publication in peer reviewed articles. These newly discovered SNP markers were combined with existing SNP markers to generate genetic baselines and for two applications of genetic stock identification (GSI). These markers may also be useful for other applications such as pedigree studies for estimating reproductive success, and evaluating adaptive divergence of populations to specific environments.

References

Bromaghin, J.F. 2008. BELS: backward elimination locus selection for studies of mixture composition or individual assignment. Molecular Ecology Resources 8:568-571.

Campbell, N. R., and S. R. Narum. 2008. 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., K. Overturf, and S. Narum. 2009. Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9:318-322.

Melton, L. 2003. Pharmacogenetics and Genotyping: On the trail of SNPs. Nature 422;917-923.

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

Sanchez, C.C., Smith, T.P., Wiedmann R.T., Vallejo, R.L., Salem, M., Yao, J., and C.E. Rexroad III. 2009. Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. BMC Genomics 10;559.

Wiedmann, R.T., Smith, T.P., and D.J. Nonneman. 2008. SNP discovery in swine by reduced representation and high throughput pyrosequencing. BMC Genetics 9;81.

Table 1: List of targets and primer sequences used for PCR amplification and sequencing.

	Target name:	Abbreviation:	Acc#:	Forward	Reverse
1	Na-K ATPase alpha3	NaKATPa3	AY319388	AAAAAGAACATCGGCGACAT	AAACCCACCTACAGCCCTTT
2	acidic ribosomal protein	arp	EU682504	CGAGGGAAGACAGGACCAC	AGCAGCATGTCCCTGATCTC
3	Growth Hormone	gh	EU682503	AGTCCTGAAGCTGCTCCATA	TTCCTCCTGACGTTTCCGTC
4	hsc71 promoter	hsc71pro	S85730	TTTCTATGCTTCCGGGTCAC	TTACCCCTCTAGGCCCAACT
5	hsc71 3' UTR	hsc71-3p	S85731	ACCAACCATTGAGGAAGTCG	ATGCAGTGGCAGTCCTTGTA
6	hsf1b 3' UTR	hsf1b-3p	AB062549	AGCAGCTGGTCCAGTACACCTC	CCATGGGCTGAATAAACCATGC
7	hsf2 3' UTR	hsf2-3p	AJ488177	AACCCAGAAGGTGACACCAC	CAAAGCATGCATTTTCTTGA
8	hsp27b (hspb1)	hsp27b	AB255361	GCGCTTGCTAGGAACTGAAA	TACGTGGGTACAGCTTTCCA
9	hsp47 3' UTR	hsp47-3p	AB196463	ACCCCAAGAACACATTCCAA	GGTTGGCAAATGGCATAGAT
10	hsp70a promoter	hsp70aPro	AB062281	TTTTCATAATGGCCGTGACA	AACAACCAGACCTGGGTTCA
11	hsp90 beta A	hsp90BA	AB196457	CCCTCAAACTCACTCCAACCGT	GGTCTGAACTCAGGCAAAGG
12	MAP kinase 4 promoter	MK4pro	EU069828	TTGGTTCCCCTTCACTCATC	CCTTGGGCTGTCCTTTCTTA
13	MAP kinase 4 3' UTR	MK4-3p	EU069829	CATGGACCACAACACTCCTG	GCATTCTCAGACAGCCACAA
14	Aldolase B (fructose 1,6 bis-phosphate aldolase)	aldB	TC18597 (TIGR)	AAACGACCACCACGTCTACC	GGGAATGTCAACATGGAAGG
15	Aspartate aminotransferase	aspat	TC9251 (TIGR)	TTCTGATCCGGCCAATCTAC	ACCTTTGACATGGTGGCTTC
16	cyclooxygenase 1 (cox1)	cox1	AJ299018	CCCATCTGTTCCCCTGAGTA	CGAAACCACACTGCTTTCAA
17	glutamic acid dehydrogenase (gdh)	gdh	AJ556997	TGCCTGTAGTCATTCCATCG	GCCCTAAGGATTCAGCACAA
18	myostatin 1a	myo1a	AF273035	TCATGGAGGTGACGATTTCA	TCTATTGCACCGTGTTCTGC
19	myostatin 1b	myo1b	AF273036	CTTTGGCTGGGACTGGATTA	AGCGAATGGACAGGTGTTTC
20	myoD	myoD	Z46924	CCATGACCCCATCTACCAAG	AAGCGCCTCAGGCTACATTA
21	natural killer efficiency factor (nkef)	nkef	AF250195	TCCAAGCAGCAGTAAGACGA	GGAGTTTTCGCCACATGTTT
22	natrual resistance-associated macrophage protein (nramp)	nramp	AF048760	AACGACTTCCTCAACGTGCT	ACCACTCCTCCAACGTGTTC
23	Ogo4 (uSat)	Ogo4	AF009796	GTCGTCACTGGCATCAGCTA	TCCAGGGTATTGTTCTAGACTGG
24	Ots208 (uSat)	Ots208	AF393187	CCAAGTGCAAGATCTGACCTC	CCCATGTTTCATGCTCACAC
25	Ots249 (uSat)	Ots249	AF393192	CTCGCTTGGTTATGGAGGAG	TGCTGACCTGTGAGTCCAAG
26	Ots474 (uSat)	Ots474	AF393200	AAGGGCTTTGAGGATCAGAA	AAACGGAGAGGCGAATATGA
27	peptidylarginine deiminase	pad	FM999001	CTGGGACTCAGCTGCACATA	TGGATTACCTCCAGTGGCTC
28	succinate-CoA ligase	suCoA	FJ607869	AGAGAAGATTGCCGCCCTAC	AGCCAGTGAGAACAGACATTTT
29	nucleoside diphosphate kinase	ndk	FJ607868	AGGAGACACCAACCCATCTG	AAATGACAGGTCAAGGGCAG
30	thymidine kinase	thyK	FJ607866	GGTGCAGGAAGTGTTACGGT	TGATGAGGCAAATCAGACCA

31	C type lectin receptor B	clrB	FJ607865	ACTGGAGATGGGTGGACAAC	GGTGACTGATTGGCAGCTTT
32	branched-chain alpha-keto acid lipoamide acyltransferase	bcAKala	NM_001124203	AGACCACCGGATCATCGAC	TGCAATGGTGCTTTCCATAG
33	carbonic anhydrase 1	carban1	NM_001124220	CAGGATCACGGCAGCTTAAT	GGTTGATTCTTCCTTGGCTG
34	dehydrodolichyl diphosphate synthase	dds	FM212438	TCTGGAATCTGTGTGATGCC	TAGCCCCATAGGGGTAAAAT
35	NILT4 leukocyte receptor precursor	NILT4	FM180058	TGGTTACTATGCCACCGTCA	GGGGCTGATTAGCATGTTGT
36	signal sequence receptor delta	ssrd	FJ591155	ATCCAGCCCCTTTTCTCTGT	CATTGGGATCGATTGAGATGT
37	programmed cell death 6	pcd6	FJ591154	AACTGTGCCAAAGCCAAATC	CAGTTTAAGGCATCATGCACA
38	GHS-R gene for growth hormone secretagogue receptor	GHS-R	AB479381	AAAGGCACAGAATTGGAGACA	TCGGAGGAATGGGACATAAG
39	brain creatine kinase b	bckB	FJ548753	TTCTCTGAGGTGGAGCTGGT	CAATGGGGCAAAGAGACTTTTA
40	glucocorticoid receptor	gluR	EU084718	CACCGGCTCTTTTCTTTCTG	TGGATGACAGCTTTGCACAT
41	gadd45 beta ii	gadd45b2	EU084728	GTTACAGGACGTGGGCAACT	CACGTCGGATAGCATAGGGT
42	Tumor protein p53	p53	NM_001124692	TTCGGTTCCATGTCATTTCA	CAAGCATCAACAGTTCACCG
43	metallothionein B gene	metB	M22487	TTTTATCGATGATCAACGTGGT	TTTAACGTGCCACCAAGTCA
44	metallothionein A gene	metA	BV725496	GGATCCTTGTGAATGCTCCA	GCTGGTATCACAAGTCTTGCC
45	defender against cell death 1	dacd1	FJ849061	ATGTGTTGGAGCCTTCATCC	TCAAGTGGTTTGCAACAATCA
46	Vesicle-trafficking protein SEC22b-B	SEC22b	BT074367	AGAGGAGATAGGGGAGGGGT	TGTTATTAACGGCAAAGCCC
47	Ictacalcin	icta	BT074357	AAACACTGACCAGGCAAAGG	ACCAAAGAGCTTCCTCTCCA
48	SUMO-1-conjugating enzyme UBC9	SUMO1	BT074321	CGAGCACAAGCCAAAAAGTT	TTGAATGGGTTCATTGAGGC
49	OtsG311	OtsG311	AF393194	GTTCTTCCCTTGATGCAGGA	CCCCATCCATTGTTCTATCG
50	insulin-like growth factor II	igf2	NM_001124697	GGACAACTACGTCAGCCACA	CCGCTAAGGATCCACCTAAA
51	thyrotropin beta-subunit	tshb	NM_001124543	GATAACGCTCTCCCTTCCCT	TCCCATTCATGATATTGGTTCA
52	Ntl T-box protein	ntl	GQ241688	CTCCTCAACCACCATCCTGT	ACATAAGGATTGGCCCCTTT
53	Pax7 gene	pax7	FJ713022	CCCATGGTAATGTGGGTAGC	TGATAACGTCTGCTTGCTGG
54	archain 1	arcn1	FM993912	CACCTGCAAGAAGGAATGGT	TTTCTTTGAATGACCTCCCG
55	adenine nucleotide translocator 2	ant2	NM_001160491	TGGTACCATTGACTGCTGGA	CATACCCACTGGAGGACAGG
56	perforin	prf1	NM_001134847	GAGATTTGAGGTGTGGGACG	GATCTGGGGTTTTGACATGG
57	atrial natriuretic peptide	anp	NM_001124211	ATTTCAGCCTGTTTTGGTGG	AGTCTTGCACATCACCAACG
58	Dolichyl-diphosphooligosaccharide-	dad1	NM_001160576	ATATCATGTGTTGGCGCCTT	TCGGATTGCAACAATCATACA
59	Peptidyl-prolyl cis-trans isomerase E	ppie	NM_001160524	TGCAGAATTGTTAAAAGGGCA	TGGGATCTTTTATTTCAGCTCAT
60	Type-4 ice-structuring protein precursor (afp4)	afp4	FN396363	GCTGATGGACCAGACCAAGT	TTGCATGTCAACAACATTCG
61	apolipoprotein A-II	apoa2	NM_001161448	TGGGTCTGTACGGACACTATG	ATACATGCACAATGTTGAGTGAA
62	Reticulon-3	rtn3	NM_001160516	CAGTCCTACACAGCGTCGAA	AGTGCCCCAGAAACATCAAC
63	Ribonuclease P/MRP protein subunit POP5	pop5	NM_001160634	CAGTACTGAGCTGCACCCTG	CACATGTCCAATGTTTTATTTGC

64	Mitogen-activated protein-binding protein-interacting protein	mapip	NM_001160633	AGACTGTGGGCTTTGGAATG	GCGTGACAGGAACAGACAAC
65	Transcription initiation factor TFIID subunit 12	taf12	NM_001160631	TTGTTTCACCCACGTCTTCA	CAGTCTGGCGATACAAGCAA
66	Crystallin J1A	crj1a	NM_001160629	TGAAGGACAACTTATCCAAGCA	GCAGCTTTGACAGAGCAACA
67	CA050 protein	ca050	NM_001160628	GCACAAACCTTTAGGCAAGC	ACAATCATTTTGCAGGTCCC
68	Eukaryotic translation initiation factor 5	if5	NM_001160627	CCAAAGAGATCCATGCGAAG	TGCAGTCAAGTAAATCCCCA
69	Vesicle-associated membrane protein 5	vamp5	NM_001160626	TGTCAGAGCAAAGGACATGG	TCCTATGGTAAGGCTGAAGCA
70	ADAM 10	ada10	NM_001160624	CGATCCCTCCCACATTTACA	AAAGCTTGCTGGATTTTTGG
71	Isopentenyl-diphosphate delta-isomerase 1	idi1	NM_001160619	CTAGGGGCCCTGGTAAAAAG	AAACGCAACAAAAATTTGGC
72	Signal recognition particle 9 kDa protein	srp09	NM_001160617	GGCTTATGGTGTCCAAGGAA	GGAGGCATGTTTCTTCCACT
73	methionine sulfoxide reductase	msra	NM_001160616	CCGGAGTCTCCTGTCCAATA	GTCAACACACCGTTTCATCA
74	SYS1 homolog	sys1	NM_001160611	CTGGCTCATTCCCTGTGAAC	GGAAAGGACATAACCGAAAGC
75	RNA-binding protein 4B	rbm4b	NM_001160608	GTAATGCATGACCCCTCGTC	AAACAAGCACGTTACCCACC
76	Prefoldin subunit 4	pfd4	NM_001160606	AGGAACAACTGCAGCAGGAA	CCATTGCTGAGAAAAATCATGT
77	Normal mucosa of esophagus-specific gene 1	nmes1	NM_001160604	GAGGGAGTGGACCCATCTAA	TGTTCTGTTGGTGGCATTGT
78	Poly synthetase 3	parp3	NM_001160600	TGCAGACACCTCCAGATGAG	GCATTTGTGGAGTGTTTATTGAA
79	Zinc finger protein 706	zn706	NM_001160599	GGGGGAATATTTGTGTGACG	TACAGGATTTGGCCTCGTTT
80	NipSnap2	nips2	NM_001160597	GACCTCGCCTCTGAAGTAACC	CACGATCCTTAATAAAGCACTGGA
81	S-acyl fatty acid synthase thioesterase	sast	NM_001160595	GCATGGAAAGACATATCATCTGG	TTTGATTTGAGGACTAGTCAGGC
82	Anaphase-promoting complex subunit 7	apc7	NM_001160593	AGCTGTTCTCACCATCCCAG	CCATTTTTCAAAACGTGTGTGT
83	Gastrula zinc finger protein XLCGF57.1	zg57	NM_001160590	GTGGCGACTGTGGGAAAAG	CACAGGCTTTACAAAAGTAAATGAA
84	26 proteasome complex subunit DSS1	dss1	NM_001160587	AAGAATTCCCAGCCGAGG	CCGTGAACTGAACCACCTTT
85	Phosphatidylinositol N-acetylglucosaminyltransferase subunit H	pigh	NM_001160586	GCATACCGACGCTTCAAACT	TCGAGCTTATCAAGTAGGGACA
86	NTF2-related export protein 2	nxt2	NM_001160583	ATGTGACCCCCAGAACTGAA	TTCACCAACAAACAAACGGA
87	Inositol monophosphatase	impa1	NM_001160581	ATGGTGGACCATTGGATCTC	CATACATGACTGGAGACACATATACA
88	Molybdenum cofactor synthesis protein cinnamon	cin	NM_001160580	CAGAAGACTTGAGAGGAGGGC	TCAGTCACGGAGCAGTTGAT
89	Guanine nucleotide-binding protein alpha-11	gna11	NM_001160579	TTGTTACACGGTTGGCTCTG	TGCTGATGAAAGTGAGGCAC
90	DNA-directed RNA polymerase II subunit J	rpo2j	NM_001160575	TGGTGCATAATGTTGGGAAA	AAGCATTTAAGCGGAATGGA
91	Checkpoint protein HUS1	hus1	NM_001160574	AATTGGCCATCCCTCCTACT	CATGAAGGCTGAGCGTGTAG
92	Hepatitis B virus X-interacting protein	xip	NM_001160573	GATTCGGGGAATATTCTAGTGAGA	TTCTACATGGATTTGAAAAATCAGTC
93	S-methyl-5-thioadenosine phosphorylase	mtap	NM_001160570	ACTATCATCCCACAGCCACC	ATGCCCTAAGCACATTCCAC
94	Betaine aldehyde dehydrogenase	badh	NM_001160550	TGAACAACGGGAAGTCAAGG	ACTGCTGCTTGAAATGTGAGG
95	Survival of motor neuron-related-splicing factor 30	spf30	NM_001160549	CCCCTGTAATGGATCAGCAC	TACGTCGCCAGCAGTGATAC
96	Inosine triphosphate pyrophosphatase	itpa	NM_001160547	CTCTGACCAATCGCTCAACA	CTTGGGGACTCCACACACTC
97	Vacuolar ATP synthase subunit F	vatf	NM_001160545	AGGGGGACTTTGGAGACTGT	TCACATGCATGCTCTGACAA

98	Thioredoxin Reducatse Interacting Protein	txnip	*TC145083*	TTTGAGAACACCTGCTCACG	TCAGAGTAGGCAGGCAGGTT
99	Thioredoxin Reductase	trdnx	CA057296	TTTATTGAACGCACCCACAC	ACGTGGTTTTCCTCAAGGTG
100	Glutathione Peroxidase	gshpx	AF281338	ACGAGCTCCATTCGCAGTAT	TCCTTAATATCTGCCTCAATGTCA
101	Glutathione S-transferase A	gsta	BT073173.1	CCCAAACTGGGAGCGTACTA	CACCAGCTATGTGGGTCTGT
102	calreticulin	crt	AY372389.1	GGGAAGAGGAAACTCCCAAG	TGCTTGACACTGAAGGGATG
103	DNA Damage Inducible Factor 4	redd-1	DQ400410	TAGCAGGGGGTCAAGTATGG	TGATTTGCATTTGAAGCCAG
104	Target of rapamycin	tTOR	EU179853.1	TGCATCAGGACCTCTTCTCA	AAGGACCAGGGTCTTGTGG
105	Asparagine Synthetase	asns	Not Submitted	TCCAGAACCAAAGAGGCGTA	GAATGTGGTGAAGGGGAGGT
106	Leptin	lpl	AB354909.1	CTCCACTATGAGGGGTCTGC	GGAGATGGTGACAGTGGGAT
107	liver X receptor	lxr	FJ470291.1	AGGACCATGAACTGGTGGAG	CTGCCAAACCACACAGAAGA

Table 2: Ascertainment panel for each species screened for SNPs.

	O. mykiss		***O. kisutch***		***O. nerka***	
Sample	Population	Region	Population	Region	Population	Region
1	Touchet R.	Washington	Sandy R.	Oregon	Birkenhead	Canada (BC)
2	Touchet R.	Washington	Sandy R.	Oregon	Birkenhead	Canada (BC)
3	Sopachnaya R.	Kamchatka,	Sandy R.	Oregon	Ualik Lake	AK - Bristol
4	Sopachnaya R.	Kamchatka,	Sandy R.	Oregon	Ualik Lake	AK - Bristol
5	Omak Cr.	Washington	Babine R.	British	Hapiza R. early	Kamchatka,
6	Omak Cr.	Washington	Babine R.	British	Hapiza R. early	Kamchatka,
7	Satus R.	Washington	Egegik R.	Alaska	McDonald	SE Alaska
8	Satus R.	Washington	Egegik R.	Alaska	McDonald	SE Alaska
9	Yakima R.	Washington	Clackamas	Oregon	Saltery Lake	Kodiak Island,
10	Yakima R.	Washington	Clackamas	Oregon	Saltery Lake	Kodiak Island,
11	Omak Cr.	Washington	Clackamas	Oregon	Suttle Lake	Oregon-
12	Omak Cr.	Washington	Clackamas	Oregon	Suttle Lake	Oregon-
13	Sashin Cr.	Alaska	Soos Cr.	Washington	Suttle Lake	Oregon-
14	Sashin Cr.	Alaska	Soos Cr.	Washington	Suttle Lake	Oregon-
15	Sashin Cr.	Alaska	Soos Cr.	Washington	Redfish Lake	Idaho
16	Shitike	Oregon	Soos Cr.	Washington	Redfish Lake	Idaho
17	Shitike	Oregon	Smith R.	California	Fishhook Cr.	Idaho Kokanee
18	Shitike	Oregon	Smith R.	California	Fishhook Cr.	Idaho Kokanee
19	Spring Cr.	Kamloops	Theodore R.	Alaska	Okanogan	WA, Kokanee
20	Spring Cr.	Kamloops	Theodore R.	Alaska	Okanogan	WA, Kokanee
21	Upper Summit	Washington	Cowlitz	Washington	Okanogan	WA, Kokanee
22	Lower Summit	Washington	Cowlitz	Washington	Okanogan	WA, Kokanee
23	Goldendale Hat.	N. California	Cowlitz	Washington	Okanogan	Washington
24	Goldendale Hat.	N. California	Cowlitz	Washington	Okanogan	Washington
25	Upper Malad R.	Idaho, Redband	Lewis R.	Washington	Okanogan	Washington
26	Upper Malad R.	Idaho, Redband	Lewis R.	Washington	Okanogan	Washington
27	Pahsimeroi	Idaho	Lewis R.	Washington	Lake	Washington
28	Pahsimeroi	Idaho	Lewis R.	Washington	Lake	Washington
29	Dworshak	Idaho	Queets R.	Coastal Wash.	Wells Dam	Washington
30	Dworshak	Idaho	Queets R.	Coastal Wash.	Wells Dam	Washington
31	Arlee Str.	Hatchery Str.	Wenachee	Col. River	Tumwater Dam	Washington
32	Donaldson Cr.	Hatchery Str.	Wenachee	Col. River	Tumwater Dam	Washington

Table 3: Validated Taqman™ assays designed from screen. Assay names contain a species identifier, a locus identifier, and the SNP position on the contig (i.e. One_aldB-152). All forward and reverse primers are unlabeled and probes contain a 5' fluorophore and a 3' quencher and minor groove binder.

Assay name	Primers	Probes
Omy_dacd1-131	GGCAGCCACCATTATTGTGAAATG	VIC-CCTGCCTACAGTCTG
	GGAAGTCTCCTTTGTTATCTGGATTGA	6FAM-CTGCCAACAGTCTG
Omy_gadd45-	AGAGAAGACTCACTGCTGTTTGC	VIC-TTGCTCCAAAATGG
	AAATCAGTTCCCACGCTATGCT	6FAM-TTGCTCCGAAATGG
Omy_GHSR-121	CTGTGTATAAGTTTATACAGTCAGCACAGT	VIC-CCTAATAACCATGATAACAGC
	TTCAGAGAGAGAAATGGCAGAAAGG	6FAM-AATAACCATGGTAACAGC
Omy_gluR-79	GACTGTCTATAGCTATTCTTCTCAAACTGT	VIC-CAAGTATTTTGCGTAGGAAT
	AGAAACTACCATTGTGATTAACAGATAGAAAATACAT	6FAM-CAAGTATTTTGCATAGGAAT
Omy_metA-161	CGCATGCACCAGTTGTAAGAAAG	VIC-CAAGTAAGTGGTTATATTCT
	AGTGCCACCAGCGATAAGAAAA	6FAM-CAAGTAAGTGGTTCTATTCT
Omy_metB-138	TCTGTCCCTGACGCTATAAAAACG	VIC-TTCGCCAAAGAGAAAT
	GAAGTATTTCAGCTTAATTTCACTGTTGAGTT	6FAM-TTCGCCAAAGTGAAAT
Omy_ndk-152	AAGAATTGAGGGATAAAAACAAAATAATATATAAACATGA	VIC-CACCCACTTTCAAAAC
	CAAACCTACATTCATTAAAGTCCAGTTTTGT	6FAM-ACCCACTCTCAAAAC
Omy_p53-262	CCCCAACATCCAGTATACAGTTTCA	VIC-CAAGTAGTATGGAGCTCTAT
	CCCAAATTGGCAATTTTAATAGGATTCAGA	6FAM-AAGTAGTATGGTGCTCTAT
Omy_pad-196	CAAACAACCACAGTAGTCCTCCAAT	VIC-AAGACAAAGGTGTAATACC
	GCTTTTCACCCTTTTGTAAATTAAGCCAAA	6FAM-AAGACAAAGGTATAATACC
Omy_SECC22b-	GGATCCCTCCTTTTAACACAAGACT	VIC-CTGTCTGTCCATATATC
	CTACAGGATGACTACCTAATTGCTAATAAAACA	6FAM-CTGTCTGTCCGTATATC
Oki_afp4-10	CTTATCACCTCGGACCGAATCAA	VIC-TTGTAAGCCCATAGAGCTG
	GCTCTTGTTGGATTGGTTTTGACT	6FAM-AAGCCCATGGAGCTG
Oki_arp-105	GCCATTGACAAAGCAGGTTAAGTTA	VIC-AAGAGCAGTACTGCACAAG
	ACAGGGATGGTAGGAAACTAGTTGA	6FAM-AAGAGCAGTACTACACAAG
Oki_aspAT-273	GAATCCACGCTGTCACAAAGTAAAA	VIC-CACCGTGCCCCACTG
	GGGTGGAGTGGGCATTATTATTTTG	6FAM-CCACCGTACCCCACTG
Oki_bcAKal-274	CCTTTCTTCCTCCGAACATGGAT	VIC-CCACAAGTTCCCTTACTAC
	TCAAACTAAGGCCACACACAAAGAT	6FAM-CACAAGTTCCCATACTAC
Oki_ca050-17	GCTCAGTAAATCTGGTCCCATTAGG	VIC-ATGTGAAACTTTAGTCATACAGAA
	GGGCAGTTTGAAGACCATACTCA	6FAM-TGTGAAACTTTAGTCCTACAGAA
Oki_carban-140	CATGCATACAAGCACACACACA	VIC-CTGTAACCTCGAATTAT
	GACAGAGCACAATAGCTAACAGTTTG	6FAM-CTGTAACCTCCAATTAT
Oki_gdh-189	CCTGTGTTGAAGTGGAGTAGGTT	VIC-CTGCTACAGAAAAAA
	GCTTTATACTGTAAGTGGACTGACCTT	6FAM-TTCTGCTACAGATAAAA
Oki_gh-183	TGTTTATGTGGTACTGGCTCAAAACT	VIC-CACATCAATACAATTTT
	GCATGCTTGGTAAGAGTGAATTTGT	6FAM-CATCAATGCAATTTT
Oki_gshpx-152	CTGCGAGCCCAAATGTAGTTTT	VIC-AGTTGTTGCAACATTGTAG
	AGTCCCAAAAGTATTATAAAGACAAGAGTTTAACA	6FAM-AGTTGTTGCAACCTTGTAG
Oki_hsc713-56	TTTGAAATGTGTTCTACCCTCTTGTGT	VIC-CAAGTCACGAGACTCA
	CTGTTGCGTAGCAAACAAAGTTG	6FAM-ACAAGTCACAAGACTCA
Oki_hsc71p-313	TGTGCCATGTAGTAGGTAGTTTGTAAC	VIC-AGTCATTGTAGTTAATAACC

	GGGCAGGCATAAAAGTAGATAAGCA	6FAM-AGTCATTGTAGTTCATAACC
Oki_hsf1b-85	CCCAGGACCAGGACCCT	VIC-AGATGGTAGGGTCCTCT
	GCTGTGTGGTAGTCGGTAGTG	6FAM-AGATGGTAGCGTCCTCT
Oki_hsp90B-83	GGGCGAATGAAGGTCATTAATAAAGT	VIC-ATTATATCAGCATACATTTC
	GATCACTTCGTCATCATCGATTCCT	6FAM-ATCAGCAAACATTTC
Oki_itpa-85	CCACACTCCGTCCTCCATAATT	VIC-TGTCTTAAATAAGTTGTTTTTG
	GACTCTGAACAGAGAATAAACCAAATCCT	6FAM-TGTCTTAAATAATTTGTTTTTG
Oki_mapK3p-93	CATTTCCTAAATATTTTCTCAACTTTCCTCATGTT	VIC-CTCCTGCCATCTCAGTC
	GGGTCACGTCACCACTGA	6FAM-CCTGCCACCTCAGTC
Oki_metA-220	AAGCAAGTAAGTGGTTCTATTCTAAATCCAA	VIC-TGGCACTTTAAAGCC
	GCCACCTAATGAGGCAATTGAGATT	6FAM-TGGCACTTTTAAAGCC
Oki_nips-159	CTGTGTCGGCATCAAATGTCTTATT	VIC-ACGCATGTCATTTTGA
	CCAAGTCTGGTTAGAGCCCTACTAT	6FAM-ACGCATGTAATTTTGA
Oki_p53-20	GGCCATTTTTTCATTGATAGGATTCAGAT	VIC-CCCAACATCAGATTTG
	GGGATTCATCATGTCTTACTCATTGGT	6FAM-ACCCAACATCTGATTTG
Oki_parp3-19	ACTTGTGTTGAATTCCAGAGACTAATGT	VIC-TCCTGGCTTTAAATTG
	GCTTCGTTATGATCTACCACCTTGA	6FAM-TGGCTTTAAAATTG
Oki_pigh-33	GCATACCAAGTCATGCAACATGTTT	VIC-CCATTGGGAAATGTTAAC
	AGACAAAAGTAGATAAAACAATAATAGGTGTATCATCAG	6FAM-CCATTGGGAATTGTTAAC
Oki_pop5-265	CGGACCTCGCAGGTTAAGAG	VIC-TTGAATACTCCTGGAATAGA
	ATTCACACTCAAACAGCTGCATTG	6FAM-TTGAATACTCCTTGAATAGA
Oki_rbm4b-129	CCAGAAGCCTTTTAGTCAGTCGTTT	VIC-CGTGTCATGATACAAAC
	GCTGCAGCTTCCATTCATCAAATTT	6FAM-AACGTGTCATACAAAC
Oki_rpo2j-235	GGAAATTGGTGATTTGAAACAGATTCTGA	VIC-TGATTCTGTTTTGAATTAACCAA
	CAGCACATACCTAACAAACTACAGTATTGA	6FAM-CTGTTTTGAATGAACCAA
Oki_SECC22-67	GGATCCCTCCTTTTAACACAAGACT	VIC-ATGGACAGACAGTTGGGA
	GGATGACTACCTAATTGCTAATAAAACATAACATG	6FAM-TGGACAGACCGTTGGGA
Oki_spf30-119	TGCATGTTTCAATGTTACACATTTATTATAGCTT	VIC-TGAACTGGGCTATATTAGACAT
	TGTTTTACTGTAAATGTACAAAACTGAAATGCT	6FAM-CTGGGCTATAGAAGACAT
Oki_srp09-107	CGGAGCTGGAATGACGACAT	VIC-TTTCATTTTTGATTTCTCCTCTC
	CAGCGAAGGTTATGCTCTCCAT	6FAM-CATTTTTGATTTCCCCTCTC
Oki_sys1-141	GCACACCTCTCAATAGATTCAAGGA	VIC-CATTTCAGTTAATTTCC
	TCGAGGAAAAAGTTAGCATTCACTGT	6FAM-CATTTCAGTTCATTTCC
Oki_taf12-40	GTCAGCTCCACTATCACACAAGTAT	VIC-TCTTGTCCTTTTTAATGCT
	GATTTTGGCCAGTTTGTGGTGAA	6FAM-CTTGTCCTTTTTTAATGCT
Oki_txnip-35	AATGGCAGGACCAAGGTCTTC	VIC-CAGAAGCTCTCCTCTGTG
	TCGCACATGCCGGAGAT	6FAM-AGAAGCTCTCGTCTGTG
Oki_vatf-363	AATCTTACGCAGGTACTGTATCTTGTG	VIC-TCTCGGCTAAGAATTG
	ACTCAATGAATTGCAGATGACTATCCA	6FAM-TCGGCAAAGAATTG
One_aldB-152	CGATCAGGTGACGCTAAAATTAACTC	VIC-CTCAGGCATTACCTTC
	GTGGCTTCCTCTTCACTCTGA	6FAM-CAGGCATCACCTTC
One_cin-177	CCTCAGACTAGTGACCGTACCTA	VIC-TCACGCACGGGACAG
	CGCTCACCGTGGTTACGT	6FAM-CACGCACGGAACAG
One_dds-529	CATAATGCTCCCCATCTTGAATTGG	VIC-AGCAATCCCATCTCTC
	CACTCAGCCCTTTAGGGAAGA	6FAM-AGCAACCCCATCTCTC
One_gdh-212	CCTGTGTTGAAGTGGAGTAGGTTAA	VIC-ATCTGTTACCAGAATGTTT
	GCTTTATACTGTAAGTGGACTGACCTT	6FAM-ATCTGTTACCATAATGTTT
One_metA-253	TTCTTATCGCTGGTGGCACTTT	VIC-AGGCAATTGAGGTTAAT
	GACCAAAGACTATTTAGTTGCCACCTA	6FAM-AGGCAATTGACGTTAAT

One_Mkpro-129	TGACGTATGTGCAATGCATGTCTAT	VIC-ATGCATATACATGTAATATAT
	AGATGAAGGACATGGCTGAAAACAT	6FAM-TGCATATACATGTAACATAT
One_Ots208-234	CAGCCGACATGCATCAGTTA	VIC-CACACGTTACATCAGATAACT
	TGACCCCATGTTTCATGCT	6FAM-CACACAATGTTACATCAGATAAC
One_pax7-248	AGTAAAGGTAGTGATGCAATGATGCA	VIC-AATTCAAAACGAAATGTG
	AACCGCATAGGACGTAAAGCA	6FAM-TGAATTCAAAACTAAATGTG
One_redd1-414rd	GTTGGCTACATCCTAAAACACAATGG	VIC-CCTAAGTCAGTCACTGTAG
	CAGCCCTGGAGTACTGAATCAG	6FAM-CCCAAGTCAGTCACTGTA
One_rpo2j-261	GATTCTGAGATCATACAGTGGATTGGT	VIC-CACATGTTTTACTCATTTGA
	GCTTGTCATCTTTCAGCACATACCTA	6FAM-CACATGTTTTACTAATTTGA
One_sast-211	TGTACTTAGTCCAATAAGCATTTCAACAGT	VIC-CATCATTTGCATTATTG
	TGGCTAGATTCACATGGTCAACAAA	6FAM-CATCATTTGAATTATTG
One_spf30-207	AGCATTTCAGTTTTGTACATTTACAGTAAAACA	VIC-AGGGACATCTTACCTCAAAA
	ACCTACTCGTAATTTCAGGGCAAAA	6FAM-AGGGACATCTTACCTAAAAA
One_srp09-127	CGGAGCTGGAATGACGACAT	VIC-CAGCGAAGGATATGCT
	AGGTTCAGCAAATCCCTCTTTAGAG	6FAM-CAGCGAAGGTTATGCT
One_ssrd-135	TGGAAACTCCTAGTGTACTTCATTCTCA	VIC-CTGCGGCTTTGTCTTG
	CGTTCCACGCTCCCTAGAATAGA	6FAM-TGCGGCTTTTGTCTTG
One_SUMO1-6	GCACAAGCCAAAAAGTTTTCTCCAT	VIC-CAAGATTGAAATTGGTTTGC
	GGACATAGTTGGAGGCAGACAAAA	6FAM-CAAGATTGAAATTTGTTTGC
One_sys1-230	CTACCTGTCTAACAGTGAATGCTAACTT	VIC-CAAAGCAAGTGATATATTAGTG
	TGAAACCATTAAGCTCTTTGTAGGACAA	6FAM-AAAGCAAGTGATATCTTAGTG
One_taf12-248	ACCTTCAATATGGTGGTGGTTACC	VIC-CCAGACAAAATCAAATTA
	ACTAAACGCACAACAGCAAACG	6FAM-CCAGACAAAATAAAATTA
One_tshB-92	GCATTGTCGTACTCGTGTGTTTG	VIC-ACCACCCTGTAGCTCA
	CACAACAGCAACAATACATGTCACA	6FAM-CACCCTGGAGCTCA
One_txnip-401	GCCAGATCCCTTCAGTTGGA	VIC-TGACTGCACTAGTTTAGAC
	GGCCATTTCAAAAGGCTGCAT	6FAM-TGACTGCACTAATTTAGAC
One_vamp5-255	GGTTGACTTTTCTTAACTTTTTAATCTGTGATATTGT	VIC-TAGGCTCCGTGCTCAGT
	GCTGAGCTAGTGATGGTACCATTT	6FAM-TAGGCTCCGTACTCAGT
One_vatf-214	TCATTCCTTTGCCTGGAGCATT	VIC-TGGTATTACTGTGCATTGAC
	GGCATACAGCAAAACAATTCAACCA	6FAM-ATGGTATTACTGTTCATTGAC

Section 2: Genetic Baseline Expansion

Introduction

Reproductively distinct aggregations of Chinook salmon (*Oncorhynchus tshawytscha*) and steelhead trout (*Oncorhynchus mykiss*) have evolved among naturally reproducing populations through the cumulative effects of selection and genetic drift (Waples 1991; Nielsen et al. 2009). Philopatry (homing to spawn) among these salmonid species is well documented (Hasler and Scholz 1983; McIssac and Quinn 1988; Quinn *et al.* 1991), and is a significant force shaping populations and defining units of productivity. Although it is easier to resolve population distinctions on a large geographic scale where gene flow and reproductive restrictions are more defined, the distribution of suitable spawning habitat and the homing behavior of locally adapted population segments may produce fine scale genetic structure between adjacent stream sections or watersheds (Beacham *et al.* 2006). While homing miscues (strays) are thought to be necessary to buffer loss of genetic diversity in salmon (Milner and Bailey 1989), particularly in small populations, the rate of straying among wild fish is generally low (Quinn 1993; Heard *et al.* 1995), and genetic structure between populations may persist despite moderate gene flow (Neville *et al.* 2007).

Objective two of the BPA project 2008-907-00 (Genetic Assessment of Columbia River Stocks) involves the collection and distribution of single nucleotide polymorphism (SNP) genotypic data and the use of that data to both expand on existing microsatellite genetic data baselines, and to characterize Chinook and steelhead population structure throughout the Columbia River Basin (CRB). Here we describe our efforts to create and build onto existing baselines founded by multi-agency consortiums called GAPS (Genetic Analysis of Pacific Salmonids; Seeb et al. 2007) for Chinook salmon, and SPAN (Stephen Phelps Allele Nomenclature; Stephenson et al. 2009) for steelhead trout. The use of SNPs is becoming increasingly popular for population analyses that were previously dominated by microsatellite markers. Many studies have been conducted that compare the relative utility of both marker types (Liu et al. 2005; Morin et al. 2009, Smith et al. 2007), and in most observations SNPs perform on par with microsatellite's for genetic distance (and nearly as well for genetic diversity analyses), though often a larger number of SNP loci are necessary to reach the same level of resolution of the more polymorphic microsatellites. The use of SNPs offers many advantages over microsatellite's and when used in concert, there is great potential to identify fine scale or localized population differentiation that is valuable for monitoring and evaluation for conservation concerns (Ryynanen et al. 2007; Narum et al. 2008). For example, SNPs are far more prolific in the genome, offering greater coverage for linkage analyses (Moen et al. 2008). Moreover, because SNPs may be located within functional genes, they are candidates for detecting positive selection or selective divergence shaping population differences. SNPs are relatively easily amplified and scored, particularly with poor quality tissue source or DNA extract (Campbell and Narum 2008), and with advances in analysis platforms, they are currently amenable to superior high throughput capabilities. Although Chinook salmon in the CRB have been fairly intensely studied (Narum et al. 2004; Waples et al. 2004; Beacham et al. 2006; Narum et al. 2007a; Narum et al. 2007b), and steelhead to a lesser degree (Narum et. al 2008; Nielson et al. 2009), our efforts are likely to provide additional information that will benefit and expound on the characterization and management of these species.

Methods

Sample Collection

In the effort to compile SNP genotypic baselines for Chinook salmon and steelhead trout, populations were included from throughout the CRB and were chosen to complement or overlap (when available) collections submitted to the standardized GAPS and SPAN consortium baselines for Chinook salmon and steelhead trout, allowing evaluation of multiple marker data sets once SNP complete (and later contributed to the identified consortiums). Some samples among our SNP baseline populations are not represented in existing microsatellite data sets due to tissue and DNA quality issues and the unavailability of genotypes solicited from outside sources (Table 1). Our goal for coverage of Chinook salmon was to include all three lineages (lower Columbia, ocean-type, and stream-type) from all primary drainages and major subbasins in the CRB. Similar goals for summer-run, winter-run and resident *O. mykiss* were also pursued (Figure 1). To date, collections have been partitioned into 52 and 54 groups for analysis of Chinook salmon and steelhead trout respectively (Table 1). Further, priority collections for both species were identified as those relevant to Tribal fishery interests (particularly in the middle Columbia and Snake River regions). Our primary focus is on robust CRB coverage to accommodate annual genetic stock identification analyses that will be used for in-season monitoring of runs and fisheries through the migratory corridor; this includes harvest GSI in the lower Columbia River, and fish passage GSI at Bonneville dam (see sections three and four of *BPA project 2008-907-00*). Biological tissues for genetic analysis were sampled from rayed fins (juveniles), and either caudal fin, opercle punch or carcasses of adult fish. Tissue samples were originally stored in individually labeled vials containing either 95% non-denatured ethanol or a lysis buffer (0.5 M EDTA, pH 8.0, 2 M Tris, pH 7.5, 5 M NaCL, 20% SDS).

Laboratory Protocol

Genomic DNA was extracted from digested tissue samples using a standard Qiagen® DNeasy™ protocol. Prior to amplification of SNP loci using primer-probe sets (fluorescent tags), an initial polymerase chain reaction (PCR) "pre-amp" step was implemented using whole genomic DNA to jumpstart SNP amplification via increased copy number of target DNA regions. The cycling regime and PCR conditions for the pre-amp step were as follows: one initial cycle of 95^{o} C for 15 min, 14 cycles of 95^{o} C for 15 seconds, 60 ^{o} C for four minutes, and a final dissociation step. For each data collection run, a panel of 96 SNP loci were arrayed with 96 samples using a Fluidigm® microfluidic 96.96 chip (including one genotype indicator and one no-template control sample) to generate high throughput genotyping. Sample cocktails included: 2.5µl Taqman[1], 0.05µl Gold taq polymerase, 0.25µl GT load buffer, 0.2µl H_2O and 2.0µl pre-amp DNA template. Single SNP assays were prepared in a 5.3µl reaction mix (per sample), containing the following reagents: 2.5µl DA load buffer, 0.25µl Rox dye, 1µl H_2O, and 1.25µl primer/probe. Microfluidic chips were loaded with assay cocktail dispensed at 4.5µl per well,

---

[1] *Single nucleotide polymorphism (SNP) loci were genotyped using TaqMan (Applied Biosystems) 5'-nuclease assays as described in Campbell and Narum (2008).*

and sample cocktail dispensed at 5.0µl per well. Chip loading and amplification was completed following standard manufacturers protocol on a Fluidigm IFC controller, and chips were imaged and scored on a Fluidigm EP1 imager using Fluidigm SNP Genotyping Analysis Software version 2.1.1. Carcass samples often provide poor quality and/or quantity of viable DNA relative to fresh tissue, and our final sample sizes were pared based on individual genotyping success. Successful genotyping for a given sample was defined proportionally as less than 10% missing data (i.e. fewer than nine SNP assays).

Statistical Analysis

Allele frequencies were generated with the program CONVERT (Glaubitz 2004). Descriptive statistics including number of samples analyzed per collection per locus, number of observed alleles, the unbiased heterozygosity (H_E), observed heterozygosity (H_O) and index of inbreeding (F_{is}) were generated using the analysis program GDA version 1.0 (Lewis and Zaykin 2001). The Markov Chain Monte Carlo approach implemented in GENEPOP version 3.4 (Raymond and Rousset 1995) was used to test for deviations from Hardy-Weinberg equilibrium (HWE) expectation evaluated across SNP loci and collections; this was completed primarily as a test for non-random mating within collections or possible marker amplification problems (e.g. null alleles). Linkage disequilibrium was tested for all pairs of loci across collections using a simulated exact test in GENEPOP. For all pairs of loci with significant non-random association (linkage) we reconstructed the (unknown) gametic phase of multi-locus genotypes (creating a phased constituent genotype) for each locus pair using the ELB algorithm implemented in ARLEQUIN version 3.5. (Excoffier et al. 2005). Statistical significance (α) was adjusted for the number of simultaneous tests k (α/k for $\alpha = 0.05$) for both HWE and linkage tests by a sequential Bonferroni correction (Rice 1989).

The program LOSITAN (Antao *et al*. 2008) was used to evaluate the relationship between F_{st} and H_e (expected heterozygosity) for all loci in an island model, to identify outlier loci (candidates for selection) having excessively high or low F_{st} compared to neutral expectations. We used data simulations based on 25,000 replicates, a mean F_{st} of 0.005, and a 0.99 confidence interval for all SNP loci under an infinite alleles model. Loci lying above or below these quantiles (outliers) may be under directional or balancing selection (respectively) in some populations. Interpreting patterns of neutral genetic diversity and drawing biological inferences from such comparisons (quantified by F_{st}) is confounded by the need to discriminate between selected diversity and the underlying demographic processes (e.g. gene flow) that influence diversity among populations; we therefore excluded significant candidate loci under selection from further analyses of population structure.

Table 1.) Inclusive Chinook salmon and steelhead trout collections in the 2009 SNP baselines. In the origin column, NOR is natural-origin and HAT is hatchery-origin. Note some populations (samples) from GAPS and SPAN were currently unavailable for genotyping; those without complementary microsatellite data are identified (*). In the Lineage column for steelhead, "AN" is anadromous and "RE" is resident. In the Run column for steelhead, "Win" is winter-run and "Sum" is summer-run. The origin of both species is identified as "NOR" – natural origin, and "HAT" – hatchery origin.

CHINOOK SALMON

Population (Collection ID)	Subbasin/ source	Region	(n)	Lineage	Run	Origin	Year	Age
Cowlitz R.	Cowlitz R.	Lower Columbia	86	Lower	Fall	HAT	2004	Adult
Lewis R.	Lewis R.	Lower Columbia	93	Lower	Fall	NOR	2003	Adult
Kalama R.	Kalama R.	Lower Columbia	81	Lower	Spring	HAT	2004	Adult
McKenzie R.	Willamette R.	Lower Columbia	84	Lower	Spring	HAT	2004	Adult
North Santiam R.	Willamette R.	Lower Columbia	87	Lower	Spring	HAT	2004	Adult
Spring Cr. (Tule)	Big White Salmon R.	Middle Columbia	88	Lower	Fall	HAT	2006	JUV.
*CherryLane	Clearwater R.	Clearwater	213	Ocean	Fall	NOR	2008	Adult
*Clearwater R.	Clearwater R.	Clearwater	118	Ocean	Fall	NOR	2008	JUV.
Clearwater R.	Clearwater R.	Clearwater	73	Ocean	Fall	NOR	2000	Adult
Nez Perce Tribal H.	Clearwater R.	Clearwater	86	Ocean	Fall	HAT	2003	Adult
Deschutes R. (lower)	Deschutes R.	Middle Columbia	90	Ocean	Fall	NOR	1999	Adult
Deschutes R. (Upper)	Deschutes R.	Middle Columbia	90	Ocean	summer	HAT	1998	JUV.
Hanford Reach	Columbia mainstem	Middle Columbia	93	Ocean	Fall	NOR	2000	Adult
Klickitat R.	Klickitat R.	Middle Columbia	92	Ocean	Fall	NOR	2004	Adult
Little White Salmon R.	Little White Salmon R.	Middle Columbia	91	Ocean	Fall	HAT	2006	JUV.
lyons ferry	Snake R. mainstem	Middle Columbia	90	Ocean	Fall	HAT	2000	Adult
Umatilla R.	Umatilla R.	Middle Columbia	82	Ocean	Fall	HAT	2006	Adult
Priest Rapids	Columbia mainstem	Upper Columbia	85	Ocean	Fall	HAT	2001	JUV.
Wells Dam	Columbia mainstem	Upper Columbia	89	Ocean	Fall	HAT	1993	Adult
Methow R.	Methow R.	Upper Columbia	88	Ocean	summer	NOR	1993	JUV.
Little White Salmon R.	Little White Salmon R.	Middle Columbia	92	Stream	Spring	HAT	-?-	Adult
Carson-NFH	Wind R.	Middle Columbia	91	Stream	Spring	HAT	2007	JUV.
John Day R.	John Day R.	Middle Columbia	84	Stream	Spring	NOR	2005	Adult
Klickitat R.	Klickitat R.	Middle Columbia	157	Stream	Spring	HAT	2002	Adult

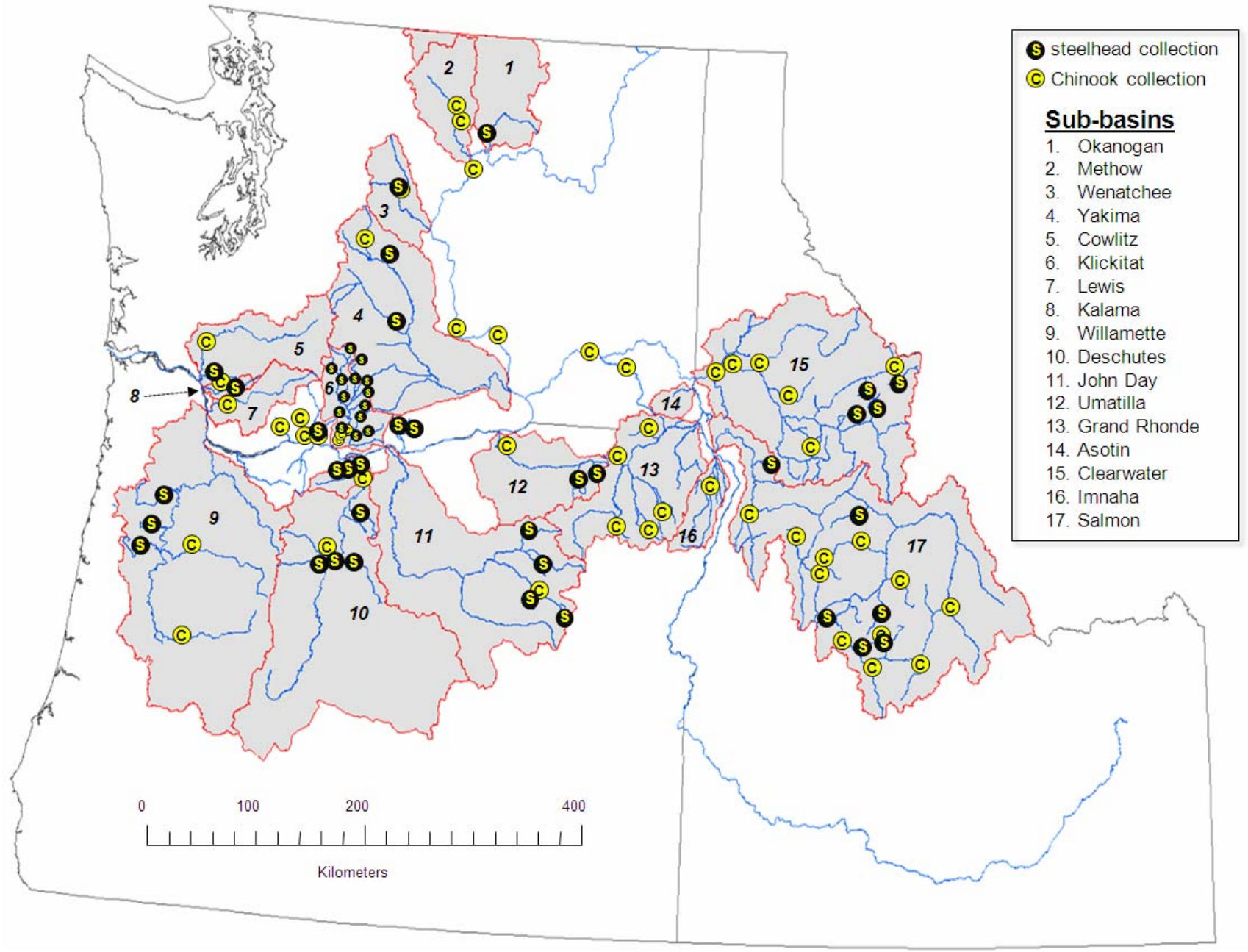
Klickitat R.	Klickitat broodstock	Middle Columbia	129	Stream	Spring	HAT	2006	Adult
Klickitat R.	Klickitat R.	Middle Columbia	187	Stream	Spring	NOR	2006	Adult
Cle Elum R.	Yakima R.	Middle Columbia	88	Stream	Spring	-?-	2006-07	-?-
Methow R.	Methow R.	Upper Columbia	93	Stream	Spring	HAT	1998	JUV.
Wenatchee R.	Wenatchee R.	Upper Columbia	85	Stream	Spring	-?-	1993	-?-
Winthrop-NFH	Carson stock	Upper Columbia	84	Stream	Spring	HAT	2001	Adult
Lolo Cr.	Clearwater R.	Clearwater	89	Stream	Spring	NOR	2001	JUV.
Dworshak-NFH	Clearwater R.	Clearwater	88	Stream	Spring	HAT	2005	Adult
Lochsa R.	Clearwater R.	Clearwater	77	Stream	Spring	HAT	2005	Adult
Newsome Cr.	S.F. Clearwater R.	Clearwater	90	Stream	Spring	NOR	2001	Adult
Looking-Glass Cr.	Grand Rhonde	Grand Rhonde	89	Stream	Spring	HAT	1994	JUV.
Lostine R.	Grand Rhonde	Grand Rhonde	82	Stream	Spring	-?-	2003	-?-
Minam R.	Grand Rhonde	Grand Rhonde	82	Stream	Spring	NOR	2002	JUV.
Wenaha R.	Grand Rhonde	Grand Rhonde	44	Stream	Spring	NOR	2002	JUV.
Big Cr.	Grand Rhonde	Grand Rhonde	92	Stream	Spring	NOR	2001	Adult
Catherine Cr.	Grand Rhonde	Grand Rhonde	85	Stream	Spring	NOR	2003	Adult
Tucannon R.	Lower Snake R.	Snake	87	Stream	Spring	NOR	2003	Adult
Imnaha R.	Snake R.	Snake	92	Stream	Spring	NOR	1998	Adult
CapeHorn Cr.	Snake R.	Snake	88	Stream	Spring	NOR	2005	JUV.
Rapid R.	Lower Salmon R.	Salmon	93	Stream	Spring	HAT	1999	Adult?
Camas Cr.	Middle Fork Salmon R.	Salmon	47	Stream	Spring	NOR	2006	JUV.
East Fork Salmon R.	Upper Salmon R.	Salmon	94	Stream	Spring	NOR	-?-	Adult
Johnson Cr.	South Fork Salmon R.	Salmon	88	Stream	Spring	HAT	2002	Adult
Secesh R.	South Fork Salmon R.	Salmon	81	Stream	Spring	NOR	2001	JUV.
Johnson Cr.	South Fork Salmon R.	Salmon	92	Stream	Spring	NOR	2002	Adult
Pahsimeroi R.	Upper Salmon R.	Salmon	93	Stream	Spring	-?-	2004	-?-
Sawtooth Hatchery	Upper Salmon R.	Salmon	91	Stream	Spring	HAT	2003	Adult
*West Fork	Yankee Fork	Salmon	75	Stream	Spring	NOR	2005	JUV.

STEELHEAD

Population (Collection ID)	Subbasin/ source	Region	(n)	Lineage	Run	Origin	Year	Age
Canyon Creek	Willamette	Lower Columbia	30	AN	Unk	NOR	1997	JUV.
Coweeman R.	Coweeman R.	Lower Columbia	47	AN	Win	NOR	2006	Adult
Kalama R.	Kalama R.	Lower Columbia	46	AN	Win	NOR	2005	Adult
Luckiamute R.	Willamette	Lower Columbia	31	Unk	Unk	NOR	1997	JUV.
Willamina Creek	Willamette R.	Lower Columbia	31	Unk	Unk	NOR	1997	JUV.
Big White Salmon R.	Big White Salmon R.	Middle Columbia	85	Unk	Unk	NOR	1990	JUV.
Klickitat R.	Klickitat R.	Middle Columbia	46	RE	Sum	NOR	2005	JUV.
Lower Little Klickitat R.	Klickitat R.	Middle Columbia	45	AN	Sum	NOR	2005	JUV.
Middle Little Klickitat R.	Klickitat R.	Middle Columbia	47	RE	Sum	NOR	2005	JUV.
Upper Little Klickitat R.	Klickitat R.	Middle Columbia	29	RE	Sum	NOR	2005	JUV.
Lower Summit Creek	Klickitat R.	Middle Columbia	46	AN	Sum	NOR	2005	JUV.
Upper Summit Creek	Klickitat R.	Middle Columbia	47	RE	Sum	NOR	2005	JUV.
Lower Trout Creek	Klickitat R.	Middle Columbia	48	AN	Sum	NOR	2005	JUV.
Upper Trout Creek	Klickitat R.	Middle Columbia	46	RE	Sum	NOR	2005	JUV.
Lower White Creek	Klickitat R.	Middle Columbia	40	AN	Sum	NOR	2005	JUV.
Upper White Creek	Klickitat R.	Middle Columbia	42	Unk	Sum	NOR	2005	JUV.
Bowman Creek	Klickitat R.	Middle Columbia	44	AN	Sum	NOR	2005	JUV.
Brush Creek	Klickitat R.	Middle Columbia	44	RE	Unk	NOR	2006	JUV.
Dead Canyon Creek	Klickitat R.	Middle Columbia	36	AN	Sum	NOR	2005	JUV.
Diamond Fork	Klickitat R.	Middle Columbia	45	RE	Sum	NOR	2005	JUV.
Fish Lake Stream	Klickitat R.	Middle Columbia	22	RE	Sum	NOR/HAT	2005	JUV.
Goldendale Hatchery	Klickitat R.	Middle Columbia	48	RE	Sum	HAT	2005	JUV.
Piscoe Creek	Klickitat R.	Middle Columbia	48	RE	Sum	NOR	2005	JUV.
Snyder Creek	Klickitat R.	Middle Columbia	46	AN	Sum	NOR	2005	JUV.
Surveyors Creek	Klickitat R.	Middle Columbia	39	RE	Sum	NOR	2005	JUV.
Swale Creek	Klickitat R.	Middle Columbia	48	AN	Sum	NOR	2005	JUV.
Teepee Creek	Klickitat R.	Middle Columbia	28	Unk	Unk	NOR	2005	JUV.
Fifteenmile Creek	Fifteenmile Creek	Middle Columbia	90	Unk	Unk	NOR	2005	JUV.
Buckhollow Creek	Deschutes R.	Middle Columbia	62	Unk	Sum	NOR	2005-06	JUV.
Trout Creek	Deschutes R.	Middle Columbia	57	Unk	Sum	NOR	2007	JUV.

Upper Mainstem	Deschutes R.	Middle Columbia	61	Unk	Sum	NOR	2005-06	JUV.
Camp Creek	Middle Fork John Day R.	Middle Columbia	46	Unk	Sum	NOR	2006	JUV.
Clear Creek	Middle Fork John Day R.	Middle Columbia	48	Unk	Sum	NOR	2005	JUV.
Camas Creek	North Fork John Day R.	Middle Columbia	22	Unk	Sum	NOR	2006	JUV.
Desolation Creek	North Fork John Day R.	Middle Columbia	25	Unk	Sum	NOR	2007	JUV.
*Rock Creek	Rock Creek	Middle Columbia	76	AN	Sum	NOR	2008	JUV.
*Squaw Creek	Rock Creek	Middle Columbia	43	AN	Sum	NOR	2008	JUV.
Meacham Creek	Umatilla R.	Middle Columbia	40	Unk	Sum	NOR	2005	JUV.
South Fork Umatilla R.	Umatilla R.	Middle Columbia	52	Unk	Sum	NOR	2005	JUV.
Nason Creek	Wenatchee R.	Upper Columbia	34	Unk	Sum	NOR	2006	JUV.
Omak Creek	Okanogan R.	Upper Columbia	57	AN	Sum	NOR/HAT	2005	Adult
Little Bear Creek	Clearwater R.	Snake R.	54	AN	Sum	NOR	2000	JUV.
North Fork Moose Creek	Selway R.	Snake R.	46	AN	Sum	NOR	2000	JUV.
Three Links Creek	Selway R.	Snake R.	46	AN	Sum	NOR	2000	JUV.
Colt Creek	Lochsa R.	Snake R.	47	AN	Sum	NOR	2000	JUV.
Lake Creek	Lochsa R.	Snake R.	47	AN	Sum	NOR	2000	JUV.
Loon Creek	Middle Fork Salmon R.	Snake R.	44	AN	Sum	NOR	2000	JUV.
Sulphur Creek	Middle Fork Salmon R.	Snake R.	46	AN	Sum	NOR	2000	JUV.
Chamberlain Creek	Salmon R.	Snake R.	46	AN	Sum	NOR	2000	JUV.
White Bird Creek	Salmon R.	Snake R.	46	AN	Sum	NOR	2000	JUV.
Lemhi R.	Upper Salmon R.	Snake R.	51	AN	Sum	NOR	2000	JUV.
Valley Creek	Upper Salmon R.	Snake R.	46	Unk	Sum	NOR	2005	JUV.
West Fork Yankee Fork R.	Upper Salmon R.	Snake R.	89	AN	Sum	NOR	2000	JUV.
Captain John Creek	Snake R.	Snake R.	55	AN	Sum	NOR	2000	JUV.

Figure 1.) Map of study area showing location of collections, and major sub-basins sampled in the CRB.



GENEPOP (Raymond and Rousset 1995) was used to calculate global F_{st} (Weir and Cockerham 1984), which indicates the proportion of total variation attributed to differences among collections and a matrix of pairwise F_{st} among all pairs of collections. Significance was determined by permutation tests based on 1000 individuals using ARLEQUIN version 3.5 (Excoffier *et al.* 2005) and critical values were adjusted for multiple tests following the method of Narum (2006). A pairwise matrix of Cavalli-Sforza and Edwards (1967) genetic chord distances (CSE) and an un-rooted neighbor-joining (NJ) phylogram were generated using PHYLIP version 3.68 (Felsenstein 1992). The NJ tree indicates similarities (clusters) among groups identified by branch associations in the genetic distance topology of the tree. The SEQBOOT option was implemented to generate 1000 simulated data sets, and a consensus topology with bootstrap support was generated using the CONSENSE option in PHYLIP. The analysis program GenAlEx version 6.2 (Peakall and Smouse 2006) was used to conduct principle components analysis (PCA) to identify clusters or patterns of genetic similarity among populations. The multivariate PCA analysis reduces redundant variables into a smaller number of principal components that will account for most of the variance in the data without much loss of information. The first principal component accounts for the greatest amount of variability in the data, and each succeeding component accounts for as much of the remaining variability as possible.

Results

Descriptive Statistics

The mean expected heterozygosity (allelic variability) per locus ranged from 0.0041-0.4772 and 0.0250-0.4994 across collections for Chinook and steelhead respectively (appendices1 & 2). For Chinook salmon, SNP loci *Ots_zP3b-1* and *Ots_SERPC1* were fixed (a single observed allele across all collections). Fixed loci are assumed to be uninformative and were therefore removed from the data set; two fixed loci in the steelhead data were cutthroat trout (*O. clarkii*) hybrid indicators, and were retained to demonstrate the absence of cutthroat trout introgression in the data set. No loci were diagnostic for any particular lineage or life history type (anadromous or resident) in the data sets for both species. Among 4742 total HWE tests for Chinook we observed 26 departures from expected genotypic proportions across groups. There were no apparent collection-specific deviations; the maximum observed was 2 deviations for any given collection, and two loci had 4 deviations across 52 collections. Among 5130 total HWE tests for steelhead we observed 58 departures from expected genotypic proportions across groups. There was at least one collection-specific deviation (Goldendale hatchery with 8 deviations); five loci had at least 3 deviations across 54 collections. A plot of expected heterozygosity and genetic distance (F_{st}) generated in LOSITAN identified several SNP outlier loci, or candidate loci under directional selection (Figure 2a & 2b). Eliminating non-neutral outlier loci is necessary before computing most population genetic parameters (e.g., F_{st}, Nm, Ne) that require neutral loci (Luikart *et al*. 2003) in order to evaluate divergence as a function of demographic differences or similarities. Ultimately only neutral loci will be retained for subsequent population structure analyses. At this point in our analyses we have not determined the appropriate confidence level threshold for assuming positive selection since some loci fall near the line of neutrality and many among those shown (Figures2a & 2b) are lineage specific. The plots indicate the number of

potential loci that may be valuable for landscape genetic analyses (based on selective divergence). One of the three potential candidate loci within both the lower and ocean-type lineages is a mitochondrial marker (C3N3) and therefore not of concern as an outlier. The stream-type lineage may have as many as nine loci that are candidates for positive selection. We will compare the effect of removing loci deemed as candidates for positive selection in an analysis to follow. However, the current population structure results were generated without excluding any loci based on this selection test.

Population Genetic Structure Analysis

For the 52 collections of Chinook salmon evaluated, we observed among-group variation (F_{st}) that ranged from 0.0115 to 0.8326 across all loci and groups.

The overall (global F_{st}) estimate for each of the three lineages was significantly greater than zero (Figure 3). Patterns of pairwise among-group variation (F_{st}) were mixed for comparisons among 14 ocean-type, 5 lower Columbia type and 32 stream-type collections (Figure 3). The least amount of among-group variation occurred within the ocean-type lineage, while the lower Columbia and stream-type lineages were comparable (note the exception in Spring Creek Tule).

For the 54 collections of steelhead evaluated, we observed among-group variation (F_{st}) that ranged from 0.0196 to 0.4932 across all loci and groups. The overall (global F_{st}) estimate was significantly greater than zero (Figure 4). Patterns of pairwise among-group variation (F_{st}) were mixed for comparisons between winter-run, summer-run, interior type, coastal type, resident and/or anadromous. There was far less variability in Fst values for steelhead than was observed among Chinook salmon. Two collections from the Klickitat River drainage (Goldendale Hatchery stock and Brush Creek) exhibited a relatively high amount of among-group variation, as did the last three collections from the right in Figure 4 (all from Williamette River subbasin). These three were the only confirmed winter-run collections in the data set.

Figure 2a.) Chinook salmon plot of expected heterozygosity vs. among-group variation.

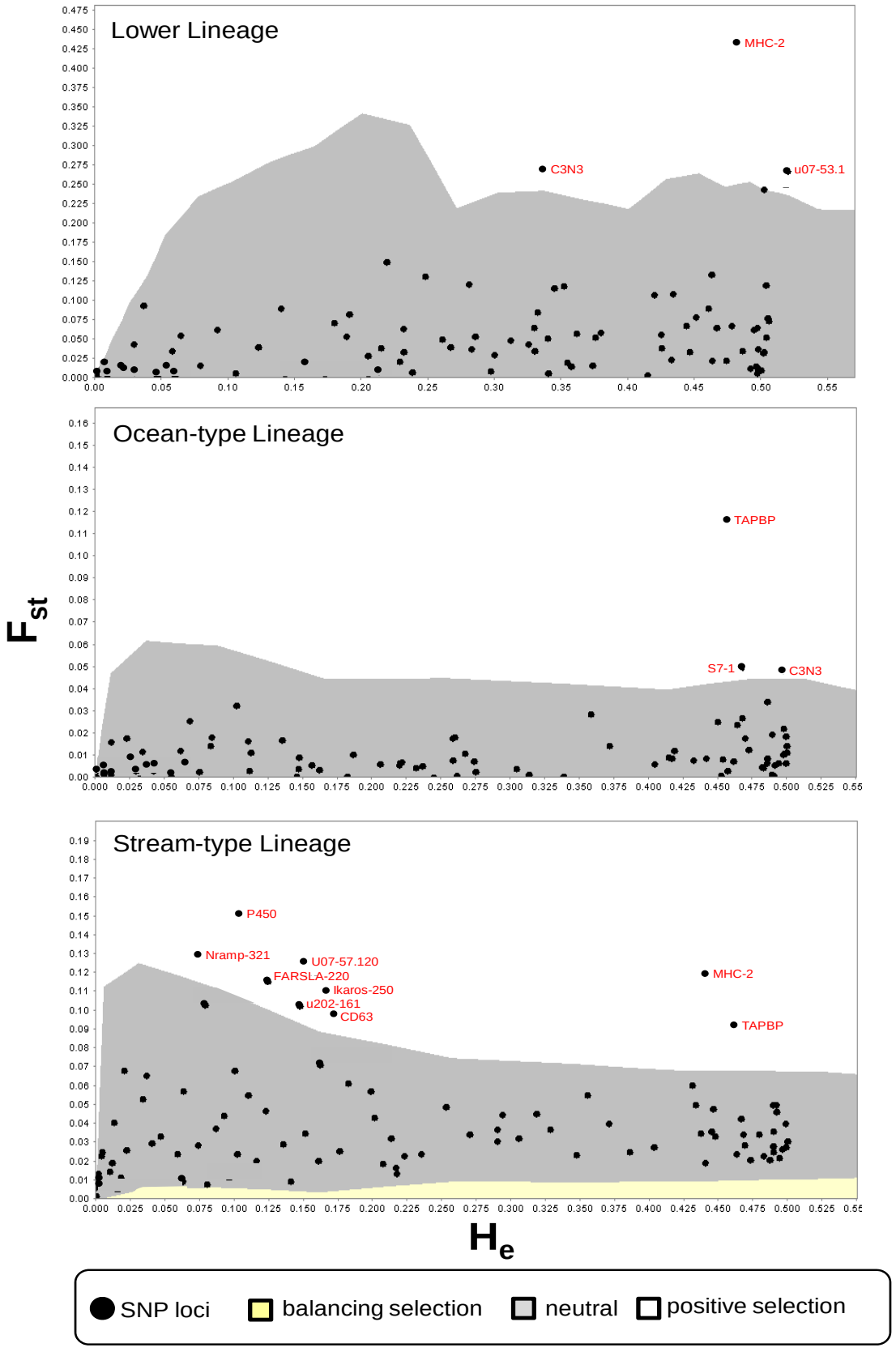
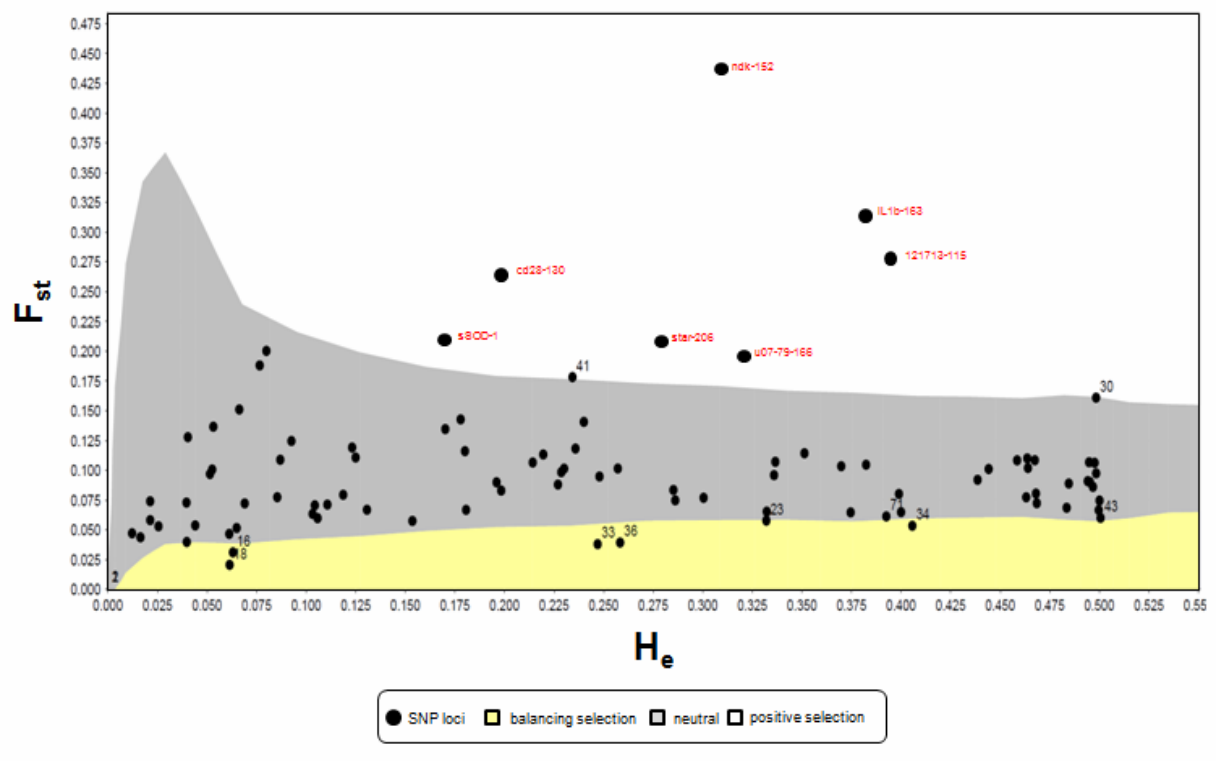


Figure 2b.) Steelhead trout



We demonstrated genetic similarity among populations within each species through phylogenetic relationships in the topology of an unrooted NJ phylogram (Figure 5) and in PCA cluster analyses to graphically display the relationship between collections based on the data variables that explain the greatest amount of total variation or differences (Figure 6). The confidence or concordance (>50%) of the NJ topology is indicated with bootstrap values at the nodes from a consensus tree. Results of the two analyses were complimentary, and revealed defined clustering of the most genetically similar collections for Chinook salmon (the three major lineages) and an interesting

association of Klickitat stream-type collections intermediate to both the stream lineage and ocean lineage (Figures 5a & 6a). Two collections of fall run (ocean-type) Chinook salmon in the Clearwater River stand out in PCA plots relative to Axis3. These collections were two of the most abundant in sample size leading to the assumption of no relatedness type of bias. The distinction may be an artifact of a hatchery-origin influence; these collections will be subject to further scrutiny. Steelhead collections clustered accurately into geographic regions or major tributary (Figure 5b). There is an indication of notable clusters based on life history origin in PCA analysis (Figure 6b)[2].

[2] *Currently we have only looked at this aspect of association to test the utility of these analyses for inferring identity of samples having unknown origin (e.g. resident vs. anadromous, coastal vs. interior). This same type of analysis will also be conducted based on location or tributary of origin to get a glimpse at distance associations.*

Figure 3.) Pairwise F_{st} comparisons between Chinook collections; the amount of total variation attributable to among-group differences.

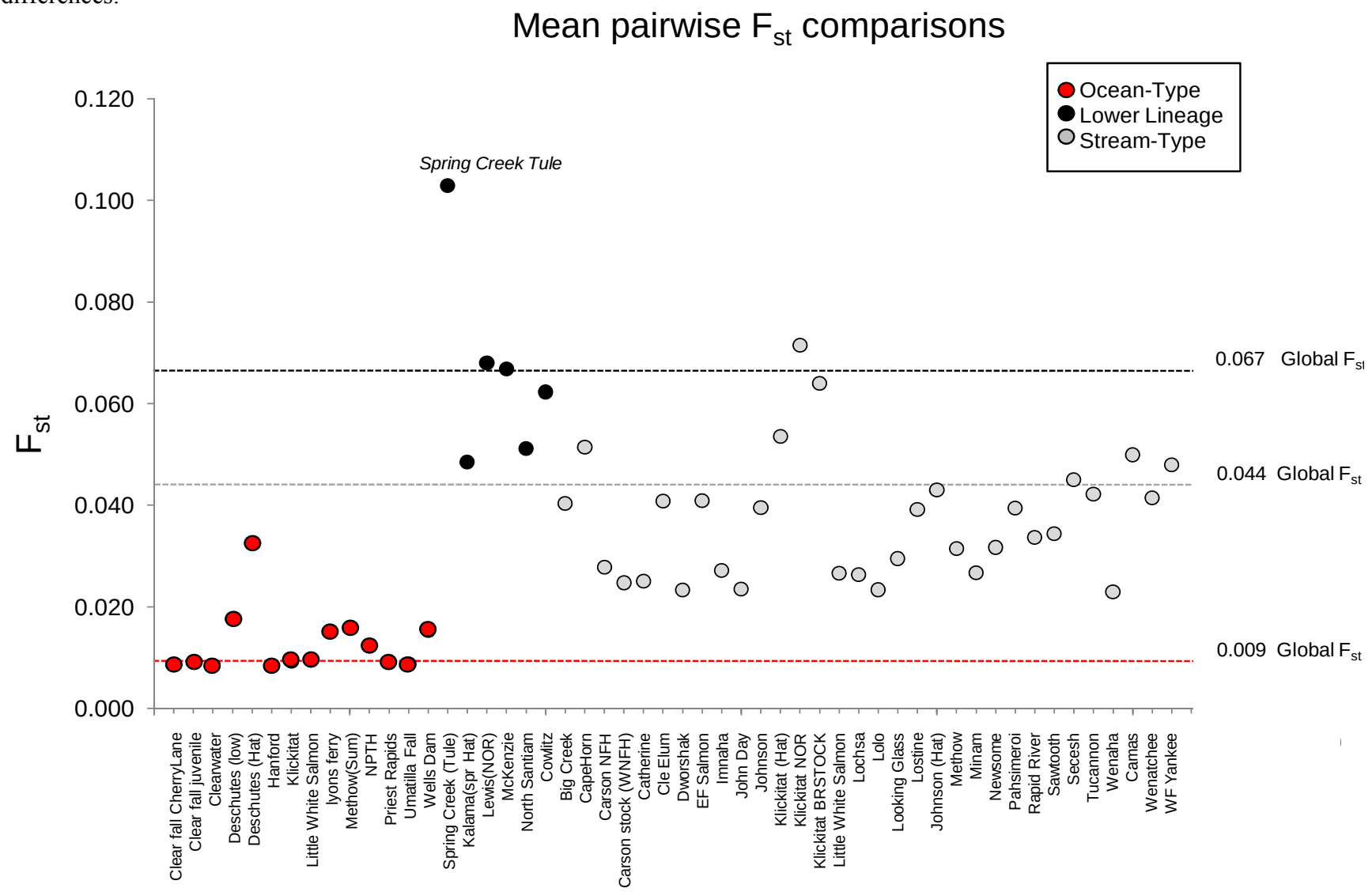
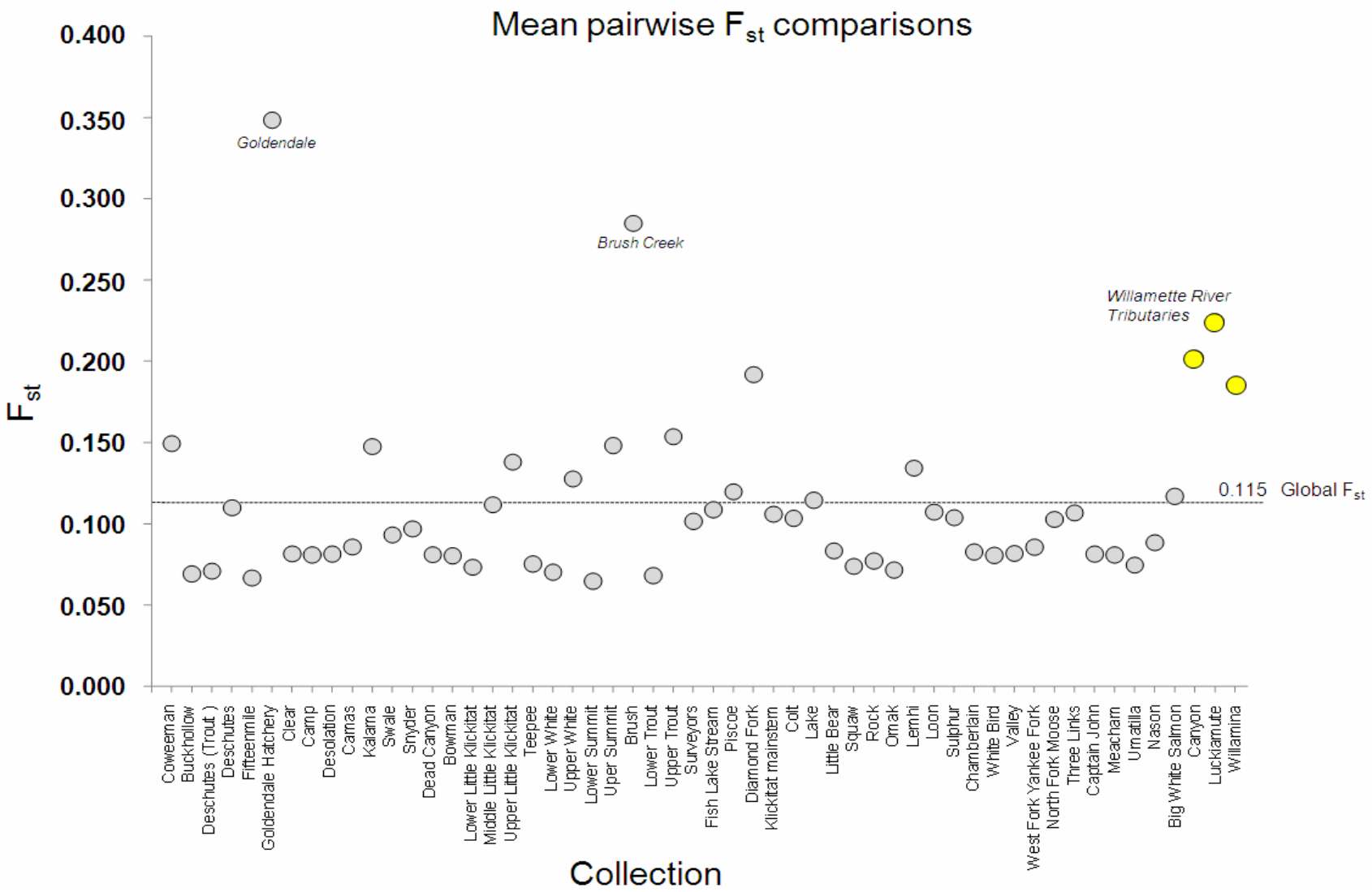


Figure 4.) Pairwise F_{st} comparisons between steelhead collections; the amount of total variation attributable to among-group differences.



Discussion

We have compiled extensive data sets of SNP genotypes for Chinook salmon and steelhead trout covering most of the Columbia River Basin (including the Snake River Basin) in our first year effort. Our goal is to construct SNP baselines of genotypes that will be amended annually to provide continued evaluation of these species that is both spatial and temporally stratified. This strategy assures the greatest likelihood of discerning reproductively distinct aggregations of Chinook salmon and steelhead trout (Waples 1991) through time, while monitoring population variability related to demographic trends that occur locally and/or regionally. Philopatry (Quinn *et al.* 1991, Hendry et al. 2003) and hatchery supplementation activities (Ford et al. 2006; Hard & Heard 1999) will play a major role in how genetic divergence and differentiation is distributed geographically, and it will be important to evaluate such impacts on the ability to differentiate populations both qualitatively and quantitatively (e. g. genetic stock identification)

The results presented in this report substantiate and complement differentiation of groups of Chinook salmon (Waples et al. 2004; Narum 2008) and steelhead trout (SPAN data set; unpublished results). Result further suggest SNPs are a class of markers that perform at least as well as microsatellites with great potential for monitoring population distinctions and composition in regard to migration and in-season fisheries.

We will begin to examine more closely those populations that display unique attributes or differences in contrast to expectations, including clustering of Klickitat River spring Chinook in comparison to overall stream-type. In addition, the nature of SNPs as candidates for detecting positive selection (e.g. locations within functional genes) should provide more clarification of how population differences are shaped across landscapes, such as the differentiation of Klickitat River anadromous *O. mykiss* populations in relation to migratory barriers (Narum et al. 2008). In fact, aside from our primary focus on GSI implementation, we will begin to investigate landscape genetics in greater detail by looking for correlations between environmental variables (e.g. temperature, migratory distance, elevation etc.) and genetic differences among populations for both Chinook salmon and steelhead trout. These additional and ongoing efforts will require further scrutiny of the genotypic data. Although loci under positive selection have great utility for landscape genetics evaluations, they are a confounding element that violate basic population genetics assumptions when conducting structure analyses. Our future efforts include adding collections to the baseline to increase basin wide coverage of both species and to provide a temporally stratified view of populations (accounting for natural inter-annual population variation).

.

Figure 5a.) Chinook salmon neighbor joining phylogram based on Cavalli-Sforza & Edwards (1967) chord-distances.

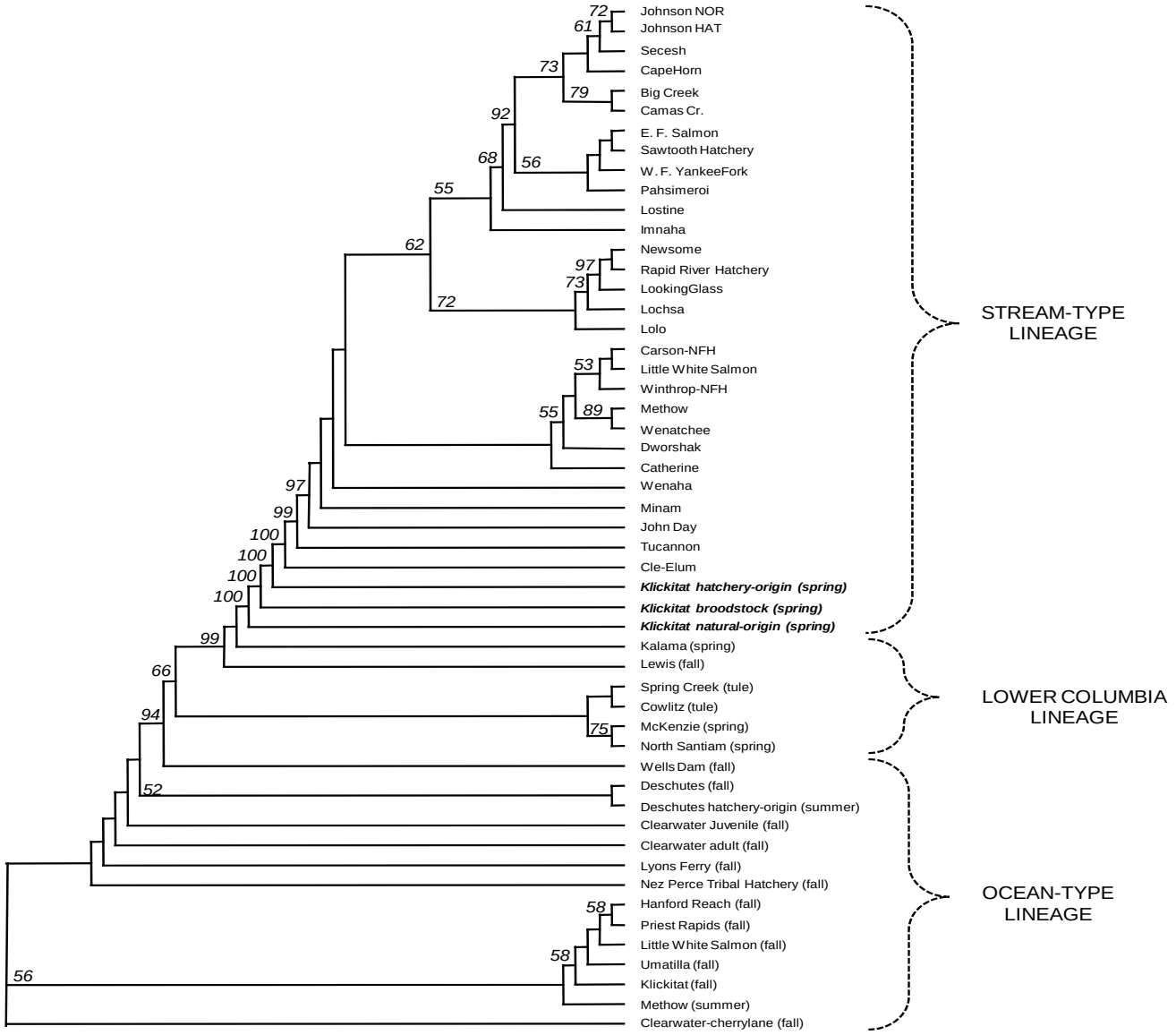


Figure 5b.) Steelhead neighbor joining phylogram based on Cavalli-Sforza & Edwards (1967) chord-distances.

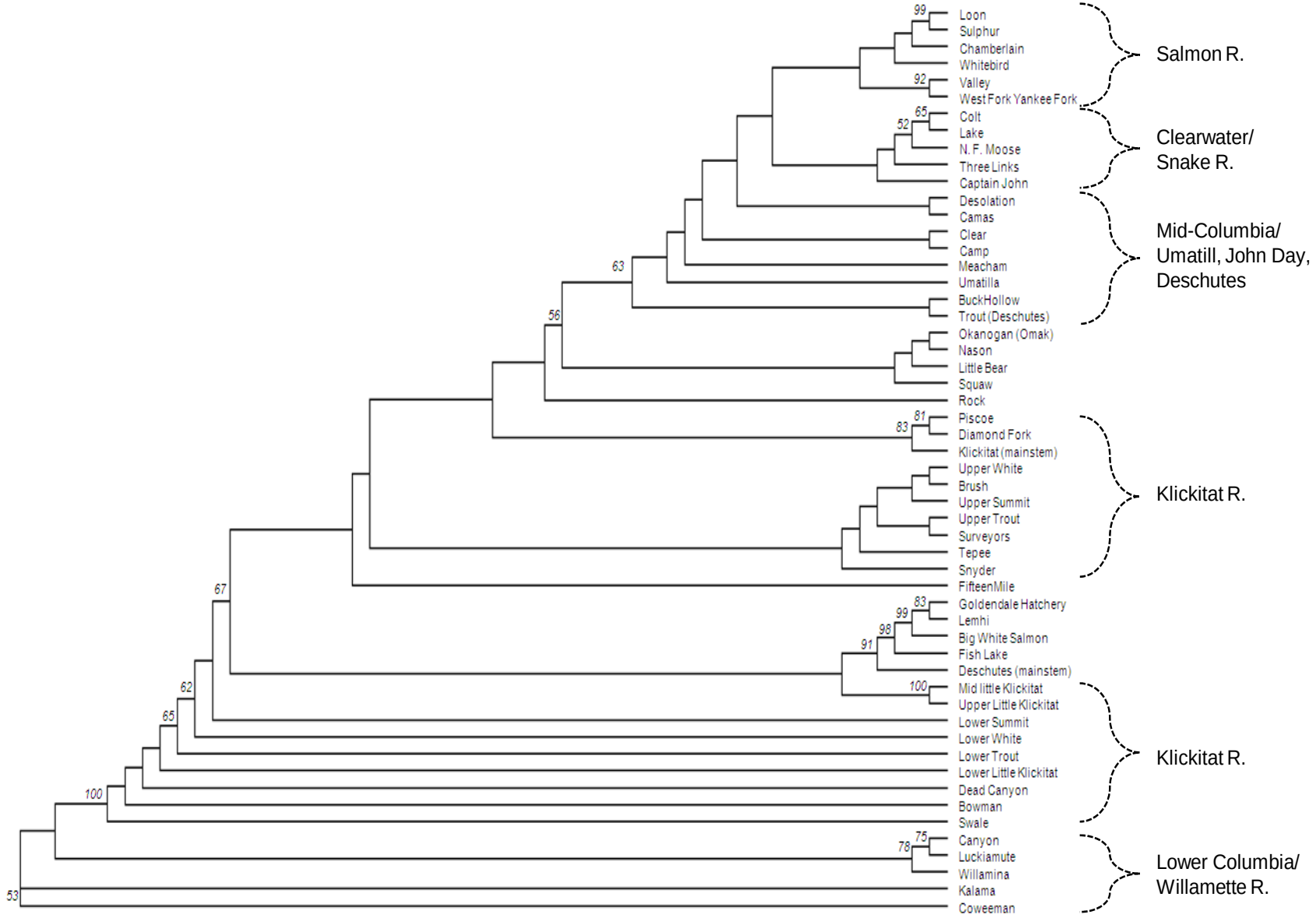
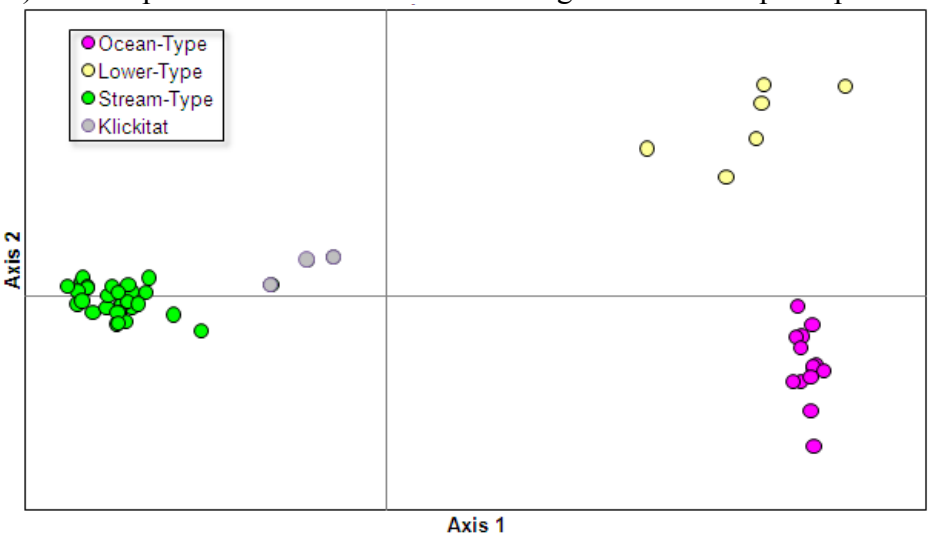
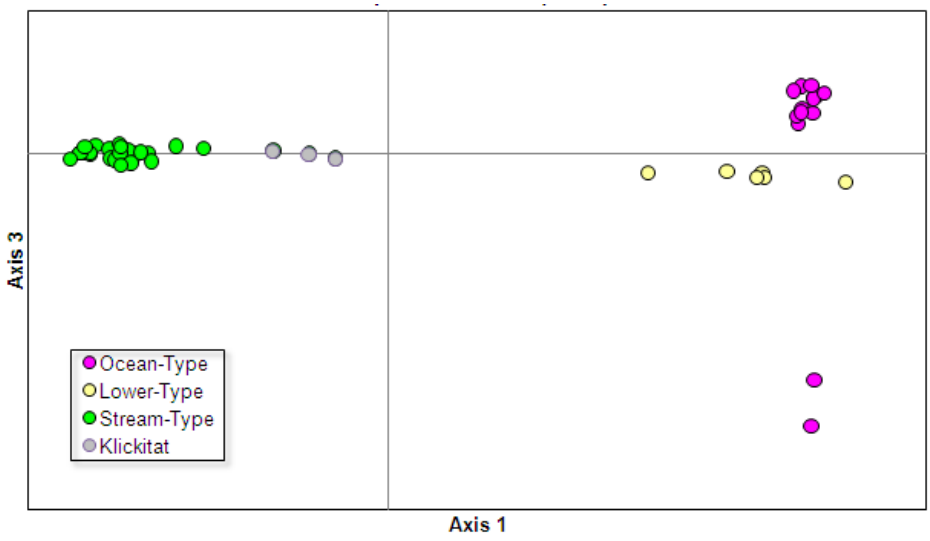


Figure 6a.) A PCA plot for Chinook salmon showing the first three principle coordinate axes.





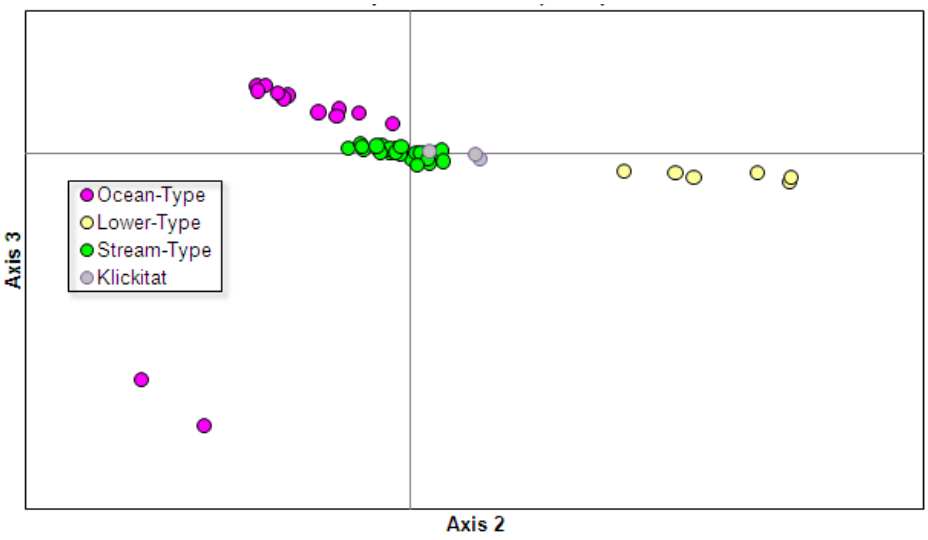
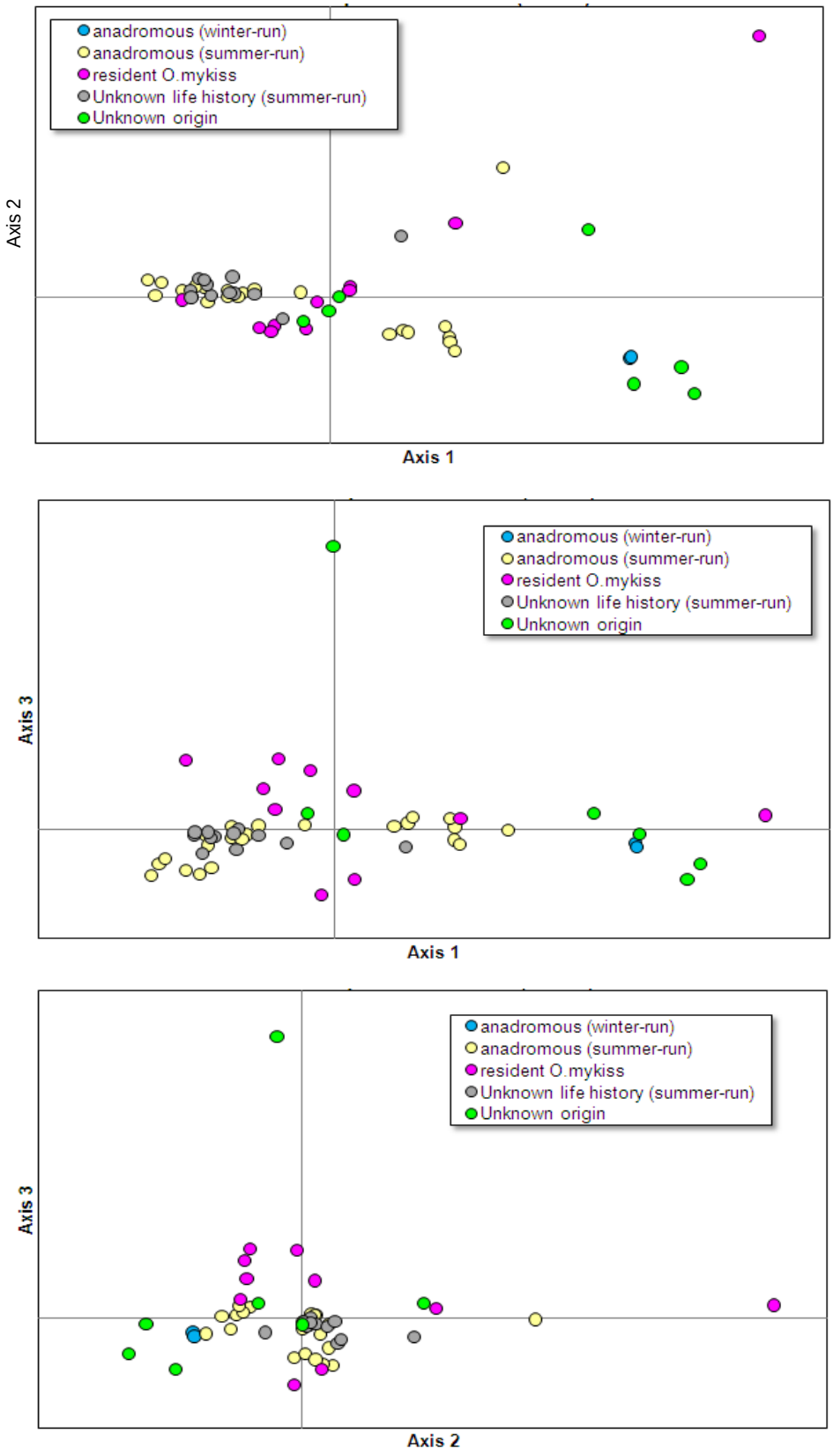


Figure 6b.) A PCA plot for steelhead showing the first three principle coordinate axes.



References

Antao, T., A. Lopes, R. J. Lopes, A. Baja-Pereira and G. Luikart. 2003. LOSITAN: A workbench to detect molecular adaptation based on an *Fst*-outlier method. BMC Bioinformatics 9:323.

Beacham, T. D., K. L. Jensen, J. Supernal, M. Wetly, L. Deng and N. Varnavskaya. 2006. Pacific rim population structure of Chinook salmon as determined from microsatellite analysis. Transactions of the American Fisheries Society 135:1604-1621.

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

Cavalli-Sforza, L.L., and W. W. F. Edwards. 1967. Phylogenetic analysis; models and estimation procedures. Evolution 32:550-570.

Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online **1**:47-50.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package), version 3.5c. Department of Genetics, University of Washington, Box 357360, Seattle, WA 98105, USA.

Ford, M. J., H. Fuss, B. Boelts, E. LaHood, J. Hard and J. Miller. 2006. Changes in run timing and natural smolt production in a naturally spawning coho salmon (*Oncorhynchus kisutch*) population after 60 years of intensive hatchery supplementation. Canadian Journal of Fisheries and Aquatic Sciences 63:2343–2355.

Glaubitz, J. C. 2004. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Molecular Ecology Notes 4 (2), 309–310.

Hard, J. J., and W. R. Heard. 1999. Analysis of straying variation in Alaskan hatchery Chinook salmon (*Oncorhynchus tshawytscha*) following transplantation. Canadian Journal of Fisheries and Aquatic Sciences 56:578-589.

Hasler, A. D., and A. T. Scholz. 1983. Olfactory imprinting and homing in salmon: investigations into the mechanism of imprinting process. Zoophysiology, Volume 14. Springer-Verlag, New York.

Heard, W. R., R. Burkett, F. Thrower, and S. McGee. 1995. A review of Chinook salmon resources in Southeast Alaska and development of an enhancement program designed for minimal hatchery-wild interaction. American Fisheries Society Symposium 15:21-37.

Hendry, A. P., V. Castric, M. T. Kinnison, and T. P. Quinn. 2003. The evolution of

philopatry and dispersal: homing vs. straying in salmonids. *In* Evolution illuminated: salmon and their relatives. *Edited by* A. P. Hendry and S. C. Stearns. Oxford Univ. Press, New York, NY. pp. 52-91.

Lewis, P. O., and Zaykin, D. 2001. Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

Liu, N., L. Chen, S. Wang, C. Oh and H. Zhao. 2005. Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. BMC Genetics 6(Suppl 1):S26.

Luikart, G., P. R. England, D. Tallmon, S. Jordan and P. Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Reviews Genetics 4(12):981-994.

McIssac, D. O. and T. P. Quinn. 1988. Evidence for a hereditary component in homing behavior of Chinook salmon (*Oncorhynchus tshawytscha*). Canadian Journal of Fisheries and Aquatic Sciences 45:2201-2205.

Milner, A. M. and R. G. Bailey. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. Aquaculture and Fisheries Management 20:179-192.

Moen, T., B. Hayes, M. Baranski, P. R. Berg, S. Kjoglum et al., 2008. A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. BMC Genomics 9: 223.

Morin, P. A., K. K. Martien and B. L. Taylor. 2009. Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources 9:66-73.

Narum, S. R., M. S. Powell and A. J. Talbot. 2004. A distinctive microsatellite locus that differentiates ocean-type from stream-type Chinook salmon in the interior Columbia River Basin. Transactions of the American Fisheries Society 133:1051-1055

Narum, S. R., C. Contor, A. Talbot and M. S. Powell, 2004 Genetic divergence of sympatric resident and anadromous forms of Oncorhynchus mykiss in the WallaWalla River, U.S.A. Journal of Fish Biology 65: 471–488.

Narum, S. R. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. Conservation Genetics 7:783-787.

Narum, S. R., J. J. Stephenson and M. R. Campbell. 2007a. Genetic variation and structure of Chinook salmon life history types in the Snake River. Transactions of the American Fisheries Society 136:1252-1262.

Narum, S. R., W. D. Arnsberg, A. J. Talbot and M. S. Powell. 2007b. Reproductive isolation following reintroduction of Chinook salmon with alternate life histories. Conservation Genetics 8:1123-1132.

Narum, S. R., J. S. Zendt, D. Graves and W. R. Sharp. 2008. Influence of landscape on resident and anadromous life history types of Oncorhynchus mykiss. Canadian Journal of Fisheries and Aquatic Sciences 65:1013-1023.

Neville, H., D. Isaak, R. Thurow, J. Dunham and B. Rieman. 2007. Microsatellite variation reveals weak genetic structure and retention of genetic variability in threatened Chinook salmon (*Oncorhynchus tshawytscha*) within a Snake R. watershed. Conservation Genetics 8:133-147.

Nielsen, J. L., A. Byrne, S. L. Graziano and C. C. Kozfkay. 2009. Steelhead genetic diversity at multiple spatial scales in a managed basin: Snake River, Idaho. North American Journal of Fisheries Management 29:680-701.

Peakall, R. O. D., and P. E. Smouse. 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295.

Quinn, T. P. 1993. A review of homing and straying of wild and hatchery-produced salmon. Fisheries Research 18:29-44.

Quinn, T. P., R. S. Nemeth and D. O. McIsaac. 1991. Homing and straying patterns of fall Chinook salmon in the lower Columbia R.. Transactions of the American Fisheries Society 120:150-156.

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

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Ryynanen, H. J., A. Tonteri, A. Vasemagi and C. R. Primmer. 2007. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic Salmon (*Salmo salar*). Journal of Heredity 98(7):692-704.

Seeb, L. W., A. Antonovich, M. A. Banks, T. D. Beacham, M. R. Bellinger et al. 2007. Development of a standardized DNA Database for Chinook salmon. Fisheries 32:541-549.

Smith, C., A. Antonovich, W. D. Templin, C. M. Elfstrom, S. R. Narum and L. W. Seeb. 2007. Impacts of Marker Class Bias Relative to Locus-Specific Variability on Population Inferences in Chinook Salmon: A Comparison of Single-Nucleotide Polymorphisms with Short Tandem Repeats and Allozymes. Transactions of the American Fisheries Society 136:1674–1687.

Stephenson, J. J., M. R. Campbell, J. E. Hess, C. Kozfkay, A. P. Matala et al. 2009. A centralized model for creating shared, standardized, microsatellite data that simplifies inter-laboratory collaboration. Conservation Genetics 10:1145-1149.

Waples, R. S. 1991. Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review 53:11-22.

Waples, R. S., D. J. Teel, J. M. Myers, and A. R. Marshall. 2004. Life-history divergence in Chinook salmon: historical contingency and parallel evolution. Evolution 58:386-403.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Appendix 1.) Chinook salmon descriptive statistics from analysis of the Chinook salmon SNP baseline of 52 collections). Column headings are: (n) sample size with complete genotype, (A) mean number of observed alleles, (He) Expected Heterozygosity, (Ho) Observed Heterozygosity, (F_{is}) Fixation Index, ($F_{st\ (mean)}$) among-collection variation per locus, and (%P) percentage of polymorphic loci. Values in bold identify numbers of significant HWE deviations.

Chinook salmon by Locus

SNP locus	(n)	A	He	Ho	F_{is}	$F_{st\ (mean)}$
Ots-102414-395	4719	2.0	0.4636	0.4655	-0.0060	0.0727
Ots-105105-613	4725	2.0	0.3978	0.3919	0.0132	0.1980
Ots-106747-239	4673	2.0	0.4462	0.4573	-0.0240	0.0694
Ots-110064-383	4729	2.0	0.4180	0.4210	-0.0030	0.1637
Ots-nramp-321	4738	1.8	0.0792	0.0757	0.0639	0.8326
Ots_113242-216	4761	2.0	0.3217	0.3276	-0.0230	0.1836
Ots_113457-40R	4802	2.0	0.2809	0.2696	0.0302	0.3488
Ots_123048-521	4808	1.6	0.1052	0.1077	-0.0180	0.0883
Ots_128757-61R	4808	2.0	0.3041	0.2984	0.0123	0.1783
Ots_94857-232R	4800	2.0	0.4772	0.4817	-0.0110	0.0452
Ots_94903-99R	4808	2.0	0.4722	0.4853	-0.0260	0.0487
Ots_96222-525	4805	2.0	0.2465	0.2462	-0.0070	0.2405
Ots_96500-180	4776	2.0	0.4265	0.4231	0.0091	0.1372
Ots_96899-357R	4812	1.8	0.1385	0.1387	-0.0060	0.0601
Ots_97077-179R	4816	1.9	0.1870	0.1766	***(4)*** 0.0799	0.2096
Ots_AldB1-122	4805	2.0	0.1746	0.1679	0.0358	0.0273
Ots_aldb-177M	4736	2.0	0.2400	0.2522	-0.0340	0.1197
Ots_ARNT	4745	2.0	0.2363	0.2296	0.0170	0.5272
Ots_arp-436	4810	1.8	0.1957	0.1957	-0.0070	0.3047
Ots_AsnRS-60	4815	2.0	0.2855	0.2807	0.0145	0.0407
Ots_aspat-196	4805	1.5	0.0999	0.0989	-0.0007	0.1110
Ots_C3N3	4817	1.8	0.2127	0.0000	1.0000	0.4284
Ots_Cath_D141	4773	1.6	0.0414	0.0424	-0.0250	0.0412
Ots_CCR7	4819	1.3	0.0310	0.0304	-0.0060	0.1390
Ots_CD59-2	4804	2.0	0.4584	0.4598	-0.0020	0.0310
Ots_CD63	4797	2.0	0.2811	0.2644	0.0500	0.2293
Ots_cox1-241	4805	2.0	0.2543	0.2499	0.0150	0.4848
Ots_CRB211	4776	1.5	0.0444	0.0463	-0.0240	0.0565
Ots_E2-275	4797	2.0	0.3534	0.3603	-0.0080	0.2917
Ots_EndoRB1-486	4760	1.9	0.2038	0.2105	-0.0270	0.1306
Ots_EP-529	4811	1.9	0.1099	0.1151	-0.0380	0.0452
Ots_ETIF1A	4710	2.0	0.3668	0.3659	-0.0020	0.2928
Ots_FARSLA-220	4795	1.9	0.1075	0.1082	0.0214	0.7764

Ots_FGF6A	4808	2.0	0.3742	0.3796	-0.0140	0.1292
Ots_FGF6B_1	4784	2.0	0.4536	0.4554	0.0047	0.0927
Ots_GDH-81x	4778	2.0	0.3992	0.3961	0.0056	0.0978
Ots_GH2_1	4823	1.7	0.0403	0.0418	-0.0260	0.0239
Ots_GnRH-271	4822	1.5	0.0285	0.0282	0.0082	0.0451
Ots_GPDH-338	4815	1.5	0.0404	0.0413	-0.0190	0.0458
Ots_GPH-318	4815	2.0	0.2060	0.2093	-0.0040	0.0859
Ots_GST-207	4814	1.6	0.0956	0.0966	-0.0180	0.0968
Ots_GST-375	4826	1.3	0.0116	0.0121	-0.0210	0.0372
Ots_GTH2B-550	4804	2.0	0.3413	0.3488	0.0179	0.2658
Ots_hsc71-3'-488	4786	2.0	0.2806	0.2831	-0.0120	0.4327
Ots_hsc71-5'-453	4793	2.0	0.2475	0.2520	-0.0030	0.2288
Ots_hsp27b-150	4806	2.0	0.2281	0.2327	-0.0090	0.2223
Ots_HSP90B-100	4813	2.0	0.2734	0.2739	-0.0030	0.4463
Ots_IGF-I.1-76	4817	1.7	0.1724	0.1690	0.0165	0.1121
Ots_Ikaros-250	4811	2.0	0.1615	0.1604	0.0114	0.6635
Ots_IL11	4819	1.8	0.1175	0.1168	-0.0040	0.0544
Ots_IL8R_C8	4808	2.0	0.3125	0.3012	0.0304	0.3224
Ots_mapK-3'-309	4801	2.0	0.4445	0.4469	-0.0020	0.1097
Ots_mapKpr-151	4795	2.0	0.3002	0.2992	-0.0030	0.0804
Ots_MHC1	4800	2.0	0.2160	0.2295	-0.0380	0.5159
Ots_MHC2	4784	2.0	0.3953	0.4177	-0.0520	0.1232
Ots_mybp-85	4809	2.0	0.3030	0.3073	-0.0010	0.2622
Ots_Myc-366	4788	1.3	0.0063	0.0064	-0.0120	0.0115
Ots_myo1a-384	4813	2.0	0.1495	0.1528	-0.0220	0.1069
Ots_myoD-364	4807	2.0	0.2694	0.2732	-0.0160	0.3509
Ots_nkef-192	4798	1.9	0.2972	0.3085	-0.0100	0.3499
Ots_NOD1	4786	2.0	0.3003	0.2957	0.0397	0.3926
Ots_OPLW173_1	4816	1.3	0.0280	0.0269	0.0175	0.1026
Ots_Ots311-101x	4810	1.7	0.1343	0.1346	0.0080	0.1950
Ots_P450	4814	1.9	0.1311	0.1245	0.0505	0.7196
Ots_P53	4803	2.0	0.3708	0.3672	0.0132	0.1016
Ots_PGK-54	4811	1.9	0.2038	0.2032	-0.0060	0.4742
Ots_Prl2	4793	2.0	0.4073	0.4147	-0.0180	0.1381
Ots_RAG3	4796	2.0	0.2818	0.2779	0.0202	0.4030
Ots_RAS1	4824	1.2	0.0041	0.0042	-0.0120	0.0193
Ots_RFC2-558	4811	1.7	0.2079	0.2074	-0.0040	0.3317
Ots_S7-1	4788	2.0	0.4477	0.4549	-0.0150	0.0535
Ots_SClkF2R2-135	4805	2.0	0.4474	0.4534	-0.0130	0.0366
Ots_SL	4821	1.8	0.1342	0.1283	0.0190	0.7065
Ots_SWS1op-182	4793	2.0	0.4205	0.4095	0.0270	0.1532

Ots_TAPBP	4801	2.0	0.3756	0.3654	0.0251	0.2318
Ots_TGFB	4805	2.0	0.1988	0.2010	-0.0080	0.0758
Ots_TLR3	4798	2.0	0.3617	0.3733	-0.0250	0.2766
Ots_TNF	4820	1.4	0.0168	0.0173	-0.0210	0.0208
Ots_Tnsf	4816	2.0	0.3097	0.3135	-0.0120	0.1328
Ots_u07-07.161	4793	2.0	0.4566	0.4619	-0.0120	0.0867
Ots_u07-17.135	4804	2.0	0.1526	0.1499	0.0207	0.0459
Ots_u07-18.378	4763	2.0	0.3116	0.3148	-0.0150	0.2300
Ots_u07-20.332	4777	1.4	0.0392	0.0397	-0.0060	0.0817
Ots_u07-25.325	4783	1.9	0.3187	0.3140	0.0148	0.2462
Ots_u07-49.290	4785	2.0	0.4368	0.4334	0.0084	0.1092
Ots_u07-53.133	4800	2.0	0.2654	0.2638	0.0092	0.3051
Ots_u07-57.120	4785	2.0	0.1215	0.1209	0.0021	0.7492
Ots_u07-64.221	4823	1.3	0.0109	0.0107	0.0446	0.0214
Ots_u202-161	4804	2.0	0.2461	0.2487	-0.0040	0.4122
Ots_u211-85	4807	2.0	0.3142	0.3184	-0.0120	0.3208
Ots_u4-92	4803	2.0	0.1015	0.1034	-0.0170	0.0351
Ots_u6-75	4817	2.0	0.1677	0.1661	-0.0080	0.0513
Ots_unk526	4810	2.0	0.2530	0.2497	0.0135	0.0300
Ots_ZR-575	4778	2.0	0.2022	0.1830	***(4)*** 0.0895	0.5737
----------	----------	----------	----------	----------	----------	----------
All	2515.8	1.9	0.2426	0.2408	0.0116	0.2142

Chinook salmon by population

Population	(n)	A	He	Ho	F_{is}	%P
Clear (CherryLane)	207.6	2	0.2654	0.2745	-0.0249	0.9574
Clearwater Juv.	117.6	1.9	0.2638	0.2511	0.0288	0.9362
Clearwater Adult	72.2	1.9	0.2640	0.2607	0.0015	0.9362
Deschutes	89.6	1.9	0.2644	0.2573	0.0154	0.9468
Deschutes (summer Hat)	89.7	1.9	0.2692	0.2593	0.0501	0.9468
Hanford Reach	92.4	1.9	0.2549	0.2547	-0.0115	0.8936
Klickitat	91.6	1.9	0.2603	0.2543	0.0120	0.9468
Little White Salmon (fall)	90.7	1.9	0.2603	0.2561	0.0187	0.9362
lyons ferry	89.6	1.9	0.2590	0.2548	0.0182	0.9149
Methow (summer)	87.7	1.9	0.2547	0.2502	0.0105	0.9149
Nez Perce Tribal Hat	85.4	1.9	0.2585	0.2567	0.0146	0.9255
Priest-Rapids	84.3	1.9	0.2511	0.2471	0.0064	0.9149
Umatilla	79.0	1.9	0.2527	0.2548	0.0021	0.9149
Wells Dam	88.5	1.9	0.2567	0.2508	0.0418	0.9362

Spring Creek (Tule)	87.2	1.9	0.2231	0.2268	-0.0053	0.8936
Kalama	80.3	1.9	0.3177	0.3067	0.0390	0.9468
Lewis	92.3	1.9	0.2919	0.2874	0.0039	0.9362
McKenzie	83.0	1.9	0.2602	0.2632	-0.0015	0.8723
North-Santiam	86.6	1.9	0.2549	0.2572	-0.0024	0.8617
Cowlitz	85.2	1.9	0.2821	0.2838	-0.0068	0.9255
Big Creek	91.5	1.7	0.2033	0.2037	0.0198	0.7447
CapeHorn	87.4	1.8	0.1956	0.1921	-0.0003	0.7553
Carson-NFH	90.7	1.9	0.2308	0.2282	0.0144	0.8617
Winthrop-NFH (Carson stock)	83.7	1.8	0.2279	0.2269	0.0182	0.8404
Catherine	84.4	1.9	0.2313	0.2288	0.0086	0.8511
Cle-Elum	87.2	1.9	0.2733	0.2680	0.0277	0.8936
Dworshak	87.8	1.9	0.2273	0.2285	0.0031	0.8617
East Fork Salmon	93.5	1.8	0.2087	0.2132	-0.0118	0.7872
Imnaha	91.7	1.9	0.2224	0.2253	0.0042	0.8511
John Day	82.7	1.9	0.2426	0.2399	0.0234	0.8936
Johnson	91.9	1.8	0.2055	0.2049	0.0105	0.8191
Klickitat (Hat)	155.6	2	0.2947	0.2907	0.0157	0.9574
Klickitat	186.3	2	0.3159	0.3027	0.0429	0.9574
Klickitat (broodstock	128.5	2	0.3082	0.2967	0.0400	0.9574
Little White Salmon (spring)	91.7	1.8	0.2338	0.2302	0.0197	0.8404
Lochsa	76.4	1.9	0.2242	0.2226	0.0160	0.8511
Lolo	88.6	1.9	0.2223	0.2236	0.0112	0.8511
LookingGlass	88.9	1.9	0.2244	0.2249	0.0033	0.8511
Lostine	81.7	1.9	0.2147	0.2198	-0.0044	0.8511
Johnson (Hat)	87.4	1.8	0.2056	0.2067	-0.0090	0.8085
Methow	92.7	1.9	0.2263	0.2166	0.0679	0.9149
Minam	81.6	1.9	0.2421	0.2370	0.0315	0.8830
Newsome	89.8	1.8	0.2226	0.2193	0.0250	0.8085
Pahsimeroi	92.4	1.8	0.2112	0.2144	-0.0151	0.7872
Rapid River (Hat)	92.8	1.8	0.2198	0.2184	0.0301	0.8404
Sawtooth	90.6	1.8	0.2068	0.2083	-0.0012	0.8191
Secesh	80.4	1.8	0.2027	0.2012	0.0070	0.7979
Tucannon	86.3	1.9	0.2483	0.2574	-0.0150	0.8723
Wenaha	43.9	1.8	0.2381	0.2420	-0.0087	0.8298
Camas	46.9	1.7	0.1937	0.1998	-0.0112	0.6702
Wenatchee	84.4	1.8	0.2205	0.2180	0.0072	0.8404
West Fork Yankee Fork	74.9	1.8	0.2071	0.2041	0.0124	0.7766
---	---	---	---	---	---	---
Mean	92.2	1.9	0.2426	0.2408	0.0114	0.8727

Appendix 2.) Steelhead descriptive statistics from analysis of the SNP baseline of 54 collections). Column headings are: (n) sample size with complete genotype, (A) mean number of observed alleles, (He) Expected Heterozygosity, (Ho) Observed Heterozygosity, (F_{is}) Fixation Index, ($F_{st\ (mean)}$) among-collection variation per locus with standard deviation (SD), and (GD) gene diversity. Values in bold identify numbers of significant HWE deviations.

Steelhead by Locus

SNP locus	(n)	A	He	Ho	F_{is}	$F_{st\ (mean)}$	$F_{st\ (SD)}$
Omy_myclarp404-111	2514	1.0	0.0000	0.0000	0.0000	---	---
Omy_Omyclmk436-96	2528	1.0	0.0000	0.0000	0.0000	---	---
Omy_113490-159	2508	2.0	0.4975	0.4406	0.1144	0.1201	0.0233
Omy_114315-438	2523	2.0	0.4678	0.4201	0.1019	0.0825	0.0184
Omy_121006-131	2530	1.6	0.0950	0.0842	0.1140	0.1144	0.0285
Omy_121713-115	2489	2.0	0.3909	0.2282	***(5)*** 0.4163	0.2727	0.0452
Omy_123044-128	2516	1.8	0.1026	0.0839	0.1830	0.0550	0.0127
Omy_123048-119	2525	1.5	0.0805	0.0547	0.3207	0.3293	0.1888
Omy_127236-583	2528	1.4	0.0709	0.0625	0.1183	0.1540	0.0534
Omy_128693-455	2530	1.6	0.0568	0.0474	0.1650	0.1712	0.1017
Omy_130295-98	2522	2.0	0.4091	0.3695	0.0966	0.1010	0.0395
Omy_130524-160	2521	2.0	0.4719	0.4276	0.0938	0.0832	0.0236
Omy_187760-385	2528	1.3	0.0331	0.0305	0.0788	0.1526	0.0660
Omy_95489-239	2529	1.5	0.0594	0.0502	0.1549	0.2073	0.1334
Omy_96222-125	2447	1.8	0.1766	0.1271	***(5)*** 0.2805	0.1100	0.0179
Omy_97077-73	2530	1.7	0.0686	0.0680	0.0094	0.0421	0.0128
Omy_97660-230	2520	2.0	0.4421	0.3937	0.1097	0.0835	0.0210
Omy_97865-196	2523	1.7	0.0625	0.0598	0.0429	0.0196	0.0041
Omy_97954-618	2520	1.9	0.2875	0.2210	0.2313	0.1642	0.0753
Omy_aldB-165	2497	2.0	0.4696	0.4197	0.1062	0.0977	0.0258
Omy_aldB-414	2515	2.0	0.3189	0.2867	0.1010	0.1165	0.0304
Omy_ALDOA_1	2528	1.3	0.0323	0.0297	0.0815	0.1628	0.0885
Omy_aromat-280	2501	2.0	0.3357	0.2819	***(6)*** 0.1603	0.0518	0.0162
Omy_arp-630	2523	2.0	0.4959	0.4586	0.0752	0.0789	0.0213
Omy_aspAT-123	2526	2.0	0.3782	0.3416	0.0967	0.1056	0.0281
Omy_aspAT-413	2529	1.7	0.1206	0.1083	0.1018	0.1045	0.0348
Omy_b1-266	2467	2.0	0.4546	0.4114	0.0950	0.1034	0.0327
Omy_b9-164	2527	1.8	0.1272	0.1041	0.1820	0.1078	0.0423
Omy_BAC-B4-126	2486	2.0	0.2963	0.2796	0.0565	0.0776	0.0262
Omy_BAC-B4-324	2513	2.0	0.4962	0.4250	0.1436	0.1619	0.0303
Omy_cd28-130	2505	1.8	0.1851	0.1417	0.2345	0.2472	0.0583
Omy_cd59-206	2519	2.0	0.4627	0.4097	0.1146	0.1145	0.0268
Omy_cd59b-112	2528	2.0	0.2552	0.2306	0.0964	0.0402	0.0106
Omy_colla1-525	2508	2.0	0.4098	0.4043	0.0133	0.0467	0.0107
Omy_cox1-221	2504	2.0	0.4814	0.4329	0.1008	0.0826	0.0217

Omy_cox2-335	2436	1.9	0.2594	0.2504	0.0347	0.0371	0.0083
Omy_crb-106	2460	2.0	0.4895	0.3858	0.2119	0.1149	0.0250
Omy_CRBF1-1	2505	1.9	0.2422	0.2100	0.1329	0.1483	0.0375
Omy_cxcr-169	2527	1.6	0.0955	0.0625	0.3452	0.4062	0.2009
Omy_dacd1-131	2524	1.5	0.1108	0.0931	0.1595	0.1801	0.0511
Omy_e1-147	2529	2.0	0.2237	0.1902	0.1500	0.1810	0.0637
Omy_g1-103	2515	1.7	0.1181	0.1125	0.0474	0.0716	0.0141
Omy_g12-82	2522	2.0	0.4994	0.4742	0.0504	0.0619	0.0154
Omy_gadd45-332	2525	1.7	0.1205	0.1097	0.0893	0.0757	0.0180
Omy_gdh-271	2523	1.9	0.1773	0.1657	0.0656	0.0647	0.0169
Omy_gh-334	2514	1.4	0.0534	0.0438	0.1806	0.1478	0.0465
Omy_gh-475	2524	1.9	0.2140	0.1961	0.0837	0.1185	0.0474
Omy_GHSR-121	2519	1.8	0.2096	0.1513	0.2783	0.2728	0.1174
Omy_gluR-79	2518	2.0	0.4994	0.4654	0.0680	0.0615	0.0208
Omy_hsc715-80	2520	2.0	0.4994	0.4425	0.1140	0.0736	0.0210
Omy_hsf1b-241	2518	2.0	0.2313	0.2041	0.1175	0.0952	0.0222
Omy_hsf2-146	2515	1.9	0.3434	0.2915	0.1512	0.1102	0.0209
Omy_hsp47-86	2524	2.0	0.3741	0.3546	0.0521	0.0585	0.0120
Omy_hsp70aPro-329	2526	1.8	0.1082	0.1021	0.0564	0.0571	0.0126
Omy_hsp90BA-193	2515	1.9	0.2347	0.2048	0.1277	0.1320	0.0242
Omy_hsp90BA-229	2522	1.8	0.1681	0.1431	0.1483	0.1322	0.0262
Omy_IL17-185	2525	2.0	0.4956	0.4772	0.0370	0.0901	0.0210
Omy_IL1b-163	2503	1.9	0.3681	0.2385	0.3521	0.2790	0.0564
Omy_IL6-320	2519	2.0	0.3502	0.3303	0.0569	0.0797	0.0200
Omy_inos-97	2527	1.6	0.0624	0.0590	0.0556	0.0460	0.0115
Omy_LDHB-1_i2	2484	1.8	0.1313	0.1083	0.1756	0.0539	0.0248
Omy_LDHB-2_e5	2520	2.0	0.3499	0.3083	0.1188	0.1214	0.0286
Omy_LDHB-2_i6	2527	1.6	0.1040	0.0910	0.1247	0.1256	0.0418
Omy_mapK3-103	2524	1.9	0.2294	0.2025	0.1174	0.1532	0.0471
Omy_mcsf-268	2502	1.7	0.1380	0.0987	0.2846	0.1821	0.0957
Omy_mcsf-371	2524	1.6	0.0834	0.0792	0.0497	0.0977	0.0244
Omy_metA-161	2522	2.0	0.4491	0.4064	0.0951	0.1064	0.0249
Omy_metB-138	2507	2.0	0.3155	0.2712	0.1404	0.0927	0.0244
Omy_myoD-178	2522	1.8	0.1524	0.1423	0.0658	0.0552	0.0105
Omy_nach-200	2528	1.6	0.0930	0.0684	0.2638	0.2512	0.0684
Omy_NaKATPa3-50	2507	2.0	0.3976	0.3781	0.0489	0.0604	0.0138
Omy_ndk-152	2517	1.8	0.2998	0.1828	0.3904	0.3910	0.0690
Omy_nkef-241	2522	2.0	0.4942	0.4588	0.0718	0.1026	0.0221
Omy_nkef-308	2525	2.0	0.4936	0.4598	0.0685	0.1154	0.0265
Omy_nramp-146	2519	1.3	0.0439	0.0401	0.0859	0.2542	0.1493
Omy_Ogo4-212	2511	2.0	0.4840	0.4656	0.0382	0.0565	0.0152
Omy_OmyP9-180	2502	1.9	0.2375	0.2146	***(3)*** 0.0964	0.1048	0.0239
Omy_Ots208-138	2481	2.0	0.4524	0.3873	***(3)*** 0.1439	0.1014	0.0209

Omy_Ots249-227	2514	2.0	0.4532	0.4073	0.1012	0.0908	0.0200
Omy_oxct-85	2518	1.9	0.2431	0.2069	0.1490	0.1273	0.0426
Omy_p53-262	2527	2.0	0.2994	0.2723	0.0905	0.0969	0.0282
Omy_pad-196	2516	1.4	0.0405	0.0366	0.0969	0.0782	0.0195
Omy_PEPA-i6	2528	1.7	0.0731	0.0601	0.1773	0.0730	0.0193
Omy_R0917-230	2530	1.2	0.0250	0.0198	0.2089	0.1943	0.1109
Omy_R1175-137	2529	1.3	0.0319	0.0261	0.1820	0.1922	0.1153
Omy_rapd-132	2525	1.5	0.0702	0.0570	0.1880	0.1931	0.0608
Omy_rapd-167	2505	2.0	0.3728	0.3269	0.1229	0.1006	0.0243
Omy_SECC22b-88	2527	1.5	0.1134	0.0700	0.3826	0.4132	0.1747
Omy_sSOD-1	2524	1.7	0.1601	0.1351	0.1564	0.1792	0.0405
Omy_star-206	2528	1.9	0.2917	0.2271	0.2216	0.2326	0.0483
Omy_stat3-273	2523	2.0	0.3937	0.3690	0.0626	0.0766	0.0209
Omy_tgfb-207	2522	2.0	0.2377	0.2169	0.0874	0.1222	0.0363
Omy_tlr3-377	2527	1.9	0.2257	0.2078	0.0793	0.1097	0.0542
Omy_tlr5-205	2526	1.9	0.1935	0.1694	0.1241	0.0879	0.0206
Omy_u07-79-166	2525	1.9	0.3288	0.2650	0.1941	0.2108	0.0491
----------	----------	----------	----------	----------	----------	----------	----------
All	2515.8	1.8	0.2541	0.2224	0.1248	0.2108	0.0490

Steelhead by population

Population	(n)	A	He	Ho	F_{is}	GD
Coweeman R.	46.7	1.9	0.2544	0.2470	0.0293	0.260
Buckhollow Creek	61.8	1.9	0.2225	0.2139	0.0388	0.227
Deschutes (Trout Creek)	56.8	1.9	0.2193	0.2133	0.0272	0.224
Deschutes	60.8	1.9	0.2485	0.2301	0.0746	0.254
Fifteenmile Creek	89.6	2.0	0.2501	0.2448	0.0213	0.256
Goldendale Hatchery	47.8	1.8	0.2948	0.3083	***(8)*** -0.0463	0.301
Clear Creek	47.8	1.8	0.1980	0.2016	-0.0188	0.202
Camp Creek	45.8	1.8	0.2016	0.2048	-0.0163	0.206
Desolation Creek	24.9	1.8	0.2002	0.1966	0.0184	0.205
Camas Creek	22.0	1.8	0.2104	0.2107	-0.0011	0.215
Kalama R.	45.9	1.8	0.2612	0.2498	0.0443	0.267
Swale Creek	47.9	1.8	0.2542	0.2671	-0.0512	0.260
Snyder Creek	45.7	1.8	0.2334	0.2188	0.0634	0.239
Dead Canyon Creek	35.7	1.8	0.2537	0.2523	0.0056	0.259
Bowman Creek	43.8	1.9	0.2559	0.2504	0.0217	0.261
Lower Little Klickitat R.	44.8	1.9	0.2626	0.2578	0.0188	0.268
Middle Little Klickitat R.	46.4	1.9	0.2616	0.2520	0.0373	0.267
Upper Little Klickitat R.	28.7	1.9	0.2314	0.2250	0.0277	0.236

Teepee Creek	27.4	1.8	0.2223	0.2073	0.0686	0.227
Lower White Creek	39.1	1.9	0.2576	0.2459	0.0461	0.263
Upper White Creek	41.6	1.6	0.1992	0.1972	0.0104	0.204
Lower Summit Creek	45.8	1.9	0.2572	0.2527	0.0179	0.263
Uper Summit Creek	46.8	1.8	0.2123	0.2109	0.0067	0.217
Brush Creek	43.5	1.4	0.1303	0.1427	-0.0961	**0.133**
Lower Trout Creek	47.9	1.9	0.2480	0.2460	0.0084	0.253
Upper Trout Creek	45.7	1.7	0.1936	0.1992	-0.0292	0.198
Surveyors Creek	38.7	1.7	0.2140	0.2138	0.0009	0.219
Fish Lake Stream	21.7	1.9	0.2915	0.2587	0.1148	0.299
Piscoe Creek	47.7	1.7	0.1844	0.1718	0.0692	0.189
Diamond Fork	44.8	1.5	0.1401	0.1402	-0.0008	0.143
Klickitat R.	45.8	1.8	0.2025	0.1915	0.0550	0.207
Colt Creek	46.9	1.7	0.1860	0.1842	0.0094	0.190
Lake Creek	46.9	1.7	0.1828	0.1820	0.0045	0.187
Little Bear Creek	53.7	1.8	0.2085	0.2170	-0.0412	0.213
Squaw Creek	42.8	1.9	0.2202	0.2242	-0.0184	0.225
Rock Creek	75.7	1.9	0.2232	0.2235	***(4)*** -0.0009	0.228
Omak Creek	56.8	1.9	0.2250	0.2118	0.0593	0.230
Lemhi R.	50.8	2.0	0.3181	0.3003	0.0564	0.325
Loon Creek	43.9	1.6	0.1862	0.1796	0.0361	0.190
Sulphur Creek	45.7	1.6	0.1908	0.1954	-0.0245	0.195
Chamberlain Creek	45.9	1.9	0.2075	0.2051	0.0117	0.212
White Bird Creek	45.9	1.8	0.2076	0.2028	0.0232	0.212
Valley Creek	45.9	1.9	0.2126	0.2133	-0.0033	0.217
West Fork Yankee Fork R.	88.6	1.9	0.2101	0.2103	-0.0010	0.215
North Fork Moose Creek	45.9	1.6	0.1964	0.1910	0.0282	0.201
Three Links Creek	45.9	1.7	0.1990	0.1990	0.0002	0.203
Captain John Creek	54.9	1.8	0.2096	0.2151	-0.0262	0.214
Meacham Creek	39.8	1.8	0.2090	0.2062	0.0136	0.214
Umatilla R.	51.8	1.9	0.2187	0.2154	0.0152	0.223
Nason Creek	33.8	1.9	0.2270	0.2271	-0.0005	0.232
Big White Salmon R.	84.3	2.0	0.3196	0.3051	0.0458	**0.327**
Canyon Creek	29.7	1.7	0.2248	0.2515	***(4)*** -0.1209	0.229
Luckiamute R.	30.7	1.7	0.2062	0.2246	-0.0907	0.210
Willamina Creek	30.9	1.8	0.2309	0.2360	-0.0225	0.236
---	---	---	---	---	---	---
Mean	46.6	1.8	0.2238	0.2212	0.0121	0.229

Section 3: Genetic Stock Identification of Chinook Salmon Mixtures in the mainstem Columbia River

Introduction

Columbia River Basin Chinook salmon consist of three major genetic lineages, which can be further broken into populations that are genetically structured on a relatively fine spatial scale (e.g., Waples et al. 2004). Genetic Stock Identification (GSI) methods have proven to be effective in determining the proportion of stock origin in several mixed stock applications of Chinook salmon (Shaklee et al. 1999, Beacham et al. 2006).
Chinook salmon fisheries in the mainstem Columbia River provide an ideal and important application of GSI because the fish harvested consist of mixtures of stocks from the entire Columbia River Basin. In addition, this represents a majority of the Columbia River Basin harvest of this species for commercial, sport, and tribal fisheries. In order to help establish sustainable fisheries, it is important to address two primary questions: 1) how are these stocks temporally and spatially distributed in the mainstem Columbia River; and 2) how are stocks distributed in the temporal and spatial harvests of sport, commercial, and tribal fisheries.
Thus, our study had two primary objectives: 1) Utilize GSI techniques to estimate stock composition of Chinook salmon passing Bonneville Dam; and 2) determine stock composition of Chinook salmon harvested in sport, commercial, and tribal fisheries in the mainstem Columbia River. To achieve this, samples of fish harvested from these fisheries over various temporal and spatial scales were necessary, as well as a representative sample of the total fish that are returning to the Columbia River Basin. Specifically for this study, we sampled fish at a fixed point at Bonneville Dam to represent the total fish present in the mainstem Columbia and also fish harvested in sport, commercial, and tribal fisheries of spring and fall-run Chinook salmon.
In this study, we utilized newly developed single nucleotide polymorphism (SNP) loci to genotype unknown mixture samples from five different sources to estimate stock composition of these mixtures, by fishery and strata. The use of SNP markers could make in-season GSI applications possible in future years.

Methods

Tissue collection
Tissues were sampled from a total of five different mixture sources: 1) Bonneville Dam (entire run), 2) Sport Fishery (spring and fall-run), 3) Test Fishery (spring-run), 4) Commercial Fishery (spring-run), and 5) Tribal Fishery (fall-run). Chinook salmon were non-lethally sampled at the

Bonneville AFF throughout the run from April until October of 2009 (Table 1, 2).
Biological data such as species, date, length, presence/absence of adipose fin, were recorded for each individual that was tissue-sampled for genetic analysis. We collected tissue samples, in the form of fin punches from each Chinook salmon. After non-lethal sampling is completed, all fish were released to a recovery pond and then to the fish ladder to continue upstream migration.
Tissues samples were preserved in ethanol or dried on whatman filter paper (Lahood et al. 2008) before being shipped to the Hagerman Fish Culture Experiment Station for GSI and estimation of stock composition. This sampling effort is covered under Scientific Research Permit #1379 under Section 10 of the ESA (permit included in PISCES attachments).

Table 1. Details of strata of fall run Chinook salmon mixture sources. Week 32 equals dates 8/3/2009-8/9/2009

		Week																			
Mixture Source	N	32	33	34	35	36	37	38	39	40	41	42	43	Male	Female	Hatchery	NonMarked	Male%	Female%	Hatchery%	Unmarked%
Buoy10	81		50	30	1									56	26	27	24	68.3	31.7	52.9	47.1
Zone6_Region01	1225			101	185	117	182	186	175	233	12	29	5	434	656	520	688	39.8	60.2	43.0	57.0
Zone6_Region02	1572			177	313	301	420	181	155	25				590	844	257	1301	41.1	58.9	16.5	83.5
Bonnneville Dam	1111		1	46	136	120	193	210	180	109	94	22		1692	1165	1508	1433	59.2	40.8	51.3	48.7

Table 2. Details of strata of spring run Chinook salmon mixture sources. Week 9 equals 2/13/2009-3/1/2009.

		Week																		
Mixture Source	N	9	10	11	12	13	14	15	16	17	18	Male	Female	Hatchery	NonMarked	Male%	Female%	Hatchery%	Unmarked%
Commercial	655						150	324	181			-	-	600	55	-	-	91.6	8.4
Sport	1287		22	102	195	219	272	182	253	42		-	-	1283	4	-	-	99.7	0.3
Test	271			4	22	15	50	11	3	71	95	-	-	215	56	-	-	79.3	20.7
Bonnneville Dam	102									39	63	39	45	85	17	46.4	53.6	83.3	16.7

Tissues were sampled in 2009 from Chinook salmon fisheries with existing programs in place with Washington Department of Fish and Wildlife (WDFW), Oregon Department of Fish and Wildlife (ODFW), and Yakama Nation. Representative tissue samples were obtained from commercial, sport, and test fisheries of spring-run Chinook, and tribal and sport fisheries from fall-run Chinook. The spring-run fisheries were all sampled below Bonneville Dam (Figure 1) and the fall-run fisheries were sampled both above Bonneville Dam (Zone 6 Tribal fishery) andbelow Bonneville Dam (Buoy 10 Sport fishery). The tribal fall-run and the sport spring-run Chinook fisheries were sampled in large enough numbers over multiple weeks that made it possible to further subdivide the locations for a finer spatial analysis- i.e. the Sport and Tribal fisheries were divided into Regions A and B and Region01 and Region02, respectively. Regions A and B, correspond to our grouping of pre-existing Oregon and Washington state Sport fishing zones 1-4 (or commercial zones 4-5) and a grouping of Sport zones 5-10 (or commercial zones 1-3), respectively. Region01 and Region02 in the Zone 6 fishery correspond to pre-existing Oregon and Washington state fishing zone 61 and a grouping of zones 62 and 63, respectively. These sets of groupings were established for this study in order to achieve balanced sampling for analysis of these fishery datasets, as well as to set an appropriate spatial scale of analysis to minimize variance of our estimates of stock proportions over temporal strata.

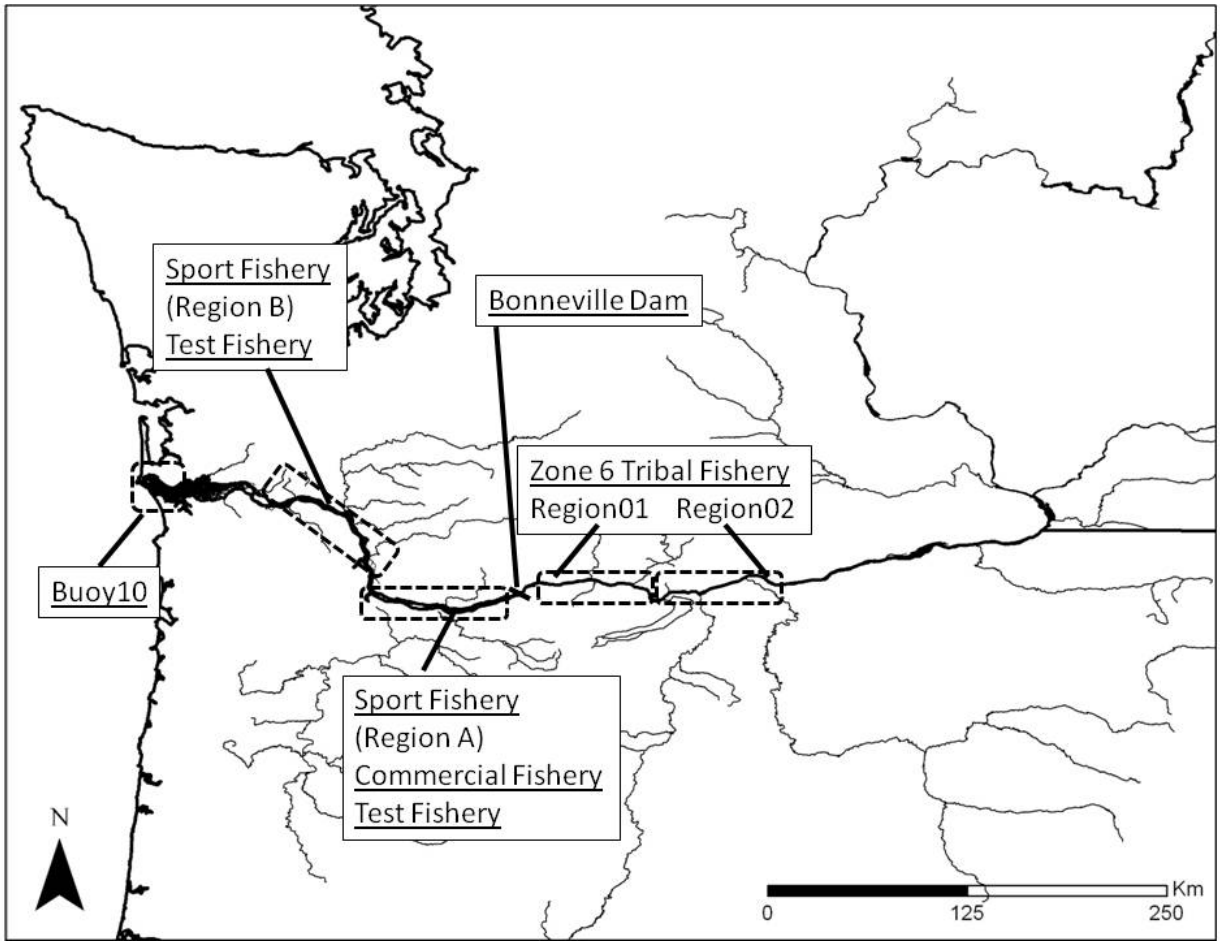


Figure 1. Map of sources of Chinook salmon mixtures

Fisheries on spring-run Chinook salmon have the advantage of widespread hatchery markings to distinguish hatchery fish from natural origin fish based on the absence or presence of the adipose fin, respectively. These adipose markings make it possible to have a mark selective sport and commercial fishery, in which only fish with missing adipose fins (hatchery-origin) are legally retained. In addition to sampling hatchery-origin fish from the mark selective commercial and sport fisheries, we were able to obtain samples from natural-origin fish before they were released from commercial and sport fishing vessels via the monitoring program of WDFW and ODFW. Samples of natural-origin fish were also obtained from the Test fishery operated by WDFW and ODFW. We compared results from GSI analyses of mixtures that included both natural-origin fish and hatchery-origin fish to the results from the same mixtures that contained only natural-origin fish to determine whether there was a significant effect of "origin" on stock proportions.

Molecular data

The following 94 SNP loci were used for genotyping fish for GSI analyses: Ots_102414-395, Ots_105105-613, Ots_106747-239, Ots_110064-383, Ots_nramp-321, Ots_113242-216,

Ots_113457-40R, Ots_123048-521, Ots_128757-61R, Ots_94857-232R, Ots_94903-99R, Ots_96222-525, Ots_96500-180, Ots_96899-357R, Ots_97077-179R, Ots_AldB1-122, Ots_aldb-177M, Ots_ARNT, Ots_arp-436, Ots_AsnRS-60, Ots_aspat-196, Ots_C3N3, Ots_Cath_D141, Ots_CCR7, Ots_CD59-2, Ots_CD63, Ots_cox1-241, Ots_CRB211, Ots_E2-275, Ots_EndoRB1-486, Ots_EP-529, Ots_ETIF1A, Ots_FARSLA-220, Ots_FGF6A, Ots_FGF6B_1, Ots_GDH-81x, Ots_GH2_1, Ots_GnRH-271, Ots_GPDH-338, Ots_GPH-318, Ots_GST-207, Ots_GST-375, Ots_GTH2B-550, Ots_hsc71-3'-488, Ots_hsc71-5'-453, Ots_hsp27b-150, Ots_HSP90B-100, Ots_IGF-I.1-76, Ots_Ikaros-250, Ots_IL11, Ots_IL8R_C8, Ots_mapK-3'-309, Ots_mapKpr-151, Ots_MHC1, Ots_MHC2, Ots_mybp-85, Ots_Myc-366, Ots_myo1a-384, Ots_myoD-364, Ots_nkef-192, Ots_NOD1, Ots_OPLW173_1, Ots_Ots311-101x, Ots_P450, Ots_P53, Ots_PGK-54, Ots_Prl2, Ots_RAG3, Ots_RAS1, Ots_RFC2-558, Ots_S7-1, Ots_SClkF2R2-135, Ots_SL, Ots_SWS1op-182, Ots_TAPBP, Ots_TGFB, Ots_TLR3, Ots_TNF, Ots_Tnsf, Ots_u07-07.161, Ots_u07-17.135, Ots_u07-18.378, Ots_u07-20.332, Ots_u07-25.325, Ots_u07-49.290, Ots_u07-53.133, Ots_u07-57.120, Ots_u07-64.221, Ots_u202-161, Ots_u211-85, Ots_u4-92, Ots_u6-75, Ots_unk526, and Ots_ZR-575. In addition, a single locus linked to the Y chromosome (Ots_SEXY1) was used to determine the sex of fish from the fall-run Chinook fishery mixtures as well as Bonneville Dam. See previous sections on SNP marker and baseline development for details of laboratory protocols involved in using these DNA markers for genotyping fish.

Statistical Analyses

SNP genotype data was utilized to estimate stock composition using the most recent version of the Chinook salmon SNP baseline. We grouped 51 baseline populations into reporting groups that were determined primarily by the relative genetic similarity among populations according to a phylogenetic analysis. Genetic distances were computed from allele frequencies according to Cavalli-Sforza and Edwards (1967), with the PHYLIP v 3.69 (Felsenstein 1989) executable "Gendist", after processing of dataset for bootstrapping (1000 replicates), using PHYLIP executable "Seqboot". Distances were clustered using the Neighbor – Joining method (Saitou and Nei, 1987), and a consensus tree was built with the PHYLIP executable "Consense"). The Phylip package is available at the following link:
http://evolution.genetics.washington.edu/phylip/

Mixture simulations were examined with the program ONCOR v1.0 (available at http://www.montana.edu/kalinowski) using the "100% simulations" feature to evaluate the power of the baseline to analyze mixture samples at the reporting group level (Anderson et al. 2008). A threshold of 90% correct assignment has been demonstrated to be an appropriate criterion for assessing the power to assign back to reporting groups (Beacham et al. 2006; Seeb et al. 2007). For these 100% simulations we set the parameters of mixture sample size and number of iterations to the values of 200 and 1000, respectively. Genotypes from fisheries mixtures were also analyzed in ONCOR to estimate stock composition by fishery and strata. We analyzed all strata that had $n > 20$ samples. These mixture proportions were generated with 95% confidence intervals using 1000 bootstraps.

Table 3. Sample sizes and reporting groups of baseline populations. See section 2.

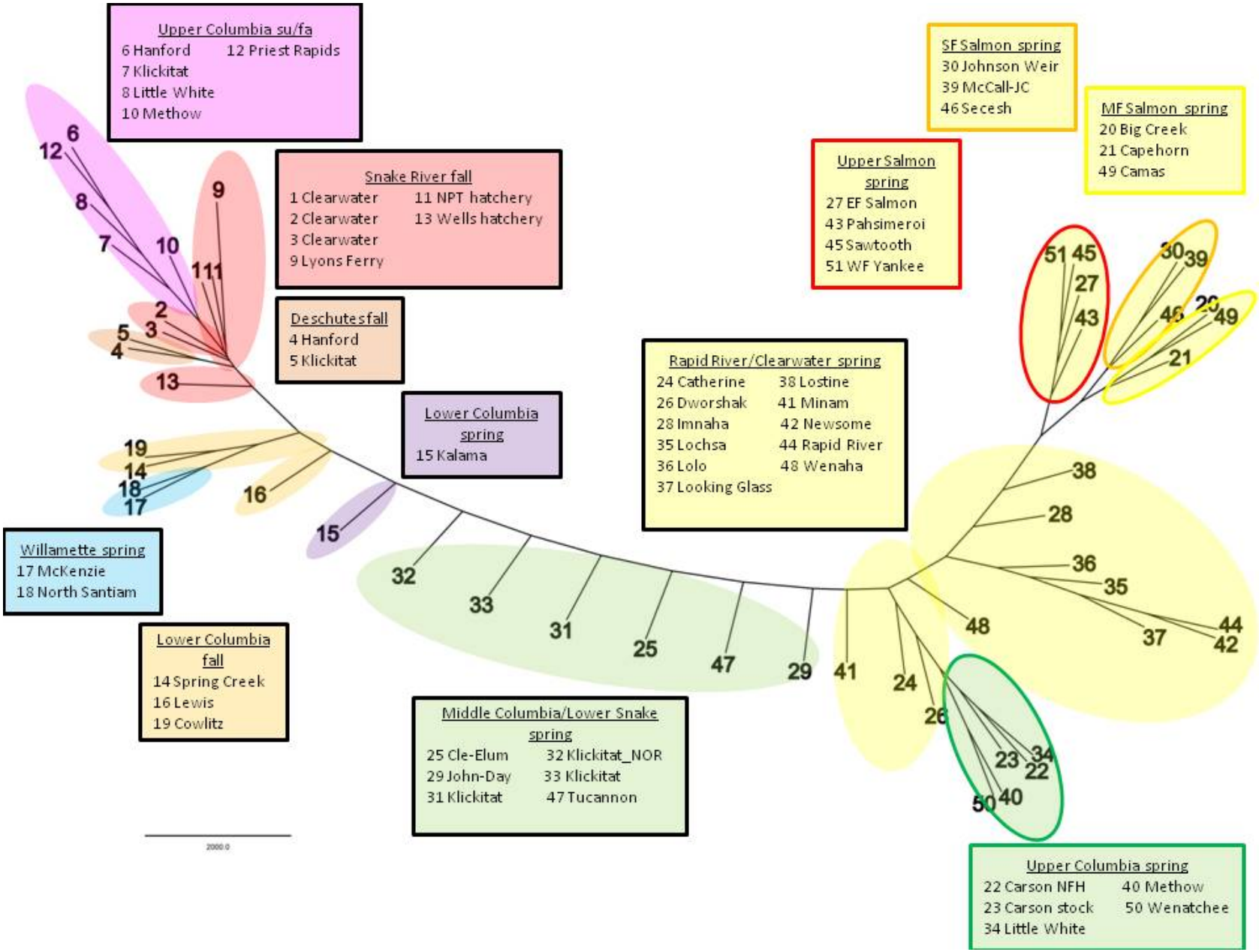
Baseline Population	Sample N	Reporting Group
Kalama_H_sp	81	L_Columbia_R_sp
McKenzie_sp	84	Willamette_R
NSantiam_sp	87	Willamette_R
CleElum_sp	88	Mid_Columbia_R_sp
JohnDay_sp	84	Mid_Columbia_R_sp
Klickitat_H_sp	157	Mid_Columbia_R_sp
Klickitat_sp	187	Mid_Columbia_R_sp
Klickitat_brdstk_H_sp	129	Mid_Columbia_R_sp
Tucannon_sp	87	Mid_Columbia_R_sp
CarsonNFH_sp	91	U_Columbia_R_sp
Carson_stock	84	U_Columbia_R_sp
LittleWhite_sp	92	U_Columbia_R_sp
Methow_sp	93	U_Columbia_R_sp
Wenatchee_sp	85	U_Columbia_R_sp
Catherine_sp	85	RapidR_Clearwater_sp
Dworshak_H_sp	88	RapidR_Clearwater_sp
Imnaha_sp	92	RapidR_Clearwater_sp
Lochsa_sp	77	RapidR_Clearwater_sp
Lolo_sp	89	RapidR_Clearwater_sp
LookingGlass_sp	89	RapidR_Clearwater_sp
Lostine_sp	82	RapidR_Clearwater_sp
Minam_sp	82	RapidR_Clearwater_sp
Newsome_sp	90	RapidR_Clearwater_sp
RapidRiver_sp	93	RapidR_Clearwater_sp
Wenaha_sp	44	RapidR_Clearwater_sp
JohnsonWeir_sp	92	SF_Salmon_sp
McCall_JC_sp	88	SF_Salmon_sp
Secesh_sp	81	SF_Salmon_sp
BigCreek_sp	92	MF_Salmon_sp
CapeHorn_sp	88	MF_Salmon_sp
Camas_sp	47	MF_Salmon_sp
EF_Salmon_sp	94	Upper_Salmon_sp
Pahsimeroi_sp	93	Upper_Salmon_sp
Sawtooth_sp	91	Upper_Salmon_sp
WF_Yankee_sp	75	Upper_Salmon_sp
Hanford_fa	93	U_Columbia_R_su/fa
Klickitat_fa	92	U_Columbia_R_su/fa
LittleWhite_fa	91	U_Columbia_R_su/fa
Methow_su	88	U_Columbia_R_su/fa
PriestRapids_fa	85	U_Columbia_R_su/fa
Wells_fa	89	U_Columbia_R_su/fa
SpringCreek_fa	88	L_Columbia_R_fa
Lewis_fa	93	L_Columbia_R_fa
Cowlitz_H_fa	86	L_Columbia_R_fa
L_Deschutes_fa	90	Deschutes_R_fa
Deschutes_H_fa	90	Deschutes_R_fa
ClearwaterCherry_fa	213	Snake_R_fa
Clearwaterjuv_fa	118	Snake_R_fa
Clearwater_fa	73	Snake_R_fa
LyonsFerry_fa	90	Snake_R_fa
NPTH_fa	86	Snake_R_fa

Results

Power analysis of baseline

The 51 collections were grouped into 12 reporting groups based on the clustering we observed in the phylogenetic analysis (Table 3, Figure 2). The ONCOR 100% simulations power analysis revealed that the majority (45 of 51; 88.2%) of baseline collections were found to correctly assign to their expected reporting group above the 90% threshold (Figure 3). However, six collections were below 90% correct assignment and of those, two were significantly below 90% (as indicated by their upper 95% confidence interval). These two collections were John-Day R. (middle Columbia R. sp) and Dworshak Hatchery (Rapid R./Clearwater R. sp). These collections may be affected by high straying rates (e.g. John-Day R.) or out-of-basin hatchery stock sources (e.g. upper Columbia R. fish in Dworshak Hatchery).

Figure 2. Neighbor Joining tree of Chinook salmon baseline populations based on Cavali-Sforza genetic distance of SNP loci. The boxes indicate the twelve reporting groups that were used to group populations based on genetic similarity, life history, and geographic proximity.



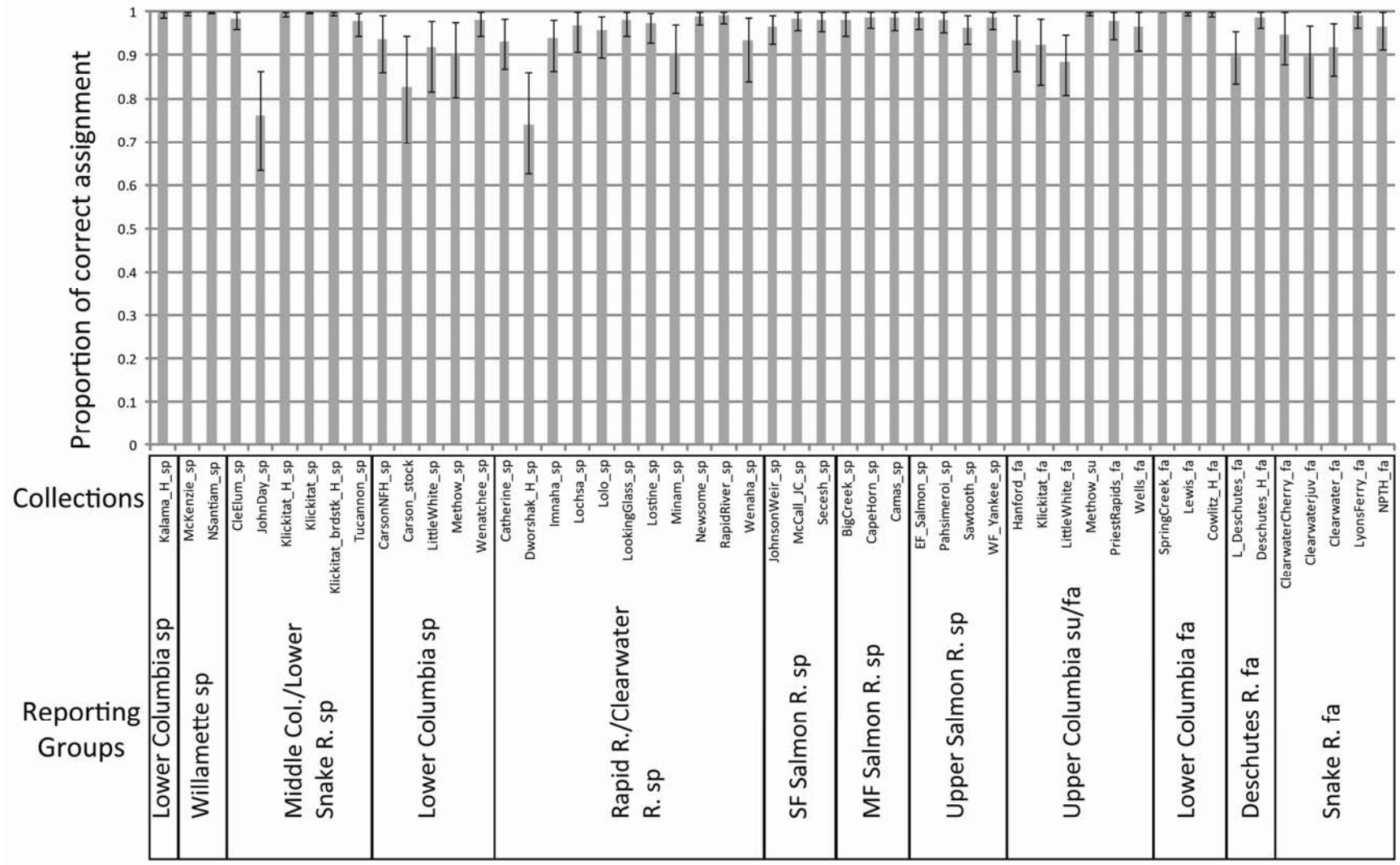


Figure 3. Results from correct assignment of 100% simulations for reporting groups.

Estimation of stock composition

Comparisons of total stock composition across the different sources of fishery mixtures revealed that these various mixtures are distinguished from each in the following two ways: 1)Some of the mixture sources contain significant proportions of stocks that are absent from other sources, or 2) A specific stock was present in all mixture sources but in significantly different proportions.

Stock proportions of the spring-run Chinook mixture sources

For the spring-run Chinook salmon collected from four mixture sources (three fisheries and Bonneville Dam), we found that the Willamette River spring-run was one of the stocks present in some mixtures but lacking in others. This stock was present in the Test and Sport fisheries, but absent in the Commercial fishery and Bonneville Dam (Figure 4).

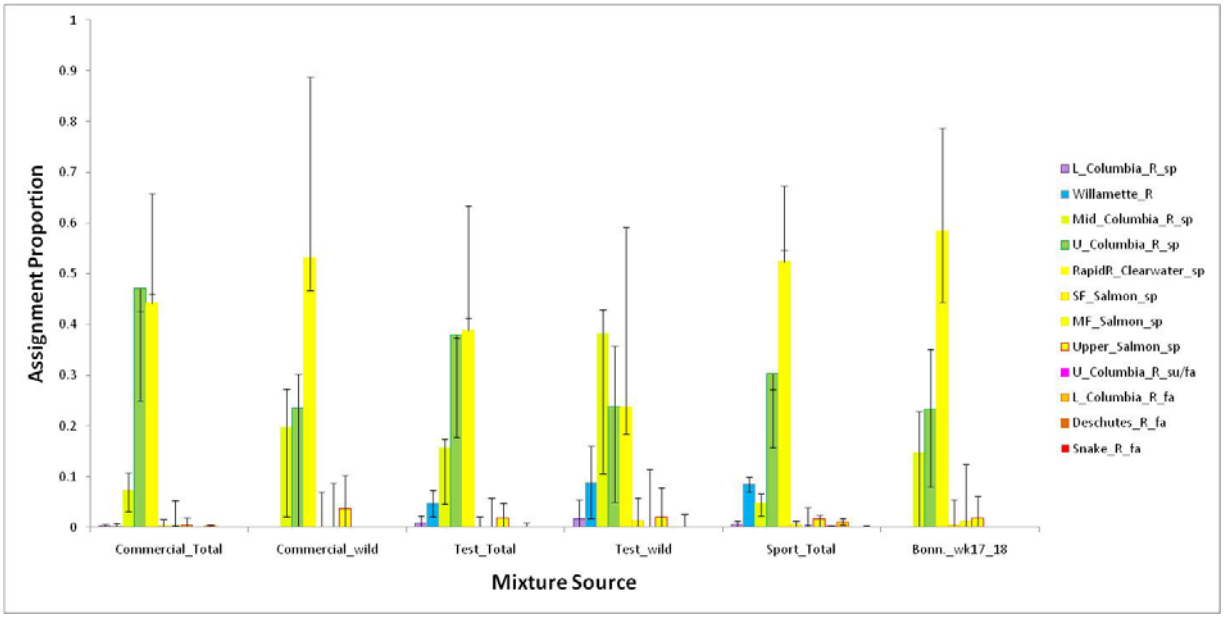
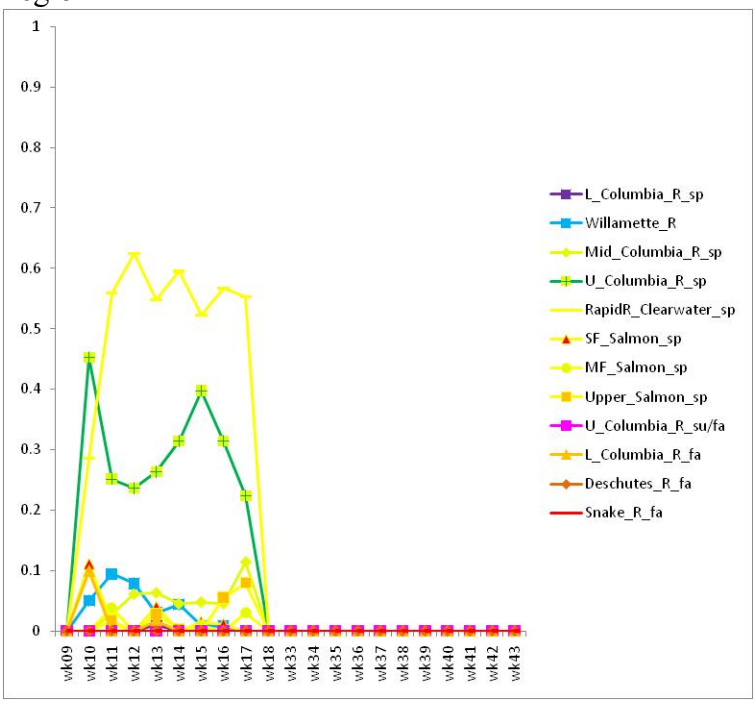


Figure 4. Comparison of stock proportions across all spring-run Chinook salmon mixture sources. Only comparable weeks (weeks 17-18) from Bonneville Dam (Bonn.) were included.

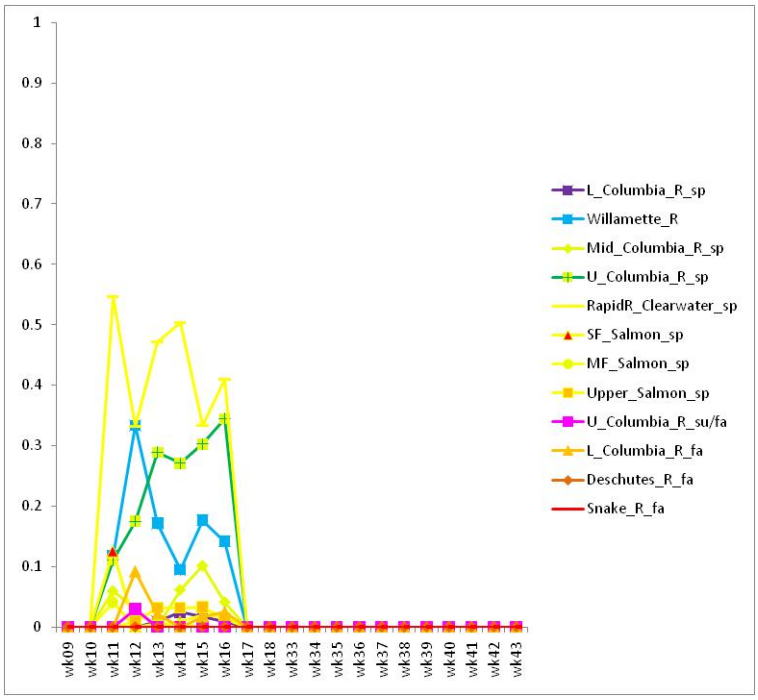
The run-timing and spatial distribution of the Willamette River spring stock could be examined using the sport fishery, which was sampled across the most weeks of any of the spring-run Chinook mixture sources. We found that the Willamette R. stock was highest in abundance in Region B of the sport fishery (Figure 5) and that the stock appears to peak early around week 12. This temporal and spatial distribution explains the lack of this stock's presence in the commercial fishery because this fishery harvested from only Region A and later in the season during weeks 14-16. On the other hand, the Test fishery mixture did show a presence of the Willamette R. stock because it harvested fish from both Region A and B, as well as spanning a wide range of weeks 11-18.

Figure 5. Sport spring-run Chinook salmon fishery by time strata and region. Region A has a smaller proportion of Willamette R. stocks than Region B due to the relative location of each region.

Region A



Region B



Hatchery and natural origin effects in spring-run Chinook salmon fisheries

We examined whether mixtures composed entirely of natural-origin fish from the spring-run commercial and test fisheries had significantly different stock composition compared to those fishery mixtures when both hatchery and natural-origin fish were included (Figure 4). There were only enough samples to compare a natural-origin (released fish) from the Commercial Fishery and Test Fishery. In both, cases comparisons between natural-origin composition and the total harvest showed significant differences in stock proportions, however each fishery was affected in opposite ways. In the Commercial Fishery, the Rapid R./Clearwater stock became significantly greater in proportion for the natural-origin-only mixture compared to the other two major stocks, upper Columbia and middle Columbia R. In the Test Fishery, the Rapid R./Clearwater stock moved from being significantly greater in proportion for the total mixture to being equal proportion with the other two major stocks, upper Columbia and middle Columbia R, in the natural-origin mixture.

Stock proportions of the fall-run Chinook salmon mixture sources

Proportions of the Lower Columbia R. fall-run stock were some of the main characteristics that distinguished all the fall-run Chinook salmon mixture sources (Figure 6).

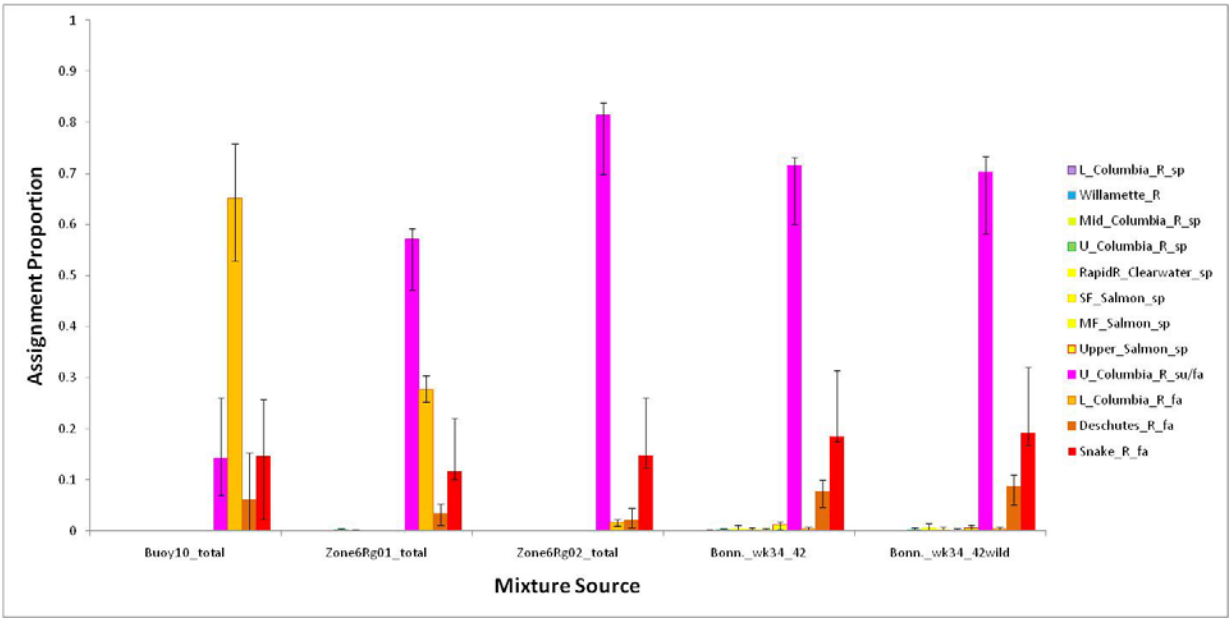
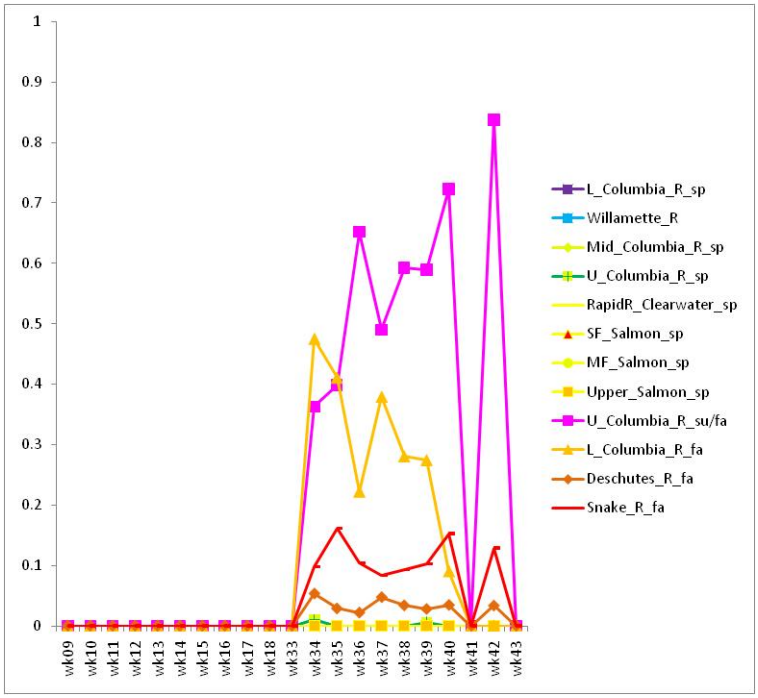


Figure 6. Comparison of stock proportions across all fall-run Chinook salmon mixture sources. Zone 6 was divided into two regions (Rg01 and Rg02). Only comparable weeks (weeks 34-42) from Bonneville Dam (Bonn.) were included.

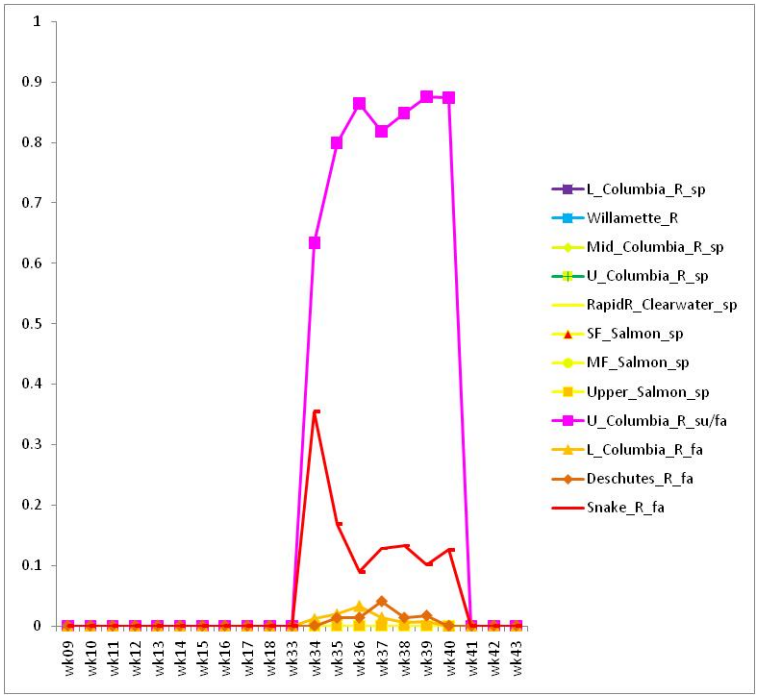
This stock represented a relatively large proportion of the mixtures from the Buoy 10 (0.65, CI 0.53-0.76) and Region 01 of the Zone 6 Tribal Fishery (0.28, CI 0.25-0.30), however it represented a much smaller or insignificant proportion of the mixtures from Bonneville Dam (0.00, CI 0.00-0.01) and Region 02 of the Zone 6 Tribal Fishery (0.02, CI 0.01-0.02). All four of the fall-run Chinook reporting groups (Upper Columbia summer/fall, Lower Columbia fall, Deschutes fall, and Snake River fall stocks) were represented in proportions with confidence intervals significantly greater than zero in all the mixture sources, except for the Deschutes fall-run stock, which was not significant in the Buoy 10 mixture source. In addition, the proportion of Upper Columbia summer/fall stock was significantly greater than all other stocks for all mixture sources, except for Buoy 10, in which the Lower Columbia fall stock represented a significantly greater proportion. This difference is likely attributed to run-timing. The Buoy 10 mixture was obtained during just two weeks (weeks 33 and 34) early in fall which is when the Lower Columbia fall stock appears to approach its peak level. This trend is best seen in Region 01 of the Zone 6 Tribal Fishery, in which we examined the temporal trends of using a weekly time strata series (Figure 7). We found that the Lower Columbia fall stock represents a larger proportion of the mixtures during weeks 34 to 35, and then subsides, while the Upper Columbia summer/fall stock dominates the rest of the time series.

Figure 7. Stock proportion trends across weekly strata in Region 01 and Region 02 of the Zone 6 Tribal fall Chinook fishery. In Region 01, the Lower Columbia fall Chinook stock appears to peak early in the fall and become less significant by week 40. In Region 02, Snake R. fall Chinook peak in 34 and decline in subsequent weeks. For reference, Week 09 equals dates 2/13/2009-3/1/2009 and Week 43 equals dates 10/19/2009-10/25/2009.

Region01



Region02



Sex-ratio bias in fall-run Chinook salmon mixture sources

The sex ratios in the Buoy 10 sport fishery mixture as well as the Bonneville Dam mixture were found to be skewed toward males (68.3% and 59.2% males in these sources respectively), whereas both Region01 and Region02 of the Zone 6 Tribal Fishery were skewed toward females (60.2% and 58.9% females in these sources, respectively). The Buoy10 mixture was not representative of the entire fall-run because it was only sampled during the weeks 33-34 and so this mixture's sex bias may be affected by run-timing differences among sexes. However, this may not be the only reason for detecting sex bias, because the Bonneville Dam and Zone 6 fishery mixtures were sampled over a much longer, more representative portion of the fall-run. Perhaps most surprising is the fact that although Bonneville Dam is geographically very close to Region01 of the Zone 6 fishery, the mixtures from these two areas were sex-biased in opposite ways. This is likely to due sampling bias for or against jacks (3 year old males) at Bonneville Dam and the Tribal Fishery, respectively. The mixture from Bonneville Dam is assumed to represent the total number of fish present in the river which is typically biased towards males and trap configuration in 2009 at the Adult Fish Facility may have disproportionately sampled jacks. Conversely, the Zone 6 Fishery is likely biased against males due to factors in the gill net fishing gear such that net size may be selecting for few jacks.

Discussion

Management implications

This study demonstrates great potential for the application of genetic stock identification in the management of mainstem Columbia River Chinook salmon fisheries evidenced by high accuracy of stock assignment, and the ability to discriminate Chinook salmon fishery mixtures by their relative stock proportions. The genetic methods were also able to detect differences in stock composition of natural-origin versus hatchery/natural-origin fishery mixtures and possibly indicate sex-selective properties of these different fisheries.

Spring-run Chinook salmon harvested in commercial, sport, and test fisheries were primarily composed of three adipose-clipped stocks (in descending order of stock composition): Rapid River Hatchery/Clearwater R., Upper Columbia R. (i.e., Carson stock), and Mid-Columbia R and these Chinook salmon stocks were also the most strongly represented at Bonneville Dam. During the spring Chinook test and sport fisheries, a fourth stock, Williamette R., was found because these fisheries include harvests spanning an earlier part of the season and locations closer to the mouth of the Columbia R.

For fall Chinook fisheries, the sport fishery at Buoy10 had predominantly Lower Columbia fall stocks (>60% composition), and less than 20% composition of the following stocks (in descending order): Snake R. fall, upper Columbia R. summer/fall, and Deschutes R. fall. The entire Zone 6 tribal Chinook fishery was heavily comprised of Upper Columbia R. summer/fall stock (60-80% depending on region), but Region 1 (closest region to Bonneville Dam) of Zone 6 contained more Lower Columbia R. fall stock (~30%) than Region 2 (< 5%), whereas Snake R. fall stock was similar in both regions (12-15%). The most variance in weekly composition of Chinook salmon stocks was observed in the Lower Columbia R. and Snake R. stocks which peaked early in the fall in Region01 and Region02 of Zone 6, respectively.

Future Directions

Known origin coded-wire tag (CWT) fish will be used to provide a measure of accuracy for GSI estimates. The GSI estimates may help refine CWT based estimates of stock composition used in fishery management. In addition, it may be possible to make in-season genetic stock composition estimates in future years.

References

Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for estimating the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65:1475-1486.

Beacham, T. D., J. R. Candy, K. L. Jonsen, J. Supernault, M. Wetklo, L. T. Deng, K. M. Miller, R. E. Withler, and N. Varnavskaya. 2006. Estimation of stock composition and individual identification of Chinook salmon across the Pacific Rim by use of microsatellite variation. Transactions of the American Fisheries Society 135(4):861-888.

Cavalli-Sforza, L. L. and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550-570.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics **5:** 164-166.

Lahood, E. S., J. J. Miller, C. Apland, and M. J. Ford. 2008. A rapid, ethanol-free fish tissue collection method for molecular genetic analyses. Transactions of the American Fisheries Society 137(4):1104-1007.

Narum, S. R., M. Banks, T. D. Beacham, et al. 2008. Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology 17(15):3464-3477.

Saitou, N., and M. Nei .1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425

Seeb, L. W., A. Antonovich, M. A. Banks, T. D. Beacham, M. R. Bellinger, S. M. Blankenship, M. R. Campbell, N. A. Decovich, J. C. Garza, C. M. Guthrie, T. A. Lundrigan, P. Morgan, S. R. Narum, J. J. Stephenson, K. J. Supernault, D. J. Teel, W. D. Templin, J. K. Wenburg, S. F. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32(11):540-552.

Shaklee, J. B., T. D. Beacham, L. Seeb, and B. A. White. 1999. Managing fisheries using genetic data: case studies form four species of Pacific salmon. Fisheries Research 43(1-3)45-78.

Section 4: Trends in Steelhead run-timing across multiple years of migrations of Columbia River using genetic stock identification

Introduction

The Columbia River Basin supports ESA listed wild stocks of steelhead as well as hatchery supplemented populations. Steelhead, like other salmonid species, have been declining in the Columbia River Basin for several reasons including climate change, hatchery practices, habitat degradation, over harvesting, and hydropower. Along with abundance estimates, basic information related to the way in which stocks of salmonids are distributed both spatially and temporally are needed by fisheries managers to achieve sustainable fisheries.

Steelhead populations have been found to be highly genetically structured in several river systems of the Columbia River Basin (Nielsen et al. 2009, Narum et al. 2006a,b), due to the high site fidelity of adults that spawn in their stream of origin. This high level of structuring makes it possible to use genetic stock identification to assign unknown fish to the population they originated, which has been successfully performed within the Columbia River Basin (e.g. the Snake River basin in Narum et al. 2008). Population genetic methods and statistical assignment models have advanced dramatically in recent years, and estimating stock composition is now possible using either Bayesian or Maximum Likelihood methods (Anderson et al. 2008). Coordination of sampling and a large effort to standardize genetic markers among a number of state, federal, and tribal agencies (Winans et al. 2004; Stephenson et al. 2009,) has now made it possible to perform GSI on a broad geographic scale that includes the entire Columbia River Basin.

The aim of this study was to determine whether genetic stock identification (GSI) analysis could accurately distinguish among stocks of steelhead within the entire Columbia River Basin and discriminate these stocks according to their peak run-timing. Specifically, we estimated stock proportions of unknown mixtures of steelhead, including both hatchery and wild fish, over five years at a fixed point (Bonneville Dam) that is located on the mainstem Columbia River. Since Bonneville Dam is the most downstream dam on the Columbia River, the mixtures obtained here represent a majority of Columbia River Basin stocks. Our study offers a rare opportunity to monitor a broad geographic scale of salmonid populations over several years. This long-term study will allow us to characterize trends in run timing and provide information regarding the consistency of anadromous steelhead across years to fisheries managers.

Methods

Sample Collection

Tissue samples were obtained from adult steelhead during migration runs at Bonneville Dam in each of five years between 2004 to 2008 (n = 715, 570, 1429, 1931, and 2545 respectively). Samples were pooled into 14 biweekly strata (mean n = 174, range 23-584 per strata) spanning the majority of the run-year from April to October (Table 1). The three years spanning 2006-2008 had nearly complete sampling across all 14 biweekly strata, however 2004-2005 were only sampled for the last 6 biweekly strata. Biological data such as species, date, length, presence/absence of adipose fin, were recorded for each individual that was tissue-sampled for genetic analysis. We collected tissue samples, in the form of fin punches from each steelhead.

After non-lethal sampling was completed, all fish were released to a recovery pond and then to the fish ladder to continue upstream migration. Tissues samples were preserved in ethanol or on Whatman filter paper (Lahood et al. 2008) before being shipped to the Hagerman Fish Culture Experiment Station for GSI and estimation of stock composition. This sampling effort is covered under Scientific Research Permit #1379 under Section 10 of the ESA (permit included in PISCES attachments).

Table 1. Sample numbers by biweekly strata for all years of sampling. Gray cells indicate sample size of n > 20.

Biweekly strata	2004	2005	2006	2007	2008	Week_totals
week16_17			16	4	7	27
week18_19			11	8	3	22
week20_21			24	42	23	89
week22_23			27	61	55	143
week24_25			25	69	38	132
week26_27			155	208	175	538
week28_29			144	186	521	851
week30_31			108	126	584	818
week32_33	83		82	118	275	558
week34_35	124	149	44	102	79	498
week36_37	162	130	358	204	418	1272
week38_39	198	168	282	437	238	1323
week40_41	117	118	120	303	125	783
week42_43	31	5	33	63	4	136

Molecular markers

The following thirteen standardized microsatellite loci were used to genotype steelhead mixtures and had been previously incorporated into a genetic baseline for performing genetic stock identification: Ogo4, Oke4, Oki23MMBL, Omy7, Omy1001, Omy1011, Oneu14, Ots3M, Ots4, Ots100, Ssa289, Ssa407, and Ssa408 (Stephenson et al. 2009).

Statistical analyses

Microsatellite genotype data was utilized to estimate stock composition using the Steven Phelps Allele Nomenclature (SPAN) version 1 baseline (map locations, Figure 1). We grouped 147 baseline collections into reporting groups that were determined primarily by the relative genetic similarity among populations according to a phylogenetic analysis (Table 2, Figure 2).

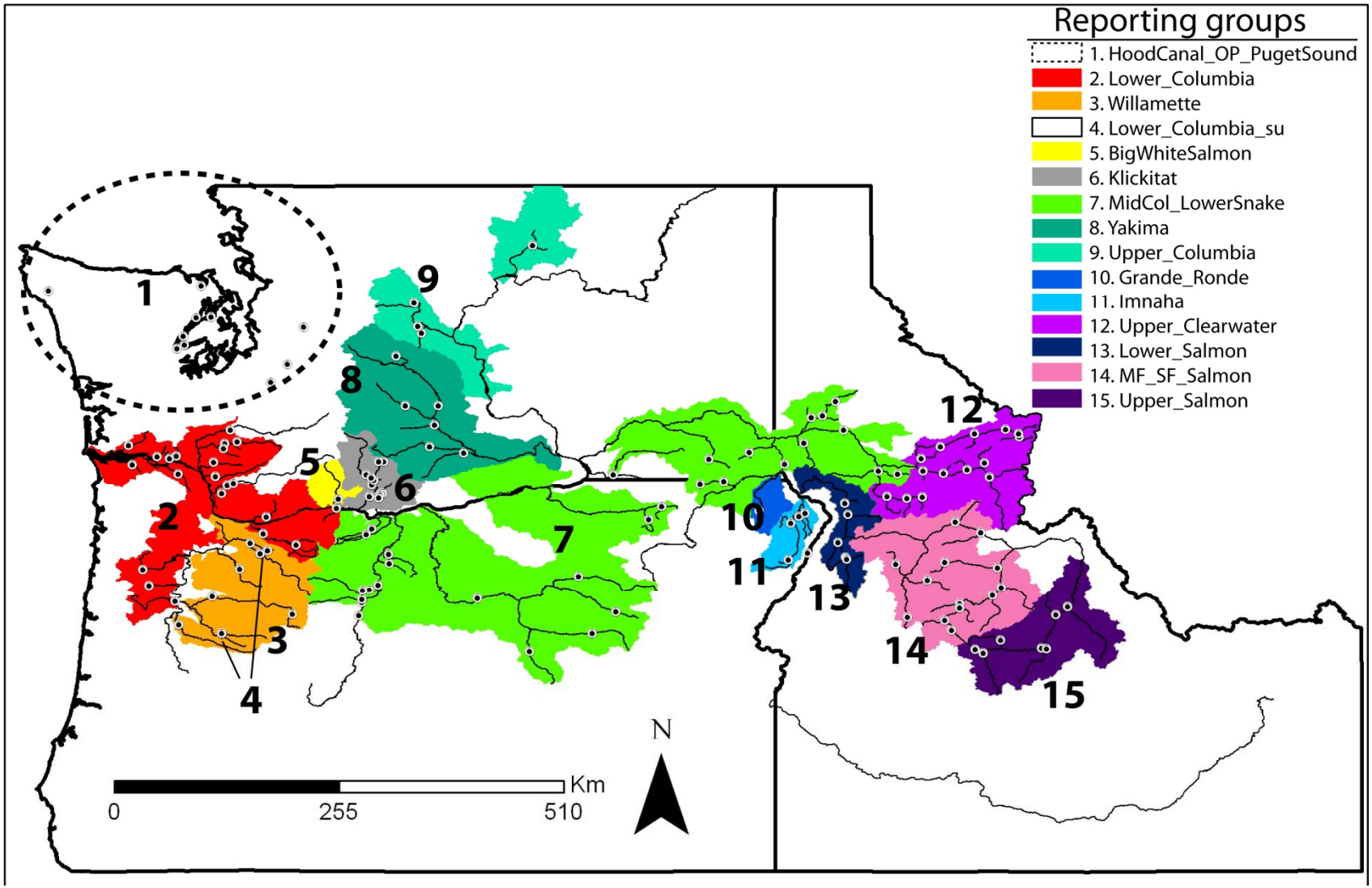


Figure 1. Map of the Columbia River Basin and collection locations of baseline samples.

ID#	Collection	N	Reporting Group	100% Avg. Correct Assignment	ST DEV	95% Confidence
1	Tahuya_R	84	HoodCanal_OP_PugetSound	0.9939	0.0055	(0.9800, 1.0000)
2	Skokomish_R	113	HoodCanal_OP_PugetSound	0.998	0.0027	(0.9925, 1.0000)
3	Duckabush_R	50	HoodCanal_OP_PugetSound	0.9519	0.0185	(0.9140, 0.9819)
4	Dewatto_R	114	HoodCanal_OP_PugetSound	0.9986	0.0024	(0.9923, 1.0000)
5	Big_Beef_Cr	83	HoodCanal_OP_PugetSound	0.9963	0.0044	(0.9875, 1.0000)
6	Bogachiel_H	89	HoodCanal_OP_PugetSound	0.9908	0.0085	(0.9720, 1.0000)
7	Snow_Cr	30	HoodCanal_OP_PugetSound	0.9887	0.0091	(0.9659, 1.0000)
8	Tokul_Cr_H	81	HoodCanal_OP_PugetSound	0.9882	0.0085	(0.9695, 1.0000)
9	Puyallup_R	64	HoodCanal_OP_PugetSound	0.9956	0.0045	(0.9842, 1.0000)
10	White_R	413	HoodCanal_OP_PugetSound	0.9971	0.0032	(0.9892, 1.0000)
11	Elochoman_R	80	Lower_Columbia	0.9949	0.0053	(0.9828, 1.0000)
12	Germany_Cr	80	Lower_Columbia	0.9853	0.0099	(0.9643, 1.0000)
13	Grays_R	77	Lower_Columbia	0.9944	0.0054	(0.9787, 1.0000)
14	Mill_Cr	85	Lower_Columbia	0.9928	0.0064	(0.9793, 1.0000)
15	Clatskanie_R	28	Lower_Columbia	0.9936	0.0066	(0.9755, 1.0000)
16	Coweeman_R	75	Lower_Columbia	0.9794	0.0104	(0.9564, 0.9946)
17	Cowlitz_R	91	Lower_Columbia	0.9976	0.0031	(0.9909, 1.0000)
18	Hood_R	79	Lower_Columbia	0.9223	0.0232	(0.8654, 0.9597)
19	Kalama_R_Su	113	Lower_Columbia	0.9669	0.0185	(0.9212, 0.9921)
20	Kalama_R_Wi	81	Lower_Columbia	0.9923	0.0055	(0.9813, 1.0000)
21	EF_Lewis_R	57	Lower_Columbia	0.9826	0.0115	(0.9603, 0.9999)
22	NF_Lewis_Cedar	50	Lower_Columbia	0.9877	0.0096	(0.9673, 1.0000)
23	NF_Lewis_Merwin	68	Lower_Columbia	0.9883	0.01	(0.9660, 1.0000)
24	Sandy_R_Marmot	80	Lower_Columbia	0.9634	0.0185	(0.9229, 0.9921)
25	Sandy_R_Still	20	Lower_Columbia	0.9532	0.0186	(0.9069, 0.9800)
26	Green_R	75	Lower_Columbia	0.9928	0.0058	(0.9782, 1.0000)
27	NF_Toutle_R	79	Lower_Columbia	0.999	0.0019	(0.9945, 1.0000)
28	SF_Toutle_R	70	Lower_Columbia	0.9951	0.0055	(0.9795, 1.0000)
29	Washougal_R	57	Lower_Columbia	0.9764	0.0136	(0.9438, 0.9948)
30	Big_Cr_H	37	Lower_Columbia	0.9683	0.0137	(0.9401, 0.9892)
35	Eagle_Cr_H	52	Lower_Columbia	0.9967	0.0044	(0.9822, 1.0000)
42	Canyon_Cr	34	Lower_Columbia	0.9848	0.0102	(0.9605, 0.9995)
43	Luckiamute_R	25	Lower_Columbia	0.9951	0.006	(0.9758, 1.0000)
44	Willamina_Cr	26	Lower_Columbia	0.9903	0.0071	(0.9730, 1.0000)
36	Clackamas_R_H	43	Lower_Columbia_su	0.9188	0.0236	(0.8749, 0.9582)
45	S_Santiam_H	36	Lower_Columbia_su	0.82	0.0367	(0.7363, 0.8939)
31	Calapooia_R	24	Willamette	0.9982	0.0029	(0.9911, 1.0000)
32	Clackamas_R_wild	59	Willamette	0.9976	0.0029	(0.9906, 1.0000)
33	Eagle_Cr_wild	52	Willamette	0.9887	0.0086	(0.9658, 1.0000)
34	NF_Dam	71	Willamette	0.9372	0.0221	(0.8913, 0.9729)
37	NF_Molalla	37	Willamette	0.987	0.0102	(0.9578, 1.0000)
38	Bennet_Dam	71	Willamette	0.9933	0.0059	(0.9780, 1.0000)
39	Marion_Forks_H	32	Willamette	0.9998	0.0009	(0.9959, 1.0000)
40	Foster_Dam	92	Willamette	0.9968	0.0034	(0.9880, 1.0000)
41	Wiley_Cr	33	Willamette	0.9917	0.0061	(0.9800, 1.0000)
75	Big_White_Salmon	54	BigWhiteSalmon	0.9779	0.0114	(0.9473, 0.9930)
62	Bowman_Cr	68	Klickitat	0.9519	0.0189	(0.9091, 0.9810)
63	Dead_Canyon_Cr	72	Klickitat	0.9416	0.0192	(0.9000, 0.9756)
64	Little_Klickitat_R_Low	31	Klickitat	0.9435	0.0206	(0.8929, 0.9714)
65	Snyder_Cr	44	Klickitat	0.9941	0.0056	(0.9814, 1.0000)
66	Summit_Cr_Low	41	Klickitat	0.9658	0.0141	(0.9351, 0.9860)
67	Swale_Cr	87	Klickitat	0.9797	0.0093	(0.9591, 0.9932)
68	Tepee_Cr	31	Klickitat	0.993	0.0058	(0.9780, 1.0000)
69	Trout_Cr	66	Klickitat	0.969	0.0128	(0.9372, 0.9861)
70	White_Cr_Low	32	Klickitat	0.9679	0.0144	(0.9318, 0.9903)
71	White_Cr_Up	30	Klickitat	0.9918	0.0062	(0.9785, 1.0000)
46	BakeOven_Cr	100	MidCol_LowerSnake	0.9684	0.0148	(0.9274, 0.9925)
47	BuckHollow_Cr	84	MidCol_LowerSnake	0.919	0.0249	(0.8696, 0.9535)
48	DeschutesR_H	43	MidCol_LowerSnake	0.9736	0.0143	(0.9445, 0.9944)
49	DRmainstem_N	136	MidCol_LowerSnake	0.9985	0.003	(0.9882, 1.0000)
50	DRmainstem_S	43	MidCol_LowerSnake	0.9916	0.0059	(0.9780, 1.0000)
51	Shitike_Cr_H	49	MidCol_LowerSnake	0.8833	0.0345	(0.8150, 0.9584)
52	Shitike_Cr_Wild	138	MidCol_LowerSnake	0.9058	0.0309	(0.8395, 0.9559)
53	Warm_Springs_R	121	MidCol_LowerSnake	0.9602	0.0168	(0.9206, 0.9880)
54	Warm_Springs_H	109	MidCol_LowerSnake	0.9003	0.0322	(0.8308, 0.9545)
55	John_Day_LowerMain	131	MidCol_LowerSnake	0.9292	0.028	(0.8644, 0.9721)
56	John_Day_MF	154	MidCol_LowerSnake	0.9714	0.0135	(0.9385, 0.9914)
57	John_Day_NF	138	MidCol_LowerSnake	0.9572	0.0182	(0.9159, 0.9873)
58	John_Day_SF	43	MidCol_LowerSnake	0.9485	0.0185	(0.9095, 0.9839)
59	John_Day_UpperMain	56	MidCol_LowerSnake	0.9232	0.0287	(0.8625, 0.9694)
60	Eightmile_Cr	39	MidCol_LowerSnake	0.8801	0.031	(0.8116, 0.9307)
61	Fifteenmile_Ramsey	44	MidCol_LowerSnake	0.9073	0.0247	(0.8510, 0.9461)
72	Meacham_Cr	63	MidCol_LowerSnake	0.9385	0.0216	(0.8926, 0.9758)
73	NF_Umatilla	20	MidCol_LowerSnake	0.9726	0.0136	(0.9411, 0.9924)

Table 2. List of collections in the genetic baseline and their reporting units. The ID#s of the collections correspond with population numbers in the phylogeny (Figure 2). The proportion of correct assignment and 95% confidence intervals were generated in ONCOR.

ID#	Collection	N	Reporting Group	100% Avg. Correct Assignment	ST DEV	95% Confidence
74	SF_Umatilla	30	MidCol_LowerSnake	0.9732	0.0157	(0.9370, 0.9960)
76	Touchet_R	47	MidCol_LowerSnake	0.9615	0.0176	(0.9204, 0.9897)
83	Cedar_Cr	46	MidCol_LowerSnake	0.9827	0.0097	(0.9585, 0.9977)
85	Cottonwood_Cr	77	MidCol_LowerSnake	0.9389	0.022	(0.8949, 0.9737)
86	EF_Potlatch_R	36	MidCol_LowerSnake	0.9003	0.0313	(0.8274, 0.9422)
87	Little_Bear_Cr	73	MidCol_LowerSnake	0.9483	0.0174	(0.9122, 0.9758)
88	Mission_Cr	34	MidCol_LowerSnake	0.9225	0.0284	(0.8647, 0.9692)
105	Crooked_Cr	84	MidCol_LowerSnake	0.9405	0.0247	(0.8858, 0.9795)
106	Wenaha_R	73	MidCol_LowerSnake	0.8926	0.0337	(0.8212, 0.9510)
123	Asotin_Cr	72	MidCol_LowerSnake	0.8039	0.0417	(0.7239, 0.8750)
124	Captain_John_Cr	40	MidCol_LowerSnake	0.9373	0.0239	(0.8900, 0.9777)
125	U_Tucannon	53	MidCol_LowerSnake	0.903	0.0281	(0.8403, 0.9442)
77	Ahtanum_Cr	65	Yakima	0.9197	0.019	(0.8835, 0.9555)
78	Naches_R	69	Yakima	0.8887	0.0307	(0.8260, 0.9437)
79	Satus_Cr	97	Yakima	0.948	0.0179	(0.9086, 0.9749)
80	Toppenish_Cr	64	Yakima	0.9821	0.0088	(0.9650, 0.9969)
81	Roza_Trap	19	Yakima	0.8141	0.0332	(0.7363, 0.8634)
82	Teanaway_R	37	Yakima	0.9094	0.0267	(0.8532, 0.9510)
142	Omak_Cr_H	198	Upper Columbia	0.9321	0.0215	(0.8866, 0.9709)
143	Omak_Cr_Wild	22	Upper Columbia	0.8484	0.0355	(0.7779, 0.9070)
144	Peshastin_Cr	82	Upper Columbia	0.6923	0.0462	(0.6107, 0.7824)
145	Nason_Cr	15	Upper Columbia	0.8948	0.033	(0.8335, 0.9517)
146	Wells_H	50	Upper Columbia	0.9559	0.0167	(0.9235, 0.9820)
104	Elk_Cr	90	Grande_Ronde	0.9497	0.0198	(0.9022, 0.9814)
107	Camp_Cr	92	Imnaha	0.961	0.0168	(0.9268, 0.9866)
108	Gumboot_Cr	74	Imnaha	0.9553	0.0176	(0.9142, 0.9817)
109	Horse_Cr	67	Imnaha	0.8738	0.0337	(0.7982, 0.9301)
110	Lightning_Cr	41	Imnaha	0.7957	0.0429	(0.6993, 0.8717)
84	CR_Clear_Cr	36	Upper_Clearwater	0.856	0.0321	(0.8002, 0.9087)
89	Dworshak_H	47	Upper_Clearwater	0.9845	0.0091	(0.9682, 1.0000)
90	CR_Canyon_Cr	63	Upper_Clearwater	0.9799	0.0115	(0.9532, 0.9974)
91	Colt_Cr	51	Upper_Clearwater	0.996	0.0048	(0.9826, 1.0000)
92	Fish_Cr	58	Upper_Clearwater	0.987	0.0075	(0.9724, 0.9989)
93	Lake_Cr	51	Upper_Clearwater	0.9961	0.004	(0.9850, 1.0000)
94	Lochsa_R	47	Upper_Clearwater	0.9938	0.0055	(0.9775, 1.0000)
95	Storm_Cr	26	Upper_Clearwater	0.9958	0.0041	(0.9861, 1.0000)
96	Bear_Cr	45	Upper_Clearwater	0.9899	0.0068	(0.9738, 1.0000)
97	Gedney_Cr	125	Upper_Clearwater	0.982	0.0098	(0.9598, 0.9960)
98	NF_Moose_Cr	80	Upper_Clearwater	0.9904	0.0076	(0.9718, 1.0000)
99	Three_Links_Cr	44	Upper_Clearwater	0.9929	0.0063	(0.9780, 1.0000)
100	Crooked_R_H	101	Upper_Clearwater	0.994	0.0058	(0.9802, 1.0000)
101	Crooked_R_Wild	80	Upper_Clearwater	0.9927	0.0073	(0.9750, 1.0000)
102	Johns_Cr	29	Upper_Clearwater	0.9162	0.0223	(0.8692, 0.9584)
103	Tenmile_Cr	47	Upper_Clearwater	0.9839	0.0102	(0.9616, 1.0000)
119	Ohara_Cr	47	Upper_Clearwater	0.9794	0.0112	(0.9574, 0.9979)
131	Boulder_Cr	46	Lower Salmon	0.8813	0.032	(0.8071, 0.9375)
132	Hazard_Cr	44	Lower Salmon	0.5224	0.044	(0.4320, 0.5995)
133	Rapid_R	256	Lower Salmon	0.9606	0.0151	(0.9273, 0.9865)
134	Slate_Cr	47	Lower Salmon	0.7672	0.0379	(0.6967, 0.8354)
135	White_Bird_Cr	86	Lower Salmon	0.8729	0.0298	(0.8131, 0.9205)
111	SR_Big_Cr	47	MF_SF_Salmon	0.9955	0.0052	(0.9849, 1.0000)
112	Camas_Cr	34	MF_SF_Salmon	0.9105	0.0219	(0.8633, 0.9463)
113	Loon_Cr	33	MF_SF_Salmon	0.9728	0.0154	(0.9371, 0.9947)
114	Lower_Big_Cr	27	MF_SF_Salmon	0.9437	0.0188	(0.8985, 0.9754)
115	Marsh_Cr	30	MF_SF_Salmon	0.9921	0.0064	(0.9789, 1.0000)
116	Pistol_Cr	15	MF_SF_Salmon	0.9405	0.0237	(0.8851, 0.9780)
117	SR_Rapid_R	45	MF_SF_Salmon	0.9644	0.018	(0.9304, 0.9942)
118	Sulphur_Cr	30	MF_SF_Salmon	0.9786	0.0103	(0.9592, 0.9957)
120	Johnson_Cr	44	MF_SF_Salmon	0.9387	0.0205	(0.8975, 0.9727)
121	Stolle_Meadows	27	MF_SF_Salmon	0.9735	0.0127	(0.9427, 0.9946)
122	Secesh_R	59	MF_SF_Salmon	0.9776	0.0128	(0.9502, 0.9967)
126	Bargamin_Cr	85	MF_SF_Salmon	0.899	0.0284	(0.8414, 0.9392)
127	Chamberlain_Cr	56	MF_SF_Salmon	0.794	0.0397	(0.7173, 0.8693)
128	EF_Salmon_R	39	Upper Salmon	0.97	0.0132	(0.9376, 0.9900)
129	Morgan_Cr	44	Upper Salmon	0.8398	0.0354	(0.7730, 0.8967)
130	Squaw_Cr	44	Upper Salmon	0.7965	0.041	(0.7233, 0.8745)
136	Pahsimeroi_R_Wild	85	Upper Salmon	0.8679	0.0413	(0.7955, 0.9363)
137	Pahsimeroi_R_H	85	Upper Salmon	0.9558	0.0194	(0.9157, 0.9848)
138	Valley_Cr_UpLow	41	Upper Salmon	0.889	0.0317	(0.8342, 0.9424)
139	WF_Yankee_Fork	88	Upper Salmon	0.9549	0.0164	(0.9198, 0.9857)
140	Sawtooth_Wild	25	Upper Salmon	0.8742	0.0332	(0.8010, 0.9250)
141	Sawtooth_H	47	Upper Salmon	0.9152	0.0293	(0.8535, 0.9609)
147	Oxbow_H	43	Upper Salmon	0.8391	0.0387	(0.7526, 0.9047)

Table 2 (continued). List of collections in the genetic baseline and their reporting units.

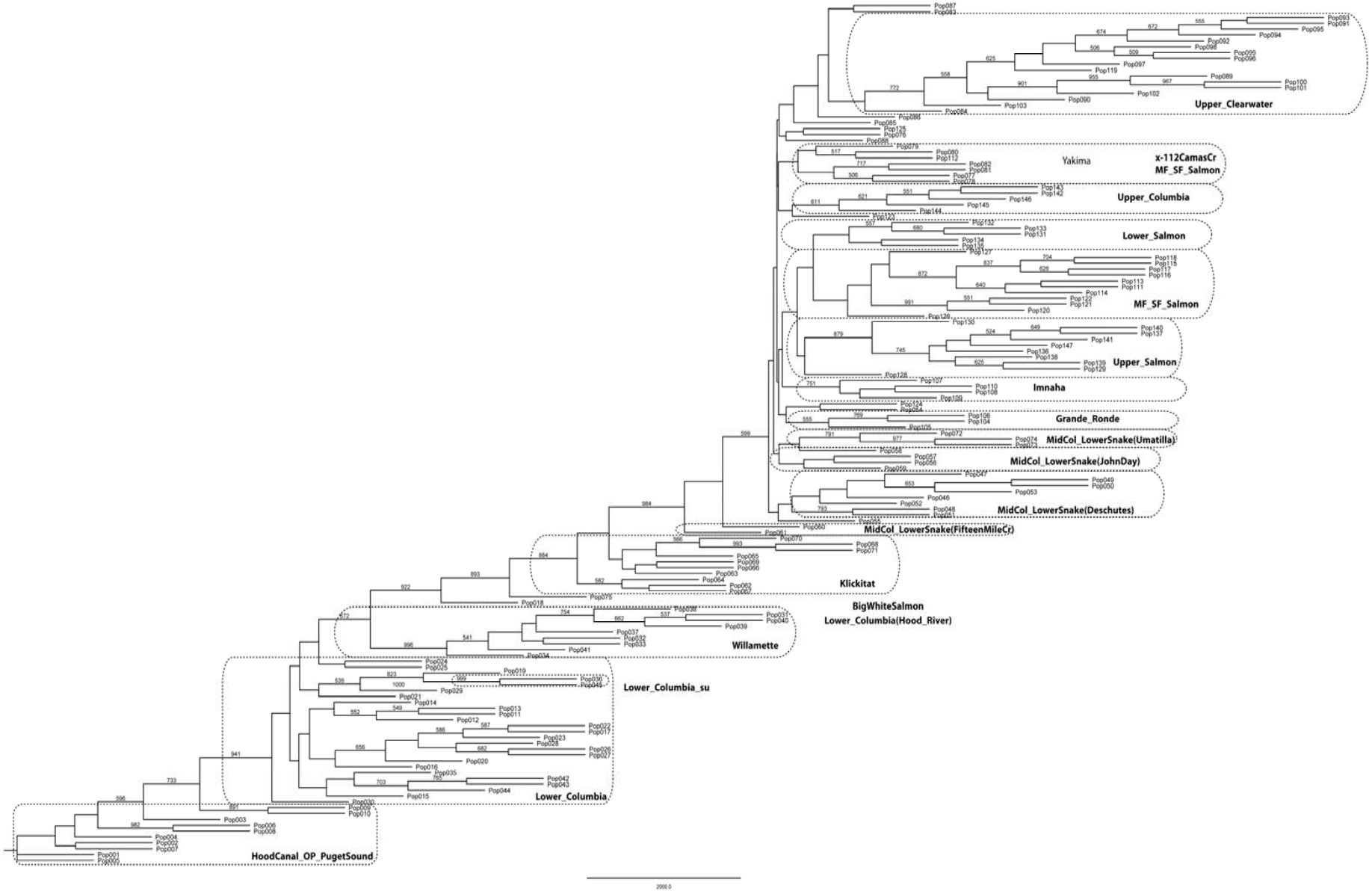


Figure 2. Neighbor-joining dendrogram of baseline collections. The bootstrap values above 50% are indicated on the branches. The dotted circles enclose reporting groups.

Genetic distances were computed from allele frequencies according to Cavalli-Sforza and Edwards (1967), with the PHYLIP v 3.69 (Felsenstein 1989) executable "Gendist", after processing of dataset for bootstrapping (1000 replicates), using PHYLIP executable "Seqboot". Distances were clustered using the Neighbor – Joining method (Saitou and Nei 1987), and a consensus tree was built with the PHYLIP executable "Consense"). The Phylip package is available at the following link: http://evolution.genetics.washington.edu/phylip/

Mixture simulations were examined with the program ONCOR v1.0 (available at http://www.montana.edu/kalinowski) using the "100% simulations" feature to evaluate the power of the baseline to analyze mixture samples at the reporting group level (Anderson et al. 2008). A threshold of 90% correct assignment has been demonstrated by Seeb et al. (2007) to be an appropriate criterion for assessing the power to assign back to reporting groups. For these 100% simulations we set the parameters of mixture sample size and number of iterations to the values of 200 and 100, respectively. Genotypes from fisheries mixtures were also analyzed in ONCOR to estimate stock composition by year and biweekly strata. We analyzed all strata that had n > 20 samples. These mixture proportions were generated with 95% confidence intervals using 1000 bootstraps.

Results

Power analysis of baseline

The 147 collections were grouped into 15 reporting groups based on the clustering we observed in the phylogenetic analysis (Figure 2). The ONCOR 100% simulations power analysis revealed that the majority (n=122) of the baseline collections were found to correctly assign back to themselves above the 90% threshold (Table 2). However, 25 collections were below 90% correct assignment and of those, 10 collections were significantly below 90% (as indicated by their upper bounding 95% confidence interval). These 10 collections were South Santiam, Asotin Creek, Roza Trap, Peshastin Creek, Lightning Creek, Hazard Creek, Slate Creek, Chamberlin Creek, Morgan Creek, and Squaw Creek. The Lower and Upper Salmon River reporting groups were found to have the lowest average correct assignment. These reporting groups may need to be adjusted in the future to account for these low values.

Run-timing of the major and minor steelhead stocks among years

Variation in stock proportions was minimal when comparing similar time strata across years, while intra-annual variation across time strata within the migrating season was relatively large. There were four "major" stocks which had above 5% average proportion of the total run (Table 3).

Table 3. Total stock proportions for mixtures from complete years of sampling

REPORTING GROUP	2006	2007	2008	Avg	SD
HoodCanal_OP_PugetSound	0.000	0.000	0.000	0.000	0.000
Lower_Columbia	0.036	0.041	0.029	0.035	0.006
Lower_Columbia_su	0.023	0.028	0.009	0.020	0.010
Willamette	0.000	0.002	0.001	0.001	0.001
BigWhiteSalmon	0.001	0.000	0.000	0.000	0.000
Klickitat	0.022	0.028	0.015	0.022	0.006
MidCol_LowerSnake	0.218	0.233	0.310	0.254	0.050
Yakima	0.008	0.008	0.013	0.010	0.003
Upper Columbia	0.119	0.142	0.142	0.135	0.013
Grand_Ronde	0.000	0.001	0.002	0.001	0.001
Imnaha	0.023	0.012	0.024	0.020	0.006
Upper_Clearwater	0.352	0.367	0.240	0.320	0.069
Lower Salmon	0.027	0.011	0.020	0.019	0.008
MF_SF_Salmon	0.007	0.013	0.009	0.010	0.003
Upper Salmon	0.164	0.113	0.186	0.154	0.037
Total	1.000	1.000	1.000	1.000	

These major stocks were the upper Columbia R. (0.135 ± 0.013), middle Columbia R. and lower Snake R. (0.254 ± 0.050), upper Clearwater R (0.320 ± 0.069), and upper Salmon R. (0.154 ± 0.037). The biweekly strata in which these major stocks were observed to peak were 24-25, 28-29, 38-39, and 32-33, respectively (Figure 3a-d,4a). There were an additional seven "minor" stocks that were between 1%- 5% average proportion, and they were lower Columbia R. (0.035 ± 0.006), lower Columbia R. summer (0.020 ± 0.010), Klickitat R. (0.022 ± 0.006), Yakima R. (0.010 ± 0.003), Imnaha R. (0.020 ± 0.006), lower Salmon R. (0.019 ± 0.008), and middle and south fork Salmon R. (0.010 ± 0.003). The biweekly strata in which these major stocks were observed to peak were18-19, 20-21, 22-23, 26-29, 32-33, and 34-37, respectively (Figure 3a-d,4a).

.

Figure 3a. Chart of proportion of reporting group biweekly assignment for 2004. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend.

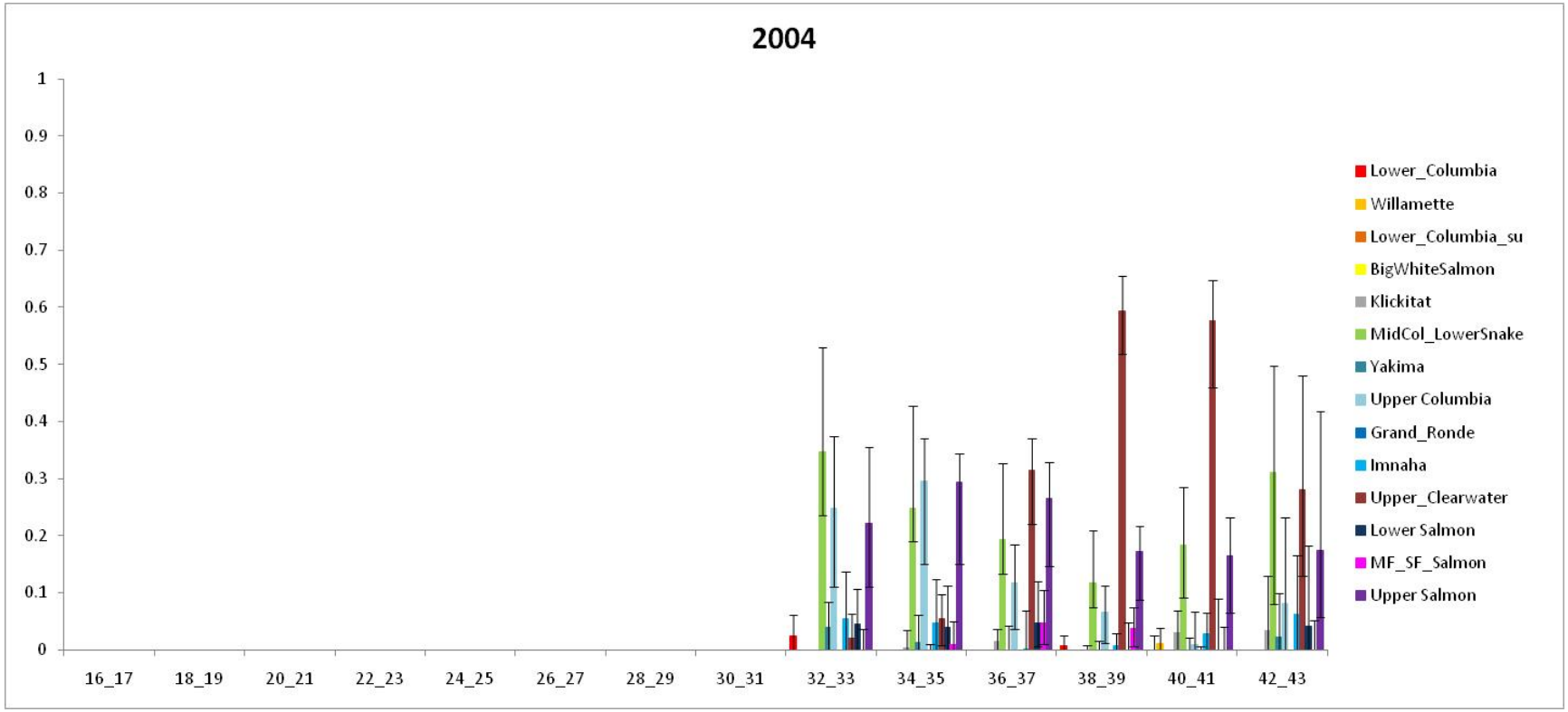


Figure 3b. Chart of proportion of reporting group biweekly assignment for 2006. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend.

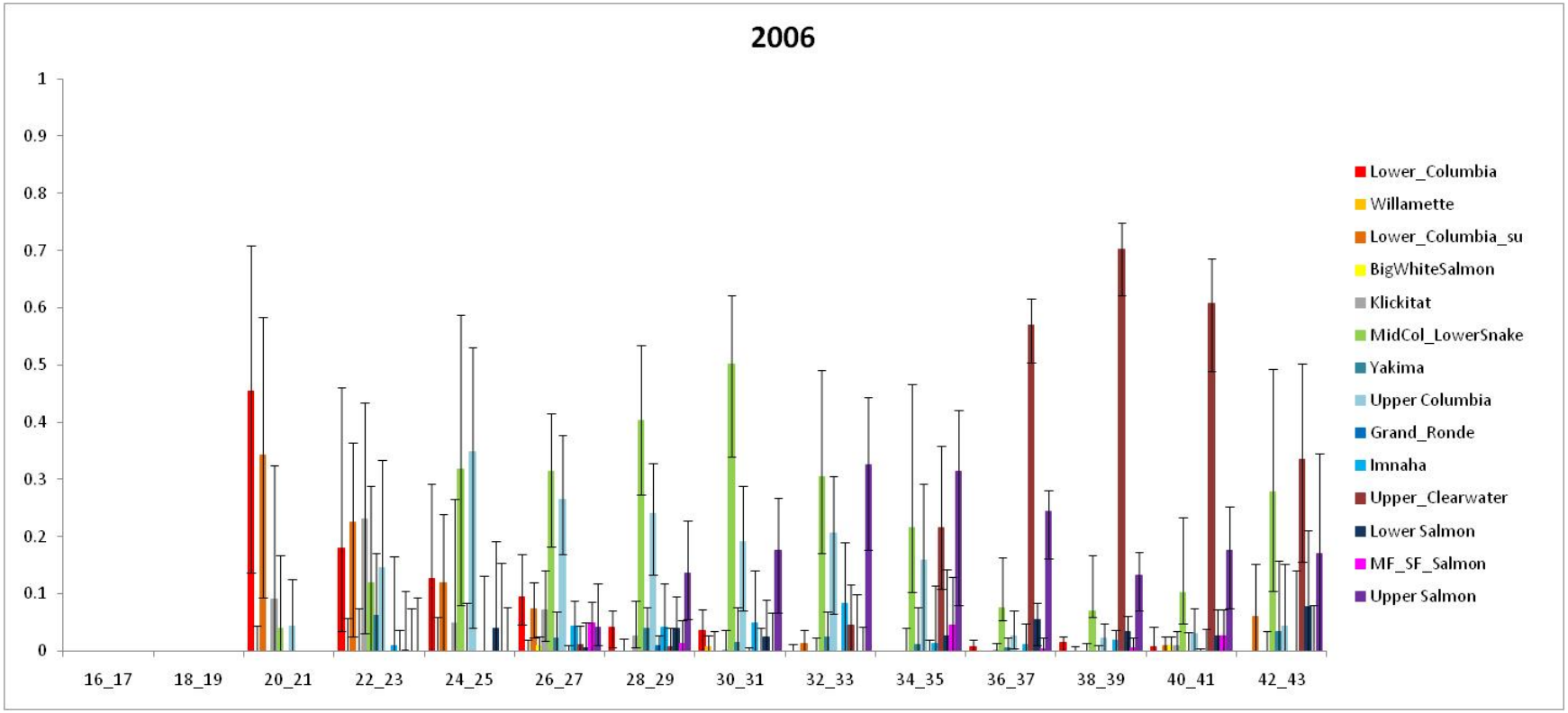


Figure 3c. Chart of proportion of reporting group biweekly assignment for 2007. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend.

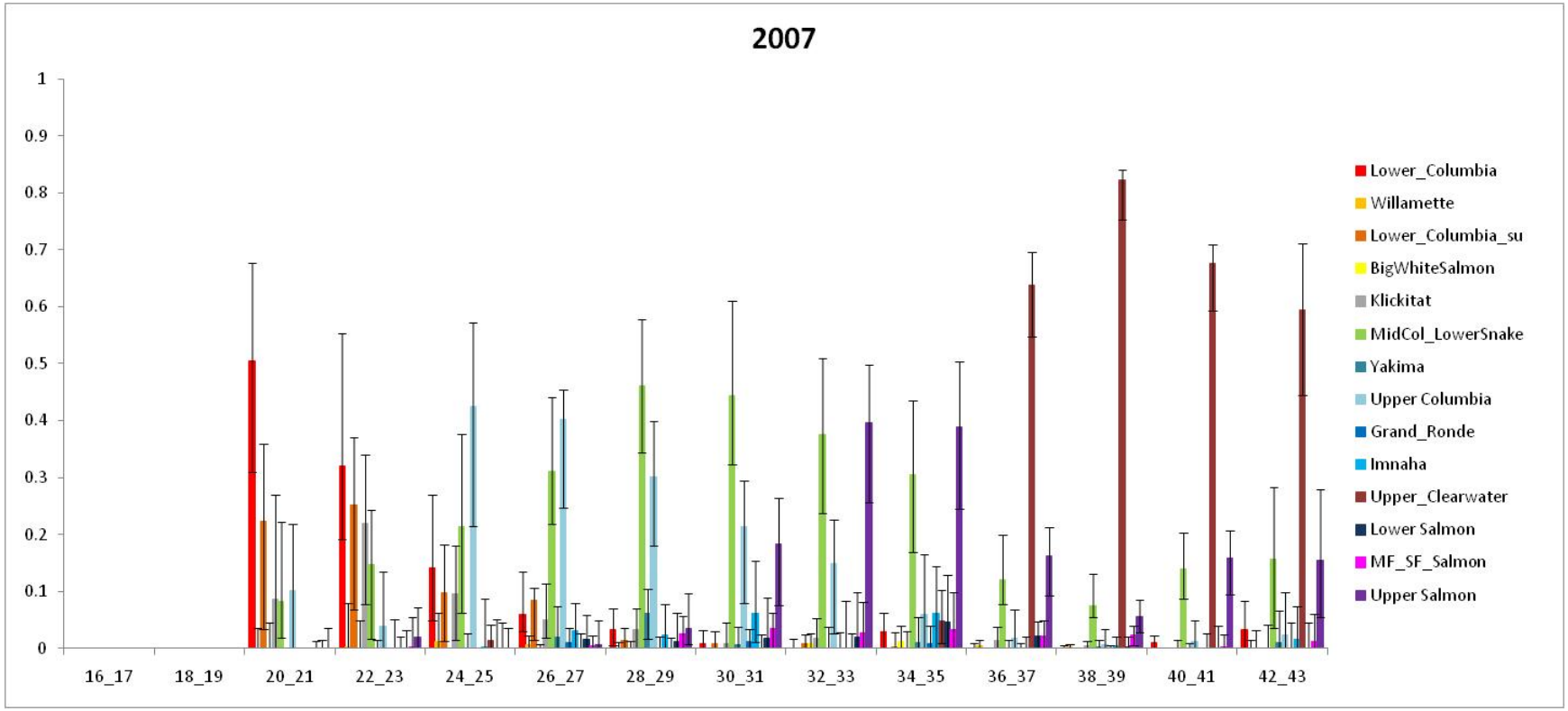


Figure 3d. Chart of proportion of reporting group biweekly assignment for 2008. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend.

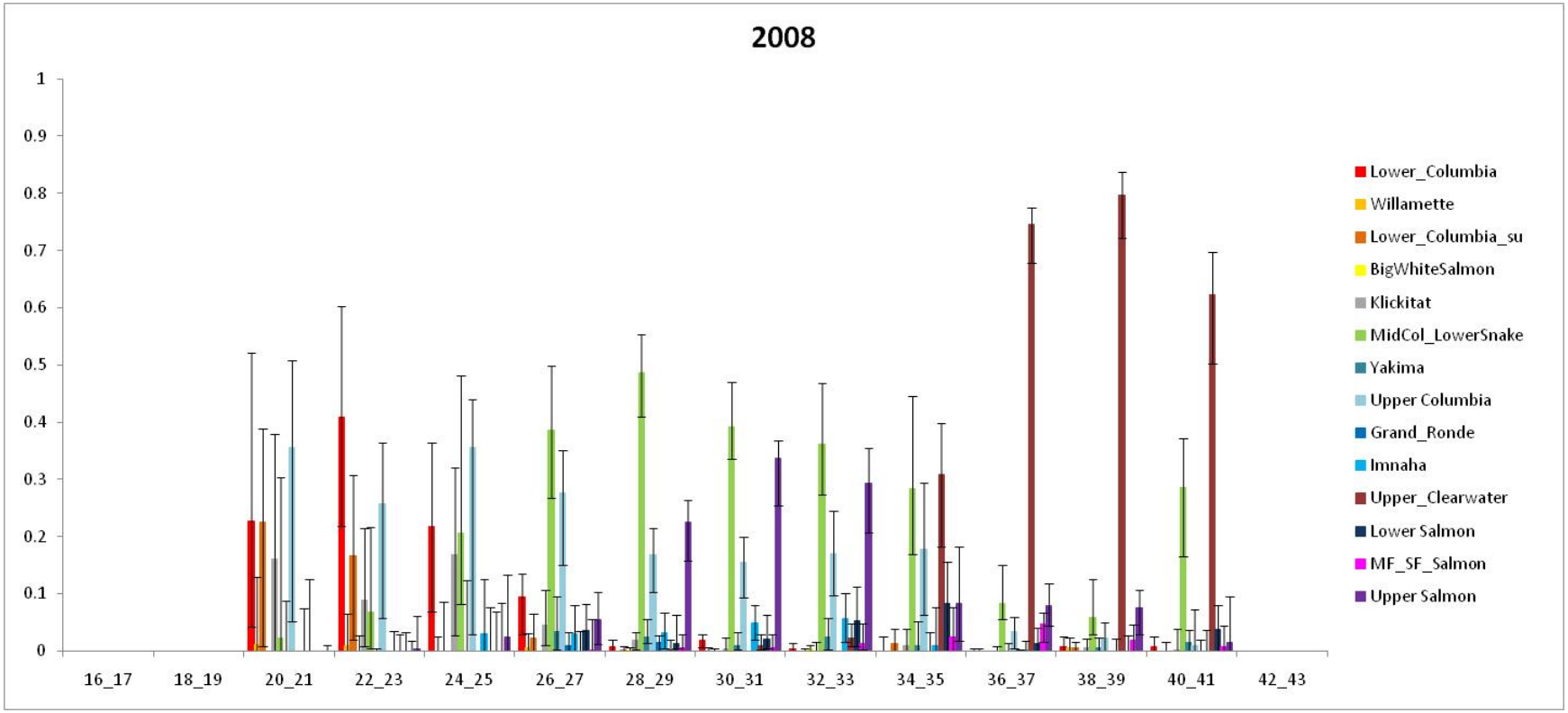


Figure 4a. Chart of proportion of reporting group biweekly assignment for total mixtures across years. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend**.**

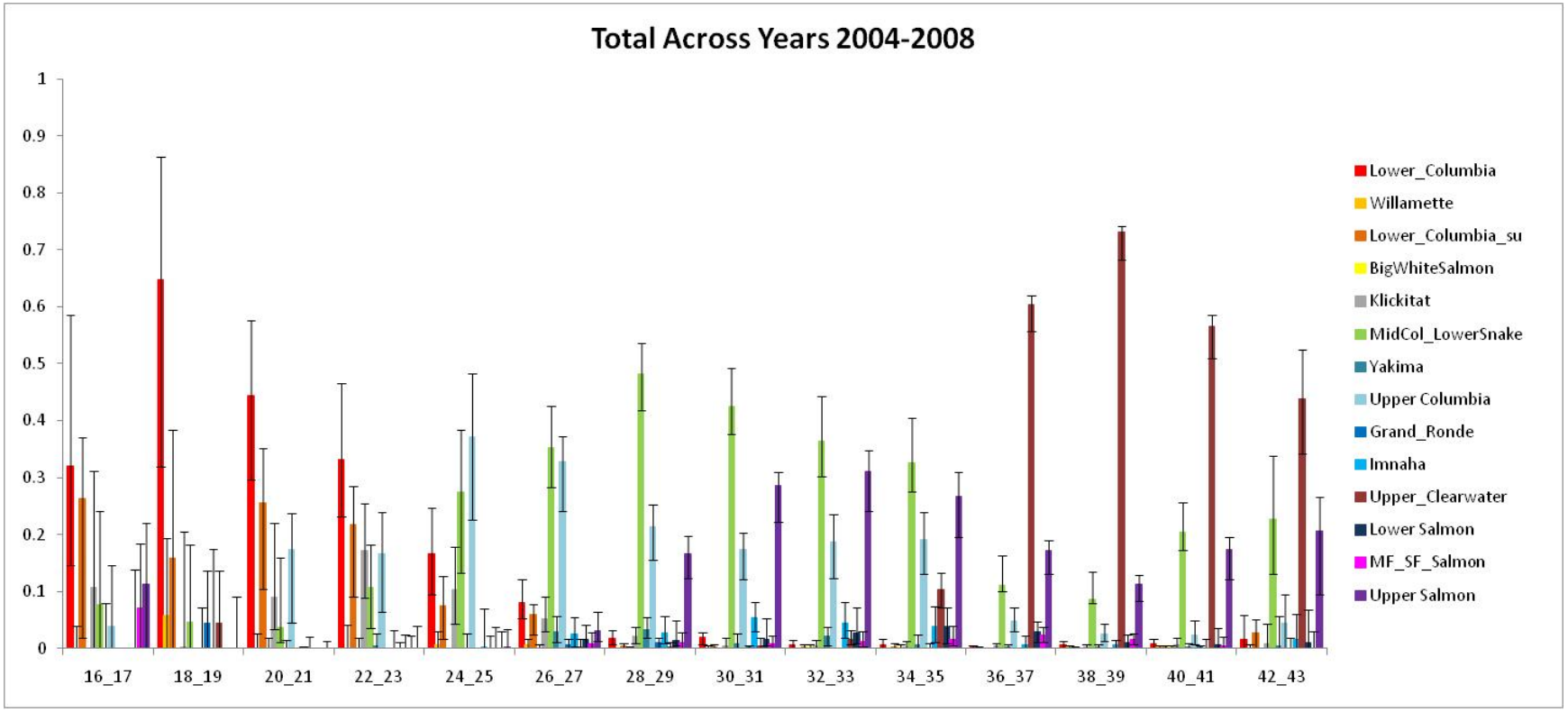
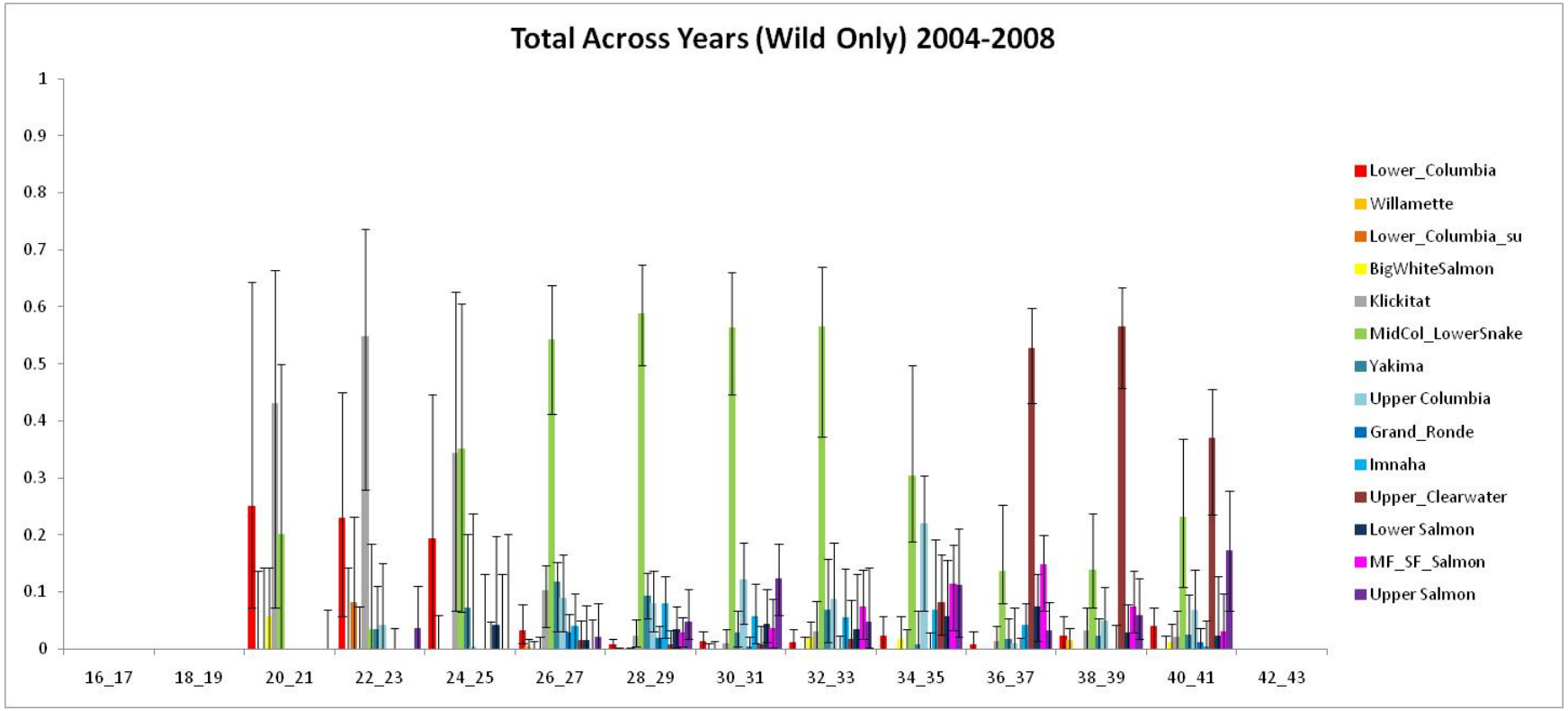


Figure 4b. Chart of proportion of reporting group biweekly assignment for total mixtures across years. The x-axis is labeled with biweekly week numbers. Week 16 begins in mid-April and week 43 ends in mid October. The y-axis indicates the proportion of assignment to the reporting groups in the figure legend. These mixtures were made up entirely of unmarked (wild only) fish.



Influence of natural origin fish on biweekly stock proportions

We removed the natural-origin fish from the total dataset across all years and analyzed these fish separately to determine whether natural-origin fish were differentiated from hatchery-origin fish based on stock proportions for biweekly strata (Figure 4b). This analysis revealed several differences between the hatchery/natural-origin (total) dataset compared with the natural-origin (wild only) dataset. First, the "major" stocks showed peaks at different biweekly strata in the two datasets: the wild-only dataset showed the upper Columbia R. stock was diminished overall and was peaking later at 34-35, middle Columbia R. and lower Snake R stock was peaking more broadly across 26-33 strata, upper Clearwater R. stock was peaking at the same time at 38-39, and upper Salmon R. was diminished and did not have a single peak until much later at 40-41. Interestingly, some of the "minor" stocks showed higher stock proportions in the wild-only dataset, and most notably these stocks were the Klickitat R., Yakima R., and Middle Fork and South Fork Salmon R.

Discussion

Management Implications

This study demonstrates great potential for the application of genetic stock identification in the management of Columbia River steelhead fisheries evidenced by high accuracy of stock assignment, and the consistency of estimated stock proportions across years combined with the ability to discriminate steelhead stocks by run-timing. Results indicate that four "major" stocks of steelhead were sampled as they migrated past Bonneville Dam: upper Columbia R. (0.135 ± 0.013), middle Columbia R./lower Snake R. (0.254 ± 0.050), upper Clearwater R (0.320 ± 0.069), and upper Salmon R. (0.154 ± 0.037). These four steelhead stocks varied considerably in peak run timing (weeks 24-25, 28-29, 38-39, and 32-33, respectively), and clear transitions occurred when each stock of steelhead was most abundant in the mainstem Columbia River. One of the advantages that genetic stock identification has over other methods for estimation of stock composition is the ability to estimate proportions for both hatchery and wild fish. We found that this advantage will help characterize run-timing of less abundant or "minor" Columbia River stocks.

References

Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for estimating the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65:1475-1486.

Lahood, E. S., J. J. Miller, C. Apland, and M. J. Ford. 2008. A rapid, ethanol-free fish tissue collection method for molecular genetic analyses. Transactions of the American Fisheries Society 137(4):1104-1007.

Narum, S. R., S. Boe, P. Moran, and M. Powell. 2006a. Small-scale genetic structure and variation in steelhead of the Grande Ronde River, Oregon, USA. Transactions of the American Fisheries Society 135(4):979-986.

Narum, S. R., M. S. Powell, R. Evenson, and B. Sharp. 2006b. Microsatellites reveal population substructure of Klickitat River native steelhead and genetic divergence from an introduced stock. Journal of Fisheries Management 26(1):147-155.

Narum, S. R., D. Hatch, A. J. Talbot, P. Moran, and M. S. Powell. 2008. Iteroparity in complex mating systems of steelhead *Oncorhynchus mykiss* (Walbaum). Journal of Fish Biology 72(1):45-60.

Nielsen, J. L., A. Byrne, S. L. Graziano, and C. C. Kozfkay. 2009. Steelhead genetic diversity at multiple spatial scales in a managed basin: Snake River, Idaho. North American Journal of Fisheries Management 29(3):680-701.

Stephenson, J. J., M. R. Campbell, J. E. Hess, C. Kozfkay, A. P. Matala, M. V. McPhee, P. Moran, S. R. Narum, M. M. Paquin, O. Schlei, M. P. Small, D. M. Van Doornik, and J. K. Wenburg. 2009. A centralized model for creating shared, standardized, microsatellite data that simplifies inter-laboratory collaboration. Conservation Genetics 10(4):1145-1149.

Winans, G. A., M. M. Paquin, D. M. Van Doornik, B. M. Baker, P. Thornton, D. Rawding, A. Marshall, P. Moran, and S. Kalinowski. 2004. Genetic stock identification of steelhead in the Columbia River Basin: An evaluation of different molecular markers. North American Journal of Fisheries Management 24(2):672-685.