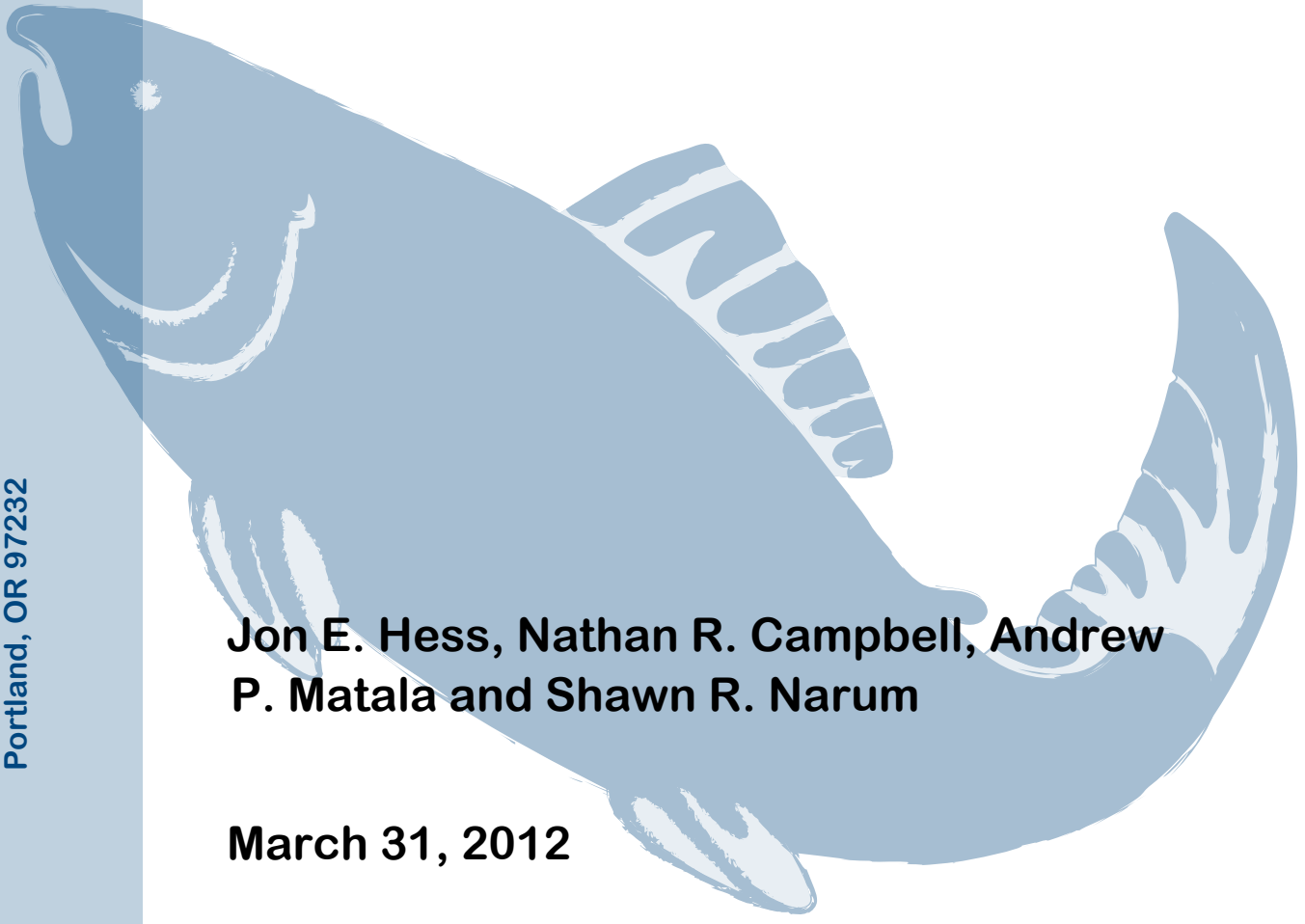
CRITFC

TECHNICAL REPORT 12-07

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

2011 Annual Report: Genetic Assessment of Columbia River Stocks



Jon E. Hess, Nathan R. Campbell, Andrew P. Matala and Shawn R. Narum

March 31, 2012

2011 Annual Report

GENETIC ASSESSMENT OF COLUMBIA RIVER STOCKS

Prepared by:

Jon E. Hess
Nathan R. Campbell
Andrew P. Matala
Shawn R. Narum

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: 4/1/11-3/31/12
Submitted: March 31, 2012

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 including Chinook salmon, steelhead, sockeye salmon and kokanee (*O. nerka*), and coho salmon; 3) implement Genetic Stock Identification (GSI) programs for mainstem Chinook salmon fisheries and 4) GSI of fish passing Bonneville Dam (steelhead and Chinook). In the third year of this project, SNP discovery and evaluation goals (Objective 1) were achieved with evaluation of 245 available assays for Chinook salmon and the results were used to select two panels of 96 SNP markers. The Chinook salmon GSI baseline will be genotyped with the combined total of these 192 SNP markers which includes a 96-SNP panel optimized for parentage based tagging (PBT) and a 96-SNP panel optimized for GSI. In addition, 126 available assays for *O. nerka* were evaluated to choose a single panel of 96 SNPs for genetic applications. For genetic baseline expansion (Objective 2), we successfully genotyped 192 SNP markers in 40 Chinook salmon collections, 192 SNP markers in 129 steelhead collections, and 96 SNP markers in 14 *O. nerka* collections 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 analyses based on selective divergence. The third year of the project included two broad applications of GSI; namely, 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 the 2011 commercial, sport, and test fisheries were primarily composed of two adipose-clipped stocks: Rapid River Hatchery/Clearwater R. and Upper Columbia R. (i.e., Carson stock). These two Chinook salmon stocks were also the most strongly represented at Bonneville Dam and upstream in the ceremonial fishery during spring weeks. A third spring-run stock, Willamette R., was found primarily in harvests spanning the earlier spring weeks and locations closer to the mouth of the Columbia R. In a rare occurrence, the sport fishery continued into summer weeks in 2011, and during this time a 20% increase in the proportion of Salmon R. stocks was observed. For fall-run Chinook salmon fisheries, the commercial fishery below Bonneville Dam contained large proportions of west cascade fall-run and Spring Creek Group Tule stocks, as well as the following stocks (in descending order): Upper Columbia R. summer/fall, Deschutes R. fall, and Snake R. fall. The entire Zone 6 tribal Chinook fishery was heavily comprised of Upper Columbia R. summer/fall stock, but Region 01 (closest region to Bonneville Dam) of Zone 6 was the only location where the Spring Creek Group Tule stock can be found. The Snake R. stock exhibited an early peak in the fall in the upstream section of Zone 6. We tested accuracy of the Chinook salmon baseline using a combination of simulated data and known-origin mixture samples based on tagging methods (e.g. coded wire tags) and observed between 2%-14% improvement in reporting group accuracy relative to what had been achieved using half (96) of the 192 SNP loci in past years. The percentage of individuals correctly assigned to a particular reporting group using a Leave-one-out method was on average 86% and ranged from greater than 95% (west cascade spring- and fall-run and Yakima R. spring stocks) to a low of 61% (Snake R. fall-run stock). For Objective 4, fish were sampled as they migrated past Bonneville Dam. We used GSI to estimate run-timing distributions and abundance of the "major" stocks of wild and hatchery Chinook salmon in 2011 and steelhead over the period 2009 to 2011. Chinook salmon abundance estimates are listed in order of median day of

peak run-timing as follows: Deschutes R. (May 9th, 18,000), Upper Columbia R. (May 10th, 111,000), Rapid R./Clearwater R. (May 11th, 96,000), John Day R. (May 12th, 16,000), and Yakima R. (May 13th, 24,000) spring-run stocks, followed by later run-timing stocks from the Middle Fork Salmon R. (May 30th, 3,000), South Fork Salmon R. (June 7th, 11,000), Klickitat R. (June 8th, 5,000), and Upper Salmon R. (June 11th, 15,000), and fall-run stocks of west cascade (Sep 8th, 3,000), Spring Cr. group tule (Sep 8th, 43,000), Upper Columbia R. summer/fall (Sep 9th, 339,000), Snake R. fall (Sep 11th, 33,000), and Deschutes R. fall (Sep. 13th, 101,000). Steelhead estimates are listed in order of median day of peak run-timing with wild and hatchery abundance separated by a semi-colon as follows: Skamania summer-run (Jul. 6th, <1,000; 7,000), Yakima R. (Jul. 22nd, 6,000; <1,000), Imnaha R. (Aug. 2nd, 5,000; 7,000), Upper Columbia R. (Aug. 3rd, 22,000; 51,000), middle Columbia R. (Aug. 4th, 48,000; 34,000), Klickitat R. (Aug. 5th, 3,000; <1,000), Grande Ronde R. (Aug. 5th, 19,000; 45,000), lower Columbia R. (Aug. 6th, 1,000; <1,000), lower Clearwater R. (Aug. 9th, 3,000; 2,000), Upper Salmon R. (Aug. 10th, 11,000; 93,000), lower/Little Salmon R. (Aug. 12th, 5,000; 12,000), Middle Fork Salmon R. (Aug. 17th, 5,000; <1,000), South Fork Salmon R. (Sep. 2nd, 2,000; <1,000), upper Clearwater R. (Sep. 7th, 6,000k; 3,000), and South Fork Clearwater R. (Sep. 11th, 9,000; 47,000). This study provided an opportunity to demonstrate the benefit of integrating the new genetic technology of parentage based tagging (PBT) into Chinook salmon and steelhead GSI applications. The challenge imposed by long histories of exogenous stock transfers from specific hatchery programs often prevents effective application of GSI in assigning hatchery fish. However, now with the prospect of expanding PBT to mass mark all hatchery fish, GSI will soon be reserved for a smaller but essential role of filling in information gaps that remain after PBT has been used to identify hatchery-origin fish.

Table of Contents

List of Figures

List of Tables

Acknowledgements
The authors of this report 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 Bobby Begay, Steve Boe, Chris Brun, Peter Cleary, Roger Dick Jr., Roger Dick Sr., Dani Evenson, Joe Hymer, Ken Keller, Chris Kern, James Kiona, Jacinda Mainord, Alfred McConnville, Jim Nagler, Terra Schultz, David Sohappy, Jason Vogel, 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, Nick Hoffman, Amanda Matala, Lori Maxwell, Megan Moore, Vanessa Morman and Jeff Stephenson. Maureen Hess contributed guidance on performing parentage based tagging analyses. Substantial data was contributed by Idaho Department of Fish and Game, and the following agencies made significant biological tissue and/or DNA sample contributions: Washington Department of Fish and Game, United States Fish and Wildlife Service, Abernathy Fish Technology Center, and NOAA fisheries

Introduction

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 salmon 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 older genetic marker technology such as microsatellite markers. Over the past few years, our lab has contributed to the increasing numbers of SNP markers that are available for salmonids, and we have reached a point where rigorous stock composition and assessment goals for timely management of fisheries and highly accurate, precise stock assignments can be achieved using one or two panels of 96 SNP markers independently of any other marker-type.

Objective 2) Baseline Expansion

Currently, genetic baselines of microsatellite markers are in place for Chinook salmon across the coastwide range (Seeb et al. 2007), steelhead (Blankenship et al. 2011), and *O. nerka* 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 (Narum et al. 2008b, Hess et al. 2011, Hess and Narum 2011). This proposal includes two GSI projects that will utilize genetic baselines: 1) GSI to Evaluate Harvest; 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, sockeye, 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), Oregon Department of Fish and Wildlife (ODFW), Yakama Nation Fisheries Program (YNFP) and Warm Springs Confederated Tribes. 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.

Finally, this is the first report in which we can begin using a new genetic technology, parentage based tagging (PBT), in combination with GSI to help augment and refine our stock identification results. PBT is an efficient approach for mass marking fish. The method is carried out by first genotyping a set of potential parents which then provides the opportunity to assign a set of genotyped offspring to their true parent pair. PBT is currently being utilized on a broad scale in the Columbia River Basin to mark all Chinook salmon and steelhead Snake River hatchery broodstock (details in Steele et al. 2011). This application has effectively marked all Snake River hatchery Chinook salmon and steelhead starting with the 2008 brood years. When parent pairs of a Snake River hatchery fish are identified with PBT, we can provide accurate information including age of the fish and the source hatchery in which its parents were spawned. In future years, we intend to use PBT in both Chinook salmon and steelhead GSI applications to identify all Snake River hatchery-origin fish, and then we will estimate stock-of-origin of all other hatchery fish that were not assigned with PBT (i.e. non-Snake River hatchery-origin) and all wild fish using GSI. In this way PBT and GSI are very complimentary, and using them in combination takes full advantage of the strengths of each method, while resolving their limitations. Exogenous stock transfers by hatcheries have made hatchery-origin fish challenging to assign with GSI and represents a main limitation that is addressed with PBT.

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, steelhead, and *O. nerka* to create genetic baselines. The third section contains GSI estimates of stock composition of Chinook salmon sampled in mainstem fisheries in 2011. The fourth section includes analysis of run-timing distributions and estimated abundance of adult Chinook salmon and steelhead stocks migrating over Bonneville Dam in 2011 and 2009-2011, respectively.

Section 1: SNP Discovery

Introduction

Informative genetic markers are a necessary tool in conservation genetics applications such as genetic stock identification (GSI), reproductive success, and selection studies. Historically, many types of genetic markers have been used but recent studies have relied on single nucleotide polymorphisms (SNPs) due to their ubiquity throughout the genome and the ability to multiplex genotype hundreds to thousands of SNPs at once using several commercially available platforms. The use of SNP markers therefore provides higher density genome coverage with lower genotyping costs than other marker types. For these reasons our laboratory has been genotyping genetic samples using panels of 96 SNP markers on the Fluidigm EP-1 system. This system uses microfluidic chips to combine up to 96 SNP genotyping assays with up to 96 DNA extracts to generate up to 9,216 genotypes per chip. In order to fully utilize the potential of this platform we have evaluated and chosen two 96 SNP panels for GSI and parentage study purposes in *O. mykiss* as previously described in Hess et al. (2011). The section below describes the evaluation of 245 *O. tshawytscha* (Chinook salmon) SNP assays and selection of the top 96 most variable for parentage studies and the creation of a second 96 SNP panel for GSI purposes by addition of 50 more SNP assays to the remaining 46 SNPs already in the baseline. All 192 SNP assays chosen for use in Chinook salmon genetics studies are reported. Similarly, a set of 96 *O. nerka* assays were evaluated and chosen from a pool of 126 assays for genetic assessment of Columbia River stocks. Methods for evaluation of SNP assays are described and assays included in each panel are reported along with minor allele frequency and G_{ST}.

Methods

O. tshawytscha parentage SNP panel

All available SNP assays for *O. tshawytscha* (N = 245) were obtained and used to genotype 92 fish from six hatchery collections slated for parentage based tagging (PBT) studies (Sawtooth, Pahsimeroi, Rapid River, McCall, Looking Glass, and Clearwater). For genotyping we used Fluidigm 96.96 chips following the manufacturer's suggested protocol modified slightly to include a pre-amplification step. For pre-amplification we used a mixture of all unlabeled primers in each assay included in 1x Qiagen multiplex PCR master mix at a concentration of 45nM and run for 14 cycles of PCR [95C for 15 min, (95C for 15 sec, 60C for 4 min x14),4C hold]. The PCR products were then diluted 1:20 in TE buffer and used as template for genotyping. SNP genotyping plots were generated using Fluidigm genotyping analysis software and manually evaluated for correct scoring. Any assay plots showing more than three clusters or poor cluster separation were noted. Minor allele frequency was calculated for each collection at each of the SNP assays (Table 1). The genotypes were then used to evaluate deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) using the program GENEPOP (Raymond and Rousset 1995). Assays showing significant deviations from HWE at two or more collections were removed from consideration in the parentage SNP panel. When pairs of SNPs showing strong LD were identified, only the one with highest allele frequency was retained. Lastly, assays noted as 'poor plot quality' were removed from consideration.

O. tshawytscha GSI SNP panel

After choosing 96 SNPs for the Chinook parentage panel we found that 50 of the assays were used previously for the genetic baseline panel as reported in Hess et al. (2011). This left a set of 46 SNPs that were previously used as part of the Columbia River genetic baseline for GSI. In order to fully utilize the Fluidigm 96.96 platform, 58 additional SNPs were evaluated and 50 were added to the 46 SNPs already in the baseline to make another 96 SNP panel for Chinook (referred to as the Chinook GSI SNP panel). Screening of the 58 potential SNP loci was done by genotyping 537 samples from 10 collections including both ocean-type and stream-type lineages in the Columbia River. Genotypes were collected using the same method as described above for the parentage SNP panel evaluation. Loci were evaluated by plot quality, minor allele frequency, population differentiation (G_{ST}), deviation from HWE, and statistically significant LD between pairs of loci. The program GENEPOP was used to evaluate deviation from HWE and to test for statistically significant LD. Allele frequencies and G_{ST} were calculated using the Excel plug in GENALEX.

O. nerka SNP panel

SNP discovery and evaluation of SNP assays was previously reported for *O. nerka* in Campbell and Narum (2010). These analyses included evaluation of plot quality, minor allele frequency, population differentiation (G_{ST}), and linkage disequilibrium (LD) derived from genotypes from five collections of *O. nerka* in the Columbia River. This information was used to choose the most informative set of 96 SNP assays for use in the Columbia River. The *O. nerka* SNP panel is reported in Table 4.

Results

O. tshawytscha parentage SNP panel

Minor allele frequencies (MAF) from each of the tested 245 SNP assays for *O. tshawytscha* were recorded for each of the six collections and averaged (Table 1). Average MAF ranged from 0.000 to 0.495 across all tested SNPs and averaged 0.137. Filtering all SNP assays that produced poor quality plots resulted in 67 assays that failed or were monomorphic within the six collections tested, leaving 178 assays for further evaluation. Genotypes from the top 100 of the remaining assays were used to evaluate deviation from HWE and statistically significant LD. No assays showed a significant heterozygote deficit and only the sex determination assay (Ots_SEXY3-1) showed significant heterozygote excess (which was expected since this assay shows all males as heterozygotes). A single pair of loci [Ots_FGF6A and Ots_FGF6B-1] showed significant LD within multiple collections and assay Ots_FGF6A was removed in favor of the higher MAF assay Ots_FGF6B-1. The top 96 of the assays remaining were selected as the Chinook parentage GSI panel and had an average MAF of 0.252 in the tested collections. The full panel with average MAF and the percentage of samples genotyped is shown in Table 2.

O. tshawytscha GSI SNP panel

After choosing the top 96 assays for parentage in Chinook salmon, an additional 50 SNP assays were needed to fill out a second SNP panel for GSI (Chinook GSI SNP panel). For this we evaluated the remaining 58 available assays that had not been previously flagged for poor plot quality or significant deviation from HWE. Each of these assays were tested by genotyping a set of 537 samples including both ocean-type and stream-type lineages from 10 collections in the Columbia River (Wenaha R., Shitike R., Lostine R., Johnson Cr., Clearwater R., Lolo Cr., Upper Deschutes R., Yakima R., Lyons-Ferry National Fish Hatchery, and Cowlitz National Fish

Hatchery). Genotyping results revealed three assays that were monomorphic in all 10 collections and four assays with poor plot quality. This left only a single assay that could be dropped for either deviation from HWE or statistically significant LD from the pool. The last assay dropped was Ots_IsoT due to deviations from HWE within several collections. The remaining assays (Table 3) were added to 46 SNPs already in the baseline but not included in the Chinook parentage panel to create the Chinook GSI SNP panel. Both Chinook SNP panels with primer/probe sequences are reported in Appendix 1.

O. nerka SNP panel

A published evaluation of all available *O. nerka* SNP assays (Campbell and Narum 2010) was used to select a set of the top 96 most informative SNPs for use in the Columbia River. After filtering SNP assays that were monomorphic, had poor plot quality, or deviations from HWE, there were 102 assays from which to choose a panel of 96. Five assays that showed statistically significant LD were then filtered leaving 97 assays. The least informative remaining assay was then dropped from the final set of 96. The final *O. nerka* SNP panel has an average MAF of 0.220 and an average G_{ST} of 0.179 indicating utility for both GSI and parentage using this panel (Table 4).

Discussion

The most expensive and time consuming aspect of genetic research in non-model organisms is developing genetic markers. Once a suitable suite of markers exists, a multitude of commercial platforms are available to increase efficiency and decrease the cost of genotyping. Sequencing efforts by many in the salmonid genetics community have now given rise to a substantial resource of SNP markers available for most Pacific salmonids. We are now in the process of tapping this resource by the creation of marker panels and genetic baselines which provide a more thorough understanding of these organisms. Currently, the *O. mykiss* marker panels are being used to track stock abundances, estimate reproductive success, and track hatchery stray rates. Similar studies can now be implemented for *O. tshawytscha* and *O. nerka* with the creation of these new marker panels. These studies provide additional information that is likely to affect conservation efforts, hatchery supplementation, and the implications of population reintroduction of these animals in the Columbia River.

References

Campbell NR, Narum SR *(*2011*)* Development of 54 novel SNP assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. *Molecular Ecology Resources,* **11 (Suppl. 1)**, 20–30.

Hess JE, Campbell NR, Matala AP, Narum SR (2011) Genetic assessment of Columbia River stocks. Annual Progress Report to Bonneville Power Adminstration, project 2008-907-00. http://fishery.critfc.org/FiSci/11_02report.aspx

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

Table 1: Evaluation of all 245 available Chinook SNP assays for inclusion in parentage SNP panel. Minor allele frequency at each of the 6 reference collections and average MAF are shown. Data are sorted by 'overall MAF'. Poorly performing assays were moved to the bottom of the table.

SNP assay:	% Geno-typed:	Saw-tooth:	Pah-simeroi:	S.F. Salmon:	Lostine:	S.F. Clear-water:	Rapid River:	overall MAF:
Ots_GTH2B-550	100.0%	43.8%	71.4%	56.3%	53.1%	37.5%	37.5%	49.5%
Ots_102414-395	100.0%	34.4%	46.4%	71.9%	62.5%	50.0%	20.8%	49.5%
Ots_OTDESMIN 19-SNP1	100.0%	34.4%	67.9%	53.1%	31.3%	59.4%	41.7%	48.4%
Ots_108820-336	100.0%	59.4%	53.6%	53.1%	28.1%	46.9%	50.0%	48.4%
Ots_101554-407	100.0%	53.1%	46.4%	50.0%	59.4%	28.1%	54.2%	47.8%
Ots_u07-07.161	100.0%	34.4%	39.3%	53.1%	59.4%	50.0%	50.0%	47.3%
Ots_IL8R_C8	100.0%	43.8%	53.6%	59.4%	43.8%	46.9%	29.2%	46.7%
Ots_96500-180	100.0%	53.1%	35.7%	50.0%	34.4%	50.0%	54.2%	46.2%
Ots_104415-88	100.0%	43.8%	60.7%	43.8%	37.5%	40.6%	58.3%	46.2%
Ots_Est740	98.9%	34.4%	35.7%	59.4%	46.9%	66.7%	33.3%	45.6%
Ots_pigh-105	96.7%	68.8%	23.1%	34.4%	40.6%	46.4%	54.2%	45.5%
Ots_94903-99R	100.0%	53.1%	32.1%	31.3%	37.5%	68.8%	41.7%	45.1%
Ots_105407-117	97.8%	40.6%	60.7%	59.4%	31.3%	36.7%	27.3%	44.4%
Ots_u07-25.325	100.0%	40.6%	42.9%	31.3%	50.0%	59.4%	29.2%	43.5%
Ots_105385-421	100.0%	25.0%	32.1%	43.8%	59.4%	53.1%	41.7%	43.5%
Ots_94857-232R	97.8%	31.3%	32.1%	37.5%	40.6%	60.7%	70.8%	43.3%
Ots_u211-85	98.9%	12.5%	28.6%	40.6%	62.5%	63.3%	50.0%	42.9%
Ots_FGF6B_1	100.0%	43.8%	46.4%	28.1%	43.8%	50.0%	41.7%	42.4%
Ots_MHC2	100.0%	56.3%	35.7%	53.1%	25.0%	25.0%	54.2%	41.8%
Ots_CD59-2	100.0%	37.5%	53.6%	40.6%	28.1%	53.1%	33.3%	40.8%
Ots_nkef-192	100.0%	40.6%	42.9%	37.5%	28.1%	50.0%	41.7%	39.7%
Ots_mapK-3'-309	100.0%	37.5%	46.4%	28.1%	34.4%	56.3%	33.3%	39.7%
Ots_106747-239	96.7%	40.6%	39.3%	37.5%	53.1%	34.6%	12.5%	38.2%
Ots_OTSTF1-SNP1	100.0%	46.9%	32.1%	50.0%	53.1%	18.8%	29.2%	38.0%
Ots_TAPBP	98.9%	28.1%	46.4%	28.1%	56.3%	16.7%	41.7%	35.7%
Ots_SClkF2R2-135	100.0%	28.1%	46.4%	28.1%	37.5%	37.5%	41.7%	35.3%
Ots_OTNAML12 _1-SNP1	100.0%	46.9%	21.4%	46.9%	15.6%	46.9%	25.0%	35.3%
Ots_FGF6A	100.0%	43.8%	39.3%	21.9%	31.3%	34.4%	37.5%	34.8%
Ots_TLR3	100.0%	50.0%	17.9%	31.3%	46.9%	21.9%	33.3%	33.7%
Ots_Tnsf	100.0%	25.0%	35.7%	21.9%	25.0%	40.6%	50.0%	33.2%
Ots_110201-363	100.0%	34.4%	21.4%	43.8%	31.3%	46.9%	16.7%	33.2%
Ots_105105-613	100.0%	18.8%	28.6%	40.6%	18.8%	31.3%	70.8%	32.6%
Ots_S7-1	98.9%	46.9%	53.6%	25.0%	18.8%	30.0%	16.7%	32.4%
Ots_ETIF1A	97.8%	34.4%	28.6%	37.5%	15.6%	40.0%	40.9%	31.7%
Ots_ntl-255	98.9%	43.8%	21.4%	34.4%	25.0%	43.3%	20.8%	31.3%
Ots_110064-383	100.0%	31.3%	28.6%	34.4%	25.0%	37.5%	20.8%	29.3%
Ots_E2-275	100.0%	25.0%	3.6%	25.0%	46.9%	31.3%	45.8%	28.8%
Ots_SWS1op-182	98.9%	31.3%	25.0%	43.8%	28.1%	20.0%	20.8%	28.6%
Ots_NOD1	100.0%	21.9%	21.4%	31.3%	37.5%	18.8%	37.5%	28.3%

Ots_115987-325	100.0%	15.6%	42.9%	21.9%	40.6%	15.6%	33.3%	27.7%
Ots_pop5-96	100.0%	31.3%	17.9%	21.9%	28.1%	28.1%	37.5%	26.6%
Ots_GPH-318	100.0%	21.9%	35.7%	25.0%	18.8%	34.4%	29.2%	26.6%
Ots_110551-64	98.9%	25.0%	50.0%	21.9%	12.5%	30.0%	20.8%	25.8%
Ots_P53	100.0%	40.6%	42.9%	12.5%	18.8%	9.4%	12.5%	23.9%
Ots_102801-308	100.0%	43.8%	17.9%	21.9%	12.5%	18.8%	29.2%	23.9%
Ots_mapKpr-151	97.8%	15.6%	28.6%	25.0%	28.1%	26.7%	18.2%	23.9%
Ots_Prl2	98.9%	25.0%	28.6%	18.8%	21.9%	34.4%	13.6%	23.6%
Ots_parp3-286	100.0%	34.4%	14.3%	21.9%	28.1%	21.9%	12.5%	22.8%
Ots_118938-325	100.0%	15.6%	28.6%	25.0%	3.1%	43.8%	20.8%	22.3%
Ots_u1002-75	98.9%	21.9%	21.4%	21.9%	15.6%	33.3%	16.7%	22.0%
Ots_redd1-187	100.0%	9.4%	17.9%	37.5%	15.6%	31.3%	16.7%	21.7%
Ots_u07-49.290	98.9%	31.3%	21.4%	15.6%	15.6%	25.0%	18.2%	21.4%
Ots_117432-409	100.0%	15.6%	32.1%	40.6%	3.1%	21.9%	12.5%	21.2%
Ots_IsoT	98.9%	9.4%	14.3%	28.1%	34.4%	9.4%	27.3%	20.9%
Ots_SEXY1	97.8%	15.6%	32.1%	46.9%	6.3%	0.0%	20.0%	20.6%
Ots_ppie-245	100.0%	18.8%	3.6%	18.8%	31.3%	28.1%	20.8%	20.1%
Ots_AsnRS-60	100.0%	28.1%	0.0%	31.3%	12.5%	37.5%	8.3%	20.1%
Ots_110689-218	100.0%	9.4%	17.9%	18.8%	37.5%	21.9%	16.7%	20.1%
Ots_112820-284	100.0%	25.0%	21.4%	12.5%	15.6%	28.1%	12.5%	19.6%
Ots_GDH-81x	100.0%	37.5%	10.7%	21.9%	9.4%	18.8%	16.7%	19.0%
Ots_Thio	98.9%	34.4%	10.7%	28.1%	9.4%	12.5%	18.2%	18.7%
Ots_100884-287	98.9%	18.8%	28.6%	28.1%	3.1%	23.3%	8.3%	18.7%
Ots_mtap-299	100.0%	37.5%	10.7%	15.6%	25.0%	9.4%	4.2%	18.5%
Ots_HMGB1-73	100.0%	9.4%	25.0%	18.8%	18.8%	21.9%	16.7%	17.9%
Ots_123921-111	98.9%	31.3%	25.0%	12.5%	3.1%	25.0%	4.5%	17.6%
Ots_u07-17.135	100.0%	31.3%	17.9%	6.3%	21.9%	15.6%	8.3%	17.4%
Ots_IGF-I.1-76	98.9%	9.4%	28.6%	21.9%	12.5%	3.3%	29.2%	17.0%
Ots_brp16-64	98.9%	21.9%	14.3%	12.5%	31.3%	16.7%	4.2%	17.0%
Ots_hsc71-3'-488	100.0%	31.3%	17.9%	9.4%	18.8%	9.4%	12.5%	16.8%
Ots_103122-180	100.0%	6.3%	25.0%	6.3%	15.6%	21.9%	29.2%	16.8%
Ots_NFYB-147	100.0%	9.4%	10.7%	21.9%	28.1%	12.5%	16.7%	16.3%
Ots_103041-52	98.9%	6.3%	17.9%	28.1%	9.4%	10.0%	29.2%	15.9%
Ots_txnip-321	100.0%	21.9%	14.3%	21.9%	21.9%	6.3%	4.2%	15.8%
Ots_112301-43	100.0%	15.6%	0.0%	15.6%	15.6%	28.1%	20.8%	15.8%
Ots_cox1-241	97.8%	21.9%	10.7%	12.5%	9.4%	23.3%	13.6%	15.6%
Ots_105132-200	98.9%	21.9%	7.1%	15.6%	13.3%	9.4%	29.2%	15.4%
Ots_mybp-85	100.0%	28.1%	0.0%	6.3%	25.0%	6.3%	29.2%	15.2%
Ots_HSP90B-100	98.9%	15.6%	17.9%	18.8%	12.5%	6.7%	12.5%	14.8%
Ots_101704-143	98.9%	0.0%	25.0%	6.3%	15.6%	9.4%	45.5%	14.8%
Ots_109525-816	100.0%	28.1%	0.0%	15.6%	3.1%	12.5%	33.3%	14.7%
Ots_112876-371	100.0%	9.4%	3.6%	18.8%	15.6%	21.9%	16.7%	14.1%
Ots_110495-380	100.0%	12.5%	10.7%	6.3%	15.6%	28.1%	12.5%	14.1%
Ots_112419-131	100.0%	15.6%	21.4%	18.8%	18.8%	3.1%	0.0%	13.6%
Ots_unk526	98.9%	3.1%	25.0%	0.0%	15.6%	23.3%	16.7%	13.2%
Ots_u07-18.378	100.0%	18.8%	0.0%	9.4%	6.3%	15.6%	33.3%	13.0%
Ots_RAG3	100.0%	3.1%	17.9%	0.0%	15.6%	28.1%	12.5%	13.0%
Ots_CirpA	98.9%	6.3%	7.1%	21.9%	15.6%	6.3%	13.6%	12.1%
Ots_124774-477	98.9%	3.1%	17.9%	12.5%	12.5%	15.6%	9.1%	12.1%
Ots_118205-61	98.9%	12.5%	7.1%	15.6%	12.5%	23.3%	0.0%	12.1%

Ots_u1008-108	100.0%	0.0%	17.9%	21.9%	6.3%	9.4%	16.7%	12.0%
Ots_nips-133	100.0%	3.1%	3.6%	6.3%	21.9%	3.1%	41.7%	12.0%
Ots_129458-451	98.9%	20.0%	21.4%	6.3%	6.3%	6.3%	8.3%	11.5%
Ots_ARNT	100.0%	9.4%	14.3%	12.5%	3.1%	15.6%	16.7%	11.4%
Ots_TGFB	100.0%	15.6%	28.6%	15.6%	0.0%	3.1%	4.2%	10.9%
Ots_96899-357R	100.0%	3.1%	10.7%	9.4%	18.8%	9.4%	12.5%	10.9%
Ots_u6-75	98.9%	18.8%	7.1%	15.6%	9.4%	0.0%	12.5%	10.4%
Ots_tpx2-125	100.0%	9.4%	3.6%	6.3%	12.5%	12.5%	16.7%	10.3%
Ots_128757-61R	100.0%	3.1%	10.7%	15.6%	9.4%	6.3%	20.8%	10.3%
Ots_OTALDBINT1-SNP1	98.9%	6.3%	14.3%	12.5%	18.8%	0.0%	4.5%	9.3%
Ots_104569-86	100.0%	6.3%	10.7%	0.0%	6.3%	25.0%	8.3%	9.2%
Ots_AldB1-122	97.8%	3.1%	0.0%	3.1%	0.0%	26.7%	25.0%	8.9%
Ots_108735-302	88.0%	6.3%	31.8%	16.7%	3.3%	0.0%	0.0%	8.6%
Ots_MHC1	100.0%	6.3%	0.0%	3.1%	9.4%	3.1%	29.2%	8.2%
Ots_aldb-177M	98.9%	6.3%	7.1%	3.1%	3.1%	13.3%	8.3%	7.1%
Ots_GCSH	100.0%	3.1%	10.7%	6.3%	6.3%	3.1%	16.7%	7.1%
Ots_AldoB4-183	98.9%	3.1%	0.0%	3.1%	0.0%	13.3%	25.0%	6.6%
Ots_IL11	100.0%	0.0%	7.1%	6.3%	15.6%	6.3%	0.0%	6.5%
Ots_126619-400	100.0%	12.5%	14.3%	3.1%	3.1%	3.1%	4.2%	6.5%
Ots_113242-216	100.0%	12.5%	21.4%	0.0%	6.3%	0.0%	0.0%	6.5%
Ots_111681-657	100.0%	0.0%	0.0%	9.4%	3.1%	3.1%	29.2%	6.5%
Ots_vatf-251	98.9%	6.3%	14.3%	0.0%	3.1%	6.7%	8.3%	6.0%
Ots_xip-130	100.0%	12.5%	3.6%	3.1%	6.3%	3.1%	8.3%	6.0%
Ots_EndoRB1-486	100.0%	0.0%	3.6%	0.0%	31.3%	0.0%	0.0%	6.0%
Ots_107285-93	98.9%	0.0%	3.6%	0.0%	12.5%	3.1%	18.2%	5.5%
Ots_u4-92	100.0%	9.4%	14.3%	0.0%	0.0%	9.4%	0.0%	5.4%
Ots_trnau1ap-86	100.0%	0.0%	0.0%	9.4%	21.9%	0.0%	0.0%	5.4%
Ots_myoD-364	100.0%	9.4%	0.0%	6.3%	9.4%	6.3%	0.0%	5.4%
Ots_hsp27b-150	100.0%	15.6%	10.7%	0.0%	0.0%	0.0%	8.3%	5.4%
Ots_130720-99	100.0%	9.4%	7.1%	6.3%	0.0%	6.3%	4.2%	5.4%
Ots_113457-40R	98.9%	9.4%	3.6%	6.3%	6.3%	0.0%	4.2%	4.9%
Ots_u07-57.120	100.0%	0.0%	7.1%	0.0%	12.5%	6.3%	4.2%	4.9%
Ots_u07-53.133	100.0%	0.0%	3.6%	9.4%	3.1%	6.3%	8.3%	4.9%
Ots_117242-136	100.0%	0.0%	7.1%	6.3%	12.5%	0.0%	4.2%	4.9%
Ots_ZR-575	95.7%	10.0%	0.0%	3.3%	6.3%	3.6%	0.0%	4.5%
Ots_GH2	98.9%	3.1%	10.7%	6.3%	3.1%	0.0%	4.2%	4.4%
Ots_itpa-79	100.0%	3.1%	10.7%	9.4%	3.1%	0.0%	0.0%	4.3%
Ots_Ikaros-250	100.0%	6.3%	0.0%	0.0%	0.0%	12.5%	8.3%	4.3%
Ots_picalm-175	98.9%	0.0%	7.1%	0.0%	6.3%	0.0%	8.3%	3.8%
Ots_hsc71-5'-453	100.0%	3.1%	0.0%	0.0%	9.4%	6.3%	0.0%	3.8%
Ots_Est1363	100.0%	0.0%	3.6%	3.1%	9.4%	0.0%	8.3%	3.8%
Ots_131906-141	100.0%	0.0%	7.1%	6.3%	6.3%	3.1%	0.0%	3.8%
Ots_131460-584	98.9%	0.0%	0.0%	6.3%	0.0%	13.3%	0.0%	3.3%
Ots_hus1-52	100.0%	0.0%	10.7%	0.0%	3.1%	3.1%	4.2%	3.3%
Ots_GnRH-271	100.0%	3.1%	3.6%	9.4%	0.0%	3.1%	0.0%	3.3%
Ots_CD63	100.0%	0.0%	10.7%	3.1%	3.1%	3.1%	0.0%	3.3%
Ots_97660_56	100.0%	0.0%	7.1%	3.1%	0.0%	9.4%	0.0%	3.3%
Ots_128693_461	100.0%	6.3%	0.0%	0.0%	12.5%	0.0%	0.0%	3.3%

Ots_102457-132	100.0%	0.0%	7.1%	0.0%	3.1%	9.4%	0.0%	3.3%
Ots_107074-284	98.9%	0.0%	10.7%	3.1%	0.0%	3.3%	0.0%	2.7%
Ots_u202-161	100.0%	0.0%	0.0%	3.1%	9.4%	3.1%	0.0%	2.7%
Ots_u1007-124	100.0%	0.0%	3.6%	6.3%	0.0%	3.1%	4.2%	2.7%
Ots_myo1a-384	100.0%	0.0%	3.6%	6.3%	3.1%	0.0%	4.2%	2.7%
Ots_111084-96	100.0%	0.0%	0.0%	0.0%	15.6%	0.0%	0.0%	2.7%
Ots_zn593-346	100.0%	0.0%	7.1%	3.1%	0.0%	0.0%	4.2%	2.2%
Ots_SL	100.0%	0.0%	10.7%	0.0%	3.1%	0.0%	0.0%	2.2%
Ots_cin-330	100.0%	3.1%	0.0%	0.0%	0.0%	9.4%	0.0%	2.2%
Ots_97077-179R	100.0%	0.0%	3.6%	0.0%	3.1%	0.0%	4.2%	2.2%
Ots_96222-525	100.0%	0.0%	0.0%	3.1%	0.0%	9.4%	0.0%	2.2%
Ots_123048-521	100.0%	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%
Ots_118175-479	100.0%	0.0%	3.6%	3.1%	3.1%	3.1%	0.0%	2.2%
Ots_FARSLA-220	98.9%	0.0%	3.6%	0.0%	0.0%	6.7%	0.0%	1.6%
Ots_Est803	98.9%	3.1%	7.1%	0.0%	0.0%	0.0%	0.0%	1.6%
Ots_P450	100.0%	0.0%	0.0%	0.0%	0.0%	6.3%	4.2%	1.6%
Ots_nramp-321	100.0%	0.0%	0.0%	3.1%	0.0%	6.3%	0.0%	1.6%
Ots_metA-199	100.0%	0.0%	3.6%	0.0%	0.0%	6.3%	0.0%	1.6%
Ots_lpl-242	100.0%	3.1%	0.0%	3.1%	0.0%	0.0%	4.2%	1.6%
Ots_EP-529	100.0%	0.0%	3.6%	0.0%	6.3%	0.0%	0.0%	1.6%
Ots_120950-417	100.0%	0.0%	0.0%	3.1%	0.0%	6.3%	0.0%	1.6%
Ots_111312-435	100.0%	0.0%	0.0%	0.0%	0.0%	9.4%	0.0%	1.6%
Ots_RFC2-558	100.0%	0.0%	0.0%	3.1%	0.0%	0.0%	4.2%	1.1%
Ots_PGK-54	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	8.3%	1.1%
Ots_Cath_D141	100.0%	0.0%	3.6%	0.0%	0.0%	0.0%	4.2%	1.1%
Ots_C3N3	100.0%	0.0%	0.0%	6.3%	0.0%	0.0%	0.0%	1.1%
Ots_131802-393	100.0%	0.0%	0.0%	6.3%	0.0%	0.0%	0.0%	1.1%
Ots_127960-181	100.0%	0.0%	0.0%	6.3%	0.0%	0.0%	0.0%	1.1%
Ots_112208-722	100.0%	0.0%	0.0%	3.1%	0.0%	3.1%	0.0%	1.1%
Ots_109693-392	100.0%	0.0%	3.6%	3.1%	0.0%	0.0%	0.0%	1.1%
Ots_108390-329	100.0%	0.0%	0.0%	6.3%	0.0%	0.0%	0.0%	1.1%
Ots_107607-315	100.0%	0.0%	0.0%	0.0%	0.0%	3.1%	4.2%	1.1%
Ots_sys1-112	100.0%	0.0%	0.0%	3.1%	0.0%	0.0%	0.0%	0.5%
Ots_OTSMTA-SNP1	100.0%	0.0%	3.6%	0.0%	0.0%	0.0%	0.0%	0.5%
Ots_nelfd-163	100.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.0%	0.5%
Ots_Hsp90a	100.0%	0.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.5%
Ots_arp-436	100.0%	0.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.5%
Ots_99550-204	100.0%	0.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.5%
Ots_129870-55	100.0%	0.0%	3.6%	0.0%	0.0%	0.0%	0.0%	0.5%
Ots_117259-271	100.0%	0.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.5%
Ots_101119-381	100.0%	0.0%	0.0%	3.1%	0.0%	0.0%	0.0%	0.5%
Ots_zP3b-215	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_USMG5-67	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u1012-34	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u1010-110	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u1006-171	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u1005-108	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u1004-117	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Ots_u07-64.221	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_u07-20.332	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_tshB-226	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_TNF	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_stk6-516	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_RAS1	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_OTSBMP-2-SNP1	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_Ots311-101x	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_OTNAML12_2-SNP2	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_ndk-167	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_Myc-366	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_msra-224	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_LWSop-638	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_HSP90BB-88	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_GST-375	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_GST-207	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_GPDH-338	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_gna11-169	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_DBLOH-73	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_CRB-211	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_CRB211	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_Chin30up-211	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_CCR7	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_casp9-99	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_bcAKal-476	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_aspat-196	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_afmid-196	98.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_98683-796	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_98409-850	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_122414-56	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_117370-471	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_107220-70	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_106313-729	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_106172-425	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_105897-124	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_104216-70	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_129170-683	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_ca050-6M	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_Carba1-147	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_Ostm1	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_u1001-110	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_u1001-73	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_u1011-76	0.0%	Null	Null	Null	Null	Null	Null	Null
Ots_107485-138	76.1%	55.6%	50.0%	50.0%	41.7%	50.0%	50.0%	49.3%
Ots_icta-101	59.8%	50.0%	50.0%	31.3%	50.0%	41.7%	50.0%	45.5%
Ots_103861-377	60.9%	45.5%	20.0%	55.0%	41.7%	72.7%	33.3%	44.6%
Ots_108210-503	51.1%	50.0%	40.0%	50.0%	25.0%	30.0%	50.0%	41.5%

Ots_128495_96	60.9%	39.3%	16.7%	28.6%	30.0%	25.0%	58.3%	32.1%
Ots_129303-499	75.0%	14.3%	25.0%	33.3%	35.7%	37.5%	31.3%	31.9%
Ots_102483-367	57.6%	27.3%	25.0%	27.8%	40.0%	40.0%	12.5%	28.3%
Ots_109243-285	31.5%	Null	75.0%	33.3%	37.5%	0.0%	33.3%	27.6%
Ots_PEMT	78.3%	5.0%	8.3%	20.8%	0.0%	15.6%	31.3%	12.5%
Ots_111666-408	75.0%	0.0%	7.7%	5.6%	0.0%	10.7%	0.0%	4.3%
Ots_103683-279	69.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Ots_105401-325	48.9%	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	2.2%
Ots_102195-157	100.0%	0.0%	17.9%	6.3%	15.6%	12.5%	8.3%	9.8%
Ots_104063-132	89.1%	6.7%	8.3%	13.6%	16.7%	26.7%	12.5%	14.0%
Ots_110594-242	85.9%	26.7%	25.0%	30.0%	36.4%	33.3%	38.9%	30.4%
Ots_apoa2-54	100.0%	25.0%	25.0%	18.8%	9.4%	12.5%	20.8%	17.9%
Ots_Phos	94.6%	10.7%	19.2%	13.3%	25.0%	21.9%	9.1%	17.8%

Table 2: Chinook salmon parentage SNP panel. Shows the percentage of fish genotyped and the average minor allele frequency at each of the 6 tested collections.

Assay Name	% Genotyped	minor allele frequency
Ots_100884-287	98.9%	0.187
Ots_101554-407	100.0%	0.478
Ots_101704-143	98.9%	0.148
Ots_102414-395	100.0%	0.495
Ots_102801-308	100.0%	0.239
Ots_103122-180	100.0%	0.168
Ots_104415-88	100.0%	0.462
Ots_105105-613	100.0%	0.326
Ots_105132-200	98.9%	0.154
Ots_105385-421	100.0%	0.435
Ots_105407-117	97.8%	0.444
Ots_108820-336	100.0%	0.484
Ots_109525-816	100.0%	0.147
Ots_110064-383	100.0%	0.293
Ots_110201-363	100.0%	0.332
Ots_110495-380	100.0%	0.141
Ots_110551-64	98.9%	0.258
Ots_110689-218	100.0%	0.201
Ots_112301-43	100.0%	0.158
Ots_112419-131	100.0%	0.136
Ots_112820-284	100.0%	0.196
Ots_112876-371	100.0%	0.141
Ots_113242-216	100.0%	0.065
Ots_115987-325	100.0%	0.277
Ots_117432-409	100.0%	0.212
Ots_118205-61	98.9%	0.121
Ots_118938-325	100.0%	0.223
Ots_123921-111	98.9%	0.176
Ots_124774-477	98.9%	0.121
Ots_128757-61R	100.0%	0.103
Ots_129458-451	98.9%	0.115
Ots_94857-232R	97.8%	0.433
Ots_94903-99R	100.0%	0.451
Ots_96500-180	100.0%	0.462
Ots_96899-357R	100.0%	0.109
Ots_ARNT	100.0%	0.114
Ots_AsnRS-60	100.0%	0.201
Ots_brp16-64	98.9%	0.170
Ots_CD59-2	100.0%	0.408

Ots_CirpA	98.9%	0.121
Ots_cox1-241	97.8%	0.156
Ots_E2-275	100.0%	0.288
Ots_Est740	98.9%	0.456
Ots_ETIF1A	97.8%	0.317
Ots_FGF6B_1	100.0%	0.424
Ots_GCSH	100.0%	0.071
Ots_GDH-81x	100.0%	0.190
Ots_GPH-318	100.0%	0.266
Ots_GTH2B-550	100.0%	0.495
Ots_HMGB1-73	100.0%	0.179
Ots_hsc71-3'-488	100.0%	0.168
Ots_HSP90B-100	98.9%	0.148
Ots_IGF-I.1-76	98.9%	0.170
Ots_Ikaros-250	100.0%	0.043
Ots_IL8R_C8	100.0%	0.467
Ots_mapK-3'-309	100.0%	0.397
Ots_mapKpr-151	97.8%	0.239
Ots_MHC1	100.0%	0.082
Ots_MHC2	100.0%	0.418
Ots_mybp-85	100.0%	0.152
Ots_NFYB-147	100.0%	0.163
Ots_nkef-192	100.0%	0.397
Ots_NOD1	100.0%	0.283
Ots_ntl-255	98.9%	0.313
Ots_OTALDBINT1-SNP1	98.9%	0.093
Ots_OTDESMIN19-SNP1	100.0%	0.484
Ots_OTSTF1-SNP1	100.0%	0.380
Ots_P53	100.0%	0.239
Ots_parp3-286	100.0%	0.228
Ots_pigh-105	96.7%	0.455
Ots_pop5-96	100.0%	0.266
Ots_ppie-245	100.0%	0.201
Ots_Prl2	98.9%	0.236
Ots_RAG3	100.0%	0.130
Ots_redd1-187	100.0%	0.217
Ots_S7-1	98.9%	0.324
Ots_SClkF2R2-135	100.0%	0.353
Ots_SEXY3-1*	100.0%	0.453
Ots_SWS1op-182	98.9%	0.286
Ots_TAPBP	98.9%	0.357
Ots_TGFB	100.0%	0.109
Ots_Thio	98.9%	0.187

Ots_TLR3	100.0%	0.337
Ots_tpx2-125	100.0%	0.103
Ots_txnip-321	100.0%	0.158
Ots_u07-07.161	100.0%	0.473
Ots_u07-17.135	100.0%	0.174
Ots_u07-18.378	100.0%	0.130
Ots_u07-25.325	100.0%	0.435
Ots_u07-49.290	98.9%	0.214
Ots_u1002-75	98.9%	0.220
Ots_u211-85	98.9%	0.429
Ots_u4-92	100.0%	0.054
Ots_u6-75	98.9%	0.104
Ots_unk526	98.9%	0.132
Ots_vatf-251	98.9%	0.060

*Sex determination marker.

Table 3: Fifty SNP loci added to Chinook genetic stock identification panel. Average allele frequencies across

Assay Name:	MAF:	Stream type G_{ST}	Ocean type G_{ST}:	All population G_{ST}:
Ots_101119-381	0.047	0.0012	0.0878	0.131
Ots_102213-210	0.046	0.0289	-0.0081	0.064
Ots_102457-132	0.198	0.0167	0.0611	0.345
Ots_102867-609	0.082	0.0044	-0.0045	0.118
Ots_104569-86	0.260	0.0124	0.0030	0.360
Ots_106499-70	0.385	0.0205	-0.0034	0.101
Ots_107074-284	0.200	0.0320	0.1813	0.375
Ots_107285-93	0.053	0.0185	0.0051	0.026
Ots_107806-821	0.362	0.0073	0.0742	0.041
Ots_108007-208	0.253	0.0121	0.0521	0.393
Ots_108390-329	0.067	0.0471	0.0004	0.080
Ots_108735-302	0.216	0.0197	0.0401	0.200
Ots_109693-392	0.099	0.0168	-0.0034	0.099
Ots_111681-657	0.173	-0.0025	0.0477	0.238
Ots_112208-722	0.256	0.0093	0.2114	0.468
Ots_117242-136	0.232	0.0034	-0.0186	0.192
Ots_117259-271	0.346	0.0244	-0.0133	0.793
Ots_118175-479	0.107	0.0300	-0.0022	0.081
Ots_122414-56	0.138	0.0498	0.1265	0.265
Ots_127236-62	0.252	0.0333	-0.0074	0.549
Ots_128302-57	0.340	0.0428	0.0392	0.737
Ots_128693-461	0.165	0.0275	-0.0038	0.155
Ots_129144-472	0.117	0.0080	0.0254	0.204
Ots_130720-99	0.283	0.0103	0.0221	0.503
Ots_131460-584	0.280	0.0444	-0.0129	0.494
Ots_131906-141	0.101	0.0208	0.0272	0.077
Ots_99550-204	0.059	0.0084	-0.0051	0.069
Ots_DDX5-171	0.238	0.0542	-0.0154	0.139
Ots_Est1363	0.401	0.0664	0.0375	0.841
Ots_HFABP-34	0.123	0.0594	0.0514	0.221
Ots_hnRNPL-533	0.359	0.0102	0.0132	0.359
Ots_Hsp90a	0.313	0.0261	0.0452	0.695
Ots_il13Ra2B-37	0.277	0.0302	0.0412	0.126
Ots_il-1racp-166	0.444	0.0147	0.0105	0.254
Ots_nelfd-163	0.288	0.0166	-0.0153	0.716

Ots_OTSMTA-SNP1	0.073	0.0064	0.0267	0.117
Ots_P450-288	0.400	0.0443	0.0106	0.394
Ots_stk6-516	0.007	0.0000	0.0091	0.023
Ots_TCTA-58	0.199	0.0321	0.0202	0.282
Ots_u1007-124	0.080	0.0714	0.0422	0.136
Ots_U2362-227	0.095	0.0437	-0.0043	0.050
Ots_U2362-330	0.318	0.1691	0.0187	0.399
Ots_U2446-123	0.359	0.1307	0.0538	0.102
Ots_unk1104-38	0.365	0.0169	0.0193	0.309
Ots_unk1832-39	0.346	0.0622	0.0413	0.175
Ots_unk3513-49	0.229	0.0339	0.0358	0.027
Ots_unk7936-50	0.142	0.0133	-0.0140	0.097
Ots_unk8200-45	0.003	0.0022	0.0000	0.007
Ots_unk9480-51	0.409	0.0492	-0.0032	0.698
Ots_zn593-346	0.017	0.0351	0.0061	0.035

Table 4: *O. nerka* SNP panel.

Assay Name:	Plot score:	Avg. MAF:	G_{ST}:
One_ACBP-79	0.933	0.220	0.013
One_agt-132	1.000	0.341	0.251
One_aldB-152	1.000	0.239	0.230
One_apoe-83	0.867	0.039	0.062
One_c3-98	0.800	0.077	0.088
One_CD9-269	0.933	0.181	0.186
One_cetn1-167	0.933	0.224	0.072
One_CFP1	1.000	0.364	0.327
One_cin-177	0.767	0.263	0.292
One_Cytb_17	0.933	0.135	0.623
One_dds-529	0.867	0.230	0.048
One_DDX5-86	0.933	0.166	0.090
One_E2-65	0.800	0.025	0.049
One_gdh-212	1.000	0.216	0.110
One_GHII-2165	1.000	0.438	0.289
One_ghsR-66	0.800	0.050	0.149
One_GPDH-201	1.000	0.408	0.051
One_GTHa	0.867	0.341	0.207
One_HGFA-49	0.867	0.260	0.474
One_HpaI-71	1.000	0.451	0.212
One_HpaI-99	0.800	0.142	0.092
One_hsc71–220	0.800	0.381	0.176
One_Hsp47	0.933	0.246	0.270
One_IL8r-362	0.867	0.054	0.119
One_ins-107	0.667	0.423	0.008
One_KCT1-453	0.867	0.153	0.314
One_KPNA-422	1.000	0.315	0.285
One_LEI-87	1.000	0.246	0.175
One_lpp1-44	1.000	0.047	0.044
One_MARCKS-241	0.800	0.049	0.087
One_metA-253	1.000	0.211	0.137
One_MHC2–190	1.000	0.330	0.194
One_MHC2–251	0.800	0.209	0.223
One_Mkpro-129	0.933	0.115	0.123
One_ODC1-196	1.000	0.251	0.041
One_Ots208-234	1.000	0.368	0.228

One_Ots213–181	1.000	0.393	0.410
One_p53–534	0.933	0.124	0.243
One_pax7-248	0.933	0.068	0.056
One_PIP_3	0.933	0.323	0.259
One_Prl2	0.933	0.391	0.458
One_psme2-354	0.800	0.045	0.063
One_rab1a-76	0.933	0.113	0.187
One_RAG3-93	0.867	0.276	0.108
One_redd1-414rd	1.000	0.490	0.348
One_RFC2-102	0.933	0.182	0.258
One_RFC2-285	0.800	0.085	0.101
One_RH2op-395	0.867	0.041	0.119
One_rpo2j-261	0.933	0.093	0.414
One_sast-211	0.867	0.084	0.073
One_spf30-207	0.800	0.088	0.046
One_srp09-127	0.933	0.156	0.365
One_ssrd-135	1.000	0.399	0.070
One_STC-410	0.867	0.082	0.120
One_STR07	0.933	0.290	0.382
One_SUMO1-6	1.000	0.082	0.069
One_sys1-230	1.000	0.323	0.065
One_taf12-248	0.933	0.022	0.027
One_Tf_ex11–750	0.933	0.089	0.104
One_Tf_in3–182	0.933	0.031	0.095
One_txnip-401	1.000	0.132	0.112
One_U1003-75	1.000	0.356	0.281
One_U1004-183	0.933	0.404	0.216
One_U1009-91	1.000	0.266	0.307
One_U1010-81	0.933	0.125	0.167
One_U1012-68	1.000	0.464	0.144
One_U1013-108	0.867	0.015	0.047
One_U1014-74	1.000	0.340	0.061
One_U1024-197	0.867	0.296	0.122
One_U1101	1.000	0.430	0.327
One_U1105	1.000	0.377	0.304
One_U1201-492	1.000	0.383	0.078
One_U1202-1052	0.800	0.033	0.024
One_U1203-175	1.000	0.359	0.334
One_U1204-53	0.933	0.186	0.176
One_U1205-57	0.867	0.043	0.184
One_U1206-108	0.933	0.251	0.370
One_U1207-231	0.933	0.308	0.041
One_U1208-67	1.000	0.238	0.144

One_U1209-111	0.933	0.072	0.062
One_U1212-106	0.933	0.339	0.345
One_U301-92	0.800	0.037	0.030
One_U401–224	0.867	0.325	0.197
One_U502–167	0.933	0.119	0.076
One_U503–170	1.000	0.376	0.166
One_U504–141	1.000	0.462	0.178
One_U508–533	0.800	0.059	0.042
One_UCA-24	0.667	0.173	0.103
One_vamp5-255	0.867	0.039	0.163
One_vatf-214	1.000	0.232	0.111
One_VIM-569	0.933	0.097	0.143
One_ZNF-61	0.933	0.318	0.100
One_Zp3b-49	1.000	0.271	0.367
One_U1214-107	1.000	0.296	0.114
One_U1215-82	0.733	0.265	0.222
One_U1216-230	0.933	0.189	0.593

Appendix 1. Details on the 96-SNP PBT and 96-SNP GSI assays utilized for the Chinook salmon 192-SNP GSI baseline.

Panel	No.	Assay:	Ref.:	V/F:	Fwd Primer	Rev Primer	VIC-mgb probe	FAM-mgb probe
OtsPBT	1	Ots_100884-287	5	T/C	CGGAAGACCAGATTCTCCAAGAGTA	CGACCAAGTAGCGGCACTT	ATAGAACTACAATTCACATATAT	AACTACAATTCGCATATAT
OtsPBT	2	Ots_101554-407	5	C/G	TGAAAGATATCAATTGTAGTAGTGGTGGTG	ACACGCCAGTCCACAAGT	ATGGAGGATTGTGGTTGT	ATGGAGGATTCTGGTTGT
OtsPBT	3	Ots_101704-143	5	T/G	ACTTCTTGAGCCAATCGGATGATG	CCAGAGATAAACTAGTGGAGGAGATCA	CTTAGACGTCAGAGGTC	CTTAGACGTCCGAGGTC
OtsPBT	4	Ots_102414-395	5	A/G	GCCTACTGATAAATGTATGACAGTAATGGA	CAATAACAAACAAGCTAGGAACAAAAGTGT	CACATAGTGTAGCTTTACTAC	CACATAGTGTAGCTCTACTAC
OtsPBT	5	Ots_102801-308	5	C/A	TGGGACAGAGGTGGGAATTGA	CCCAAAGATGCTTAACTGAAGATGTG	AGGGACAGTTTCGCAGACG	AAGGGACAGTTTCTCAGACG
OtsPBT	6	Ots_103122-180	5	T/C	CAAACGCGCACTCACACA	TCACAATGGTACGATTTTACGACTCAA	CATCAACACAATCTGC	CATCAACACGATCTGC
OtsPBT	7	Ots_104415-88	5	C/T	CCTGAGCATCCCAGTTGAACT	TGTTTTCAATACACTGCAATTTAGTTTTGGT	TCCTGAAAAACGACATCC	CTGAAAAACAACATCC
OtsPBT	8	Ots_105105-613	5	C/G	AGTACAAGTGCAGAGAATGACATCATG	GGTGTTTTATTTTCCCATATATCTTTTAACTTTAAGCT	CCGAGCTTGAGTTAGGA	CCGAGCTTGACTTAGGA
OtsPBT	9	Ots_105132-200	5	G/T	CGATGTACTGAGGGCAGTGT	GAGTGGAGTTCCTTAATAATCATTGACCTT	CAAGAGTGGCATAAAA	CAAGAGTGGAATAAAA
OtsPBT	10	Ots_105385-421	5	A/G	GACTGTCTTGGAACCGTTGCTA	TCCCGGAACACACCAATGTC	CCTCCTGGGTATATCG	CTCCTGGGCATATCG
OtsPBT	11	Ots_105407-117	5	T/A	TGTGTACATCCGCGTAAATATTGAAGATAA	CTGTGAGCTGCTGCAAACC	CAGGTTAGGAATGGTTG	CAGGTTAGGATTGGTTG
OtsPBT	12	Ots_108820-336	5	G/A	TGAAATAAATTGTTCTGTTGATATGT	CAACGACACACCAACAACGT	ATTGCCCATCTCAGAATA	AATTGCCCATCTTAGAATA

OtsPBT	13	Ots_109525-816	5	C/T	GAATTTTGGAGCCAGATAGTAGCGTACATCATGAG	CTCCCCATGTCCCTGAGTCT	CATGAGGCGTTCGGC	ATGAGGCATTCGGC
OtsPBT	14	Ots_110064-383	5	C/T	AACAAAGAATGTTAAACACCAAACAGGAA	GTGCAAGGGACCTAGCTAATCC	CTACGTAATGAACGTTAGCT	ACGTAATGAACATTAGCT
OtsPBT	15	Ots_110201-363	5	A/T	GTTTGGCTATTGAAATTATACATTAAAACATGTAGCT	CCATGGCATCCTGTAAAGAACAACA	TGGATGCCAGTTTTAAAA	TGGATGCCAGTTTAAAAA
OtsPBT	16	Ots_110495-380	5	G/C	GCCTAGGTATGTACGAAACTTCACA	AGGCTTTTTCAGATGGTCGTATGA	ATGGCCCCTGTCTATG	ATGGCCCCTGTGTATG
OtsPBT	17	Ots_110551-64	5	C/A	GAGTGGTCAAGGTTTCAGTTTCTG	GAAATGGACAGACACAAGGTCAAAC	ACGCTCGGAACATT	ACGCTCTGAACATT
OtsPBT	18	Ots_110689-218	5	T/G	GTATAAACTAGAGTCCAGTGTTATGTTAATGTCTT	CATGGCAGACAACAGTAGAGAATATGA	CACCAATCAATTAATTATT	ACCAATCAATTCATTATT
OtsPBT	19	Ots_112301-43	5	T/C	GCATGGCTGCCCTAGAACA	TCAGAACATTTCCTTCAGCTTCGT	CGTCGCATTCAGC	CGTCGCGTTCAGC
OtsPBT	20	Ots_112419-131	5	A/T	GTGGGTAATCGATGCCAAAGAGAT	TGGCAGTGTTTTCAACTAGCTTTG	AAGCGACTTGATTATC	AGCGACATGATTATC
OtsPBT	21	Ots_112820-284	5	C/T	CATAGATGTTTATATGAAAAACCTCCCACTGT	GCATCCAAAAAGACGTGTGTGTTT	ACTCACACTCGAGTGACT	ACTCACACTCAAGTGACT
OtsPBT	22	Ots_112876-371	5	C/A	GCCTACAGCAAATTCAGCTACACAT	TGGACCTTCAATCATCACAGCTT	CATCACAACGATGTGTG	CACATCACAACTATGTGTG
OtsPBT	23	Ots_113242-216	5	C/T	GAGGCCTAATGTCTCTTGTGACT	GACATCTTCAACAAGTGTTCATTCACC	ATTACCAACGGAGAACC	TTACCAACAGAGAACC
OtsPBT	24	Ots_115987-325	5	T/G	GGAGGTGTAGTGAAATGGGAAGAT	GCATTCAGTGAACCAGTAGTGCTAT	ATGCATAAAAGGTAATTGTG	ATGCATAAAAGGTCATTGTG
OtsPBT	25	Ots_117432-409	5	A/G	TCATCAAAACATGCCTCTTCTGTGT	TGTTGAACCTGTCACTCTGTCTTC	TTTAGACTTTGCTCTATAACAG	ACTTTGCTCCATAACAG
OtsPBT	26	Ots_118205-61	5	T/C	CCATACAGCCAGTCCAGGTG	ACTGGACAGGGCTGGGT	TAGTAGCCCCTACACCTC	TAGCCCCTGCACCTC

OtsPBT	27	Ots_118938-325	5	C/T	ATTTTCAAACAGGCATTTATCATTGGTGAA	GGTCTGTCCCTCATTCTTTGCA	AGAGATGCAAAGTGGAGTT	AGAGATGCAAAATGGAGTT
OtsPBT	28	Ots_123921-111	5	A/G	TCGCTAGGCAGAAATATAGGGTTCT	GAGCATGGCGCTTGCA	TGCTAAATGGCATATATTAT	CTAAATGGCACATATTAT
OtsPBT	29	Ots_124774-477	5	T/C	AGTTGTTCTTTTTATATTGTGTTTTTATTCCATTCCA	GCCAAATAAAAACAAAGCATGAACACA	CCACCGCCATCTGATA	CACCGCCGTCTGATA
OtsPBT	30	Ots_128757-61	5	A/-	CGTGTCCGGCTTCTTTTATTTCATT	GATGGGTATGTTAATCATATTACCAGCGTAA	TTGTGCATTTTCCCC	TGTGCATTTCCCCC
OtsPBT	31	Ots_129458-451	5	T/C	TGGGACCCACATAAAGCAACTG	GACATAAGACCCATTTAGCCCCTTTT	CATCTGGCAATGCCTT	CATCTGGCAGTGCCTT
OtsPBT	32	Ots_94857-232	5	T/C	GGCACTCTCCCTGGCTAGA	CCCCATCACTTCTCTGGCTTTAAAT	CAGGATAATAACAAACAAG	CAGGATAATAACGAACAAG
OtsPBT	33	Ots_94903-99	5	G/T	CCGTCTGAGTAGGAGGATCAATACA	TTTGGATCCAGCTCTCCGTATAGA	CAAACCAGCAAACAT	ACAAACCAGAAAACAT
OtsPBT	34	Ots_96500-180	5	G/T	GATCATGTCAGATAGGATGCTGAAAGT	CAGGTCTGGTCTACATCGAACAC	AAAACAAATCATTTTTCG	AAAAACAAATAATTTTTCG
OtsPBT	35	Ots_96899-357	5	T/A	TCTCCTGAACTAATTTAGACCTCTGAATGT	CCTCATATTGCTTTCATCTGAAGAGAGA	CTGAATGTTTTTTTTAATCTTT	CTGAATGTTTTTTTTTATCTTT
OtsPBT	36	Ots_ARNT	9	G/T	CCACTGGCTGTGGAGCTT	GGGTTCAGTGATAGTTGGGCAAAT	TACAGATGTCATTTTAC	CTACAGATGTAATTTTAC
OtsPBT	37	Ots_AsnRS-60	2	T/C	CCGACGCCTCACTGAGT	TGGTTTTTCAGGTCATGGTTTCCA	TGAGTCCCTGACCAGC	AGTCCCCGACCAGC
OtsPBT	38	Ots_brp16-64	15	T/C	ACTCTGGGTCCAGGAGGTTTT	CTGACGAGACCATGCACCAA	AAGTCAGCATCTTTCA	AGTCAGCGTCTTTCA
OtsPBT	39	Ots_CD59-2	9	G/A	TGTTTATCTCTGAGTGAAAAAGGTGTGT	CATGTTACCCAGCTAAAAGTCTATAGCA	CTAAAATGTCATGTAAATAT	ACTAAAATGTCATATAAATAT
OtsPBT	40	Ots_CirpA	12	C/T	GCTGTGATTGTGCTCTAAAGACATG	CTCCCACTTAGCATTCCTACCTT	AATGCATTACAGAACTGA	AATGCATTACAAAACTGA

OtsPBT	41	Ots_cox1-241	4	C/T	CACTGAACTGTAAGCCATTGTGATT	GTAAATGTAGTATACAGTATAGGCATCGTAGGT	CACTACGGTAAGACCAT	CACTACAGTAAGACCAT
OtsPBT	42	Ots_E2-275	2	A/G	GGTGCCACTTTAGTATAGCTGCTTA	CCCTACCCCCTGTGTTCCA	CCCCCATATTGCTG	CCCCACATTGCTGG
OtsPBT	43	Ots_Est740	12	T/C	GGACTCGTGCTTGAGGAAGATG	TGCATGGCTCCAACTCCTT	TCTGGATGGAACCGTTAG	CTGGATGGAGCCGTTAG
OtsPBT	44	Ots_ETIF1A	9	A/C	TCTGAACTCACCAAAGGAACACTTG	GAGAGAAAAGGAGAAATGATTGCCATT	CAACTGAAGAAAATAATATG	CTGAAGAAAAGAATATG
OtsPBT	45	Ots_FGF6B_1	7	A/C	GAGACAAAGGTTTGCAGGTTCATG	GGGAGCCATGCACTAATATATTGGA	CCTGTTATCAGACCCAAAT	CTGTTATCAGCCCCAAAT
OtsPBT	46	Ots_GCSH	12	C/T	GTTCTTTTTAATGATGACTACAGGTCTTTCAC	GCTACTTTACATAATACCATTTGAGCTGAGA	TATCTGGGCGGGCTG	CTATCTGGACGGGCTG
OtsPBT	47	Ots_GDH-81x	4	C/-	CTTTTCTGAATTAGTGCTGTGCTTGT	CCAACTTCTTCAACTCTGTCAGTGA	TGTTACGGGACATACT	TCTGTTACGGACATACT
OtsPBT	48	Ots_GPH-318	3	C/T	GGTGATAACAGGTGTTGCACCAA	TCAGGTGGTGGTGGACAAC	ATCAAGCTGACGAACCA	CAAGCTGACAAACCA
OtsPBT	49	Ots_GTH2B-550	7	C/G	TGACTACCCGTTGTACCAATGAAC	CACAGGAAGGACGTGTTTTGATG	TTAATGCTGCAGATGTTAT	ATGCTGCACATGTTAT
OtsPBT	50	Ots_HMGB1-73	15	G/T	TGCTTCAGTGAAAATAAGCGTGAGA	GTCGAGCGGTATGAATACTTTCTGA	ACTGTATATGTTACGTTTTC	ACTGTATATGTTAAGTTTTC
OtsPBT	51	Ots_hsc71-3'-488	6	C/T	TGCATCCATTCATACCTGACCAATT	TTTGGTTAGGCACACGATAATTTGC	TTTCCAATGGTATAGATATGA	TTTCCAATGATATAGATATGA
OtsPBT	52	Ots_HSP90B-100	3	C/T	CACCTTAGTTCCACGCAACATG	CTGCGTGTATTGTAGTGGTGACA	TCTATGGTGTGATTCATT	TTCTATGGTGTAATTCATT
OtsPBT	53	Ots_IGF-I.1-76	2	A/T	GGTAGGCCGTCAGTGTAAAATAAGT	GATGGAGGCCACTGTGTTCTTA	CTGCCTAGTTAAATAAAATA	CTGCCTAGTTAAATTAAATA
OtsPBT	54	Ots_Ikaros-250	2	G/A	GAGGCTGACTTGGACTTTGC	GGCCTGTCAGCCAAGGA	ACAGAAGATTTTCGGCTGC	ACAGAAGATTTTCGACTGC
OtsPBT	55	Ots_IL8R_C8	9	C/T	CGTGGTGTTCGCCTTCCT	TGTCGGCCATCACTGTCATG	CTGGACGCCGTTACA	TGGACGCCATTACA
OtsPBT	56	Ots_mapK-3'-309	11	T/G	CGTGACCCTTGTAACTGAAAAGC	GGCCACTGTCATAGAATTAGGCATT	ATGCTATTAAATGAATATTC	ATGCTATTAAATGACTATTC

OtsPBT	57	Ots_mapKpr-151	11	A/T	TGTTGTCTCGGACTGCATGAC	GAAGGCACAGAGATGAAGGACAT	CATGCATTGCACATAC	CATGCAATGCACATAC
OtsPBT	58	Ots_MHC1	1	G/A	GTCCACATTCTCCAGTACATGTATGG	CAAACCCCTCTGTCTGTTCAGT	CATCATCCCGTGAGCAG	TCATCATCCCATGAGCAG
OtsPBT	59	Ots_MHC2	1	T/G	GTCCTCAGCTGGGTCAAGAG	GTAGTGGAGAGCAGCGTTAGG	CTGGAGCGTTTCTGTA	CTGGAGCGTGTCTGTA
OtsPBT	60	Ots_mybp-85	10	C/T	CAAGGGATGTGACAAATTAATCAAACACATAA	AAGAGGTCTAATAAATCTCCAATGTAAAAACGT	AGAGCATGTAGTTTTG	AGCATGTAATTTTG
OtsPBT	61	Ots_NFYB-147	15	C/T	CCGTCCACAGCACAAGACTATAATA	CAGATGATAGCTTCAGTAAGTGGTTCA	TGTTCCAATGTAAAATGTATGC	TTCCAATGTAAAATATATGC
OtsPBT	62	Ots_nkef-192	4	C/T	CATTTAGCAGACACTCTTATCTTAGTGTCA	CGAATGTCCACCTCAGATGTTACAA	AATAGGCCGACATCAA	AAATAGGCCAACATCAA
OtsPBT	63	Ots_NOD1	7	C/G	GTGCTGCAGGAACCATGTG	CTGTGTGGACTGCTGTCTAAGG	CCAACGGCGACTTG	CCAACGCCGACTTG
OtsPBT	64	Ots_ntl-255	11	T/A	TGCAGTTACAAGCCTAAGACAATCT	CAACTAAAGTAACACACCAGCAACTG	TTGTAGAGGAAGAATATTC	TTGTAGAGGAAGTATATTC
OtsPBT	65	Ots_ALDBINT1-SNP1	5	T/C	CGCTGGGCATGGATGAGT	GGCCAACACTGCTACTTCCT	CTACTGTTGTATTTTCTC	CTGTTGTGTTTTCTC
OtsPBT	66	Ots_DESMIN19-SNP1	5	C/A	GGTCTGTCTGTCTGTCTATCTGTCA	TGTGTGTCTTTGTTCATTCCTACCA	CCAGTCATGGGTCATT	TCCAGTCATTGGTCATT
OtsPBT	67	Ots_TF1-SNP1	5	G/T	CGGACAAAGAGCTACAGAAATGC	CGTCCCTCTTCACGCATGA	CCGCCACCTTGGCT	CGCCACATTGGCT
OtsPBT	68	Ots_P53	1	G/A	GGAACTTCCTCTCCCGTTCTG	GCACACACACGCACCTCAA	CTGGGTCGGCGCT	TGGGTCGACGCTC
OtsPBT	69	Ots_parp3-286	11	A/G	AGTCAGTGTTGGTGTAGTGAAGAGA	CATTTGTGGAGTGTTTATTGAACAGTAACA	AGTTACAAGTGGTGTTTCA	ACAAGTGGCGTTTCA
OtsPBT	70	Ots_pigh-105	11	A/-	GTTTGGAATGTTTCTCTGATTGTGTTAACAA	GCATTACTAAAAACTGGTGTGTGGAA	TGACCTGAAAATATATATTTTT	ACCTGAAAATATATTTTTTT
OtsPBT	71	Ots_pop5-96	11	T/C	CTCTTGCTACTTG	AGTTTGAGGGCTC	TTCTGTTACTGGA	CTGTTACTGGGCT

OtsPBT	72	Ots_ppie-245	11	C/A	CAGTGTATCTCATGTTTTTGGTCATGTATTTTCTCTGCTATTTTT	TATTCTGTCATGGGACTGGAGCTGCTGAACATA	CTGATGATGTCTGAAATGAAAGCC	GATGAATGTCTGAAATTAAAGCC
OtsPBT	73	Ots_Prl2	1	A/G	CCTGGTCTGTTTGTGATCAAGATG	GGTTAACTCAAATAGAACATACTCTGACACA	ATGTATTGTTCATTTAATG	TGTATTGTTCGTTTAATG
OtsPBT	74	Ots_RAG3	7	C/T	CATTTCCACGAAAAGCCAGATGAC	ACAGAATAAAGTATCTTCCTCTTACATCACTACTAAT	CTCTACAGTATGAACTATG	CTCTACAATATGAACTATG
OtsPBT	75	Ots_redd1-187	11	A/G	TTCTGGGTTGCCATACTCTTTCAAT	AGTTGAGACCTTCAGTTCTTAGGGTAT	ATTCTGACAGCTGTTTTG	CTGACAGCCGTTTTG
OtsPBT	76	Ots_S7-1	7	T/C	TGCCATCATAAACAACCTAACAAGTAACT	CCTGGTTTAAAAACGGCCAACTG	TACAGGAGATAAGGTCGCA	CAGGAGATAGGGTCGCA
OtsPBT	77	Ots_SClkF2R2-135	2	A/T	CCAAATACAGACCAGCTACTTGTGT	CTTCAAGTCCCTGAATAATGGTACGT	ATTCAAAGTCAAATTTT	ATTCAAAGTCTAATTTT
OtsPBT	78	Ots_SEXY3-1	13	X/Y	GGTCTTGCAGTCAGGAGAGG	CCAGGTGGTGAAGGTAGGAA		ATCTCCACTTCGCTGA
		Ots_SEXY3-1 AC			TCCTTGTGTCTAAAGGGCTTTGAG	GGGCTTGCTAGTCCTAAACAGATC	CAGAATTAGCTTTGGACATT	
OtsPBT	79	Ots_SWS1op-182	2	T/A	TCAAAGACATCGAACACAAGAACGA	GCAGGTAAATTCAAACGTCATCATAAGAA	ATGTACTTTAACGATTCATTT	ATGTACTTTAACGTTTCATTT
OtsPBT	80	Ots_TAPBP	9	C/T	TTTCTCATCCTTCTCTCTTCCAGTCT	GGACAAACCAGCACTCCAGAA	CTGGACAGCTGGTCC	CTGGACAACTGGTCC
OtsPBT	81	Ots_TGFB	9	C/T	GCCTCACATTTTACTGATGTCACTTC	GAGCAGATCTCTTCAGTAGTGGTTT	CTTCCGAGAGCTAGGCT	CTTCCGAGAACTAGGCT
OtsPBT	82	Ots_Thio	12	T/C	TTTTAAAAATGGAGATAAACTCCTGACCTGAA	AATACCAAACCATGCCACTAATACCT	CAGTGTATTAGTCATTCTTA	CAGTGTATTAGTCGTTCTTA
OtsPBT	83	Ots_TLR3	9	C/T	TGCACCTGCGAGAGCAT	CTGGCGTTTGTTCCGTTCAG	CTGTGGTTTGTGGCGTG	CTGTGGTTTGTAGCGTG
OtsPBT	84	Ots_tpx2-125	11	C/T	TGTTGTAATCTTT	TCTTCCAAATTGA	CAGGCGGTTCTCC	CAGGCAGTTCTCC

					CTGAATATTTGCTTGCTT	GCACAAAAGCATTGCTT		
OtsPBT	85	Ots_txnip-321	11	T/C	CCTTCAAACTAACACATCATAGACATGCTT	TTATCAAACTGAAGGCGGATTTACTGA	TCTGGCGGATTTACA	CTGGCGGGTTTACA
OtsPBT	86	Ots_u07-07.161	8	C/T	GTCAACAAATGCAGGTAACATAAATGGT	GATGCAAACACCTGTGAAATTGTGA	ATCAGTGACATAAGTTGTCCA	TCAGTGACATAAATTGTCCA
OtsPBT	87	Ots_u07-17.135	8	A/G	CTCGCCTCTGTCATTGTATTACCTT	TGACACACGAGCCATTTTGATGAT	AAAATGTACCACATACTTGT	AAATGTACCACATACTCGT
OtsPBT	88	Ots_u07-18.378	8	A/T	GGAAACCAGCTAGGATTCAGGAA	CGTTATATGGTTTGCTTGTTTGCGATA	ATATGGTATGTAGAGGCTAGTTA	TATGTAGAGGCAAGTTA
OtsPBT	89	Ots_u07-25.325	8	T/C	AGACAATCATGGTGTTTTGAGTCTTTCT	GCCTAGGCTTGATGGAGTCA	CCGCTTGAAAGTTTGA	CGCTTGAAGGTTTGA
OtsPBT	90	Ots_u07-49.290	8	G/A	GCTGAGGAAGGATTCTGTATTTGCT	TCGGACAGAGCGCATCC	CTTTCCCCGTGTTGGT	ACTTTCCCTGTGTTGGT
OtsPBT	91	Ots_u1002-75	8	T/C	CCGCCTTTCCCACCTTCTC	TCAAACGAGAACACACTAAGGTTGT	ATGGCCCTTACACTATC	TGGCCCTTACGCTATC
OtsPBT	92	Ots_u211-85	2	C/T	TGGTGAGAGCAGCTTTAAATGTCTT	ACCCATTCTTCTGTCTGGTTTAAGC	TCCCAAAGTCGAGTGTG	CCCAAAGTCAAGTGTG
OtsPBT	93	Ots_u4-92	2	T/C	ATCCAAGGAGCCCCATTAAAGATTT	CGTACCAGAGTTGTAGAAGCATCT	CTGTGTTGAATTTAACATAAT	TCTGTGTTGAATTTAACGTAAT
OtsPBT	94	Ots_u6-75	2	C/T	GAAAAAGTAAAGTAAAAGTAAAGTATTATACCACTAAAGACAAT	GATCCACACTGTTGGTCTACTACAA	TTAGTCAACTGTTGTTTTT	TTAGTCAACTGTTATTTTT
OtsPBT	95	Ots_unk526	7	A/G	TCAAGACTGTGCTGTAGTTGTCTAC	CCTCCCCCTTTTCCACATCAG	CAACATTCCAGTCTGAAAC	CATTCCAGCCTGAAAC
OtsPBT	96	Ots_vatf-251	11	G/-	CTTTTCGGGTTATTCATGCTGTTGT	GCAAGCATTTGAAAAACAGACTGGAT	AGACCACAAGATACAGTACC	AGACCACAAGATA--GTACC
OtsGSI	1	Ots_101119-381	5	T/C	TTTTCTAGGACAGGTTGCTTGCA	CCAGGTTTCTTTAGCCTACTTATTCTTTACA	TGCCACATGATAATTGA	CCACATGGTAATTGA

OtsGSI	2	Ots_102213-210	5	A/G	CATTCCATGACAATGATTGAAATCTAAAAACAC	GAGTATCTCAATTGCAACACTATGGTATGT	CTGTATACAGTAAGAGTATTAAT	ACAGTAAGAGCATTAAT
OtsGSI	3	Ots_106747-239	5	C/A	ATCGAGGATGCCTCAAAGACATC	GTTAGACCCACCACCAGTCATC	CCCGCGGTGAGTAT	CCCGCTGTGAGTAT
OtsGSI	4	Ots_102457-132	5	A/G	CCAGCAGAGACTGGGTTCAC	TTCCCTACCGGCGAAACC	CAATTGTGCGTTGCCCCA	ATTGTGCGTCGCCCCA
OtsGSI	5	Ots_102867-609	5	A/G	CTCTGCCATTCATTTGGGCTTTG	GTCTAAAGTGGTCCCCTTGGAT	ACAGAGAGAAGTCCCAGGTG	AGAGAGAAGCCCCAGGTG
OtsGSI	6	Ots_arp-436	11	A/T	GCCCTGGAGAAGTACGTTTTAAACTAA	GCAACCATGTCAACATTGCACATAA	CTAGGTGAAACTTTTTTTAAA	CTAGGTGAAACTTTTTAAAAA
OtsGSI	7	Ots_104569-86	5	T/G	CCTGCATGTTGTTCACGTTGTC	CGGCCGGAGGGATCAC	TGGTCGCAGATGCC	TGGTCGCCGATGCC
OtsGSI	8	Ots_106499-70	5	C/G	ACTCTATCATCGGCAGGACCAT	ACCGTAAGTGTGGTTGTGTTCATTA	CTCATTTTTCAGAATTGTATTC	CTCATTTTTCAGAATTCTATTC
OtsGSI	9	Ots_107074-284	5	A/T	CCCACTTCCAGAGCCTGAA	TTTTCCATGGCTGTGTGTACTGT	ACCGTAGCTGCACCTG	CGTAGCAGCACCTG
OtsGSI	10	Ots_107285-93	5	T/A	GCCCTTGTGACAATGCACTGTTATA	AACATACACCAATACTTAGGTCTAGACAGT	AAGTAACGTATCAAATGGC	AAAGTAACGTATCATATGGC
OtsGSI	11	Ots_107806-821	5	T/A	CTCCCTTGCTTTTGGTCATTGG	TGCAGTGCTGAATTAGAGATTAATTTTTGTG	CAAAGAAAATCAAAATTT	CAAAGAAAATCTAAATTT
OtsGSI	12	Ots_108007-208	5	A/T	CAGGCTTGTGTTAAGTAGGGAGAAA	CATTGGACAAGACCGGGTAGTC	CAGTTTCACTTAATTTTAAAATG	TTTCACTTAATTTAAAAATG
OtsGSI	13	Ots_108390-329	5	G/C	GAGGTTTGTTACTGTCACCCATAGA	CCTGCTGTAGCAAACTGTCTCAAA	CTACTTATGTAGCATTTTAA	CTACTTATGTAGGATTTTAA
OtsGSI	14	Ots_108735-302	5	C/T	CCTTTTTCTTATTAGTTTTACTTCCCCAGAGA	CAATTCCATTCTTGATTCTGTTTAACGGT	AAACAAACAACGCCTCATG	AACAAACAACACCTCATG
OtsGSI	15	Ots_109693-392	5	T/G	TCTCCCTCATTCCCATGTCATATCA	GGGAACGTATCAGGTGAGTGT	TCCGTTAGTTCATCCTGG	TCCGTTAGTTCCTCCTGG
OtsGSI	16	Ots_Cath_D141	5	T/C	CACTTGTTCTGCACACTACTTGTC	CACACATGGATTTTGCCTGTCTAAA	TGGGAAGCAATCAA	AATTGGGAAGCAGTCAA
OtsGSI	17	Ots_111681-657	5	G/T	CTGAGCTTTTTCA	GGCGCAGCAGCA	TAGCGCAAACCC	CGCAAACACCGA

					ACTTACTTGTTGGA	ACTG	CGAACC	ACC
OtsGSI	18	Ots_112208-722	5	C/A	CTGCATGAACGTTAACTCAAATAAAAGGT	AATGAGTTCTACTGACATTGTATACTAGAATAAGTATCA	TGTGAGGGCGGTCTT	ATGTGAGGTCGGTCTT
OtsGSI	19	Ots_117242-136	5	A/G	GTGACAGGAGACAGAAAGAGACATT	TGGTCCTCCCTGTCTCTATCTACTA	CAGCACATAACTTGACCTC	AGCACATAACCTGACCTC
OtsGSI	20	Ots_117259-271	5	T/C	ACACCCACTTCAACCTCCATAAC	GCCTCAGAGCTTAGCTTGGA	CTCTCCTGATCACTCTGT	CTCTCCTGATCCCTCTGT
OtsGSI	21	Ots_118175-479	5	C/T	TGCGCGTCTCATTCAACCAT	ACCTTACGTCCTAGGTAGGAAACA	AGAATGAAGTGAAAAGAA	AGAATGAAGTAAAAAGAA
OtsGSI	22	Ots_122414-56	5	C/T	GCACCGTATCAACGAGCTCAT	TGCATGGATTTCCTTTGTGTTGTTG	TGTATGACCTCTGACCTGT	TGTATGACCTCTAACCTGT
OtsGSI	23	Ots_127236-62	5	T/A	TGGAGAACTTGCACTGAATGTGAAA	GCTGTTGGACCTTGACTTTAACAAATT	TCTCTTATCTGAGTTCTGC	CTCTTATCTGTGTTCTGC
OtsGSI	24	Ots_128302-57	5	C/T	GGTTGCAGGGCAGAACTGT	ACCCATCCAATAACCCATTTTCCTT	CCTGCAATACGACCAAC	CTGCAATACAACCAAC
OtsGSI	25	Ots_128693-461	5	C/T	TCAATGTTCATCAATGCACTTCCTGTA	GCCTGCAGGAGAAGGTAGAGTTA	CACTCAGCTGGTACCCA	ACTCAGCTGATACCCA
OtsGSI	26	Ots_129144-472	5	C/A	CTGTTAGTGCAGAAGACGTAGCT	GCAGAGCTATTGAGCCAAGTTACAA	TGGGTCTCGAGCCTGTA	TGGGTCTCGATCCTGTA
OtsGSI	27	Ots_130720-99	5	A/G	CGGTCATTGTAAATGTCAACGGTTT	TGCTTGCATGTTCTTGGTGTAGTAA	CCTGTCTCATTCCC	CTGTCCCATTCCC
OtsGSI	28	Ots_131460-584	5	T/C	CCTATTTTTGATAGGTCATAGTGAATGGGATAG	CTGTACTCCTCCATTCCTTTTCACT	CTATCAAAGCAATACATTG	CTATCAAAGCAGTACATTG
OtsGSI	29	Ots_131906-141	5	A/T	GGCTCGAACCACCCAGTTTA	TGCCCAACTGGTTTGCAATC	CACGGTTTACACTCCTATTA	ACGGTTTACACTCCAATTA
OtsGSI	30	Ots_99550-204	5	C/T	TGACAGATTTCACCTTTAACTAGCTAAGC	GCAACCTCTTTCACACTTCAGTAAC	AAGGCTTTGGTTGTTTG	AAGGCTTTGATTGTTTG

OtsGSI	31	Ots_DDX5-171	16	C/T	ATGACCAATTGAAGAGTTCTTCCGT	CAAAGCCAAACGTCACATTTACACT	TTCATAATTGAACGATTTCA	CATAATTGAACAATTTCA
OtsGSI	32	Ots_Est1363	12	A/T	GGTGATTTTGCCACAGAGTAGAGAT	AGTGTTAAATGTAACTTGCATATACAGGCAAT	CCATCCTGTCTTGTCTG	CATCCTGTCATGTCTG
OtsGSI	33	Ots_HFABP-34	16	C/T	CAAGAACACCGAGATCTCCTTCA	TCGGCGGTGGTCTCG	TCGAACTCCGCTCCTAG	TCGAACTCCACTCCTAG
OtsGSI	34	Ots_hnRNPL-533	3	A/T	TCTTTGATATTGAGCTCATAAAAGCAAGGT	TCCTTGTTCATCCATCAGGCATAAAA	CATTTACCAGTTCTCACACAC	TTTACCAGTTCACACACAC
OtsGSI	35	Ots_Hsp90a	5	G/C	ACAGTATACCGGCTGCCTATTCATA	GTCGTTTTTCATAGAAAATAGCTCACAGTT	ATTTGACTTGTCTTTTTG	TTTGACTTGTGTTTTTG
OtsGSI	36	Ots_il13Ra2B-37	16	T/G	AGGACTGGCTGCACATTCA	GAGGAGCTGTTCACACATATGTTG	CCAGGGAATCTATCCCAG	CCAGGGAATCTCTCCCAG
OtsGSI	37	Ots_il-1racp-166	2	G/T	GCCAAGAAAGTGTAGCTCCAACATA	AAGCAGAAACCCAGTAAGAAGGAAA	CCACATTCGTTTTTC	ACCACATTAGTTTTTC
OtsGSI	38	Ots_CCR7	9	C/T	CTGCTCACCTGCATCAGTGT	CCATGGTGGTCTGGACGAT	CCACGTAGCGATCG	ACCACATAGCGATCG
OtsGSI	39	Ots_nelfd-163	15	A/G	CTCACTGCAAATCCAACTTCATCAT	CCACTACATCCTCATCCAAGGTT	ACCCACCAGTGTCATT	CCACCAGCGTCATT
OtsGSI	40	Ots_CRB-211	5	A/C	CAACGCGGGAATGGCTTTTAA	GCCAGAGTCGCCAAAATAGTAGAAT	CTACCGTACTGAACTC	CCGTACGGAACTC
OtsGSI	41	Ots_MTA-SNP1	5	C/T	GCCGAAAAATAAGCGATTAGTGATGA	GCCCCATGGTAAACCTAATTAACCT	AATTGCCTCATTGGGTG	AATTGCCTCATTAGGTG
OtsGSI	42	Ots_P450-288	14	A/G	ATGTCAATATATTTCACTATAATGATTGGAAGCCA	CACTGAACTCGAAGCTGTTAGGA	CTATAAAGTTGGACAGTTGG	AAAGTTGGGCAGTTGG
OtsGSI	43	Ots_stk6-516	15	C/A	TGTGTTTAGGATTGAACTGACCATGTT	GTAAACTCCACCTGCAAGAAGGA	AACATAACGGACTCCC	TAGAACATAACTGACTCCC
OtsGSI	44	Ots_TCTA-58	16	C/T	ACCAGTACCTAAACGTTAGAAAGC	CGTTAGTTAGCTATGTCTGAAAGGC	CTGCCATGAAGTGCTAG	TGCCATGAAATGCTAG

					AA	A		
OtsGSI	45	Ots_GST-207	3	G/A	GGAGAACATGCA TCACCATTCAAG	TCAGCAAACGAA GGCTATGTAGAA T	ATGAGAGAGTCT TTCTCTGTT	ATGAGAGAGTCT TTTTCTGTT
OtsGSI	46	Ots_hsc71-5'-453	6	C/T	TTGAGAACATGT GGTAATTAACTAC AATGACTAA	GTACGAAGTTGC GCCTTGTC	CTGAGGTGGCAA AAT	TGAGGTGACAAA AT
OtsGSI	47	Ots_u1007-124	15	A/G	CGAAATAAGGGC CTGGTGTTTAAAA	TGTACCAGGTGG AAGCTTTGG	TGTCCTGTCCTCA GATCA	TCCTGTCCCCAGA TCA
OtsGSI	48	Ots_U2362-227	16	A/T	TCGTGGATTGTGG CTTACGT	GGGTGTTTAACA AGTAGTCCCTTCA	CTTAAGAAGCATT TTTTTG	AAGAAGCATTTA TTTTTG
OtsGSI	49	Ots_U2362-330	16	A/G	AATGGGTAACAA AGAAATAGCTAG CTACTT	GACAGACCACAG TGAAGGTGAAA	ACTGGGAAGATT GTTTG	CTGGGAAGACTG TTTG
OtsGSI	50	Ots_U2446-123	16	C/A	CTGGTCTGTGACG TCAAAATGATG	AGCTAGACCAGG CCATTTGAG	CTGCAACTCGAC GCAAG	ACTGCAACTCTAC GCAAG
OtsGSI	51	Ots_unk1104-38	16	C/T	TAACCATGACTTC TATCAATCACCCC	CCTCCATACATCG TCAAAGCTGTA	CCACTAAGGATT ACGTTACG	CACTAAGGATTA CATTACG
OtsGSI	52	Ots_unk1832-39	16	C/T	GAAACGTCTATG CTGTCCCCTTTAA	CTGCAGTATTAGC TCTAGTTGAATCC A	CACCACTAGAAC TCTC	CACCACTAAAAC TCTC
OtsGSI	53	Ots_unk3513-49	16	C/T	TTTGAGTGAGTCA CTGCACCAA	CAGCTCCACAGT GTCACCAT	AGTGCGAAGAAC C	AGTGCAAAGAAC C
OtsGSI	54	Ots_unk7936-50	16	C/G	ATGGGTTGGGATT ATGGTTCATTGT	CAAAATGGTTACT TGCATAGTCTTTT GT	AGACATGTAGCT ATGTAGGTAA	AGACATGTAGCT ATCTAGGTAA
OtsGSI	55	Ots_unk8200-45	16	A/G	TCAGGAGTGAAG CTGGTCTCT	TTCCATAGTAACT GACCTCAGTGTCT	CAGTTTAAAGTGT ATTCTCC	TTTAAAGTGCATT CTCC
OtsGSI	56	Ots_unk9480-51	16	G/C	CAAATCAGAACA AAACCTCCCACA A	GGAAGTCTGTCTG AATGGTTGTCTT	CTCCCACAAACCC	TCCCAGAAACCC
OtsGSI	57	Ots_hsp27b-150	6	G/A	TAGGAGTTGGAA AGACTGCACA	CCCATTGGTTCTT TGGTGTT	YGATCTGGACCA GGCT	YGATTTGGACCA GGCT
OtsGSI	58	Ots_zn593-346	15	A/T	CTACGCGAGAAA TAACACTTTTCAA AACT	GGCGAGTTTATTA CGGTGTTATGAC	TCTTGCAATCATT TTTAAC	CTTGCAATCATAT TTAAC

OtsGSI	59	Ots_113457-40	5	C/T	CCCAAGTGGTGAGTGTCAGT	ACTACAACAGGTGTTGATAATAGAATCATTCTC	ATATGGATTGGAGAATAG	CATATGGATTAGAGAATAG
OtsGSI	60	Ots_123048-521	5	A/C	CTCAACAGTGCACCTCCCTTAATT	CCAAACACACCCTTCCATAATCTCT	TCACATCCAACTCAGTACT	CATCCAACGCAGTACT
OtsGSI	61	Ots_96222-525	5	C/T	GCTCTTGCCCATCTGTAGGAT	GGCGCAACATATGTATTAAGCAACT	TGTAGCTAATTTTAAGTTCTC	AGCTAATTTTAAATTCTC
OtsGSI	62	Ots_97077-179	5	G/T	CCTGAACAAATACTTAACGCTCCAGTT	GTAATAATACTTCACACCATTGCCACTTC	TCACAAATGTATCCTAAAGC	CACAAATGTATACTAAAGC
OtsGSI	63	Ots_AldB1-122	5	C/T	GCCATGGAGGACTGGATGA	GCCACCACTACTTGCTGAGAAAATA	ACCCACTTCGCCAACA	ACCCACTTCACCAACA
OtsGSI	64	Ots_aldb-177M	4	T/A	GCGATCAGGTGACGCTAAAATGA	AGGAAGGTGATGCCTGAGAGA	CCAAATTGCTTAACCC	CCAAATTGCTTTATCC
OtsGSI	65	Ots_aspat-196	4	G/C	CCTGAACAGGTACACACAAACGA	TCCAACTGATGAATATGACCAACATGAAT	CACACCCACTCTTTAT	CACACCCAGTCTTTAT
OtsGSI	66	Ots_C3N3	1	T/G	CCGGATTCCATGGCCTACAC	GCCAAAATGATGTTCGGATGTAAAGT	CTAGAAAGGTTGATCCAATAA	AAAGGTTGAGCCAATAA
OtsGSI	67	Ots_CD63	9	A/C	TGCATGTTTTCTAACTGTGTTTTTGTGT	TGAATGCCCCCCATCAACA	AGATCATGGGAATCATAT	ATCATGGGCATCATAT
OtsGSI	68	Ots_EndoRB1-486	5	G/A	CCTTTGGGTCTGCTTGAGGTT	GGAGCCAAATCCTAATGCTGAAGTA	TCCTTCTCACGCTTCT	CTCCTTCTCATGCTTCT
OtsGSI	69	Ots_EP-529	10	A/G	GCCCTGCCTGCAACTTC	GAAACCAACGTCTTGATGTAGACCTA	CAGTGTCATTTTCGGC	ATCAGTGTCATCTTCGGC
OtsGSI	70	Ots_FARSLA-220	3	G/A	GTTCGTGGGATTGTTCAATGTTCAT	CTTGGACAGGCTCACATTACCATA	CCTTGGATGGGATGTG	CCTTGGATAGGATGTG
OtsGSI	71	Ots_FGF6A	7	G/T	TCAAAAATGTCTATCCAACAAATACTCTGAAAAATATTG	CTTGTGCGCACCTTGCA	CACGATTAGCAATGAACAA	CACGATTAGCAATTAACAA
OtsGSI	72	Ots_GH2	1	A/T	GCGTACTGAGCCT	CCCCCAGGTTCTG	TGACTCTCAGCAT	TGACTCTCTGCAT

OtsGSI	73	Ots_GnRH-271	2	C/T	GGATGACA CAGATGAAAAAT AAATAATTGGGC CATTAGGAA	GTAGTAGTTC CAGAGAGACTGA GACCATATGATGT AGT	CT CAATGAATACAA TATCTAACCTAAT	CTG AATGAATACAAT ATCTAATCTAAT
OtsGSI	74	Ots_GPDH-338	2	G/A	CACTAAATATTCC TTATCATTTCATA CTAAGTCTGAAG AA	AGCTGATACACA ATCAAAACACAA AACAT	CCACTACTTAACG TGCTTT	CCACTACTTAACA TGCTTT
OtsGSI	75	Ots_GST-375	3	C/T	CAGCCCGTCCCA AAATCAAG	CAGGAATATCAC TGTTTGCCATTGC	TTTCTTGTAGGCG TCAGAG	TCTTGTAGGCATC AGAG
OtsGSI	76	Ots_IL11	9	T/C	CCTCCAGATGAG ACCCACTCT	CAAAATGGTGCT CAAACGACTTCA	AGTCCGCATGGA GCT	TCCGCGTGGAGCT
OtsGSI	77	Ots_LWSop-638	2	T/C	CAATTACTCTTTC TCAGCCCTGTGT	GCGGTAAGATGC AGTTTTACATGGA	TTTAACAAGAAA ATTATACATTTC	CAAGAAAGTTAT ACATTTC
OtsGSI	78	Ots_Myc-366	5	T/C	CCTTAGCTGCTCT TTGAAGTTGACT	GGCTATAGAGTG TATTTACAGCATG CA	TCTCTGCTCATCT GTC	CTCTGCTCGTCTG TC
OtsGSI	79	Ots_myo1a-384	4	A/C	CTCCCCCCTGGAC TTTGG	GCTCTATTGCACC GTGTTCTG	ACAGATCCATCC ACCACT	AGATCCAGCCAC CACT
OtsGSI	80	Ots_myoD-364	4	T/G	GTGTGTGTGTGTG TGTGTCATC	TTTACACATATAC AAAAATGGTCCT CTATTGTCAT	TCATCTTTTGTTA TTTCCTTG	ATCTTTTGTTCTT TCCTTG
OtsGSI	81	Ots_Ots311-101x	4	A/-	AAATGAGGCCGT CCTTTACACT	GCAATACAAGCC CTTGATAATGAA GT	CTGAGATCACTTT GAGCAC	ACTGAGATCACT GAGCAC
OtsGSI	82	Ots_P450	1	T/A	TGAGCGAGATTT ATCAAACTGTCA AAGA	CCCAAGCGGGAG AACTTACAG	CCCCGAAGTACTT TT	CCCGAAGAACTTT T
OtsGSI	83	Ots_PGK-54	7	T/A	CTCATACTTTGTA CCTGTGTGTTCCA	CGACCCAAGTGG CTCATCAG	CCACCATCAAGC ACTG	CCACCATCATGCA CTG
OtsGSI	84	Ots_RFC2-558	2	A/-	GTAAGGTCTACTC CGGTTGTATTCG	CAATACGACAGT ACCGGTGTTAAA CT	TGCATGTAACAA ATAACAT	TGCATGTAACATA ACAT
OtsGSI	85	Ots_SL	1	A/G	AATATTGGCTTTC TGAGAATGCATTT GG	CCAAGATACTTCC TTTAACTTCTCTG TCA	TCAAAGATATGA TTCAATTAA	AAGATATGGTTC AATTAA

OtsGSI	86	Ots_Tnsf	1	A/G	GCCAATACGGGTTCTGAACTGT	CGGAATAGTCATAGTAGGGCTCGTT	TGCTCCAGATCTC	TGCTCCAGGTCTC
OtsGSI	87	Ots_u07-53.133	8	C/T	AGCTAGGCTGTAAATGCAAGGAT	CAGTGCTTTCAATTCATGCTGTCAA	TAACACATGTTGGAGGTC	AACACATGTTAGAGGTC
OtsGSI	88	Ots_u07-57.120	8	A/T	GGTTTGAGCCAATCAGTTGTGTT	CGGTCTAATGTCCATTGCTCATGTT	CAACCCCTACCTTGTCAC	CCCCTACCATGTCAC
OtsGSI	89	Ots_u202-161	2	T/A	CACTTTTGACTTTACATGGAACTTAACTCAT	GGGACTTCACTTTCTACAAACATGTCA	ATTAGCTGCTAAGCACTAG	ATTAGCTGCTATGCACTAG
OtsGSI	90	Ots_zP3b-215	2	G/T	TGCTGAGGACCATCTGCAATTC	AGGTCCATGAATAACTGAAAATGTACAAGT	CCAAATATCCTACCCGTGATG	CAAATATCCTACCAGTGATG
OtsGSI	91	Ots_ZR-575	10	G/A	GCCTACCAGAAAGTACCAATTGTGA	ACTTTTCACTGTCCTATTACAATTAGTATTTGTGATAT	CCGACACAATTTTGT	CCGACATAATTTTGT
OtsGSI	92	Ots_nramp-321	4	G/A	GGCCATCTTTCAGGACGTACAG	GCATGCTCTGCAATACGTTGAG	TCGTTCATGCCCGTTAG	TCATTCATGCCCGT
OtsGSI	93	Ots_RAS1	9	C/T	TCATAAACATGGTGTCTTTCAGTCAGTT	CTGACATGTGAAACTACTAAAGCATTTAATCAC	CAATCTATCATCGACCAGC	CAATCTATCATCAACCAGC
OtsGSI	94	Ots_TNF	9	C/T	CCAAATCCTCATCCCACACACT	CCGTTGCACTTGACCCTAAAC	CTGGCTGTAAACGAAGA	TGGCTGTAAACAAAGA
OtsGSI	95	Ots_u07-20.332	8	A/C	CGCGAGTTAGCTCGAATATTATGATTTC	TCAAGCTAGCATAGCAACTTCATCAA	ACCATTTGATATAACTGCGTTAG	CATTTGATATAACGGCGTTAG
OtsGSI	96	Ots_u07-64.221	8	G/C	GAGGATGACACTGTCCGTTTGT	CACAGTCCTTCGTATTCACCTTGAT	ATCGACCCTGTCATTAG	CGACCCTGTGATTAG

Note- Appendix references may be cited as follows:

1-Smith CT, Seeb JE, Schwenke P, Seeb LW (2005) Use of the 5'-nuclease reaction for SNP genotyping in Chinook salmon. Transactions of the American Fisheries Society 134:207-217.

2-Smith CT, Elfstrom CM, Seeb JE, Seeb LW (2005) Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Molecular Ecology 14:4193-4203.

3-Smith CT, Antonovich A, Templin WD, Narum SR, Elfstrom CM, Seeb LW (2007) Impacts of marker class bias relative to locus-specific variability on population inferences in Chinook salmon; a comparison of SNPs to STRs and allozymes. Transactions of the American Fisheries Society 136:1674–1687.

4-Campbell NR, Narum SR (2008) Identification of novel single-nucleotide polymorphisms in Chinook salmon and variation among life history types. Transactions of the American Fisheries Society 137:96-106.
5-Clemento AJ, Abadia-Cardoso A, Starks HA, Garza JC (2011) Discovery and characterization of single nucleaotide polymorphisms in Chinook salmon, Oncorhynchus tsawytscha. Molecular Ecology Resources 11(suppl. 1), 50-66.
6-Campbell NR, Narum SR (2009) Identification and characterization of heat shock response related single nucleotide polymorphisms in O. mykiss and O. tshawytscha. Molecular Ecology Resources 9, 1460-1559.
7-Unpublished Northwest Fisheries Science Center. Contact Anna Elz - Anna.Elz@noaa.gov
8-Unpublished Washinton Department of Fish and Wildlife. Contact Sewall Young - Sewall.Young@wdfw.wa.gov
9-Unpublished Washington State University - Vancouver. Contact Jennifer DeKoning - dekoning@vancouver.wsu.edu
10-Unpublished Oregon State University. Contact Renee Bellinger - renee.bellinger@oregonstate.edu
11-Unpublished Columbia River Inter-Tribal Fish Commission. Contact Nathan Campbell - camn@critfc.org
12-Miller KM, Beacham TD, Ming TJ (2008) Chinook Salmon SNP Development. Report to the Northern Endowment Fund.
13-Sex determination marker- based on GenBank sequence DQ393586.
14-Unpublished Alaska Department of Fish and Game. Contact Bill Templin - Bill.Templin@alaska.gov
15-Unpublished University of Washington & Washington Department of Fish and Wildlife. Contact Lisa Seeb - lseeb@uw.edu
16-Unpublished University of Washington. Contact Lisa Seeb - lseeb@uw.edu

Section 2: Genetic Baseline Expansion

Introduction

Reproductively distinct aggregations of Chinook salmon (*Oncorhynchus tshawytscha*), steelhead trout (*Oncorhynchus mykiss*), and sockeye or kokanee salmon (*Oncorhynchus nerka*) 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 better defined by physical distance, the patchy 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 (e.g., Beacham et al. 2006). While homing miscues (straying) 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). The challenge of detecting fine-scale genetic structure can be overcome by choosing an appropriate suite of genetic markers that can characterize broad regions of a species genome, representing both neutral and adaptive genetic variation.

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 to expand on existing genetic data baselines for the characterization of Chinook salmon, steelhead and *O. nerka* population structure throughout the Columbia River Basin (CRB). Here we describe our efforts to create and build upon existing baselines that were originally populated with microsatellite (μSAT) data and were founded by multi-agency consortiums called GAPS (Genetic Analysis of Pacific Salmonids; Seeb et al. 2007) for Pacific salmon species, and SPAN (Stephen Phelps Allele Nomenclature; Stephenson et al. 2009) specifically for steelhead trout. The use of SNPs for population analyses has become largely main stream with similar or overlapping consortium efforts underway. These include coast-wide development of SNP baselines (e.g., the expansion of GAPS under the Chinook Technical Committee LOA), establishment of standardized SNP scoring protocols among fisheries management entities, and specific directed applications including parentage based tagging (PBT). Although it has been shown that a larger number of SNP loci are necessary to reach the same level of resolution as the more polymorphic μSAT markers that have dominated past studies (e.g. ~200 SNPs needed for Chinook salmon GSI, Hess et al. 2011a), large numbers of highly informative SNP assays are indeed available, and their abundance stands to increase with the advent of new technologies such as Restriction site Associated DNA (RAD) sequencing (Miller et al. 2007; Baird et al. 2008). SNPs are highly prolific in the genome, with substantial coverage for linkage analyses (Moen et al. 2008). Moreover, because SNPs may be located within functional genes, they are candidates for detecting positive (adaptive) selection or selective divergence, a force instrumental in shaping population differences. SNPs are relatively easily amplified and scored, even with poor quality tissue source or DNA extract (Campbell and Narum 2008). With

advances in analysis platforms, they are currently amenable to superior high throughput capabilities.

Genetic analysis of *O. nerka* (hereafter intended to represent both sockeye salmon and kokanee) in the CRB has been limited in scope, and although Chinook salmon have been studied in great detail (Narum et al. 2004a; Waples et al. 2004; Beacham et al. 2006; Narun et al. 2008b; Matala et al. 2011), and steelhead to a similar degree (Narum et. al 2004b; Nielson et al. 2009; Blankenship 2011; Narum et al. 2011), our efforts are likely to provide additional information that will benefit and expound on the characterization, productivity and management of these species. The additional information that SNPs provide represents advanced potential for identifying fine scale or localized population differentiation that may prove valuable in monitoring and evaluation for conservation purposes (Ryynanen et al. 2007; Narum et al. 2008a). For example, the criteria for defining populations are somewhat ambiguously defined or applied among management agencies. They include distinct population segments (DPS), evolutionarily significant units (ESU) and major population groups (MPG). These definitions may rely heavily on geography, records of stock transfer, and basic biology or behaviors observed across locales or subbasins. Although genetic structure also influence the delineation of defined populations, such characterizations are likely to change coincident with improvements in genetic tools (particularly SNPs associated with adaptive selection). This additional information may ultimately influence criteria for identifying populations on the basis of both gene flow and adaptive potential across diverse landscapes in the CRB (Matala et al. 2011, Narum et al. 2011).

Methods

Sample Collection and baseline expansion

In 2011 we initiated an effort to create an *O. nerka* SNP baseline (see Section 1 for SNP panel information). A total of 14 discrete collections were genotyped in 2011, and the goal was to broadly include both the resident form (kokanee) and anadromous form (sockeye salmon). *O. nerka* collections were also gathered in conjunction with progress on external projects, including a study focused on population structure in the upper Deschutes River Basin. In 2011 we also expanded on existing SNP baselines for Chinook salmon and steelhead trout (Table 1). Representative collections were included from throughout the CRB (where available). For Chinook salmon our goal for coverage included all three lineages (lower Columbia, ocean-type, and stream-type). The addition of a second panel of 96 SNPs in 2011 brought the total number of Chinook salmon assays to 192 (Section 1, Appendix 1), and our primary effort was therefore directed at updating current baseline populations rather than adding new collections. In 2011 a total of 40 discrete Chinook salmon collections from the existing baseline were genotyped (Table 1). For steelhead trout we had a similar goal. Because the 2009 baseline (see 2009 annual progress report) was completed prior to finalizing our updated SNP panels of 192 assays (see 2010 annual progress report), we focused on updating the data for the 2009 baseline collections to include the PBT panel of 96 SNPs (Section 1, Appendix 1). An additional 19 new collections were genotyped and added to the baseline in 2011 (Table 1, Figure 1) to augment CRB regions with minimal or no representation. For the long-term, we intend to include coastal and inland lineages of O. mykiss, both summer-run, winter-run ecotypes where they occur, and to test the utility of the SNPs for differentiating anadromous from resident *O. mykiss* life histories (Narum et al. 2008a; Narum et al. 2011). Further, priority collections for all three species have been

identified as those relevant to basin-wide management and tribal fishery interests (particularly in the middle Columbia and Snake River regions). Priority collections include major supplementation stocks for all three species, and several Idaho *O. nerka* populations. These collections have been archived for future baseline expansion efforts.

We have three primary goals for the application of the compiled SNP genotypic data. Objective one is annual genetic stock identification (GSI) analyses that will be used for monitoring of fishery returns through the migratory corridor, including harvest GSI in the lower Columbia River, and fish passage GSI at Bonneville and Lower Granite dams (*see sections 3 & 4 of this report*). In objective two we will conduct PBT monitoring in cooperation with Idaho Department of Fish and Game (IDFG). The PBT effort entails genotyping all hatchery broodstocks on an annual basis in the Snake River Basin; this provides the ability to genetically assign all subsequently returning adult (hatchery) progeny to their respective hatcheries of origin. Lastly, our third objective is to maintain a ten year expanding baseline of SNP data for application in various analyses including population structure analyses, and investigations of landscape genetics and adaptation among populations for these and other *Oncorhynchid* species in the future.

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), or using a dry Whatman paper medium (LaHood et al. 2006?). Many samples were contributed by outside agencies including NOAA Fisheries, Washington Department of fish and Game (WDFW), Oregon Department of Fish and Wildlife (ODFW), United States Fish and Wildlife Service (USFWS) and IDFG.

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, each 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: 3.4µl GTXpress Taqman (Applied Biosystems), 0.30µl GT load buffer (inclusing taq polymerase), 0.30µl H_2O and 2.0µl pre-amp DNA template. Single SNP assays were prepared in a 5.0µl 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, and sample cocktail dispensed at 5.0µl per well. Chip loading and amplification was completed following standard manufacturers protocol on a Fluidigm IFC controller. Amplification conditions using a fast-cycling protocol were; 70^{o} C for 30 min, 25^{o} C for 10 minutes, and 95 ^{o} C for one minutes, followed by 50 cycles of 95^{o} C for 5 seconds, and 50^{o} C for 25 seconds, and a final cool down step of 25^{o} C for 10 minutes. 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 loci for Chinook salmon, and fewer than 18 SNP loci for steelhead trout).

Statistical Analysis

Allele frequencies were generated with the program GenAlEx version 6.41 (Peakall & Smouse 2006). Descriptive statistics including number of samples analyzed per collection per locus, and the unbiased heterozygosity (H_E), were generated using the analysis program GenAlEx version 6.41 (Peakall & Smouse 2006). 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). Tests for linkage disequilibrium (LD) for all pairs of loci across collections were conducted in GENEPOP using a computationally demanding simulated exact test. Linkage analysis identifies non-random association of loci, where population specific linkage indicates population admixture. For all pairs of loci with significant linkage we will select the most informative of the two, and drop the remaining locus in each pair from further analyses. However the results of the current analyses represent all available loci with the exclusion of hybrid and sex determining loci. Statistical significance (α) was adjusted for the number of simultaneous tests k (α/k for $\alpha = 0.05$) for HWE tests and LD using the modified BY-FDR method (Narum 2006).

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 50,000 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. Evaluating adaptive diversity (selection) is confounded by the need to scrutinize the underlying demographic processes that also influence diversity. It is necessary to eliminate non-neutral outlier loci for computing most population genetic parameters (e.g., F_{ST}, Nm, Ne, gene flow) in order to draw biological inferences from comparisons that reveal patterns of neutral genetic diversity (Luikart *et al*. 2003). Therefore as analyses develop and the baseline is applied for specific applications we will exclude significant candidate loci under selection from further analyses of neutral population structure. Analyses conducted for this report were conducted without excluding any candidate loci as identified by LOSITAN. Determining which candidate loci are under selection will require additional computationally demanding tests to show statistical associations (see Matala et al. 2011). 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 are lineage specific.

GENEPOP (Raymond and Rousset 1995) was used to calculate global F_{ST} (θ of Weir and Cockerham 1984) and a matrix of pairwise F_{ST} among all pairs of collections, which indicates the proportion of total variation attributed to differences among collections. Based on minimal intermediate genetic distances and similarity in clustering among steelhead collections from the

same location (e.g. major tributary or watershed), we combined collections by location to form 34 groups of steelhead collections for pairwise analysis (e.g., all collections from the Klickitat River were combined (Table 1). A pairwise matrix of Neis genetic distances (1972) 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) using genetic distance 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

Analyses were conducted using all 96 loci for *O. nerka*. After excluding potential hybrid individuals among steelhead collections, the hybrid detection markers and sex determining marker were omitted from analysis leaving 188 loci remaining. For Chinook salmon the sex marker was excluded before analyses were conducted.

Chinook salmon Descriptive Statistics

The mean expected heterozygosity (allelic variability) across loci and collections ranged from 0.2656-0.3307 for the lower Columbia lineage, 0.2741-0.3158 for the ocean-type lineage, and 0.1996-2945 for the stream-type lineage. For the lower Columbia lineage, SNP loci *Ots_GnRH-271 and Ots_SERPC1-209* were fixed. Among the ocean-type collections, *Ots_SERPC1-209* and *Ots_zP3b-1* were fixed, and for the stream-type lineage there were four fixed loci: *Ots_RAS1, Ots_SERPC1-209, Ots_u07-64.221, Ots_zP3b-1*. No loci were diagnostic for any particular lineage or life history type in the data set. Among 6,661 total HWE tests across all collections we observed 144 departures from expected genotypic proportions. We identified some population and locus specific deviations out of a total of 191 loci and 40 populations; 14 deviations in Little White Salmon spring-run, 16 in the Entiat spring-run, 23 in the Little White Salmon fall-run, and 31 in the Entiat summer-run. Deviations occurred in approximately 25% of populations for locus *Ots_EndoRB,* and approximately 15% of populations for *Ots_ZR-575* and *Ots_nramp-redesign*. A plot of expected heterozygosity and genetic distance (F_{ST}) generated in LOSITAN identified 30 SNP outlier loci, or candidate loci under directional selection across lineages (Table 2). Significant linkage disequilibrium was observed in four pair of loci: *Ots_OTALDBINT1-SNP1* and *Ots_aldb-177M, Ots_hsc71-3'-488* and *Ots_hsc71-5'-453, Ots_FGF6A* and *Ots_FGF6B_1,* and *Ots_OTSTF1-SNP1* and *Ots_Tnsf.*

Chinook salmon Genetic Structure Analysis

For the 40 collections of Chinook salmon evaluated, we observed among-group variation (F_{ST}) across loci that ranged from 0.0022-0.1268 for the lower Columbia lineage (mean 0.0557), 0.0014-0.0595 for the ocean-type (mean 0.0217), and 0.0019-0.1325 for the stream-type (mean

0.0494). The overall mean pairwise genetic distance was 0.2089 across lineages (Figure 2a). The least amount of among-group variation occurred within the ocean-type lineage, while the lower Columbia and stream-type lineages were comparabl. We demonstrated genetic similarity among populations within each lineage through phylogenetic relationships in the topology of an unrooted NJ phylogram (Figure 3a), and in PCA cluster analyses to graphically display the relationship between collections; the latter is a spatial ordination of data in the context of the two axes that explain the greatest amount of total variation or differences among collections (Figure 4a). The confidence or concordance (>50%) of the NJ topology is indicated with bootstrap values at the nodes (Figures 3a). Results of the two analyses were complimentary, and revealed defined clustering of the most genetically similar collections for Chinook salmon, but collections between the three major lineages were clearly divergent. In large part, the Chinook salmon collections clustered accurately into geographic regions or major tributary (Figure 3a & 4a). The North Santiam and Big White Salmon spring-run collections in the lower Columbia lineage were surprisingly distinct, as were the American, Warm Springs, and Klickitat collections in the stream-type lineage comparisons. Such distinctions were likely highly influential in overall results (e.g., unusually high pairwise F_{ST} in the lower lineage; Figure 2), and those results may be reflective of stray influences within and/or among lineages (see Hess et al. 2011b) and well as ecotype distinctions (e.g., summer-run vs. fall-run in the ocean-type lineage).

Steelhead Descriptive Statistics

The mean expected heterozygosity across loci and collections ranged from 0.2667-0.3627 for the coastal lineage, and 0.2284-0.3430 for the inland lineage. A single fixed locus was observed in the coastal lineage (*Omy_pad-196*), while no loci were fixed among inland lineage collections. No loci were diagnostic for either lineage or or life history type in the data set. Among 22,255 total HWE tests across all collections we observed 197 departures from expected genotypic proportions. We identified some population and locus specific deviations out of 188 loci and 129 populations; 9 deviations in the Abernathy collection, 10 in the Umatilla, and 10 in Canyon Creek in the Willamette. Deviations occurred in approximately 6% of populations for locus *Omy_crb-106* and *OMS00087*. A plot of expected heterozygosity and genetic distance (F_{ST}) generated in LOSITAN identified 23 SNP outlier loci, or candidate loci under directional selection across lineages (Table 2). Significant linkage disequilibrium was observed in one pair of loci: *Omy_GHSR-121* and *Omy_mapK3-103*

Steelhead Genetic Structure Analysis

The 34 grouped steelhead collections are identified in Figure 2b. We observed among-group variation (F_{ST}) across loci that ranged from 0.0064-0.0941 for the coastal lineage (mean 0.0349), and 0.0031-0.1092 for the inland lineage (mean 0.0286). The overall mean pairwise genetic distance was 0.0692 across lineages (Figure 2b). On average the pairwise genetic distances were lower in the coastal lineage although the average overall was higher than inland. We demonstrated genetic similarity among populations within each lineage through phylogenetic relationships in the topology of an unrooted NJ phylogram (Figure 3b), and in PCA cluster analyses to graphically display the relationship between collections (Figure 4b). The confidence or concordance (>50%) of the NJ topology is indicated with bootstrap values at the nodes (Figures 3b). Results of the two analyses were complimentary, and revealed defined clustering of the most genetically similar collections for steelhead, but populations between lineages were clearly divergent. In large part, the steelhead collections clustered accurately into geographic

regions or major tributary (Figure 3b & 4b). Similar to observations among Chinook salmon results, the Klickitat and Big White Salmon collections are among the most divergent overall, appearing intermediate between lineages.

O. nerka Descriptive Statistics

The mean expected heterozygosity across loci and collections ranged from 0.2513-0.3286. No loci were fixed 14 baseline collections. No loci appear diagnostic for either life history type in the data set (kokanee or sockeye salmon). Among 1,234 total HWE tests across all collections we observed 21 departures from expected genotypic proportions. We identified one locus specific deviations (One_UCA-24) which occurred in approximately 25% of collections, but there were no observed population specific deviations. A plot of expected heterozygosity and genetic distance (F_{ST}) generated in LOSITAN identified 2 SNP outlier loci, or candidate loci under directional selection across lineages (Table 2), however one is a mitochondrial locus. Significant linkage disequilibrium was observed between *One_MHC2-190* and *One_MHC2-251*, and *One_ODC1-196* and *One_U508-533*.

O. nerka Genetic Structure Analysis

We observed among-group variation (F_{ST}) across loci that ranged from -0.0018-0.2722 (mean 0.1047; Figure 2c). We demonstrated genetic similarity among populations within each region or major subbasin through phylogenetic relationships in the topology of an unrooted NJ phylogram (Figure 3c), and in PCA cluster analyses to graphically display the relationship between collections (Figure 4c). The confidence or concordance (>50%) of the NJ topology is indicated with bootstrap values at the nodes (Figures 3c). Results of the two analyses were complimentary, and revealed defined clustering of the most genetically similar collections for *O. nerka*. Collections within each region, particularly temporally stratified collections exhibited substantial similarity, while populations between regions (e.g., Wenatchee and Okanogan) were highly divergent (Figure 2c-4c).

Discussion

We have compiled extensive data sets of SNP genotypes for Chinook salmon, steelhead trout, and *O. nerka* covering diverse regions in the Columbia River Basin (including the Snake River Basin) in this expansion and update effort that complements our first two years of results. Our goal was to construct SNP baselines of genotypes that will be expanded annually to provide continued evaluation of these species that is both spatially and temporally stratified to account for inter-annual variation. 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 2008b; Narum et al. 2010b) and steelhead trout

(Blankenship et al. 2011). Results further suggest SNPs are a class of markers that perform at least as well as µSATs in terms of their potential for monitoring population distinctions and composition during fish migrations and fisheries harvests. In addition, we have demonstrated that SNPs offer an opportunity to characterize adaptive variation, which is often beyond the scope of most µSAT datasets that utilize neutral markers.

The expansion efforts reported here complement previously reported results. The continued expansion of SNP panels and updating of baseline collections will help us achieve a greater level of resolution (or statistical power to identify population distinctions), at least among the major tributaries and subbasins throughout the CRB. Such results will be most beneficial to the larger application of the baselines, namely genetic stock ID. The degree to which our resolution in the Chinook salmon baseline might increase is as yet unclear given that the majority of our efforts this report year were directed at updating the existing baseline populations with additional SNP loci; we were only able to achieve a portion of the overall effort necessary to complete this task. However, initial indications suggest an improved GSI quality with the new data set (see sections 3 and 4 of this report). Our steelhead baseline in 2011 is robust and representative of the majority of watersheds within the Columbia/Snake River basins. With the SNP panel updating effort that was completed this year (including contributions of a significant amount of data from IDFG) we now have over 129 separate collections, including spatially and temporally stratified coverage in many watersheds. On the basis of our SNP genotypes, steelhead appear to be a highly diverse species in the basin, with clear distinctions that corroborate previous studies and are consistent with biologically significant distinctions among life history types, lineages and ecotypes. We anticipate a peer reviewed manuscript will be submitted for publication in a scientific journal for calendar year 2012 that will examine demographic and adaptive divergence of steelhead to expand on currently published studies (e.g., Blankenship et al. 2011). In comparisons among species specific SNP panels, the panel for *O. nerka* is extremely powerful owing to the biological nature of the species and its distribution within the CRB. Our ability to differentiate among regions and in some cases within regions (the Deschutes River populations) exceeds what we have been able to ascertain with Chinook salmon and steelhead, as evidenced by comparisons of among-group variation. For example, in a related study, we have shown that a mixed sample of sockeye salmon from Okanogan and Wenatchee River systems can be differentiated with 100% accuracy (data not shown).

We will begin to more closely examine those populations that display unique attributes or differences in contrast to, or in accordance with expectations based on published information (e.g., Big White Salmon River, Little White Salmon River, and Klickitat River). 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 (e.g., Matala et al. 2011). We will continue 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, and compare these results to some of our initial or preliminary findings as these baselines develop. Our data will be implemented in current and ongoing application of PBT and GSI methods for each species. These additional and ongoing efforts will require further scrutiny of the genotypic data, as we have not yet identified significant selection candidate loci for the expanded data set. Our future efforts include adding collections to the baseline to increase basin-wide coverage of all species,

particularly those that account for stock transfer history of *O. nerka* throughout the basin, and to provide a temporally stratified view of populations (accounting for natural inter-annual population variation). We also intend to continually strive to increase the numbers of markers (SNP loci) employed for genetic applications, as evident by our updating effort in this report and with new RAD sequencing capabilities. Some HWE and linkage issues that point to problems with population admixture (e.g., straying) may persist and will require certain populations and/or loci be dropped from the baseline in the future.

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 F_{ST} -outlier method. BMC Bioinformatics 9:323.

Baird, N. A., P. d. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko and E. A. Johnson. 2008. PLos ONE 3(10): 7 pages.

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.

Blankenship, S. M., M. R. Campbell, J. E. Hess, M. A. Hess, T. W. Kassler, C. C. Kozfkay, A. P. Matala, S. R. Narum, M. M. Paquin, M. P. Small, J. J. Stephenson, K. I. Warheit and P. Moran. Major lineages and metapopulations in Columbia River Oncorhynchus mykiss are structured by dynamic landscape features and environments. Transactions of the American Fisheries Society 140:665-684.

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.

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.

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.

Hess, J. E., A. P. Matala, and S. R. Narum. 2011a. Comparison of SNP and microsatellite markers for application of genetic stock identification for Chinook salmon in the Columbia River Basin. Molecular Ecology Resources 11 (Suppl. 1):1–13.

Hess, J. E., A. P. Matala, J. S. Zendt, C. R. Frederiksen, B. Sharp and S. R. Narum. 2011b. Introgressive hybridization among major Columbia River Chinook salmon (Oncorhynchus tshawytscha) lineages within the Klickitat River due to hatchery practices. Canadian Journal of Fisheries and Aquatic Science 68:1876–1891.

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 137:1104-1107.

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.

Matala, A. P., J. Hess and S. R. Narum. 2011. Resolving adaptive and demographic divergence among Chinook salmon populations in the Columbia River Basin. Transactions of the American Fisheries Society. Transactions of the American Fisheries Society 140:783-807.

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.

Miller, M. R., J. P. Dunham, A. Amores, W. A. Cresko and E. A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17:240-248.

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.

Narum, S. R., M. S. Powell and A. J. Talbot. 2004a. 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. 2004b. 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. S. Zendt, D. Graves and W. R. Sharp. 2008a. Influence of landscape on resident and anadromous life history types of Oncorhynchus mykiss. Canadian Journal of Fisheries and Aquatic Sciences 65:1013-1023.

Narum, S. R., T. L. Schultz, D. M. Van Doornik and D. Teel. 2008b. Localized genetic structure persists in wild populations of Chinook salmon in the John Day River despite gene flow from outside sources. Transactions of the American Fisheries Society 137:1650-1656.

Narum, S. R., J. Hess and A. P. Matala. 2010. Examining Genetic Lineages of Chinook Salmon in the Columbia River Basin. Transactions of the American Fisheries Society 139:1465–1477.

Narum, S. R. J. Zendt, C. Frederiksen, N. Campbell, A. P. Matala and Bill Sharp. 2011. Candidate genetic markers associated with anadromy in *Oncorhynchus mykiss* of the Klickitat River. Transactions of the American Fisheries Society 140:843-854.

Nei, M. 1972. Genetic distance between populations. The American Naturalist 106:283-292.

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., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6, 288-295.Program note available from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1471-8286.2005.01155.x

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.

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

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

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.

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.

Table 1.) Chinook salmon, steelhead trout, and sockeye/kokanee salmon collections included in the 2011 SNP baseline expansion. Refer to the map identifier for location. The entry '---' indicates no data or unknown status. Chinook salmon lineage is LC – Lower Columbia, OT – ocean type, and ST – stream type. Steelhead lineage is I – interior, C – coastal. *O. nerka* lineage is K – kokanee, and SO – sockeye. Ecotypes is AN – anadromous, and R – resident. Origins are identified as; wild or natural-origin (N) and hatchery-origin (H). Run time is F – fall, W – winter, SP – spring and SU – summer. "Win" is winter-run and "Sum" is summer-run. Age is A – adult, J – juvenile (age unknown), and S – smolt. Collections that appeared in previous reports and were updated with revised SNP panels in 2011 do not appear on the map (Figure 1). The identifier column correlates to PCA analysis and NJ tree ID (Figures 3 & 4) for Chinook salmon, and to map locations for all new collections in 2011. The majority of the steelhead baseline collections were described in previous report years and do not appear is this table list. Some data was contributed form external projects (®).

ID	Collection	BPA subbasin	(n)	Lat	Long	Lineage	Ecotype	Run	Origin	Year	Stage
O. tshawytscha											
1	Cowlitz R.	Cowlitz	89	46.513	-122.635	LC	N/A	SP	H	2004	A
2	Kalama R.	Kalama	85	46.017	-122.733	LC	N/A	SP	H	2004	A
3	McKenzie R.	Willamette	75	44.117	-123.086	LC	N/A	SP	H	2004	A
4	Sandy R.	Sandy	46	45.563	-122.395	LC	N/A	SP	N	2006	A
5	North Santiam R.	Willamette	75	44.697	-122.983	LC	N/A	SP	H	2004	A
6	Elochoman R.	Elochoman	81	46.261	-123.298	LC	N/A	F	N	1995-97	A
7	Lewis R.	Lewis	93	45.953	-122.584	LC	N/A	F	N	2003	A
8	Sandy R.	Sandy	82	45.563	-122.395	LC	N/A	F	N	2002	A
9	North Fork Lewis R.	Lewis	81	45.867	-122.724	LC	N/A	F	N	2004	A
10	White Salmon R.	Big White Salmon	77	45.744	-121.525	LC	N/A	SP/F	N	2008	J
11	Big Cr.	Salmon	89	45.138	-115.038	ST	N/A	SP	N	2001	A
12	Chamberlain Cr.	Salmon	45	45.454	-114.933	ST	N/A	SP	N	2009	J
13	Shitike Cr.	Deschutes	93	44.764	-121.238	ST	N/A	SP	N	2004	J
14	Warm Springs R.	Deschutes	88	44.861	-121.244	ST	N/A	SP	H	2004	A
15	Entiat R.	Entiat	92	47.696	-120.321	ST	N/A	SP	N	2006	J
16	Imnaha R.	Imnaha	91	45.561	-116.834	ST	N/A	SP	N	1998	A

17	John Day R.	John Day	76	44.76	-119.65	ST	N/A	SP	N	2000	J/A
18	Little White Salmon R.	Little White Salmon	93	45.722	-121.641	ST	N/A	SP	H	2007	J
19	Rapid R.	Salmon	93	45.353	-116.394	ST	N/A	SP	H	1999	A
20	East Fork Salmon R.	Salmon	94	44.259	-114.317	ST	N/A	SP	N	---	A
21	Tucannon R.	Tucannon	82	46.526	-118.142	ST	N/A	SP	N	2003	A
22	Wenaha R.	Grande Ronde	48	45.946	-117.455	ST	N/A	SP	N	2006	J
23	Leavenworth-NFH	Wenatchee	88	47.559	-120.672	ST	N/A	SP	H	2005	A
24	Peshastin Cr.	Wenatchee	86	47.558	-120.575	ST	N/A	SP	N	2005	J
25	W. F. Yankee Fork	Salmon	75	44.349	-114.727	ST	N/A	SP	N	2005	J
26	American R.	Yakima	66	46.976	-121.158	ST	N/A	SP	N	2003	A
27	Cle Elum R.	Yakima	86	47.178	-120.999	ST	N/A	SP	H	1997	---
28	Johnson Cr.	Salmon	92	44.899	-115.492	ST	N/A	SP/SU	N	2002	A
29	Klickitat R.	Klickitat	84	45.71	-121.27	ST	N/A	SP	H	2002	A
30	Newsome Cr.	Clearwater	82	45.831	-115.608	ST	N/A	SP	N	2001	A
31	Clearwater-NPTH	Clearwater	85	46.519	-116.665	OT	N/A	F	H	2003	A
32	Upper Deschutes R.	Deschutes	90	44.878	-121.048	OT	N/A	SU	N	1998	J
33	Lower Deschutes R.	Deschutes	90	45.28	-121.02	OT	N/A	F	N	1999	A
34	Entiat R.	Entiat	51	47.696	-120.321	OT	N/A	SU	N	2008	A
35	Hanford Reach	Columbia Lower Mid.	90	46.713	-119.481	OT	N/A	F	N	2000	A
36	Little White Salmon R.	Little White Salmon	99	45.722	-121.641	OT	N/A	F	H	2007	J
37	Lyons Ferry	Tucannon	90	46.589	-118.22	OT	N/A	F	H	2000	A
38	Tumwater & Dryden	Wenatchee	93	47.542	-120.559	OT	N/A	SU	N	1993	A
39	White Salmon R.	Big White Salmon	90	45.744	-121.525	OT	N/A	F	N	2008	J
40	Lower Yakima R.	Yakima	60	46.312	-119.473	OT	N/A	F	N	1998	A

O. mykiss

2009	Canyon Cr.	Willamette	25	44.91	-123.418	C	AN	---	N	1997	J
2009	Coweeman R.	Coweeman	47	46.174	-122.758	C	AN	W	N	2006	A
2009	Kalama R.	Kalama	94	46.033	-122.87	C	AN	W	N	2005	A
2009	Luckiamute R.	Willamette	28	44.789	-123.563	C	---	---	N	1997	J
2009	Willamina Cr.	Willamette	32	45.123	-123.489	C	---	---	N	1997	J

2009	Big White Salmon R.	Big White Salmon	81	45.744	-121.525	C	---	---	N	1990	J
2009	Lower Little Klickitat R.	Klickitat	46	45.843	-121.06	I	AN	SU	N	2005	J
2009	Lower Summit Cr.	Klickitat	46	45.988	-121.125	I	AN	SU	N	2005	J
2009	Lower Trout Cr.	Klickitat	48	46.038	-121.199	I	AN	SU	N	2005	J
2009	Upper Trout Cr.	Klickitat	46	46.077	-121.212	I	R	SU	N	2005	J
2009	Lower White Cr.	Klickitat	35	46.013	-121.15	I	AN	SU	N	2005	J
2009	Bowman Cr.	Klickitat	48	45.845	-121.042	I	AN	SU	N	2005	J
2009	Dead Canyon Cr.	Klickitat	35	45.942	-121.144	I	AN	SU	N	2005	J
2009	Snyder Cr.	Klickitat	47	45.826	-121.166	I	AN	SU	N	2005	J
2009	Surveyors Cr.	Klickitat	39	46.196	-121.255	I	R	SU	N	2005	J
2009	Swale Cr.	Klickitat	48	45.809	-121.065	I	AN	SU	N	2005	J
2009	Fifteenmile Cr.	Fifteenmile	91	45.451	-121.124	I	---	---	N	2005	J
2009	Buckhollow Cr.	Deschutes	63	45.263	-121.024	I	---	SU	N	2005-06	J
2009	Trout Cr.	Deschutes	57	44.821	-121.087	I	---	SU	N	2007	J
2009	Upper Mainstem	Deschutes	61	44.724	-121.248	I	---	SU	N	2005-06	J
2009	Camp Cr.	John Day	22	44.689	-118.797	I	---	SU	N	2006	J
2009	Clear Cr.	John Day	39	44.589	-118.507	I	---	SU	N	2005	J
2009	Camas Cr.	John Day	19	45.021	-118.991	I	---	SU	N	2006	J
2009	Desolation Cr.	John Day	19	44.994	-118.928	I	---	SU	N	2007	J
2009	S. F. Umatilla R.	Umatilla	34	45.723	-118.187	I	---	SU	N	2005	J
2009	Nason Cr.	Wenatchee	21	47.802	-120.715	I	---	SU	N	2006	J
2009	Omak Cr.	Okanogan	94	48.396	-119.504	I	AN	SU	---	2005	A
A	®Rock Cr. Bickleton	Columbia Lower Mid.	51	45.707	-120.464	I	---	SU	N	2008	J
A	®Rock Cr.	Columbia Lower Mid.	43	45.748	-120.436	I	---	SU	N	2010	J
B	®Rock Cr. Hwy. 8 Bridge	Columbia Lower Mid.	33	45.748	-120.436	I	---	SU	N	2008	J
C	®Squaw Cr.	Columbia Lower Mid.	70	45.796	-120.464	I	---	SU	N	2008	J
C	®Squaw Cr.	Columbia Lower Mid.	69	45.796	-120.464	I	---	SU	N	2010	J
D	®Iskuukpa Cr.	Umatilla	440	45.699	-118.396	I	AN	SU	N	2010	J/A
E	Nile Cr.	Yakima	59	46.837	-120.956	I	---	SU	N	2005;08	---
F	Pile Up Cr.	Yakima	26	47.049	-121.186	I	---	SU	N	2005;08	---
G	Quartz Cr.	Yakima	26	47.019	-121.131	I	---	SU	N	2008	---

H	N. F. Little Naches R.	Yakima	21	47.043	-121.189	I	---	SU	N	2008	---
I	Toppenish Cr.	Yakima	46	46.324	-120.169	I	---	SU	N	2009	---
J	Rattlesnake Cr.	Yakima	36	46.818	-120.939	I	---	SU	N	2005;08	---
K	Satus Cr.	Yakima	46	46.262	-120.112	I	---	SU	N	2009	---
L	Icicle Cr.	Wenatchee	24	47.559	-120.674	I	AN	SU	N	2011	---
L	Levenworth NFH	Wenatchee	19	47.559	-120.674	I	AN	SU	H	1999	---
M	Chiwaukum R.	Wenatchee	54	47.688	-120.741	I	AN	SU	N	2007	---
M	Upper Chiwaukum R.	Wenatchee	29	47.688	-120.741	I	AN	SU	N	2007	---
N	Bonaparte Cr.	Okanogan	100	48.705	-119.446	I	---	SU	N	2010	---
O	Salmon Cr.	Okanogan	100	48.379	-119.591	I	---	SU	N	2010	---
O. nerka											
P	Suttle Lake	Deschutes	99	44.427	-121.727	K	R	N/A	N	2009	A
P	Suttle Lake	Deschutes	92	44.427	-121.727	K	R	N/A	N	2010	A
P	Suttle Lake	Deschutes	100	44.427	-121.727	K	R	N/A	N	2011	A
Q	Pelton Trap	Deschutes	24	44.726	-121.247	S	AN	N/A	N	2009-10	A
R	Lake Billy Chinook	Deschutes	98	44.603	-121.28	K	R	N/A	N	2010	S
R	Lake Billy Chinook	Deschutes	100	44.603	-121.28	K	R	N/A	N	2011	S
S	Metolius R.	Deschutes	94	44.619	-121.47	K	R	N/A	N	2009	A
S	Metolius R.	Deschutes	98	44.619	-121.47	K	R	N/A	N	2010	A
T	Meadow Cr.	Okanogan	46	50.1784	-116.938	K	R	N/A	H	2005	A
U	Tumwater Dam	Wenatchee	94	47.617	-120.723	S	AN	N/A	---	2004	A
U	Tumwater Dam	Wenatchee	138	47.617	-120.723	S	AN	N/A	---	2009	A
V	Wells Dam	Columbia Upper Mid.	94	47.945	-119.866	S	AN	N/A	---	2004	A
V	Wells Dam	Columbia Upper Mid.	138	47.945	-119.866	S	AN	N/A	---	2009	A
W	Lake Whatcom	N/A	46	48.673	-122.278	---	R	N/A	H	2005	A

Table 2. Putative outlier loci for directional (adaptive) selection; results of analysis using LOSITAN. These loci have been flagged but their correlation to adaptive selection has not been confirmed by additional tests (see Matala et al. 2011; Narum et al. 2011).

steelhead trout			Chinook salmon				*O. nerka*	
Loci	inland	coastal	Loci	LC	OT	ST	Loci	
OMS00013	✓		Ots_103122-180	✓			One_Cytb_17	✓
Omy_107031-704	✓		Ots_131460-584	✓			One_U1216-230	✓
Omy_cd28-130	✓		Ots_unk1832-39	✓			---	
Omy_hus1-52	✓		Ots_MHC2	✓			---	
Omy_IL1b-163	✓		Ots_u07-07.161	✓			---	
Omy_ndk-152	✓		Ots_Hsp90a	✓	✓		---	
Omy_sSOD-1	✓		Ots_Est1362	✓	✓	✓	---	
Omy_star-206	✓		Ots_C3N3	✓		✓	---	
Omy_u07-79-166	✓		Ots_FARSLA-220		✓	✓	---	
Omy_u09-52.284	✓		Ots_U2446-123		✓		---	
Omy_u09-56.119	✓		Ots_unk1104-38		✓		---	
Omy_vatf-406	✓		Ots_TAPBP		✓		---	
OMS00118	✓	✓	Ots_ppie-245			✓	---	
OMS00064		✓	Ots_117259-271			✓	---	
OMS00096		✓	Ots_128302-57			✓	---	
Omy_102505-102		✓	Ots_nelfd-163			✓	---	
Omy_105714-265		✓	Ots_113457-40R			✓	---	
Omy_97954-618		✓	Ots_Ikaros-250			✓	---	
Omy_bcAKala-380rd		✓	Ots_MHC1			✓	---	
Omy_cd59-206		✓	Ots_nramp-321			✓	---	
Omy_Ogo4-212		✓	Ots_P450			✓	---	
Omy_sast-264		✓	Ots_PGK-54			✓	---	
---			Ots_SL			✓	---	
---			Ots_u07-57.120			✓	---	
---			Ots_u202-161			✓	---	
Total	13	10		8	6	16		2

Figure 1.) Map of study area and baseline expansion collections for 2011. The novel steelhead and *O. nerka* expansion collections genotyped in 2011 are shown with map ID (see Table 1). Those steelhead and Chinook salmon populations that were updated with revised SNP panels (Table 1; Appendix 1) in 2011 appear in earlier reports (see 2009 and 2010 annual reports to BPA) and are not specifically identified here. The unfilled circles represent previously described collections in the 2009 and 2010 steelhead (*O. mykiss*) baseline, while black circles represent collections in the 2009 and 2010 Chinook salmon (*O. tshawytscha*) baseline.

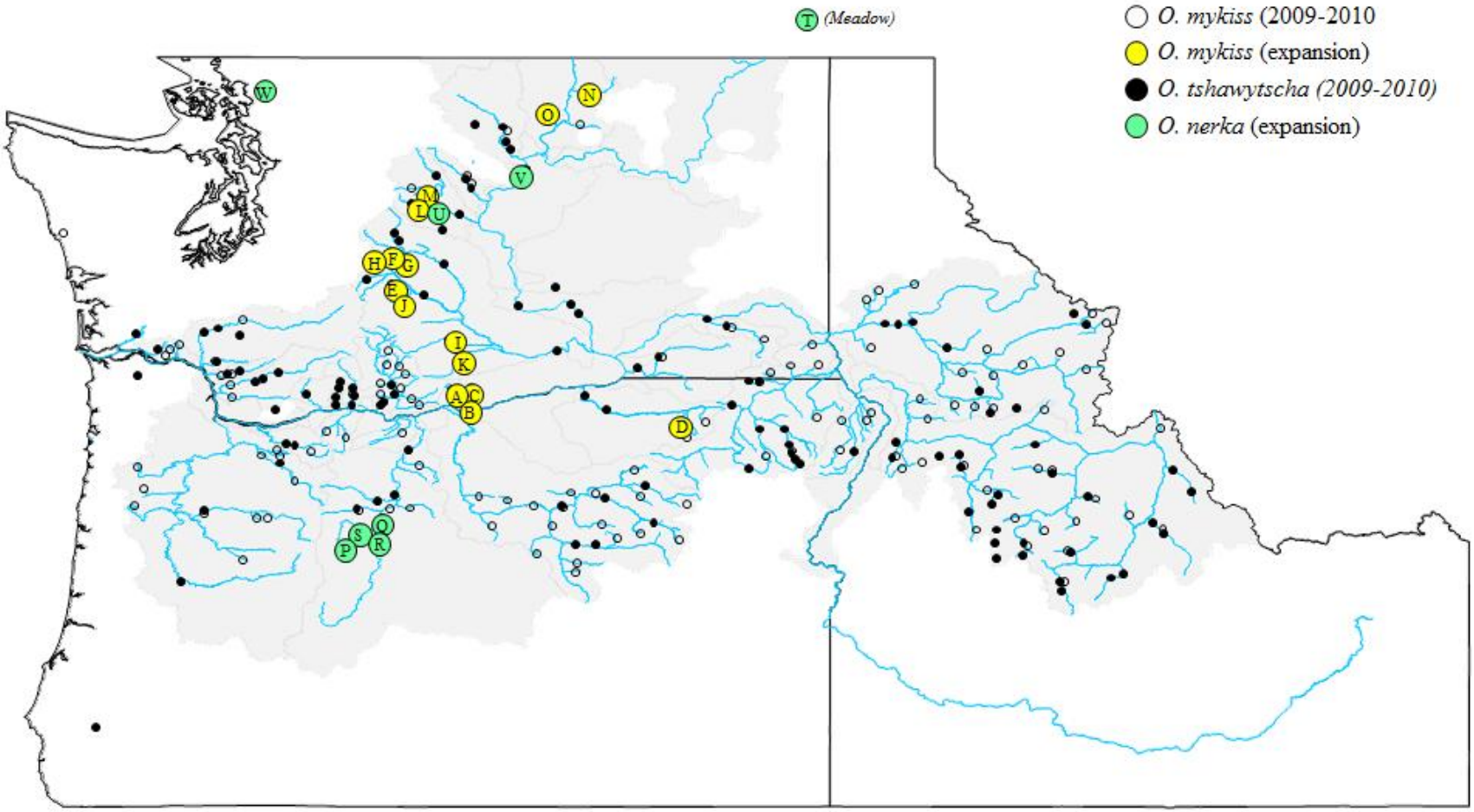


Figure 2a.) Pairwise F_{ST} comparisons (by lineage) between Chinook salmon collections, based on 191 SNP markers; the amount of total variation attributable to among-group differences. Among-lineage global F_{ST} is 0.2089. Mean overall F_{ST} by lineage is indicated by dashed lines.

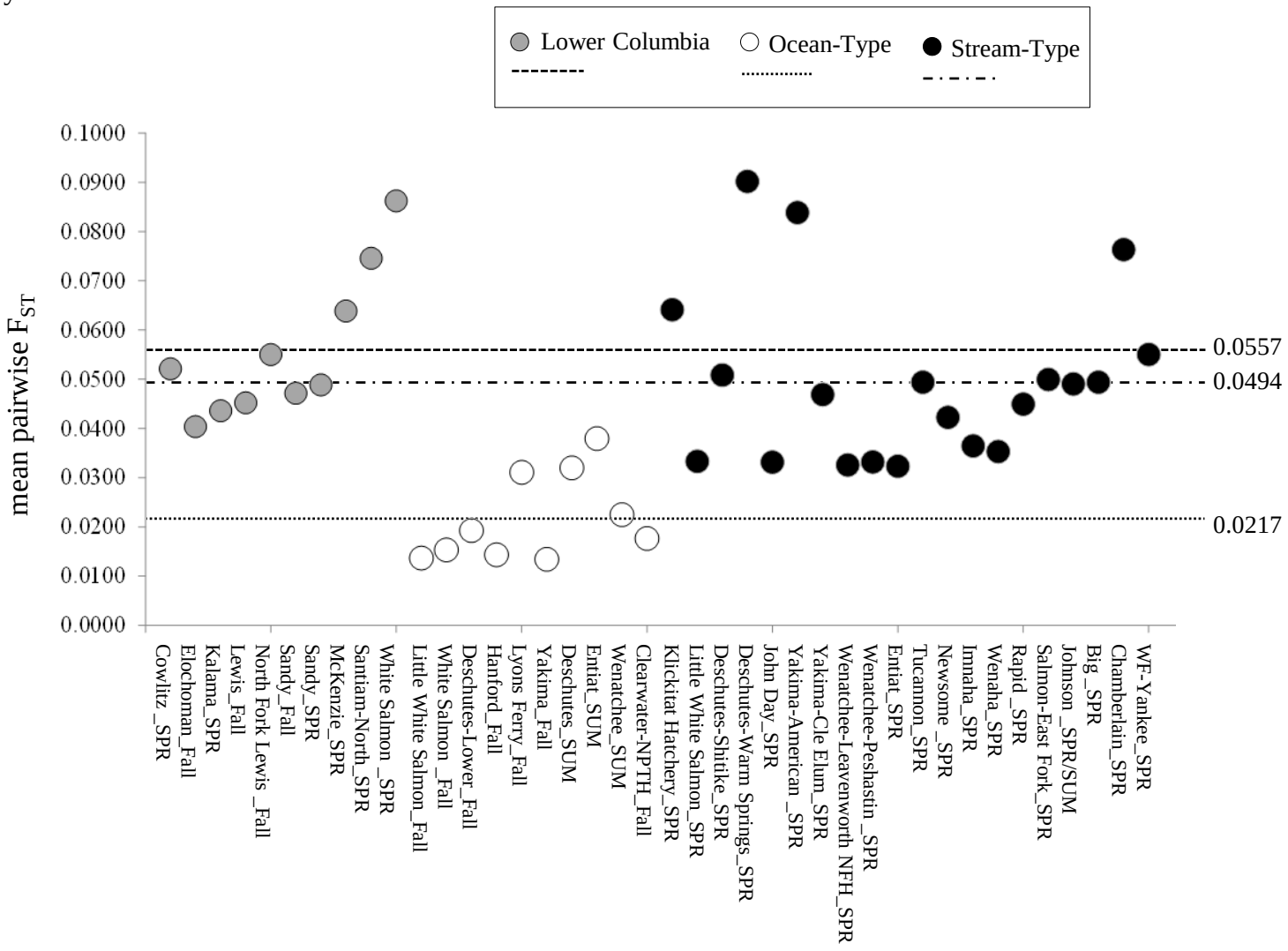


Figure 2b.) Pairwise F_{ST} comparisons (by lineage) between steelhead collections, based on 188 SNP markers. The overall mean F_{ST} among lineages is 0.069. Collections grouped by subbasin or watershed are: 1) Quinault, 2) Willamette, 3) Abernathy/Germany/Mill, 4) Cowlitz, 5) Kalama, 6) Lewis, 7) Sandy, 8) Big White Salmon, 9) Coweeman, 10) Hood, 11) Deschutes, 12) fifteenmile, 13) Klickitat, 14) Touchet, 15) Umatilla , 16) Yakima, 17) John Day, 18) Rock, 19) Entiat, 20) Methow, 21) Okanogan, 22) Wenatchee, 23) Tucannon, 24) Asotin, 25) M. F. Clearwater, 26) S. F. Clearwater, 27) Clearwater, 28) Grande Ronde, 29) Imnaha, 30) M. F. Salmon, 31) S. F. Salmon, 32) upper Salmon, 33) Bargamin, 34) lower Salmon.

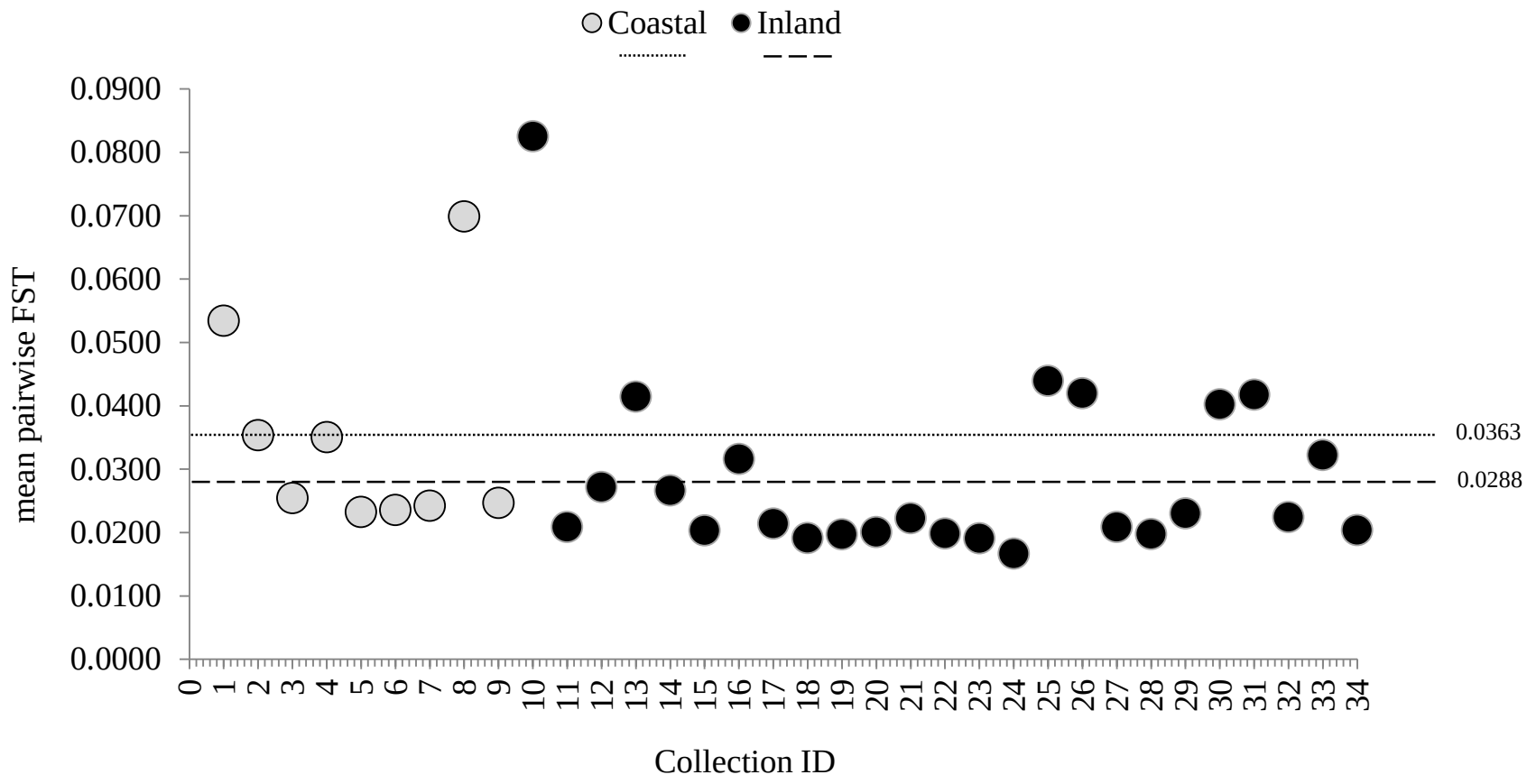


Figure 2c.) Mean pairwise F_{ST} between *O. nerka* collections based on 96 SNP markers. Means were calculated from comparisons among regions or watersheds (excluding comparisons between collections from the same subbasin). The mean pairwise values among temporal collections or collections from the same site are denoted by a red bar. The global or overall value is marked by a dashed line; in the Deschutes River region, (SL) is Suttle Lake, and (LBC) is Lake Billy Chinook.

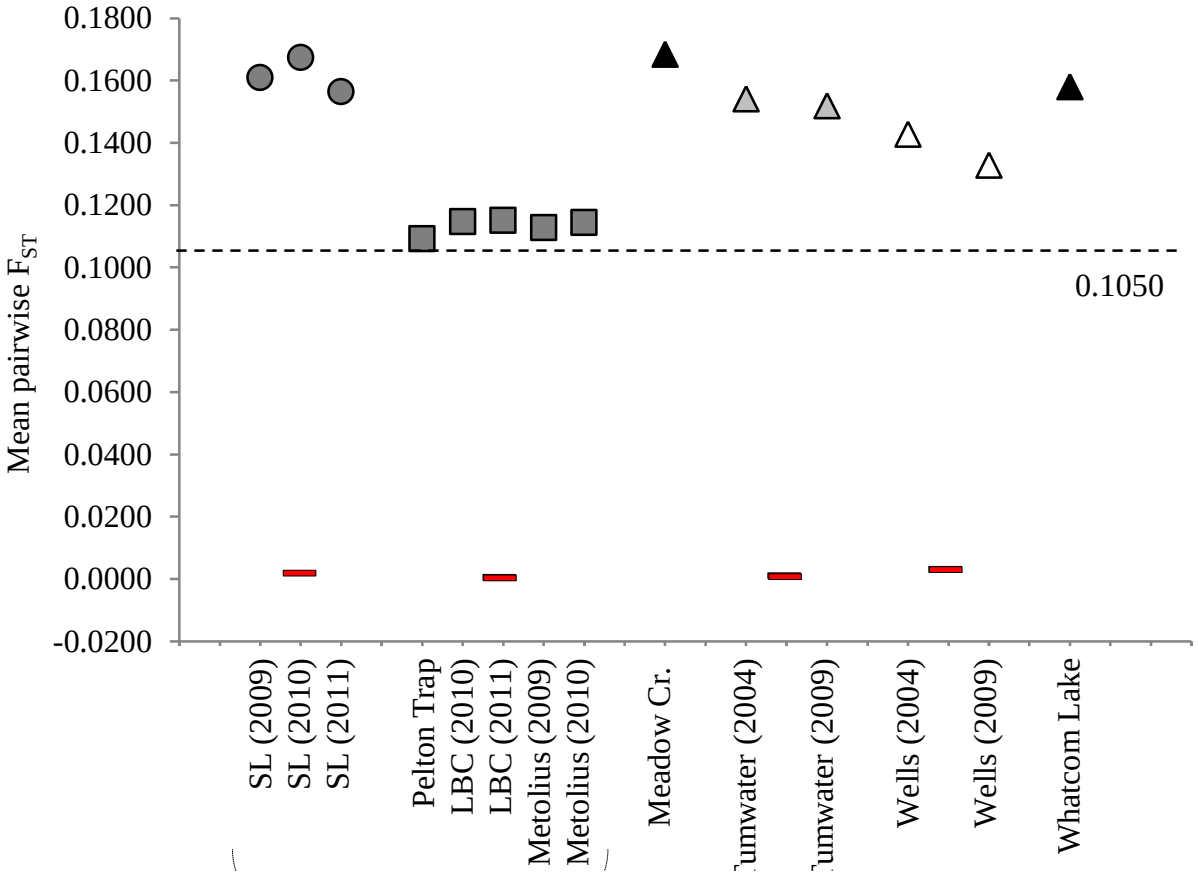


Figure 3a.) Chinook salmon NJ-phylogram based on Neis distance (1972) and 191 SNPs. Consensus in the tree topology greater than 50% is shown in red text at nodes. Note the 100% bootstrap support for the topology separating stream-type from the remaining two lineages. Numbers correlate to collection descriptions in Table 1.

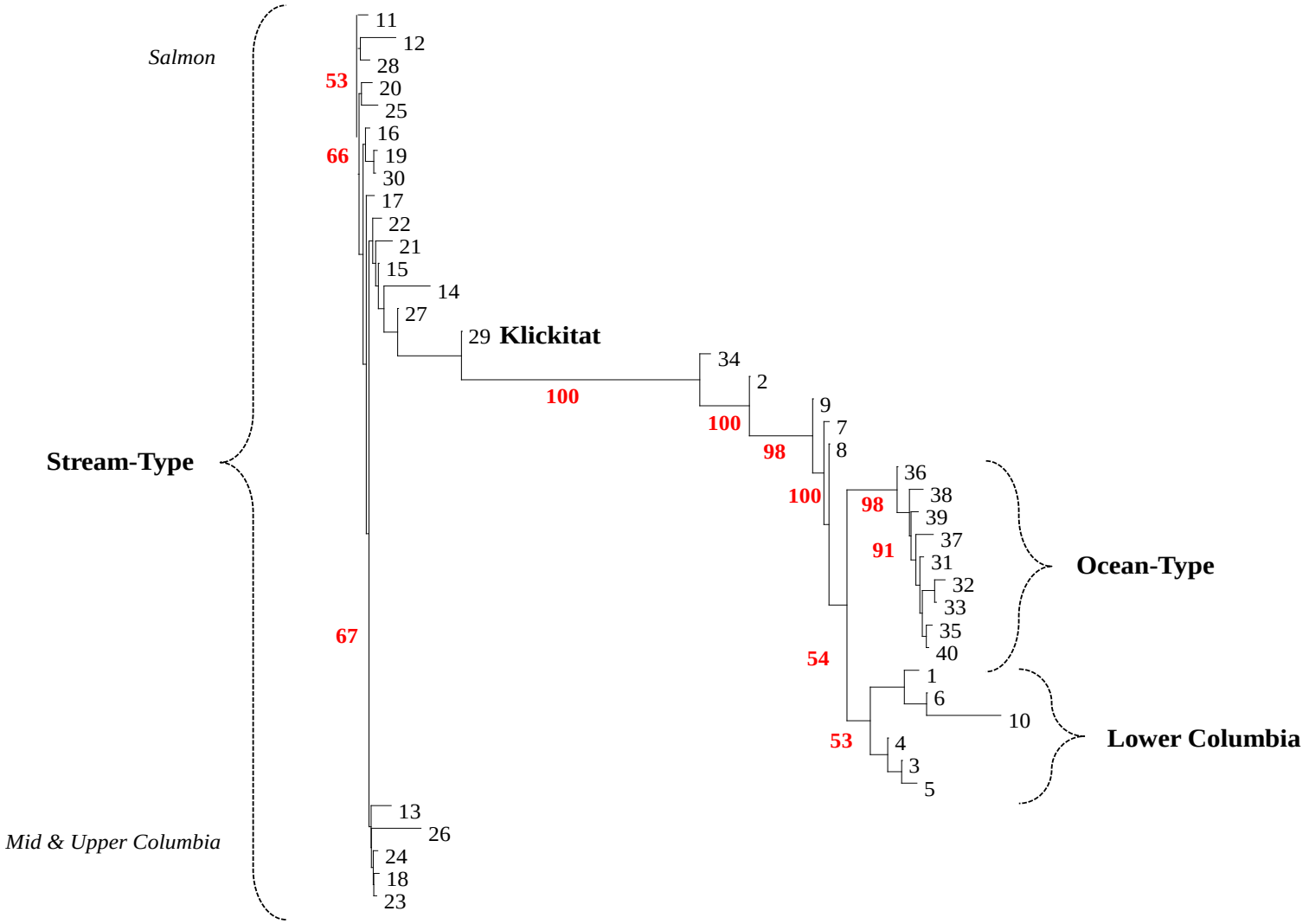


Figure 3b.) Steelhead NJ-phylogram based on Neis distance (1972) and 188 SNPs. Coastal collections are separated from inland collections with 100% bootstrap support. A key to population ID# is provided in Appendix 1.

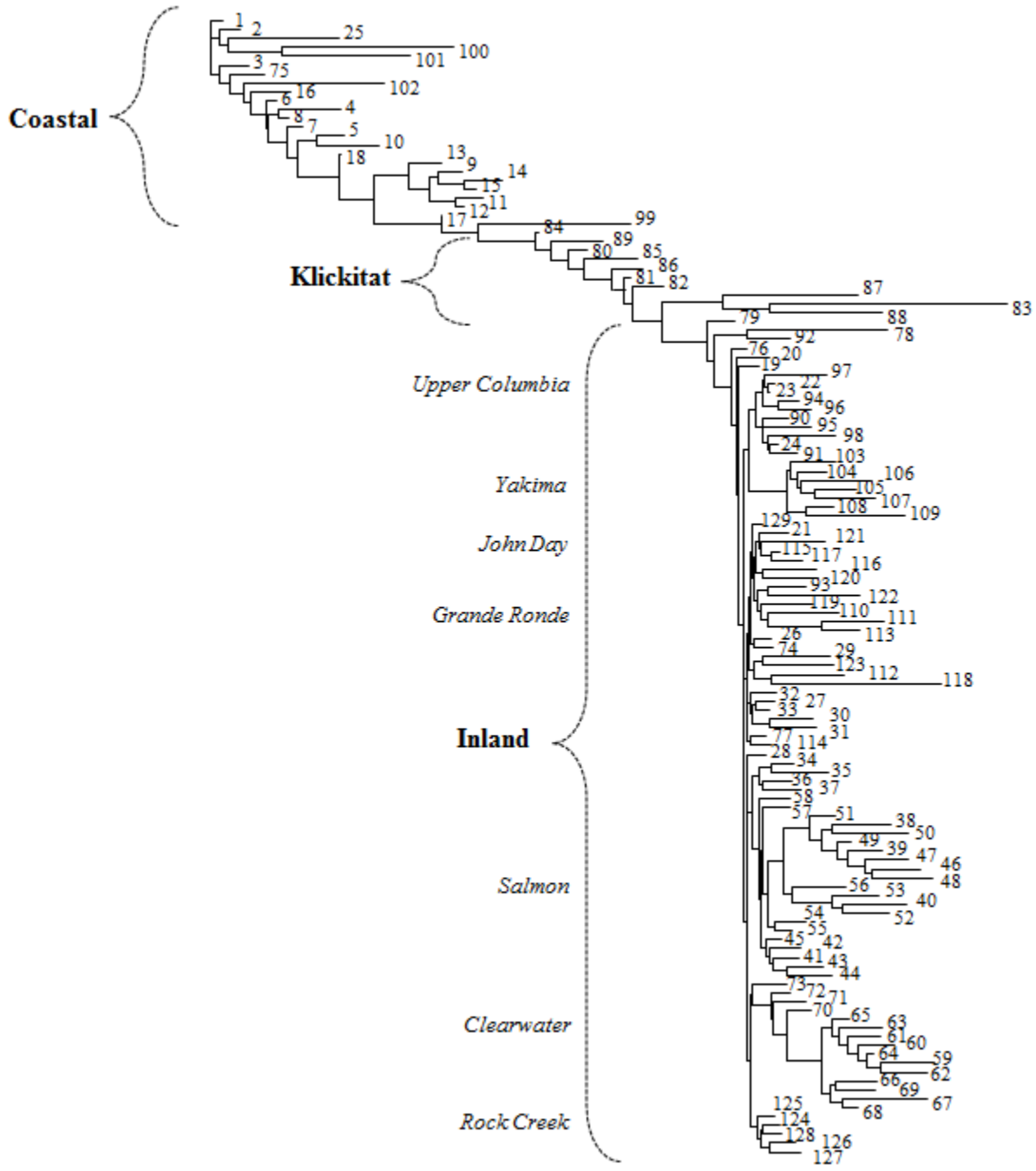


Figure 3c.) *O. nerka* NJ-phylogram based on Neis distance (1972) and 96 SNPs. Consensus in the tree topology greater than 50% is shown in red text at nodes.

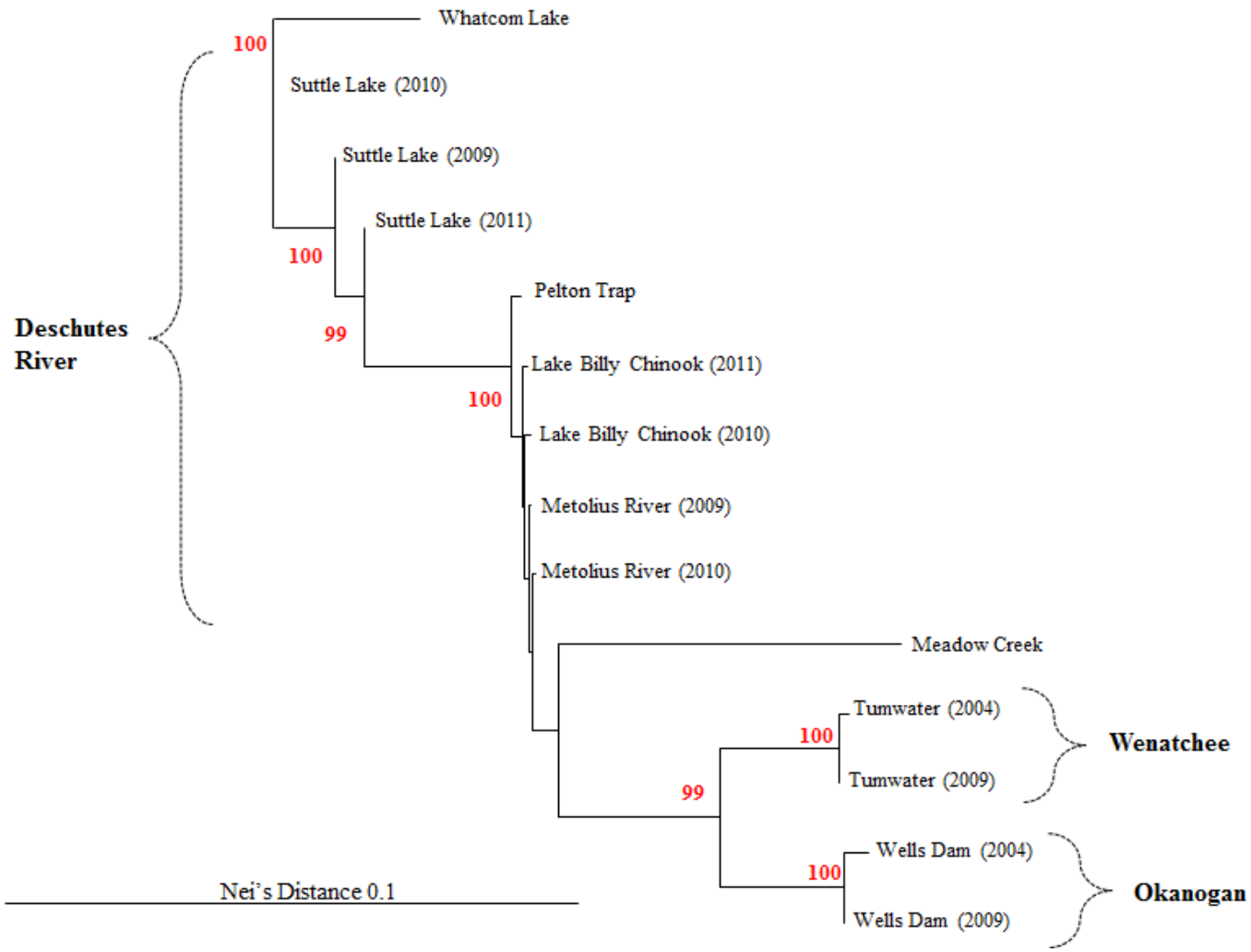


Figure 4a.) Principle coordinates analysis (PCA) plot for Chinook salmon based on 191 SNPs. Numbers correspond with collection ID#'s (see Table 1).

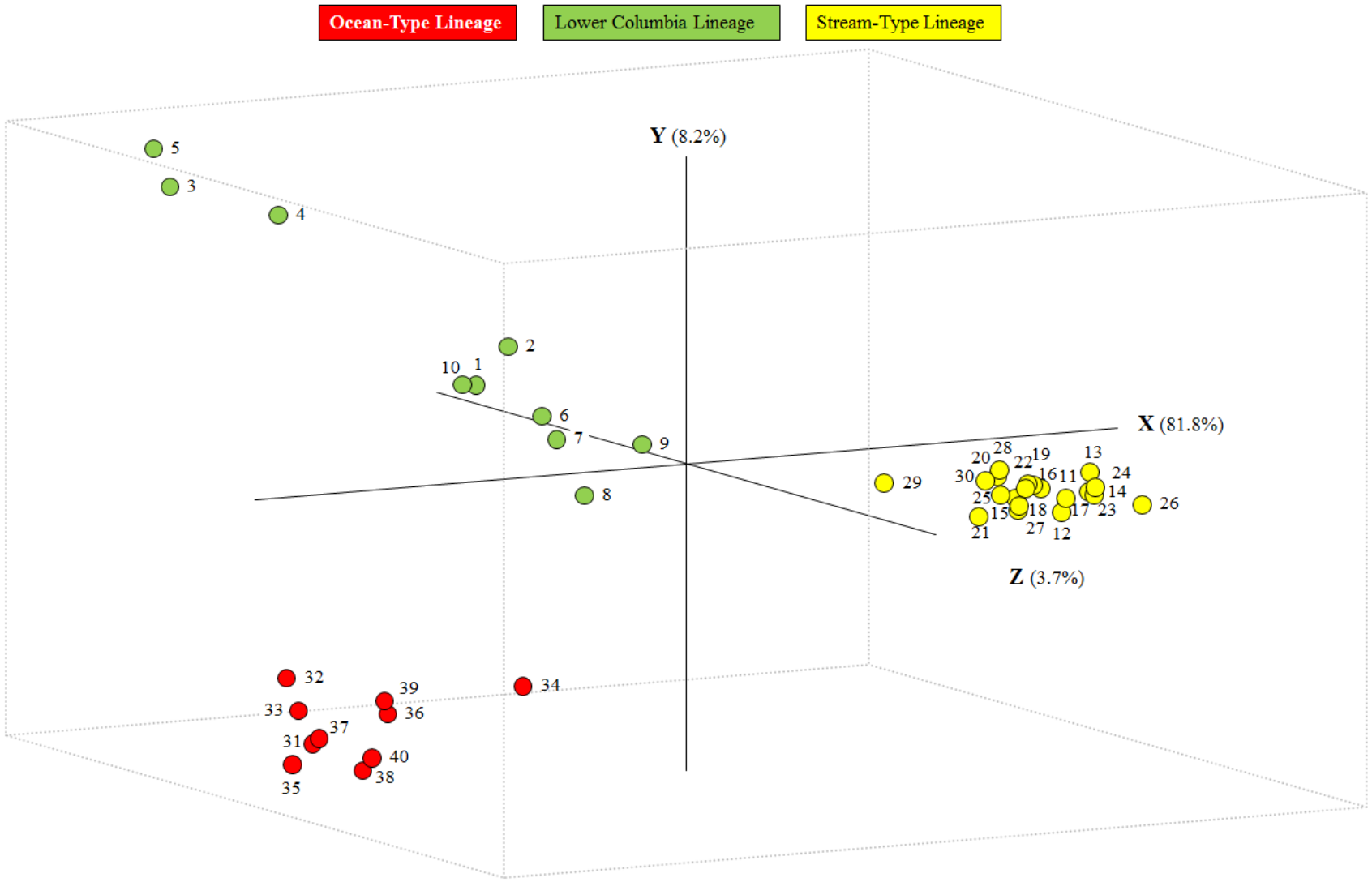


Figure 4b.) Principle coordinates analysis (PCA) plot for steelhead trout based on 188 SNPs. Collections grouped by subbasin or watershed are: 1) Quinault, 2) Willamette, 3) Abernathy/Germany/Mill, 4) Cowlitz, 5) Kalama, 6) Lewis, 7) Sandy, 8) Big White Salmon, 9) Coweeman, 10) Hood, 11) Deschutes, 12) fifteenmile, 13) Klickitat, 14) Touchet, 15) Umatilla , 16) Yakima, 17) John Day, 18) Rock, 19) Entiat, 20) Methow, 21) Okanogan, 22) Wenatchee, 23) Tucannon, 24) Asotin, 25) M. F. Clearwater, 26) S. F. Clearwater, 27) Clearwater, 28) Grande Ronde, 29) Imnaha, 30) M. F. Salmon, 31) S. F. Salmon, 32) upper Salmon, 33) Bargamin, 34) lower Salmon.

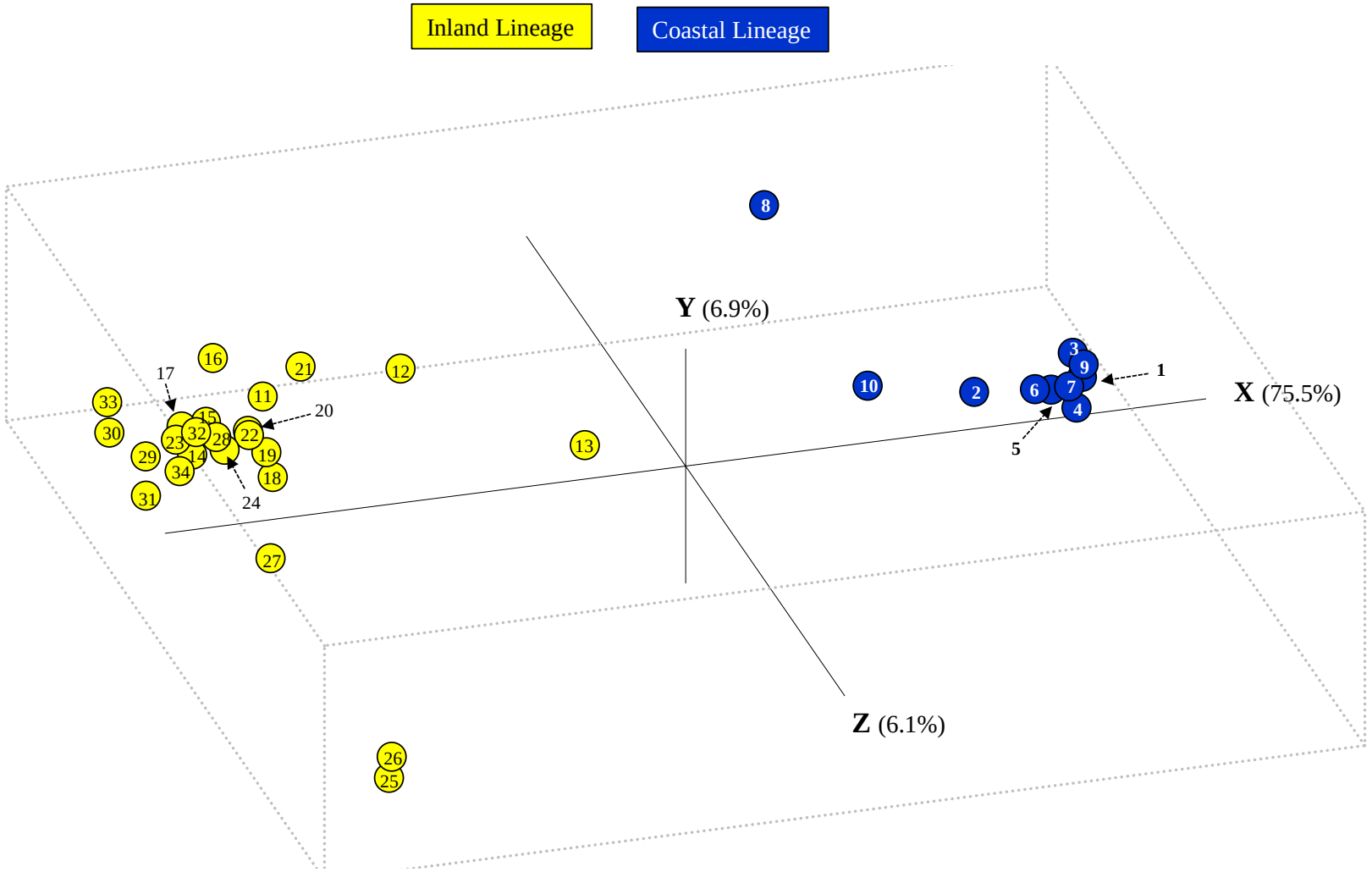
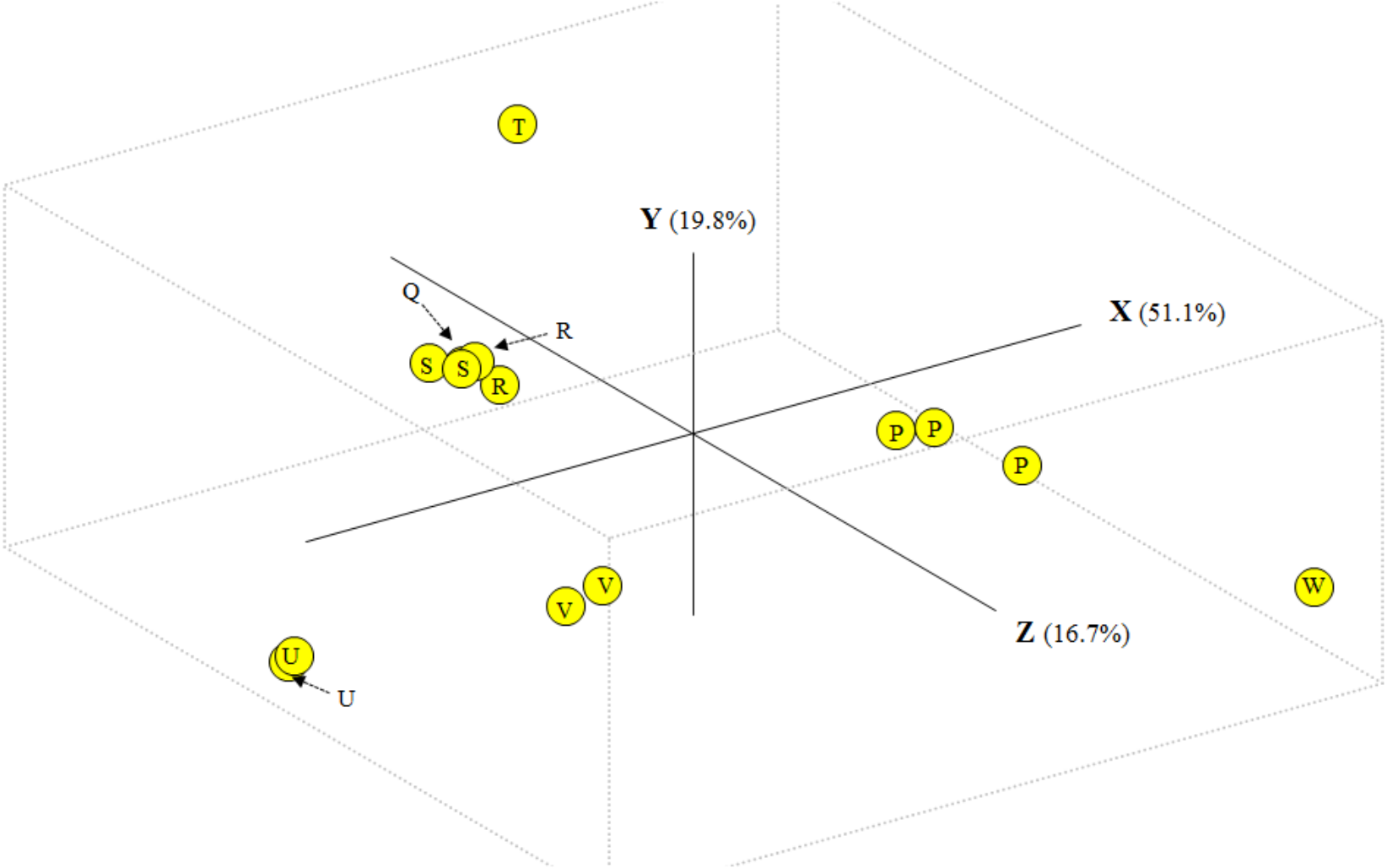


Figure 4c.) Principle coordinates analysis (PCA) plot for *O. nerka* based on 96 SNPs. Letters correspond to collection ID (see Table 1).



Appendix 1. Collection codes for all steelhead collections comprising the current 2011 baseline. Nineteen new collections were added in 2011 (±). An additional 27 collections identified in the 2009 annual report (**) were updated using the 192 complementary SNP assays (2 x 96 panels) first developed and described in the 2010 annual report; Snake River collections were updated and contributed by Idaho Department of Fish and Game.

ID	Watershed	Region/Major Trib.	Genotyper	Lineage
1	Abernathy	Lower Columbia	CRITFC	**coastal**
2	Germany	Lower Columbia	CRITFC	**coastal**
3	Mill	Lower Columbia	CRITFC	**coastal**
4	Cowlitz	Lower Columbia	CRITFC	**coastal**
5	**Kalama (summer)	Lower Columbia	CRITFC	**coastal**
6	Kalama (winter)	Lower Columbia	CRITFC	**coastal**
7	E. F. Lewis	Lower Columbia	CRITFC	**coastal**
8	N. F. Lewis	Lower Columbia	CRITFC	**coastal**
9	Clackamas	Willamette	CRITFC	**coastal**
10	Clackamas (N. F. Dam)	Willamette	CRITFC	**coastal**
11	N. F. Eagle	Willamette	CRITFC	**coastal**
12	Eagle	Willamette	CRITFC	**coastal**
13	Little Rock/Mad	Willamette	CRITFC	**coastal**
14	N. F. Santiam	Willamette	CRITFC	**coastal**
15	Wiley	Willamette	CRITFC	**coastal**
16	Still	Sandy	CRITFC	**coastal**
17	Hood (east & middle forks)	Middle Columbia	CRITFC	inland
18	Hood (west fork)	Middle Columbia	CRITFC	inland
19	Pelton	Deschutes	CRITFC	inland
20	Shitike	Deschutes	CRITFC	inland
21	Touchet	Walla Walla	CRITFC	inland
22	Entiat	Upper Columbia	CRITFC	inland
23	Entiat Trap	Upper Columbia	CRITFC	inland
24	Methow	Upper Columbia	CRITFC	inland
25	Quinault	Coast	CRITFC	**coastal**
26	Tucannon	Snake	CRITFC	inland
27	Crooked	Grande Ronde	CRITFC	inland
28	Asotin	Snake	CRITFC	inland
29	Elk	Grande Ronde	IDFG	inland
30	Lostine	Grande Ronde	IDFG	inland
31	Little Minam	Grande Ronde	IDFG	inland
32	Lower Grande Ronde	Grande Ronde	IDFG	inland
33	Wenaha	Grande Ronde	IDFG	inland

34	Big Sheep	Imnaha	IDFG	inland
35	Camp	Imnaha	IDFG	inland
36	Cow	Imnaha	IDFG	inland
37	Lightning	Imnaha	IDFG	inland
38	Big	M. F. Salmon	IDFG	inland
39	Loon	M. F. Salmon	IDFG	inland
40	Stolle	S. F. Salmon	IDFG	inland
41	Sawtooth	Upper Salmon	IDFG	inland
42	W. F. Yankee Fork	Upper Salmon	IDFG	inland
43	Morgan	Upper Salmon	IDFG	inland
44	Pahsimeroi	Upper Salmon	CRITFC	inland
45	N. F. Salmon	Salmon	IDFG	inland
46	Marsh	M. F. Salmon	IDFG	inland
47	Rapid	M. F. Salmon	IDFG	inland
48	Pistol	M. F. Salmon	IDFG	inland
49	Camas	M. F. Salmon	IDFG	inland
50	Big	M. F. Salmon	IDFG	inland
51	Bargamin	Salmon	IDFG	inland
52	E. F. South Fork	S. F. Salmon	IDFG	inland
53	Secesh	S. F. Salmon	IDFG	inland
54	Boulder	Little Salmon	IDFG	inland
55	Hazard	Little Salmon	IDFG	inland
56	Rapid	Little Salmon	IDFG	inland
57	Slate	Lower Salmon	IDFG	inland
58	White Bird	Lower Salmon	IDFG	inland
59	Storm	M. F. Clearwater	IDFG	inland
60	Loch_CF	Lochsa	IDFG	inland
61	Loch_Canyon	Lochsa	IDFG	inland
62	Sel_Bear	Selway	IDFG	inland
63	Sel_Moose	Selway	IDFG	inland
64	Sel_Gedney	Selway	IDFG	inland
65	Sel_Ohara	Selway	IDFG	inland
66	SFC_Crook	S. F. Clearwater	IDFG	inland
67	SFC_Tenmile	S. F. Clearwater	IDFG	inland
68	SFC_Johns	S. F. Clearwater	IDFG	inland
69	Clear	S. F. Clearwater	IDFG	inland
70	E. F. Potlatch	Clearwater	IDFG	inland
71	Big Bear	Clearwater	IDFG	inland
72	Little Bear	Clearwater	IDFG	inland
73	Lapwai (Mission)	Clearwater	IDFG	inland
74	Tucannon	Snake	IDFG	inland

75	**Coweeman	Lower Columbia	CRITFC	**coastal**
76	**BuckHollow	Deschutes	CRITFC	inland
77	**Tout	Deschutes	CRITFC	inland
78	**Upper Mainstem	Deschutes	CRITFC	inland
79	**Fifteenmile	Middle Columbia	CRITFC	inland
80	**lower Little Klickitat	Klickitat	CRITFC	inland
81	**Summit	Klickitat	CRITFC	inland
82	**Trout	Klickitat	CRITFC	inland
83	**Upper Trout	Klickitat	CRITFC	inland
84	**Bowman	Klickitat	CRITFC	inland
85	**Dead Canyon	Klickitat	CRITFC	inland
86	**Lower White	Klickitat	CRITFC	inland
87	**Snyder	Klickitat	CRITFC	inland
88	**Surveyor	Klickitat	CRITFC	inland
89	**Swale	Klickitat	CRITFC	inland
90	±Bonaparte	Okanogan	CRITFC	inland
91	**Omak	Okanogan	CRITFC	inland
92	±Salmon	Okanogan	CRITFC	inland
93	**Umatilla	Middle Columbia	CRITFC	inland
94	±Chiwaukum	Wenatchee	CRITFC	inland
95	±Leavenworth-NFH	Wenatchee	CRITFC	inland
96	±Upper Chiwaukum	Wenatchee	CRITFC	inland
97	**Nason	Wenatchee	CRITFC	inland
98	±Icicle	Wenatchee	CRITFC	inland
99	**Big White Salmon	Middle Columbia	CRITFC	**coastal**
100	**Canyon	Willamette	CRITFC	**coastal**
101	**Luckiamute	Willamette	CRITFC	**coastal**
102	**Willamina	Willamette	CRITFC	**coastal**
103	±Rattlesnake	Yakima	CRITFC	inland
104	±Nile	Yakima	CRITFC	inland
105	±Pileup	Yakima	CRITFC	inland
106	±Quartz	Yakima	CRITFC	inland
107	±N. F. little Naches	Yakima	CRITFC	inland
108	±Satus	Yakima	CRITFC	inland
109	±Toppenish	Yakima	CRITFC	inland
110	Beech	John Day	CRITFC	inland
111	Upper Mainstem	John Day	CRITFC	inland
112	Rock/Baldy	John Day	CRITFC	inland
113	Lower Mainstem	John Day	CRITFC	inland
114	Pine/Service/Bridge	John Day	CRITFC	inland
115	Middle Fork	John Day	CRITFC	inland

116	**Camp	John Day	CRITFC	inland
117	**Clear	John Day	CRITFC	inland
118	Granite	John Day	CRITFC	**coastal**
119	**Camus	John Day	CRITFC	inland
120	**Desolation	John Day	CRITFC	inland
121	Big/Fox	John Day	CRITFC	inland
122	Deer	John Day	CRITFC	inland
123	Murderers	John Day	CRITFC	inland
124	±Rock (2009)	Rock Creek	CRITFC	inland
125	±Squaw (2009)	Rock Creek	CRITFC	inland
126	±Bickleton	Rock Creek	CRITFC	inland
127	±Highway 8 Bridge	Rock Creek	CRITFC	inland
128	±Squaw (2008)	Rock Creek	CRITFC	inland
129	±Iskuulpa	Umatilla	CRITFC	inland

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

Introduction

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). These methods have been demonstrated to be useful even at relatively fine geographic scales within the Columbia River Basin (CRB) (Hess et al. 2011, Hess and Narum 2011). Within the CRB, Chinook salmon consist of three major genetic lineages, which can be further broken into populations that are genetically structured on a finer spatial scale (e.g., Waples et al. 2004; Narum et al. 2010). In past analyses, we aggregated CRB Chinook salmon populations into twelve reporting groups for GSI applications (e.g. Hess et al. 2011), but this current study uses an adjusted set of seventeen reporting groups made possible with a large increase in genetic markers, from 96 single nucleotide polymorphism (SNP) markers to 192 SNPs.

Chinook salmon fisheries in the mainstem of the lower Columbia River provide an ideal and important application of GSI because the fish harvested consist of mixtures of stocks from the entire CRB. In addition, fisheries in this location represent a majority of the CRB harvest of this species taken by the commercial, sport, and tribal fishermen. In order to help establish sustainable fisheries, GSI can be used to address two primary questions: 1) how are Chinook salmon stocks temporally and spatially distributed in the mainstem lower Columbia River; and 2) how are these stocks temporally and spatially distributed in the harvests of fisheries. This information has the potential to be used by fisheries managers to shape the various fisheries in a sustainable way that would protect less abundant Chinook salmon stocks, while targeting the stocks of adequate abundance.

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. Specifically for this study, we represented the total fish present in the mainstem Columbia River by sampling fish at a fixed point at Bonneville Dam (results presented in Section 4) and we estimate stock composition for spring- and fall-run Chinook salmon harvested above and below Bonneville Dam by sampling sport, commercial, and tribal fisheries in the 2011 season (results discussed in this section). To estimate temporal stock distributions, we obtained sample sizes that were large enough to estimate proportions of Chinook salmon stocks from each type of fishery across weekly strata. We characterized spatial distributions of Chinook salmon stocks by pooling fishery samples into two regions located within the downriver and upriver side of Bonneville Dam.

In this report, we utilized 192 SNP loci to genotype unknown mixture samples to estimate stock composition of these mixtures, by fishery, location, and weekly strata. Previously we have shown that 92 SNPs are nearly as powerful as a suite of thirteen microsatellite markers (established for Chinook salmon range-wide GSI applications) for performing GSI within the CRB based on results from simulations and statistical power analyses. We then extrapolated that using twice as many SNPs (~192) would surpass the power of 13 microsatellites (Hess et al. 2011). Here we examine the truth of this extrapolation by testing the accuracy of the new 192-SNP baseline. Our accuracy testing includes analyses of 100% mixtures simulated for each

baseline population and an estimation of correct assignment of baseline individuals to their reporting group of origin. We include a discussion of the identified weaknesses of the baseline and the overall power of these 192 SNP markers as well as a new strategy for improving stock identification. This new strategy takes advantage of a recently developed genetic technology, parentage based tagging (PBT), by jointly using the two genetic tools, GSI and PBT, in a tiered approach for stock identification. PBT is an efficient approach for mass marking fish. The method is carried out by first genotyping a set of potential parents which then provides the opportunity to assign a set of genotyped offspring to their true parent pair. PBT is currently being utilized on a broad scale in the Columbia River Basin to mark all Chinook salmon and steelhead Snake River hatchery broodstock (details in Steele et al. 2011). This application has effectively marked all Snake River hatchery Chinook salmon and steelhead starting with the 2008 brood years. When parent pairs of a Snake River hatchery fish are identified with PBT, we can provide accurate information including age of the fish and the source hatchery in which its parents were spawned. In future years, we intend to use PBT in this Chinook salmon harvest study to identify all Snake River hatchery-origin salmon, and then we will estimate stock-of-origin of all other hatchery fish that were not assigned with PBT (i.e. non-Snake River hatchery-origin) and all wild fish using GSI. We demonstrate the utility of PBT by assigning the 3-year old Snake River hatchery-origin spring Chinook salmon (i.e. brood year 2008) from the 2011 spring Chinook salmon harvest. Similar to coded wire tags (CWT), PBT provides source hatchery for all tagged fish. In this study, we use the known origins of subsets of harvested fish based on CWT data and assignments using parentage based tagging (PBT), in order to assess accuracy of GSI-estimation of the origins of these same fish.

Methods

Tissue collection

Tissues were sampled from Chinook salmon in 2011 from a total of five different mixture sources: 1) Bonneville Dam (results discussed in section 4), and the spring- and fall-run seasons of the following fisheries: 2) commercial, 3) sport, 4) test, and 5) tribal. The tribal harvest included both a commercial gill-net fishery in the fall (sampled by Yakama Nation monitors), as well as a ceremonial harvest conducted in the spring (sampled by the Warm Springs fishery program). These tissues were collected in coordination with existing monitoring programs led by Washington Department of Fish and Wildlife (WDFW), Oregon Department of Fish and Wildlife (ODFW), and the Warm Springs and Yakama Nation tribes. The spring-run fisheries were sampled below Bonneville Dam in the sport, commercial, and test fishery (regions A and B), and sampled above Bonneville Dam in region 01 as part of the Warm Springs tribal ceremonial fishery (Figure 1, Table 1). The fall-run fisheries were sampled above Bonneville Dam (Zone 6 Yakama Nation tribal fishery) and below Bonneville Dam (regions A and B via commercial fishery). Stock proportions were calculated for strata within each fishery source, such that stock proportions could be compared across weeks strata and regions. We use the following four main geographic regions (Figure 1): Region A corresponds to our grouping of pre-existing Oregon and Washington state sport fishing zones 1-4 (or commercial zones 4-5), Region B corresponds to our grouping of sport zones 5-10 (or commercial zones 1-3), and Region 01 and Region 02 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.

Table 1. Characteristics of Chinook salmon harvest samples by fishery source, region, tag recovery, and weekly strata.

Fishery source	Region	Type	Statistical week: Spring 9	10	11	12	13	14	15	16	17	18	20	21	Summer 22	23	24	25	Fall 32	34	35	36	37	39	40	41	42	43	Total
Commercial	A	CWT						1	2	1			86	121					1	24	84	53		49	4	5		1	432
		Regular											112	117					1	97	97	108		109	75	67	31	95	909
	B	CWT						92	80				34						19					47	10	6	5	2	295
		Regular						106	118				65						81					94	90	66	73	62	755
Sport	A	CWT	4	6	8	5	21	12	2	5	7			32	3	13	30	14											162
		Regular	11	20	49	54	79	66	26	46	36			67	14	38	61	38											605
	B	CWT					10	12	12	25	10		4	11	7	14	22	16											143
		Regular					49	37	50	75	58			23	8	31	33	15											379
Onboard	A	CWT																											
		Regular						4	2																				6
	B	CWT						2																					2
		Regular						13	13																				26
Test	B	CWT				3	3	7		6			14																33
		Regular	7	4	9	6	59	47		107			70																309
Tribal	01	Regular					4	2	22	100	99	85									60		8			42			422
	02	Regular																			39	48	16		9	52			164
																													4642

Note: a.) Region refers to the areas harvested shown in figure 1.
b.) Type refers to whether a CWT was detected and recovered from a fish that was sampled. "Regular" indicates absence of a CWT.
c.) Weekly strata highlighted in grey indicate weeks in which CWT and regular samples were both in sufficient numbers (n>19) to compare stock proportions across these sample types. Outlined weeks indicate sufficient sample sizes within collections of a certain sample type. Only strata with greater than 19 samples were used to estimate stock proportions.
d.) Statistical week 9 equals 2/21/2011-2/27/2011 and week 43 equals 10/17/2011-10/23/2011.

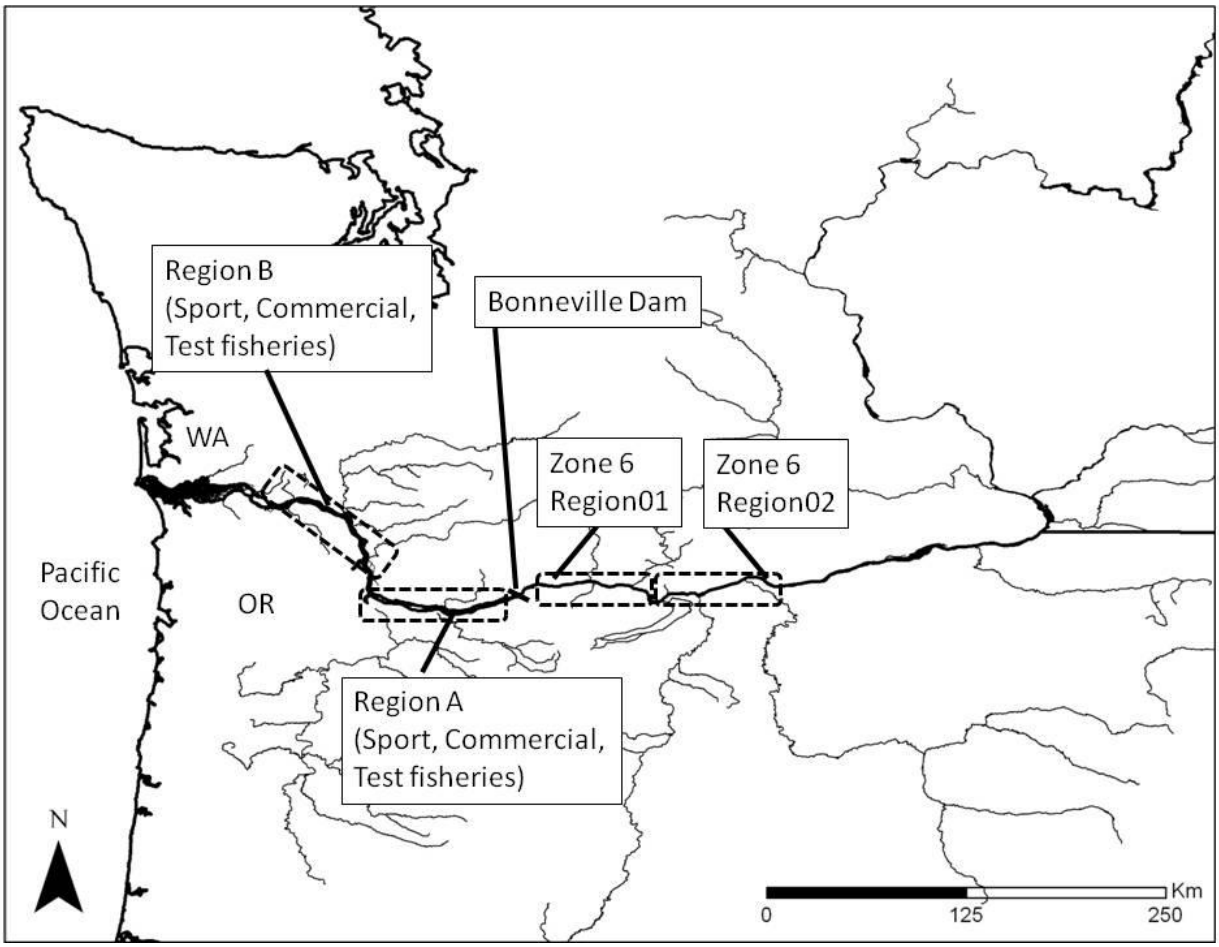


Figure 1. Map of sources of Chinook salmon mixtures.

Fisheries on spring-run Chinook salmon are mark selected based on absence or presence of the adipose fin to distinguish hatchery fish from natural origin fish, 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. This year, only the test fishery mixture contained enough representative natural-origin fish samples in sufficient numbers to compare weekly strata the mark-selective spring-run fisheries to determine whether there was any effect of "origin" on stock proportions.

Molecular data

A total of 192 SNP loci were used for genotyping Chinook salmon (Section 1, Appendix 1). However, for GSI analyses, the following four loci were excluded: Ots_SEXY1 (a marker for sex determination), Ots_FGF6A (found to be linked to another marker, Ots_FGF6B_1, which was retained), Ots_hsc71-3'-488 (found to be linked to another marker, Ots_hsc71-5'-453, which was retained), and Ots_zP3b-215 found to be monomorphic in the Columbia River Chinook salmon baseline. Therefore, we used 188 total SNP markers for GSI. See previous sections on SNP marker and baseline development for details of laboratory protocols involved in using these DNA markers for genotyping fish. For PBT analyses, we used 95 of the 192 SNP loci (see

Section 1, Appendix 1 for a list of PBT markers), which have previously been demonstrated to provide accurate parent assignments (Steele et al. 2011).

Statistical Analyses

We grouped 40 baseline populations (Table 2) into seventeen reporting groups that were primarily determined by the relative genetic similarity among populations according to a phylogenetic analysis. STRUCTURE v2.3.2 (Pritchard et al. 2000) was used to identify strays (>80% assignment to a different lineage) in each collection by setting a K value of 3 or higher; 30 strays were swapped into a more appropriate collection (e.g. in Entiat R., interior ocean-type and stream-type lineages co-occur and have slightly overlapping spring and summer run-timing and so must be genetically identified to make a pure collection) and 30 strays were removed entirely from the dataset (results not shown). Genetic distances were computed from allele frequencies based on Cavalli-Sforza and Edwards (1967) chord distance, with the PHYLIP v 3.69 (Felsenstein 1989) and 1000 bootstrap replicates were performed. Distances were clustered using the Neighbor – Joining method (Saitou and Nei, 1987), and a consensus tree was constructed (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). It has been recommended that reporting groups meet a threshold of 90% proportions in these 100% mixture simulation analyses to be useful for fishery management applications (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 100 bootstraps. ONCOR was also used to assign individuals from the baseline in a "leave-one-out" analysis to estimate correct individual assignment to reporting groups. To examine how much greater power 188 SNPs can generate compared to our previous 92-SNP baseline, we excluded all SNPs that weren't part of the 92 SNP marker set and ran all these analyses.

Aside from simulations, we examined how well GSI assignments corresponded to known origin fish based on CWT data and PBT assignments. In total, 757 CWT were recovered during the spring and summer fishery harvest of Chinook salmon, and 310 were recovered during the fall season (Table 1). The hatchery-of-origin information was fit into our reporting group categories to make it compatible for comparing to GSI results. We used PBT to identify the source hatcheries of 3-year old Snake River spring-run Chinook salmon. Genotypes of parents and offspring were analyzed with SNPPITv1.0, a software program developed for large PBT datasets based on SNP markers (Anderson 2010). Analyses with SNPPIT were performed on the 2008 Chinook salmon broodstock baseline (Steele et al. 2011), to assign any offspring that were included in the 2011 Chinook salmon harvest dataset. PBT assignments were based on marker exclusion and a false discovery rate (FDR) threshold of 1.5% was used to ensure high confidence in all parent matches.

Table 2. Sample sizes and reporting groups of Chinook salmon baseline populations.

#	Collection	N	Lineage	Reporting Group
1	Cowlitz_sp	89	LC	W_Cascade_sp
2	Kalama_sp	69	LC	W_Cascade_sp
3	Elochoman_fa	81	LC	W_Cascade_fa
4	Lewis_fa	93	LC	W_Cascade_fa
5	NFLewis_fa	81	LC	W_Cascade_fa
6	Sandy_fa	78	LC	W_Cascade_fa
7	McKenzie_sp	75	LC	Willamette_sp
8	Sandy_sp	46	LC	Willamette_sp
9	Nsantiam_sp	75	LC	Willamette_sp
10	WhiteSalmon_fa	89	LC	Spring_Cr_Group_Tule
11	Klickitat_sp	85	ST	Klickitat_sp
12	Shitike_sp	93	ST	Deschutes_R_sp
13	WarmSprings_sp	88	ST	Deschutes_R_sp
14	JohnDay_sp	76	ST	John_Day_sp
15	American_sp	66	ST	Yakima_sp
16	CleElum_sp	86	ST	Yakima_sp
17	Entiat_sp	96	ST	Upper_Columbia_R_sp
18	LWhiteSalmon_sp	93	ST	Upper_Columbia_R_sp
19	LeavenworthNFH_sp	88	ST	Upper_Columbia_R_sp
20	Peshastin_sp	86	ST	Upper_Columbia_R_sp
21	Tucannon_sp	82	ST	Lower_Snake_sp
22	Wenahasp	45	ST	RapidR_Clearwater_sp
23	Imnahasp	91	ST	RapidR_Clearwater_sp
24	NewsomeCrsp	82	ST	RapidR_Clearwater_sp
25	RapidR_sp	93	ST	RapidR_Clearwater_sp
26	BigCr_sp	89	ST	MF_Salmon_sp
27	Chamberlain_sp	45	ST	MF_Salmon_sp
28	JohnsonCr_sp	92	ST	SF_Salmon_sp
29	EFSalmon_sp	94	ST	Upper_Salmon_sp
30	WFYankee_sp	75	ST	Upper_Salmon_sp
31	Entiat_su	46	OT	Upper_Columbia_R_su/fa
32	Wenatchee_su	92	OT	Upper_Columbia_R_su/fa
33	Hanford_fa	90	OT	Upper_Columbia_R_su/fa
34	LWhiteSalmon_fa	89	OT	Upper_Columbia_R_su/fa
35	WhiteSalmon_fa	78	OT	Upper_Columbia_R_su/fa
36	Yakima_fa	60	OT	Upper_Columbia_R_su/fa
37	Deschutes_fa	89	OT	Deschutes_R_fa
38	LDeschutes_fa	90	OT	Deschutes_R_fa
39	Clearwater_fa	85	OT	Snake_R_fa
40	LyonsFerry_fa	90	OT	Snake_R_fa

Note: Chinook salmon baseline collections (n=3230). Lineages are: ST- stream type, OT – ocean type, and LC – Lower Columbia. See Figure 1, Section 2 for geographical locations. "Sp" , "su", and "fa" notation designate spring-, summer-, and fall-run-timing, respectively.

Results

Accuracy testing of 188 SNP baseline

The 40 collections were grouped into 17 reporting groups based on the clustering we observed in the phylogenetic analysis (Figure 2). Results from 100% mixture simulations from each of the baseline collections, showed a majority (38 of 40; 95%) of the simulated mixtures were estimated to be composed of greater than 90% proportion of the correct reporting group (Figure 3b). Two collections, Wenaha spring-run and Clearwater fall-run, showed less than 90% proportions, but were not significantly below 90% (as indicated by their upper 95% confidence interval). These collections may be affected by out-of-basin hatchery stock sources (e.g. Wenaha may be influenced by use of Carson stock in the Rapid River source hatchery).

Results from the leave-one-out analysis (Figure 3a), showed lower performance of baseline power, where only half of the baseline collections displayed greater than 90% correct individual assignment. Among the baseline collections with the lowest correct assignment were Wenaha spring-run (53%), Clearwater R. fall-run (61%), and Lower Deschutes R. fall (67%).

Based on CWT source hatchery data, we calculated proportions of each reporting group represented among the 757 CWT recovered from hatchery-origin fish in the spring/summer harvest, and the 310 CWT recovered from hatchery-origin fish in the fall harvest. The following ten reporting groups were represented in the subset of spring/summer CWT salmon (listed in descending order): Rapid_R/Clearwater_sp (28%), Willamette_sp (20%), Deschutes_R_sp (19%), Upper_Columbia_R_sp (13%), Yakima_sp (9%), SF_Salmon_sp (4%), W_Cascade_sp (3%), Upper_Salmon_sp (2%), Upper_Columbia_R_su/fa (1%), and Klickitat_sp (<1%) (Figure 4). The GSI analysis of the spring CWT mixture produced similar proportions for the following reporting groups (95% confidence interval is reported as a range that follows point estimate): Rapid_R/Clearwater_sp (17%, 21-34%), Willamette_sp (20%, 17-23%), Yakima_sp (9%,6-11%), W_Cascade_sp (3%,1-4%), Upper_Salmon_sp (2%,1-4%), Upper_Columbia_R_su/fa (1%,0-2%), and Klickitat_sp (1%,0-1%). However, the following three reporting groups were shown to have proportions based on CWT data that fell outside the 95% confidence intervals estimated by GSI: Deschutes_R_sp (14%,11-16%), Upper_Columbia_R_sp (27%,14-23%), SF_Salmon_sp (1%,0-2%). Further, one reporting group not represented by CWT data, John_Day_sp, but received 5% (2-8% C.I.) in the GSI estimates.

The following five reporting groups were represented in the subset of CWTs from hatchery-origin fall Chinook salmon (listed in descending order): Upper_Columbia_R_su/fa (48%), Snake_R_fa (37%), W_Cascade_fa (8%), Spring_Cr_Group_Tule (7%), and Deschutes_R_fa (<1%). None of the GSI proportions were significantly different from CWT proportions: Upper_Columbia_R_su/fa (39%, 31-72%), Snake_R_fa (21%, 4-38%), W_Cascade_fa (7%,4-10%), Spring_Cr_Group_Tule (11%, 7-15%), and Deschutes_R_fa (21%, 1-29%). However, the GSI point estimates were sometimes very different from the known CWT proportions, and the confidence intervals were large especially for Upper_Columbia_R_su/fa, Snake_R_fa, and Deschutes_R_fa.

Comparing GSI individual assignments of salmon with known hatchery origins based on CWT showed low accuracy in a group of reporting groups that were similar to those discussed above in the stock proportional estimates. For the fourteen total reporting groups represented among the 1067 CWT salmon, there were eight reporting groups that achieved greater than 75% concordance with the GSI stock-of-origin assignments (Table 3). The following six reporting groups, listed from least to most concordant, were found to be below a 75% concordance rate:

SF_Salmon_sp (17.2%), Upper_Salmon_sp (30.8%), Snake_R_fa (31.9%), RapidR_Clearwater_sp (48.3%), Deschutes_R_sp (65.8%), and Upper_Columbia_R_su/fa (70.5%).

The PBT analysis resulted in 162 hatchery-origin salmon that could be assigned back to Snake River hatchery broodstock parents that were spawned in 2008 (Table 3). These salmon were found to have originated from the RapidR_Clearwater_sp, SF_Salmon_sp, and Upper_Salmon_sp reporting groups, and GSI assignment concordance with the PBT data was below 75% and very similar to the CWT results for these particular reporting groups.

Although not part of our main objectives, we also compared the accuracy of CWT by validating with our PBT assignment results. In total, of the 162 PBT assignments and 1067 CWT (757 were recovered in the spring Chinook salmon harvests), only 18 salmon (11% of PBT tags) were found to overlap across these tagging datasets. The three reporting groups represented by these 18 PBT-assigned salmon were Rapid_R/Clearwater_sp (n=14), SF_Salmon_sp (n=1), and Upper_Salmon_sp (n=3), and the percent concordance values with CWT data for these reporting groups were 71%, 100%, and 33%, respectively. It is alarming why this source hatchery information is not 100% concordant for these methods, and we believe that errors in CWT reading and errors in filing the wrong snout bag with fish ID's is most likely the reason for these discrepancies (A few CWT source hatcheries for these 18 salmon were even outside the Snake River basin). In contrast, the PBT assignments have high confidence, and hatchery spawn records confirm most of the identified parent pairs of these assigned salmon were in fact spawned in the hatchery.

Comparing these various power metrics side by side with the two sets of marker panels (188 versus 92) showed overall improvement for many of the reporting groups (Table 3). The largest average improvement in accuracy was found in CWT concordance, which increased by an average of 14% with the 188 versus 92 SNPs. The Leave-one-out test showed a similarly high average improvement of 12% correct assignment, whereas average improvement measured by 100% simulations and PBT concordance were much smaller, 2% and 4%, respectively. Reasons for the small improvement shown by the 100% simulations appears largely due to diminishing returns, as both the 188 and 92 SNP datasets achieved high accuracy (above 90% for most reporting groups). The PBT concordance could only compare with hatchery-origin fish from Snake River reporting groups, and hatcheries from these reporting groups appear to be a challenge to accurately assign using GSI methods. Since the 100% simulations do not appear to be a very conservative test of power, we are more inclined to use the Leave-one-analysis to determine reporting group accuracy. If we were to set a threshold such that 75% average correct assignment (based on leave-one-out) must be met for each reporting group, then fourteen of the seventeen total reporting groups would qualify using 188 SNPs, but only ten would qualify using 92 SNPs (Table 3).

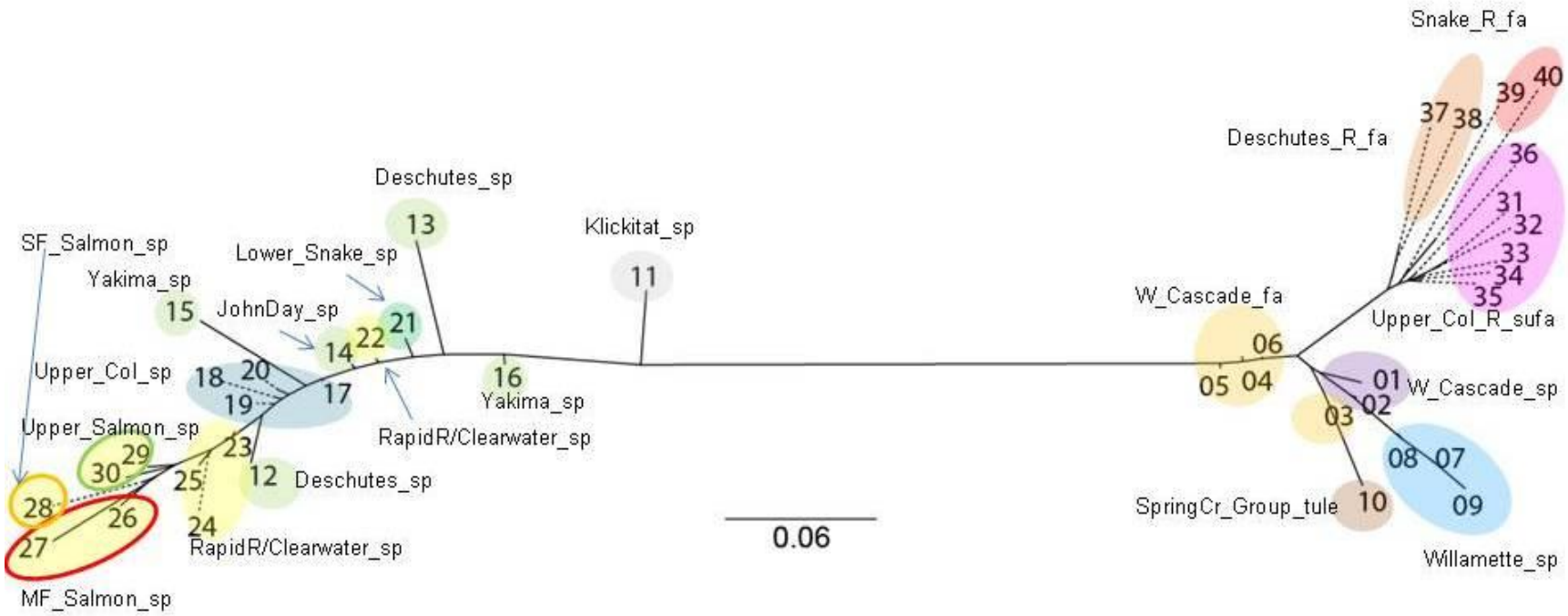


Figure 2. Neighbor-joining tree of Chinook salmon baseline populations using Cavalli-Sforza and Edwards (1967) chord distance of 188 SNP loci. The clusters are labeled with names of reporting groups used to aggregate the collections based on a combination of factors including genetic similarity, life history, and geographic proximity. We used a final set of 17 reporting groups for all GSI analyses (Table 2).

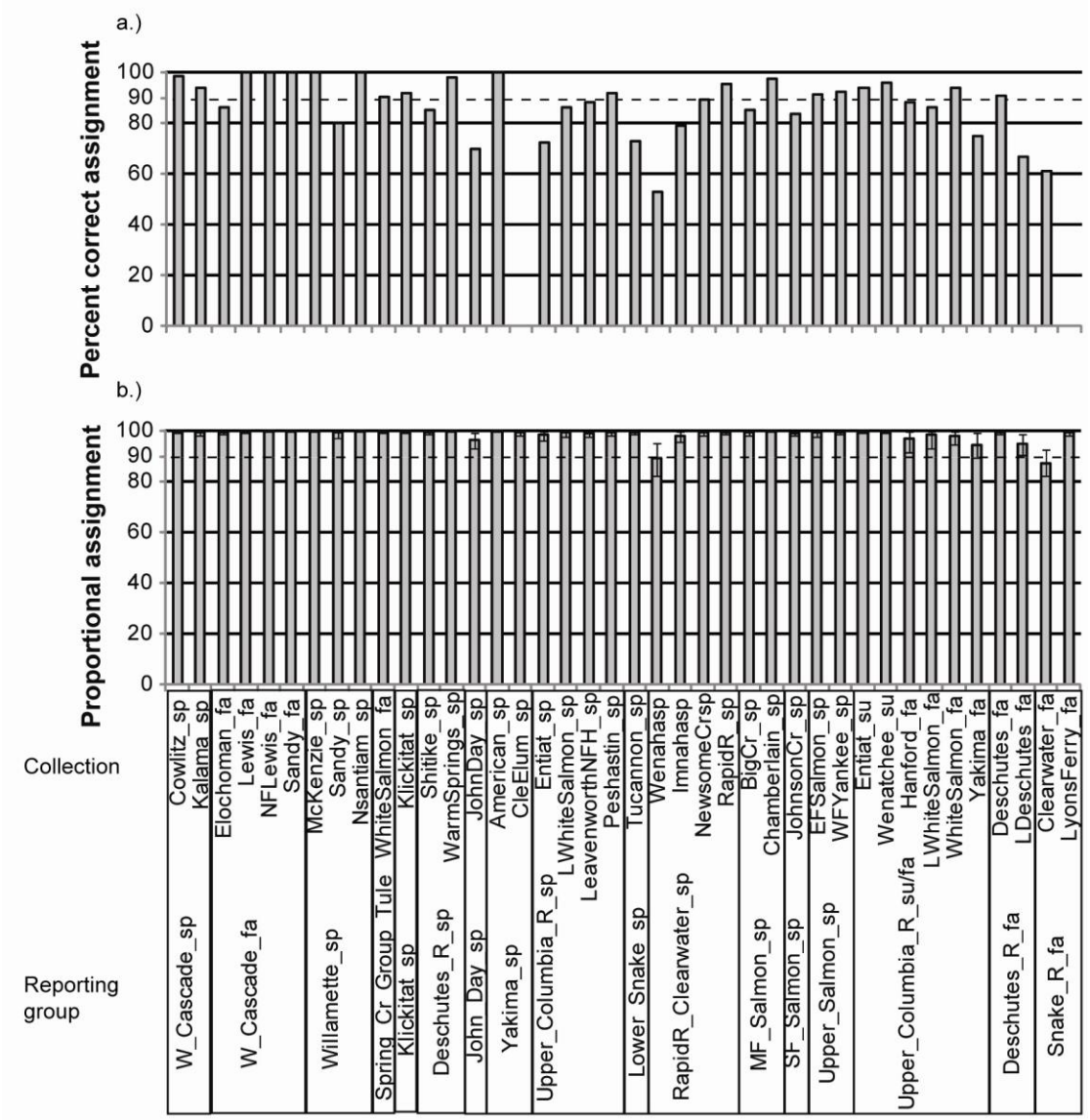


Figure 3. Accuracy of the Chinook salmon GSI baseline. Results are based on a.) correct assignment of baseline individuals back to reporting-group-of-origin from leave-one-out analysis, and b.) estimated reporting-group proportions from 100% mixture simulations using ONCOR. Estimated mixture proportions include 95% confidence intervals. Both charts show a dashed line at the 90% level, considered to be a guidance threshold for determining whether accuracy of reporting groups is sufficient for fisheries management applications. The Cle Elum and Lyons Ferry collections had too much missing genotypic data to calculate a leave-one-out correct assignment value in ONCOR.

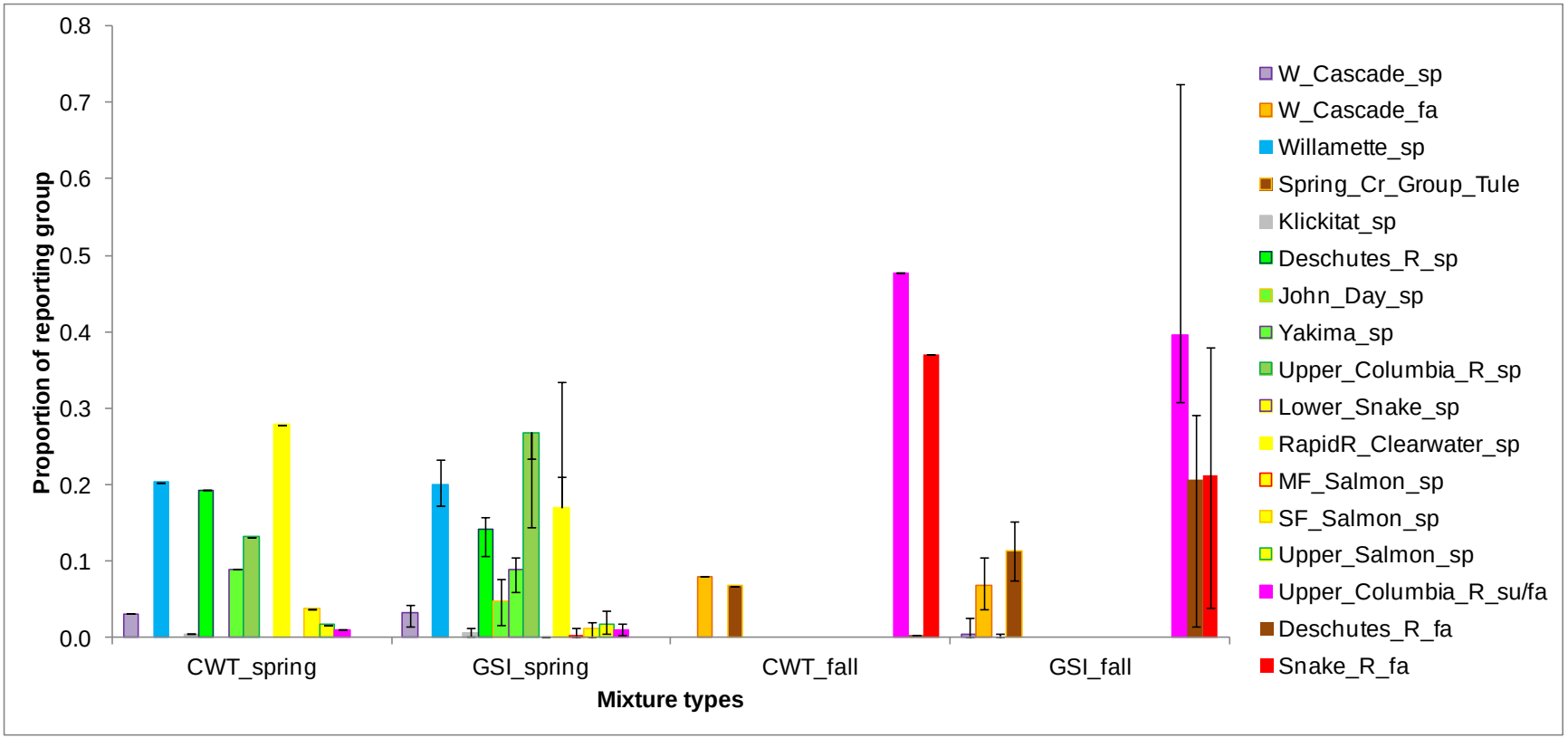


Figure 4. Comparison of expected stock proportions based on a mixture of hatchery-origin Chinook salmon with coded wire tags (CWTs) and estimated proportions of the same mixtures with genetic stock identification (GSI). The spring mixture was constructed by pooling all CWT hatchery-origin Chinook salmon harvested between Feb. 21st-June 16th in the sport, commercial and test fisheries (n=757). The fall mixture was constructed by pooling all CWT hatchery-origin Chinook salmon harvested between Aug. 5th-October 20th in the commercial fishery (n=310). The GSI proportions indicate 95% confidence intervals based on 1000 bootstraps.

Table 3. Comparison of GSI accuracy using 188 versus 92 SNPs.

	Leave-1-out		100%Sims			CWT concordance			PBT concordance	
Reporting Groups	188SNP	92SNP	188SNP	92SNP	N	188SNP	92SNP	N	188SNP	92SNP
W_Cascade_sp	96.2%	81.4%	99.8%	98.6%	25	88.0%	76.0%	0	-	-
W_Cascade_fa	96.5%	85.5%	99.9%	99.6%	24	83.3%	62.5%	0	-	-
Willamette_sp	93.3%	90.1%	99.8%	99.4%	154	94.8%	87.7%	0	-	-
Spring_Cr_Group_Tule	90.3%	90.7%	100.0%	99.9%	21	90.5%	61.9%	0	-	-
Klickitat_sp	92.1%	75.6%	99.9%	98.7%	4	100.0%	100.0%	0	-	-
Deschutes_R_sp	91.9%	84.2%	99.9%	99.4%	146	65.8%	38.4%	0	-	-
John_Day_sp	70.0%	52.6%	96.6%	89.9%	0	-	-	0	-	-
Yakima_sp	100.0%	80.6%	99.7%	98.9%	68	89.7%	66.2%	0	-	-
Upper_Columbia_R_sp	84.8%	66.7%	99.2%	96.5%	100	88.0%	75.0%	0	-	-
Lower_Snake_sp	73.0%	69.8%	99.6%	97.4%	0	-	-	0	-	-
RapidR_Clearwater_sp	79.2%	67.0%	96.7%	92.8%	211	48.3%	46.0%	105	50.5%	47.6%
SF_Salmon_sp	84.0%	51.2%	99.3%	96.4%	29	17.2%	0.0%	24	12.5%	0.0%
MF_Salmon_sp	91.5%	79.7%	99.7%	99.1%	0	-	-	0	-	-
Upper_Salmon_sp	91.9%	77.1%	99.6%	98.6%	13	30.8%	38.5%	33	36.4%	39.4%
Upper_Columbia_R_su/fa	89.0%	79.7%	98.1%	96.6%	156	70.5%	46.8%	0	-	-
Deschutes_R_fa	78.7%	67.5%	97.4%	95.2%	1	100.0%	100.0%	0	-	-
Snake_R_fa	61.3%	57.8%	93.7%	91.7%	116	31.9%	0.9%	0	-	-
Average	86.1%	73.9%	98.8%	97.0%		71.3%	57.1%		33.1%	29.0%
Improvement	12.1%		1.8%			14.2%			4.1%	

Note: Accuracy is based on results from leave-1-out and 100% mixture simulations performed with the program ONCOR. "CWT concordance" refers to correct assignment to reporting group of all CWTs recovered in the 2011 Chinook salmon harvest, and "PBT concordance" refers to correct assignment of all PBT-assigned Snake River hatchery fish from the 2011 spring Chinook salmon harvest. Numbers of total fish with CWTs or assigned with PBT analysis are shown. "Improvement" refers to the average percent increase in accuracy gained by the 188-SNP- compared to the 92-SNP-dataset. Values shaded with green, blue, orange, pink, and red correspond with ranges >90%, 80-90%, 70-80%, 60-70%, and below 60%, respectively.

Stock proportions of the spring Chinook salmon fishery sources

Comparisons of stock composition across the different fishery sources generally produced results consistent with our analyses from last year using the 2010 Chinook salmon harvest. For the spring-run Chinook salmon harvest, we continue to observe the reporting group representing the Willamette River spring-run stock with large changes in proportion that appear associated with fishery type, and associated both spatially and temporally. This stock was present in all the fisheries below Bonneville Dam (Regions A and B, Figure 5), but absent at both Bonneville Dam (Section 4) and the Ceremonial fishery in Zone 6 (not shown). This Willamette River stock shows consistently high proportions above 20% in the sport fishery in Region B, and shows a dramatic decrease in proportion in the sport fishery that occurs upstream in the same weeks in Region A (Figure 5). Further, a temporal shift in proportions of the Willamette River stock appear in both the sport and test fisheries, as the proportions are higher in earlier weeks versus latter. These spatial and temporal patterns in the Willamette stock proportions are surprisingly absent in the commercial fishery which observed high proportions of this stock throughout the spring and summer season and across regions (average 49% and 30% in region A and B, respectively). The reason for the lack of these spatial and temporal patterns in the commercial fishery may be explained by a much more concentrated fishing effort focused at the boundary delineating these two regions (the mouth of the Willamette River). This concentration of effort would result in higher yield of the Willamette stock, and is in contrast to a more even spatial representation of each region that is characteristic of the sport fishery.

The test fishery, which unlike the mark selective commercial and sport fisheries, samples both hatchery-marked and unmarked wild fish. However, this difference does not explain why the Willamette River stock is found in lower proportions (average 21%) in the test versus the commercial (30%) and sport (32%) fisheries in region B, because Willamette River stock was not underrepresented by the unmarked fish caught in the test fishery. In fact, of the three major stocks (RapidR_Clearwater_sp, Upper_Columbia_R_sp, and Willamette_sp) caught in the spring fishery, the Willamette River stock had the highest percent unmarked fish (26%) in the test fishery. The reason for the lower proportion is again due to subtle differences in concentration of fishing effort within these regions, as the test fishery was conducted slightly more downstream away from the mouth of the Willamette River than the commercial fishery in region B.

For the commercial, sport, and test fishery the other two major stocks, RapidR_Clearwater_sp and Upper_Columbia_R_sp were in the following proportions: sport fishery region B (33%, 27%) and region A (43%, 45%), commercial fishery region B (24%, 35%) and region A (13%, 20%), and test fishery region B (29%, 36%).

Not surprisingly, Chinook salmon that passed above Bonneville Dam during week 18 (Section 4) and were harvested in the ceremonial fishery were very similar to the total adult fish that were non-lethally sampled at the dam in terms of stock composition. The two reporting groups, Rapid R./Clearwater R. and Upper Columbia R, comprised on average 42% and 48% of the ceremonial harvest, respectively.

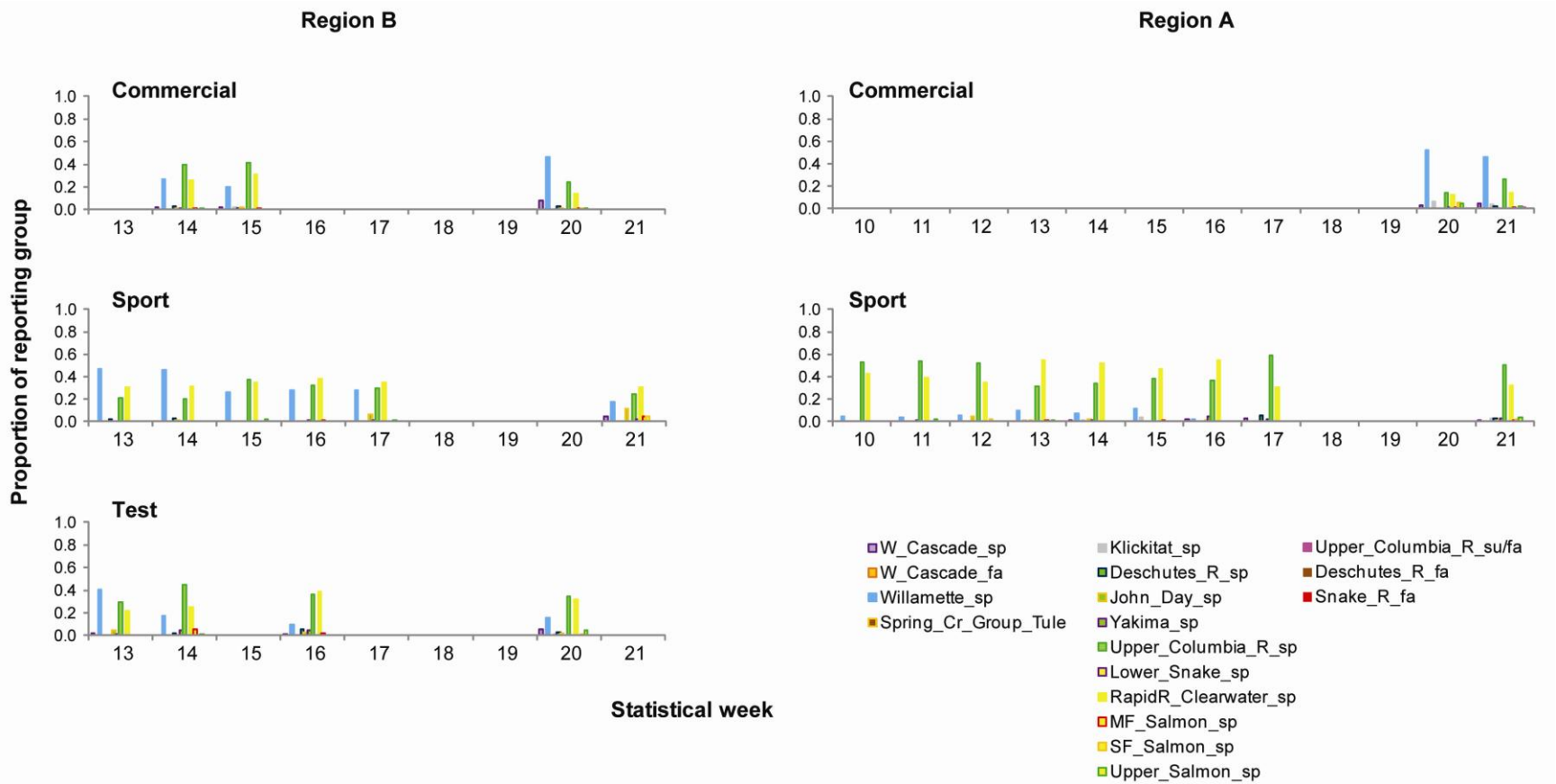


Figure 5. Comparison of stock proportions across all spring-run Chinook salmon fishery sources below Bonneville Dam. Included are commercial, sport, and test fisheries in regions B (mouth of the Columbia River to Willamette River) and A (Willamette River to Bonneville Dam). Stock proportions are indicated by different colored bars that correspond with 17 reporting groups as shown.

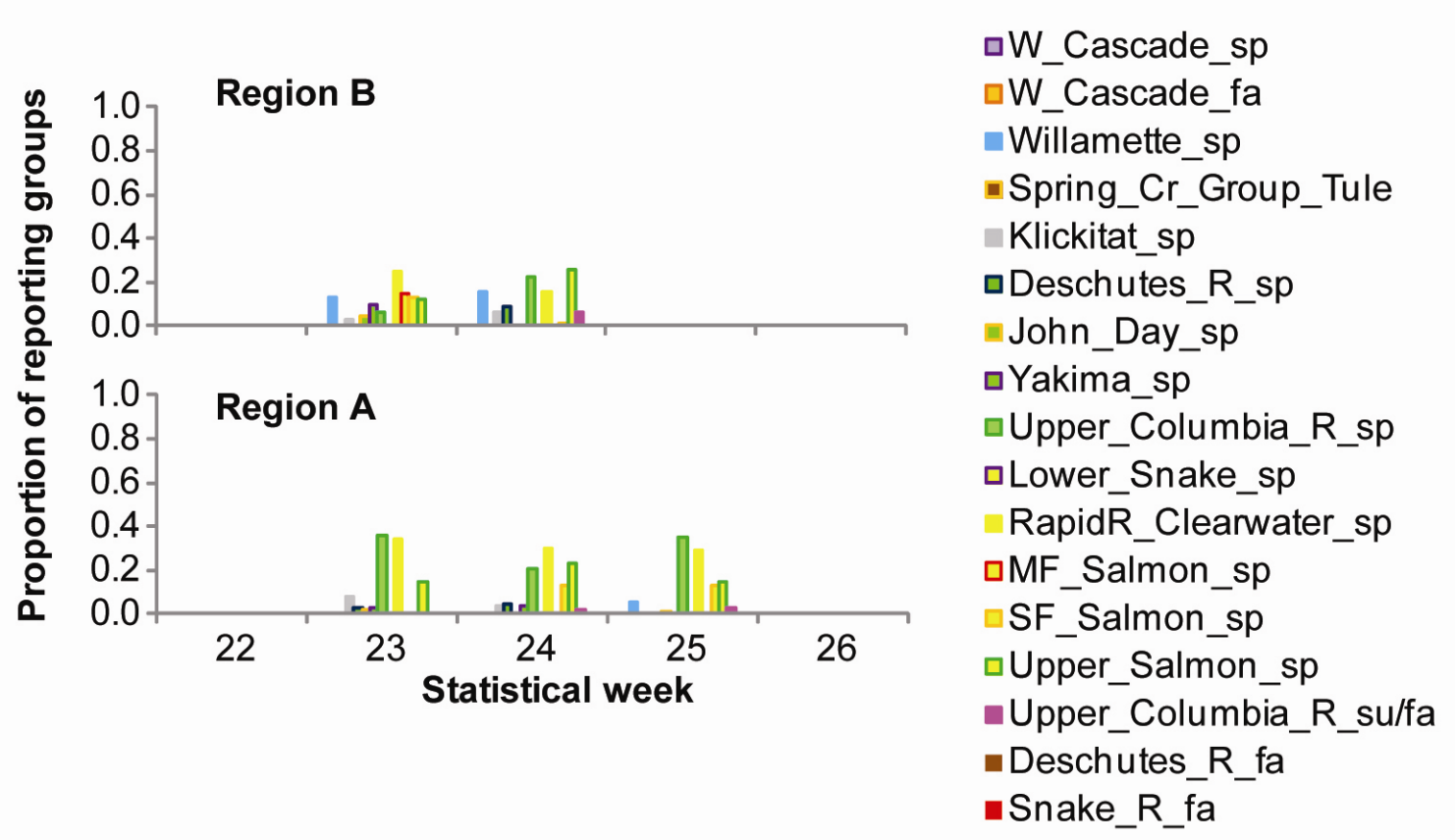


Figure 6. Comparison of region B versus A for the summer Chinook sport fishery.

Stock proportions of the summer Chinook salmon sport fishery

Due to the delayed run-timing of Chinook salmon in 2011, the sport fishing season was extended longer than usual and there were summer weeks (statistical weeks 22 to 26) in which we could estimate stock proportions of this fishery in both region A and B. The most notable difference in stock proportions was an increase in the Salmon River stocks, especially the Upper_Salmon_sp. This stock showed average proportions of 18% and 19% in regions A and B, respectively, during summer weeks versus less than 1% in these regions during all earlier weeks. These changes were expected according to Bonneville run-timing results that show June 11th is the median of the run distribution for the Upper_Salmon_sp reporting group (Section 4).

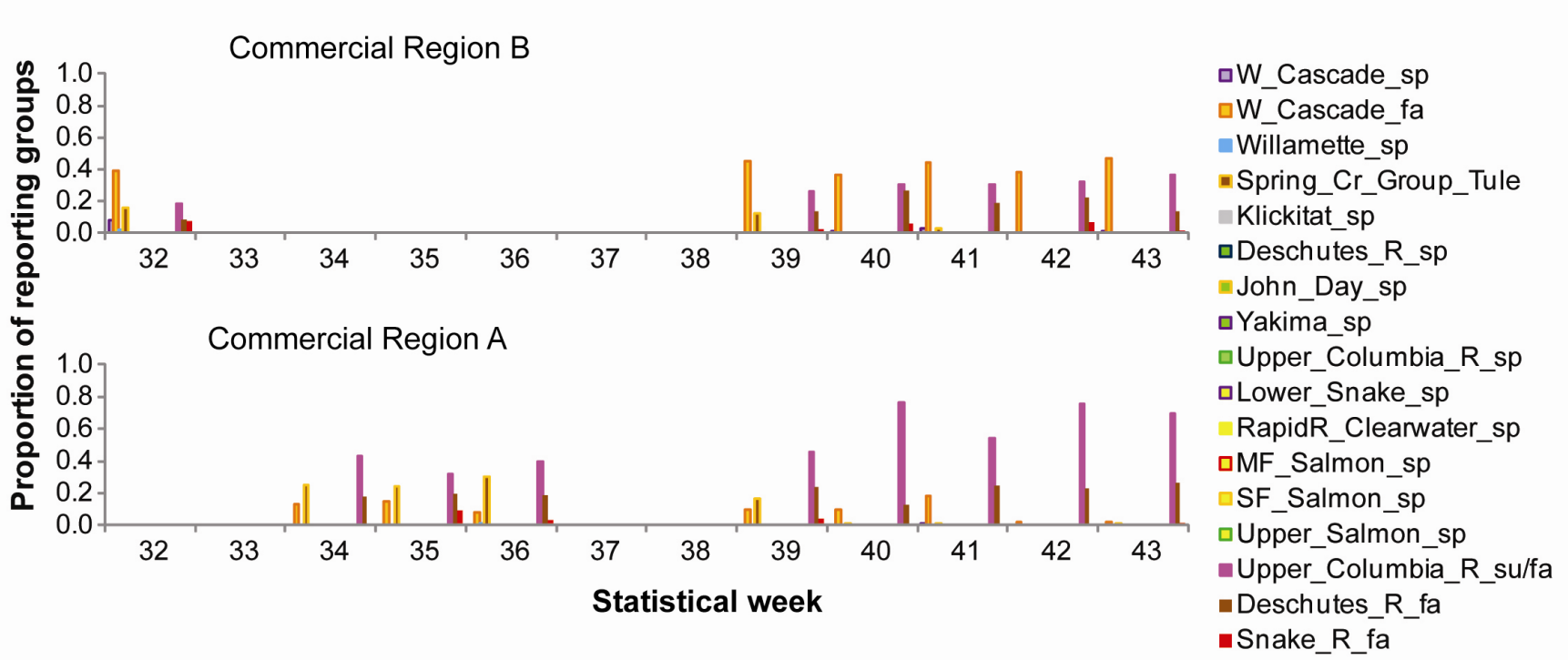


Figure 7. Comparison of regional effect on stock proportions in the commercial fall Chinook salmon fishery below Bonneville Dam.

Stock proportions of the fall Chinook salmon fishery

The 2011 fall harvest dataset, which includes both the commercial and tribal harvest, had greater representation of the commercial harvest below Bonneville Dam in order to maximize the number of CWT recoveries that were included in analyses for use in testing accuracy of the GSI baseline. Similar to our previous findings from 2010, proportions of the West Cascade fall-run and Spring Creek Group Tule stocks showed large variation across weeks and regions in the fall-run Chinook salmon commercial and tribal fishery mixtures (Figure 7). For example, in the same range of weeks (39-43) in the commercial harvest the West Cascade fall-run was estimated to be 42% proportion in region B, but only 8% proportion in regions A. The Spring Creek Group Tule stock was found in greater proportions earlier in the fall. For example, in early weeks (32-39), this stock made up 14% and 24% of regions B and A, respectively, compared to later weeks (40-43), in which both regions decreased to 1% average proportion of this stock.

In general, the following five reporting groups are represented in these fall fisheries: West Cascade fall-run, Spring Creek Group Tule, Deschutes River fall, Snake River Fall, and Upper Columbia River summer/fall.

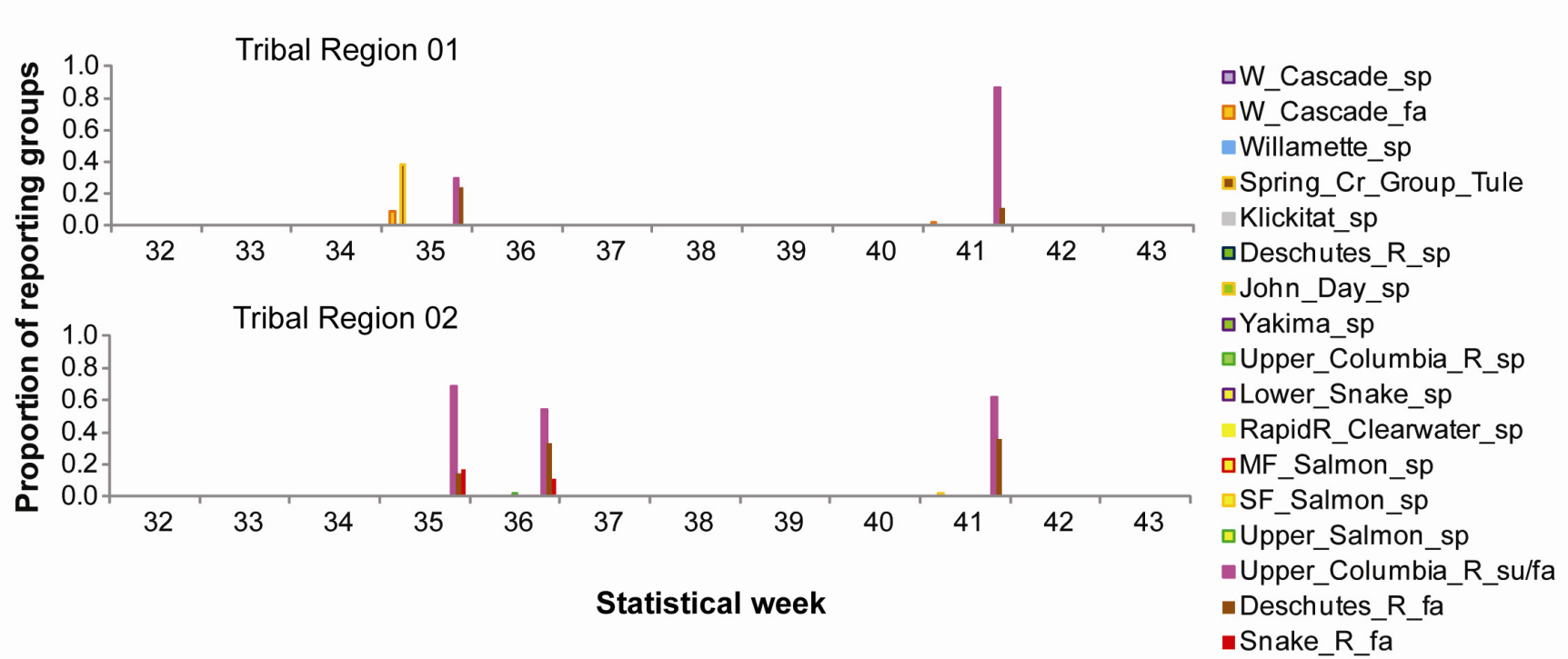


Figure 8. Comparison of regional effect on stock proportions in the tribal fall Chinook salmon fishery in zone 6.

In the tribal fall Chinook salmon fishery, the Spring Creek Group Tule stock can only be found in Region 01, and similar to regions below Bonneville Dam, this stock is no longer present by the end of the season (Figure 8). Since the fish that assign to the Spring Creek Group Tule stock and pass Bonneville Dam are most likely part of the Spring Creek National Fish Hatchery, their final destination is a tributary that empties into Region 01 of the Zone 6 area. In Region 02, an appearance of Snake River fall stock appears in the early weeks and is missing in the last week that we sampled. This result is consistent with our analysis in 2010, which showed that the Snake River has an earlier run-timing compared to the Deschutes River fall stock. This result is also consistent with the order of the median dates of the run-timing distributions of these stocks at Bonneville Dam (Section 4).

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 and our results show large improvements in accuracy of stock assignment. We are now able to discriminate 17 reporting groups without sacrificing levels of accuracy that we had achieved with our previous baseline that utilized a smaller number of SNP markers. When we compare our previous set of 92 SNP markers to our current set of 188 SNPs we observe an increase between 2%-14% based on various methods for measuring baseline accuracy.

We also observed consistent patterns in stock proportions of Chinook salmon spring and fall harvest in the lower mainstem of the Columbia River. The spring-run Chinook salmon harvested in commercial, sport, and test fisheries were primarily composed of three adipose-clipped stocks: Rapid River Hatchery/Clearwater R., Upper Columbia R. (i.e., Carson stock), and Williamette R. At least in the sport fishery, the Willamette River stock appears in highest abundance during the earlier part of the season and locations closer to the mouth of the Columbia R. For fall-run Chinook salmon fisheries, the commercial fishery below Bonneville Dam contained large proportions of West Cascade fall-run and Spring Creek Group Tule stocks, as well as the following stocks (in descending order): upper Columbia R. summer/fall, Deschutes R. fall, and Snake R. fall. The entire Zone 6 tribal Chinook fishery was heavily comprised of Upper Columbia R. summer/fall stock, but Region 01 (closest region to Bonneville Dam) of Zone 6 was the only location where the Spring Creek Group Tule stock can be found. The Snake R. stock continues to exhibit an early peak in the fall, based on comparisons of early and late weekly strata in Region 02 of Zone 6.

Future directions

Accuracy testing using known origin hatchery-Chinook salmon based on CWT and PBT information revealed some weaknesses in the GSI baseline that will need to be improved in the future. For spring Chinook salmon, misassignment between Upper Columbia R., Rapid River, and Salmon River spring-run stocks continues to represent an analytical challenge, and for fall Chinook salmon, where we experience difficulty distinguishing among Upper Columbia R. summer/fall, Deschutes R fall and Snake R. fall-run stocks. These misassignment issues likely are the main reasons for observing low concordance between GSI and CWT/PBT data and are primarily attributed to historical interbreeding of these stocks by hatchery programs. In addition, to this misassignment we are working with a relatively low number of collections (40) to represent this relatively large set of reporting groups (17). One of the factors that could contribute to the apparent low resolving power of the GSI analyses for these troublesome reporting groups may be due to the fact that we are using primarily hatchery origin fish to test the accuracy of a baseline that is in some cases composed entirely of natural origin collections that represent the reporting groups. An important test to conduct in the future would be to use mixtures of both natural-origin and hatchery-origin fish for accuracy testing of this GSI baseline. We expect that GSI is much more accurate for natural-origin fish than hatchery-origin fish due to stock transfers between hatcheries that have led to a mixed genetic signal. Further, we will continue to add collections to the baseline and this expansion will improve representation of each reporting group.

Even if these plans to expand and improve the GSI baseline achieve minimal improvement in assignment accuracy for hatchery-origin fish, we continue to integrate PBT into our analyses and this strategy will resolve accuracy issues related to all Snake River hatchery-origin Chinook salmon. In the future, stock composition analyses will start by assigning hatchery-origin fish with PBT analyses, and then GSI will be used for natural-origin fish and any hatchery-origin fish that cannot be assigned by PBT.

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.

Anderson E. C. 2010. Computational algorithms and user-friendly software for parentage-based tagging of Pacific salmonids. Final report submitted to the Pacific Salmon Commission's Chinook Technical Committee (US Section). 46 p. http://swfsc.noaa.gov/textblock.aspx?Division=FED&ParentMenuId=54&id=16021.

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.

Hess, J. E., A. P. Matala, and S. R. Narum. 2011. Comparison of SNP and microsatellite markers for application of genetic stock identification for Chinook salmon in the Columbia River Basin. Molecular Ecology Resources 11 (Suppl. 1):1–13.

Hess, J. E. and S. R. Narum. 2011. SNP loci correlated with run-timing in adult Chinook salmon from the Columbia River Basin. Transactions of the American Fisheries Society 140(3):855-864

Narum, S. R., J.E. Hess, and A.P. Matala. 2010 Examining genetic lineages of Chinook salmon in the Columbia River Basin. Transactions of the American Fisheries Society 139:1465-1477.

Pritchard**,** J.K., Stephens**,** M., Donnelly**,** P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155(2): 945–959.

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.

Steele CA, Campbell MR, Ackerman M, McCane J, Hess MA, Campbell N, Narum SR. 2011. Parentage Based Tagging of Snake River hatchery steelhead and Chinook salmon. Bonneville Power Administration. Annual Progress Report, Project number 2010-031-00. https://research.idfg.idaho.gov/Fisheries%20Research%20Reports/Res11-111Steele2010%20Parentage%20Based%20Tagging%20Snake%20River%20Steelhead%20Salmon.pdf

Waples**,** R.S., Teel**,** D.J., Myers**,** J.M., Marshall**,** A.R., 2004. Life-history divergence in Chinook salmon: historical contingency and parallel evolution. Evolution, 58(2): 386–403.

Section 4: Characterization of Chinook salmon and steelhead run-timing and abundance at Bonneville Dam

Introduction

The Columbia River Basin supports ESA listed wild stocks of Chinook salmon and steelhead as well as hatchery supplemented populations. Both Chinook salmon and steelhead have been declining in the Columbia River Basin for several reasons including climate change, habitat degradation, hydropower, hatchery practices, and over-harvesting. 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.

As evident from the genetic stock identification (GSI) analyses of Chinook salmon fisheries harvests in Section 3, certain stocks seem to have strong spatial and temporal associations. However, because the type of fishery gear, harvest regulations, and locations targeted varies considerably among fisheries, it is necessary to conduct a study that samples a representative mixture of all hatchery- and natural-origin stocks at a fixed location to accurately estimate abundance and characterize run-timing distributions of stocks. In addition to information on abundance and run-timing, biological data including fork length and age can be examined with estimated stock of origin in order to characterize life history differences among stocks. This type of examination is especially important for steelhead, which has been managed using two life-history categories, A- and B-run. These life-history categories are defined by run-timing at Bonneville Dam (e.g. B-run arrives after Aug. 25th), fork length (e.g. B-run is greater than 78 cm), and ocean age (e.g. B-run spends 2 or more years in saltwater) and all of these types of data have been collected for steelhead in this study.

Here we analyze fish across the entire run of Chinook salmon and steelhead from April to October to estimate temporally stratified proportions of stocks and extrapolate abundance using a daily census that is conducted at the Bonneville Dam fish counting window. We examine Chinook salmon and steelhead using two sets of species-specific SNP assays for a combined total of 192 loci per set. Although there are slight methodological differences between these two species-specific applications (e.g. different temporal strata), the general approach to estimating abundance and characterizing run-timing distributions was applied consistently across species. In addition, similar to the way in which we demonstrated the power of the Chinook salmon 192 SNP baseline for GSI applications (Section 3), we include a demonstration of the power of the steelhead baseline. In a recent study by Narum et al. 2010 (Section 4), microsatellite genotypes were used to apply GSI analyses and assign unknown steelhead to the population they originated. This current report represents the first analysis using 192 SNP loci (n=192) for GSI analysis of Bonneville Dam steelhead mixtures, and we expect these SNP markers to have comparable resolution of the highly structured steelhead populations that have previously been characterized by microsatellites (Narum et al. 2006a,b, Nielsen et al. 2009, Blankenship et al. 2011).

In the past two years of this study, we have had to rely entirely on GSI as the genetic tool to characterize mixtures of both hatchery and wild interior Columbia River Chinook salmon and steelhead based on stock proportions and individual assignment to their likely stock (reporting group) of origin. However, the 2011 year is the first year we are able to apply an additional genetic tool, referred to as Parentage Based Tagging (PBT), to assign a portion of Snake River hatchery-origin spring-run Chinook salmon and summer-run steelhead back to their hatchery

parents (Steele et al. 2011). This new and powerful genetic tool provides the opportunity to obtain additional types of data including accurate age of fish, quantification of the number of unmarked (non-adipose clipped) hatchery fish, and precise assignments of fish to their source hatchery. The ability of PBT to identify a fish's source hatchery is a much finer spatial distinction as compared to the stock identity provided by GSI. However, these tools can provide the greatest benefit when applied in combination, as GSI has the ability to provide information on wild fish and hatchery fish that originate from outside the Snake River basin. We include results from these PBT analyses to demonstrate the utility of this new genetic tool as applied to Chinook salmon and steelhead passing Bonneville Dam.

The aim of this study was to use GSI to discriminate Columbia River steelhead and Chinook salmon stocks according to their peak run-timing. Since Bonneville Dam is the most downstream dam on the Columbia River, the fishery 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 abundance of Chinook salmon and steelhead and provide this data to fisheries managers.

Methods

Sample Collection

Tissue samples were obtained from adult Chinook salmon in 2011 (n = 3335) and adult steelhead in 2009, 2010, and 2011 during migration runs at Bonneville Dam (n = 2286, 1671, and 1352, respectively). Based on numbers of fish collected, samples were pooled into weekly strata for Chinook salmon (Table 1) and biweekly strata for steelhead (Table 4) spanning the majority of the run-year from April to October. 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 fish. 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 on Whatman filter paper (Lahood et al. 2008) before being shipped to the Hagerman Fish Culture Experiment Station for DNA extraction and analysis. In addition, scales are collected for estimating the number of years spent in freshwater and saltwater for each fish. 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 weekly strata for Chinook salmon that were DNA sampled or tallied for abundance at Bonneville Dam.

Spring				Fall			
Statistical week	N	Abundance	Sample rate	Statistical week	N	Abundance	Sample rate
17	9	2529	0.4%	31	0	7293	0.0%
18	128	47599	0.3%	32	24	4631	0.5%
19	287	61460	0.5%	33	6	6333	0.1%
20	238	64036	0.4%	34	6	14875	0.0%
21	154	22307	0.7%	35	67	23533	0.3%
22	212	12548	1.7%	36	96	68816	0.1%
23	128	22209	0.6%	37	239	135133	0.2%
24	239	27455	0.9%	38	208	98585	0.2%
25	189	24668	0.8%	39	263	65949	0.4%
26	81	23555	0.3%	40	237	28140	0.8%
27	91	25417	0.4%	41	211	21033	1.0%
28	56	16461	0.3%	42	42	9888	0.4%
29	66	11005	0.6%				
30	58	7956	0.7%				
Grand total					3335	853414	0.4%

Note: For reference, statistical week 17 is 4/18/11-4/24/11 and 42 is 10/10/11-10/16/11. "Abundance" is based on tallies of Chinook salmon adults and jacks provided by the Fish Passage Center (http://www.fpc.org) as observed at their fish counting window. Genetic sample numbers (N) above 20 are highlighted in gray and these weeks were the only ones deemed sufficient to estimate stock abundance. The grand total is a sum that includes both spring and fall sample columns.

Molecular markers

We used both the GSI-96 and PBT-96 SNP panels (Section 1, Appendix 1) for a total of 192 SNP loci to genotype Chinook salmon mixtures, and we removed the same four loci for reasons described in Section 3 which leaves 188 for all GSI analyses. For steelhead, we also used the GSI-96 and PBT-96 SNP panels (Appendix 2, Section 2, Hess et al. 2011) for a total of 192 SNP loci. However, we removed three loci that are used to detect cutthroat hybrids and the sex determination marker resulting in a combined total of 188 SNP loci for GSI applications.

Statistical analyses

Estimation of Chinook salmon stock composition utilized the baseline and 17 reporting groups described in Section 3, Table 2. Estimation of steelhead stock composition utilized the baseline and 17 reporting groups described in Table 3. The program ONCOR v1.0 (available at http://www.montana.edu/kalinowski) was used to estimate stock composition for each temporally stratified Bonneville Dam mixture. We analyzed all strata that had $n > 20$ samples. To estimate total abundance of each stock, we multiplied stock proportions estimated for each time stratum (weekly or biweekly) with the total species abundance tallied at the fish counting window at the Bonneville Dam fish ladder (data available at the Fish Passage Center website: http://www.fpc.org). This method was utilized to minimize any sampling bias due to the uneven sample rate. For example, the sample rate for Chinook salmon in 2011 ranged from 0% to 1.7% (Table 1). To characterize each stock's run-timing distribution, we extrapolated each stock's daily abundance from weekly estimates of stock proportions and multiplied these proportions evenly to daily fish counts within a particular week and calculated the median, inter-quartile, 5^{th}, and 95^{th} percentile of the abundance distributed across ordinal days. For steelhead, each year's samples were first divided into two datasets of hatchery- and natural-origin fish based on absence and presence of an adipose fin, respectively. Analyses were conducted on each of these datasets to estimate stock abundance and run-timing distributions for all hatchery and natural-origin steelhead stocks separately.

Similar to the accuracy testing that was conducted with the Chinook salmon baseline in Section 3, we performed both Leave-one-out assignment tests and 100% mixture simulations for the steelhead baseline using ONCOR. The steelhead reporting groups were constructed based on the same reporting groups used by Ackerman et al. (2011) in the Snake River Basin, and using the reporting groups discussed in Hess et al. (2011) as a guide for the rest of the Columbia River Basin.

Parentage based tagging (PBT) analysis was used to identify the source hatcheries of 3-year old Snake River spring-run Chinook salmon and 4-year old Snake River steelhead. Genotypes of parents and offspring were analyzed with SNPPITv1.0, a software program developed for large PBT datasets based on SNP markers (Anderson 2010). Analyses with SNPPIT were performed separately on the BY2008 Chinook salmon broodstock baseline and BY2008 steelhead broodstock baseline (Steele et al. 2011), to assign any offspring that were included in the 2011 Bonneville Dam mixtures of Chinook salmon and steelhead, respectively. PBT assignments were based on marker exclusion and a false discovery rate (FDR) threshold of 1.5% was used to ensure high confidence in all parent matches. Due to this high confidence in the identity of an individual's hatchery parents, we could also accurately age these fish as 3-year old and 4-year old spawn age for Chinook salmon and steelhead, respectively. We used age and

hatchery-of-origin information inferred by PBT assignments to test the accuracy of GSI assignments, scale-aging methods, and to quantify the percent of unmarked (non-adipose clipped) fish that were hatchery-origin.

Results

Estimated abundance of Chinook salmon stocks in 2011

There were fourteen Chinook salmon stocks passing Bonneville Dam that we estimated abundance greater than 2,000 fish in the season (Table 2). The nine major stocks of the spring-run in order of magnitude were Middle Fork Salmon R. (3,000), Klickitat R. (5,000), South Fork Salmon R. (11,000), upper Salmon R. (15,000), John Day R. (16,000), Deschutes R. (18,000), Yakima R. (24,000), Rapid R./Clearwater R. (96,000), and upper Columbia R. (111,000). The five major stocks of the fall-run in order of magnitude were West Cascade (3,000), Spring Cr. group tule (43,000), Snake R. (33,000), Deschutes R. (101,000), and upper Columbia R. summer/fall (339,000). These stock abundance estimates were based on the stock proportions that were estimated in ONCOR across weekly strata (Figure 1), and were multiplied with the total abundance of Chinook salmon that was tallied on a daily basis at the Bonneville Dam fish counting window. This calculation was used to generate the stock abundance on a weekly basis to visualize peak run-timing (Figure 2) and to characterize the distributions of run-timing for each stock (Figure 3).

Table 2. Basic information on run-timing distributions of Chinook salmon stocks passing Bonneville Dam.

Reporting group	Estimated abundance	Median	1st quartile	3rd quartile	5th percentile	95th percentile	Median date	Interquartile range
W_Cascade_sp	1267	-	-	-	-	-	-	-
W_Cascade_fa	2961	251	248	255	243	261	8-Sep	7
Willamette_sp	290	-	-	-	-	-	-	-
Spring_Cr_Group_Tule	43541	251	247	255	242	264	8-Sep	8
Klickitat_sp	4911	159	149	166	129	196	8-Jun	17
Deschutes_R_sp	17966	129	123	152	118	186	9-May	29
John_Day_sp	16412	132	127	152	119	169	12-May	25
Yakima_sp	23851	133	129	154	122	167	13-May	25
Upper_Columbia_R_sp	110648	130	123	148	118	172	10-May	25
Lower_Snake_sp	1734	-	-	-	-	-	-	-
RapidR_Clearwater_sp	96234	131	123	153	118	179	11-May	30
MF_Salmon_sp	2919	150	130	154	123	162	30-May	24
SF_Salmon_sp	11020	158	134	165	121	179	7-Jun	31
Upper_Salmon_sp	15284	162	149	171	120	183	11-Jun	22
Upper_Columbia_R_su/fa	338823	252	241	261	178	277	9-Sep	20
Deschutes_R_fa	101210	256	249	266	241	280	13-Sep	17
Snake_R_fa	33325	254	246	261	237	276	11-Sep	15

Note:These summary statistics of run-timing distributions were calculated using a method to estimate abundance of each stock based on stock proportions and total numbers of Chinook salmon that were observed passing Bonneville Dam at the fish counting window. This information is also plotted in Figure 3.

Run-timing of Chinook salmon stocks in 2011

We were able to obtain sufficient sample sizes to characterize the run-timing distributions of fourteen Chinook salmon stocks. Variation in stock proportions across weekly strata within the migrating season was relatively large. The largest run-timing distributional differences for the major spring-run stocks (stocks arriving before June 30th;Ordinal day 180), were observed between relatively early peak run-timing stocks Deschutes R., upper Columbia R., Rapid R./Clearwater R., John Day R., and Yakima R. spring-run stocks (median dates May 9th, May 10th, May 11th, May 12th, and May 13th, respectively), versus the relatively late peak run-timing stocks from the Middle Fork, South Fork, and Upper Salmon R. and Klickitat R. (median dates May 30th, June 7th, June 11th and 8th, respectively). The major fall-run stocks showed minimal differences in peak run-timing, however the stocks can be ordered by median date as follows: West Cascade (Sep 8th), Spring Cr. group tule (Sep 8th),Upper Columbia R. summer/fall (Sep 9th), Snake R. fall (Sep 11th), and Deschutes R. fall (Sep. 13th).

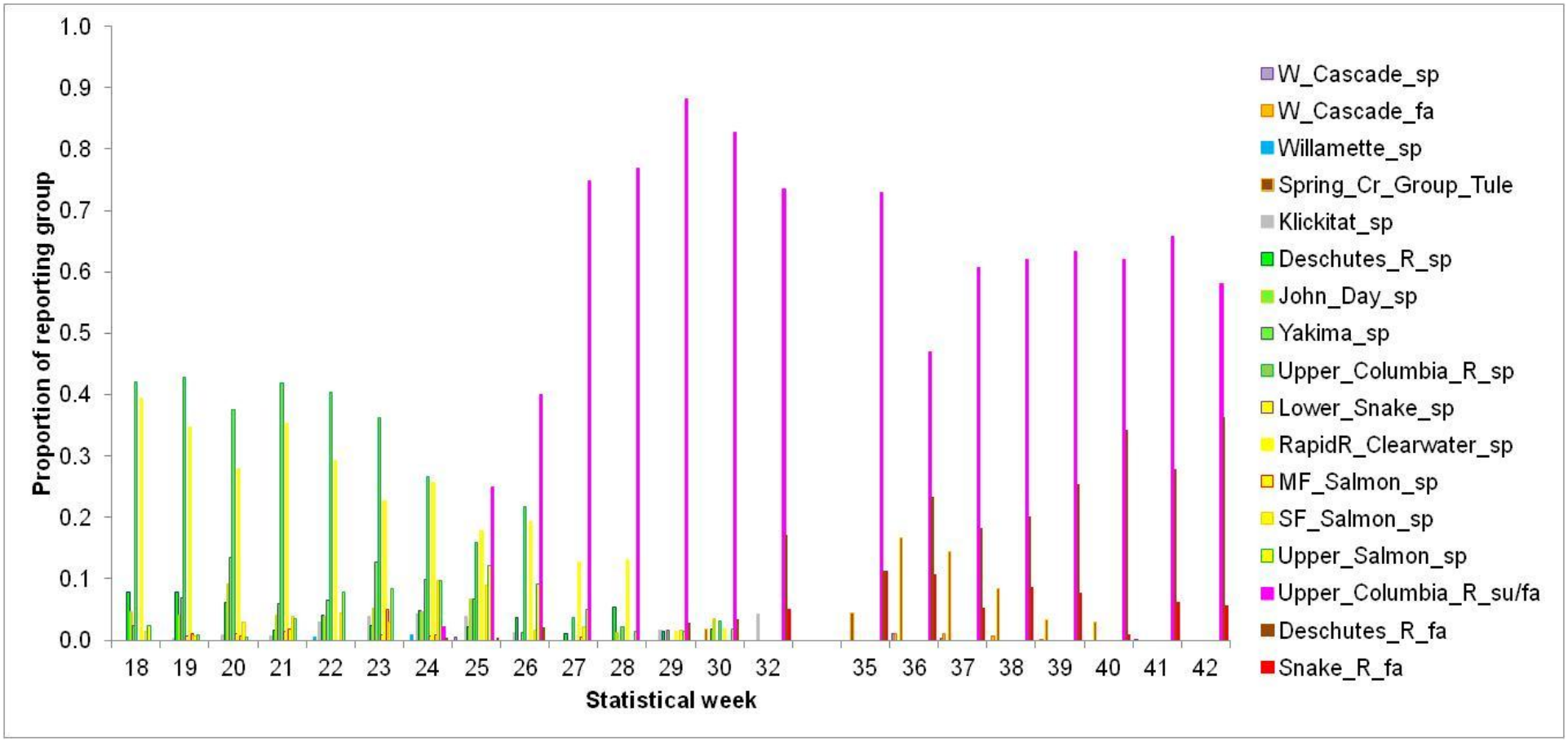


Figure 1. Estimated weekly proportions of reporting groups for Chinook salmon passing Bonneville Dam in 2011. Estimated reporting group proportions for weeks 31, 33, and 34 were unavailable due to insufficient sample numbers of Chinook salmon.

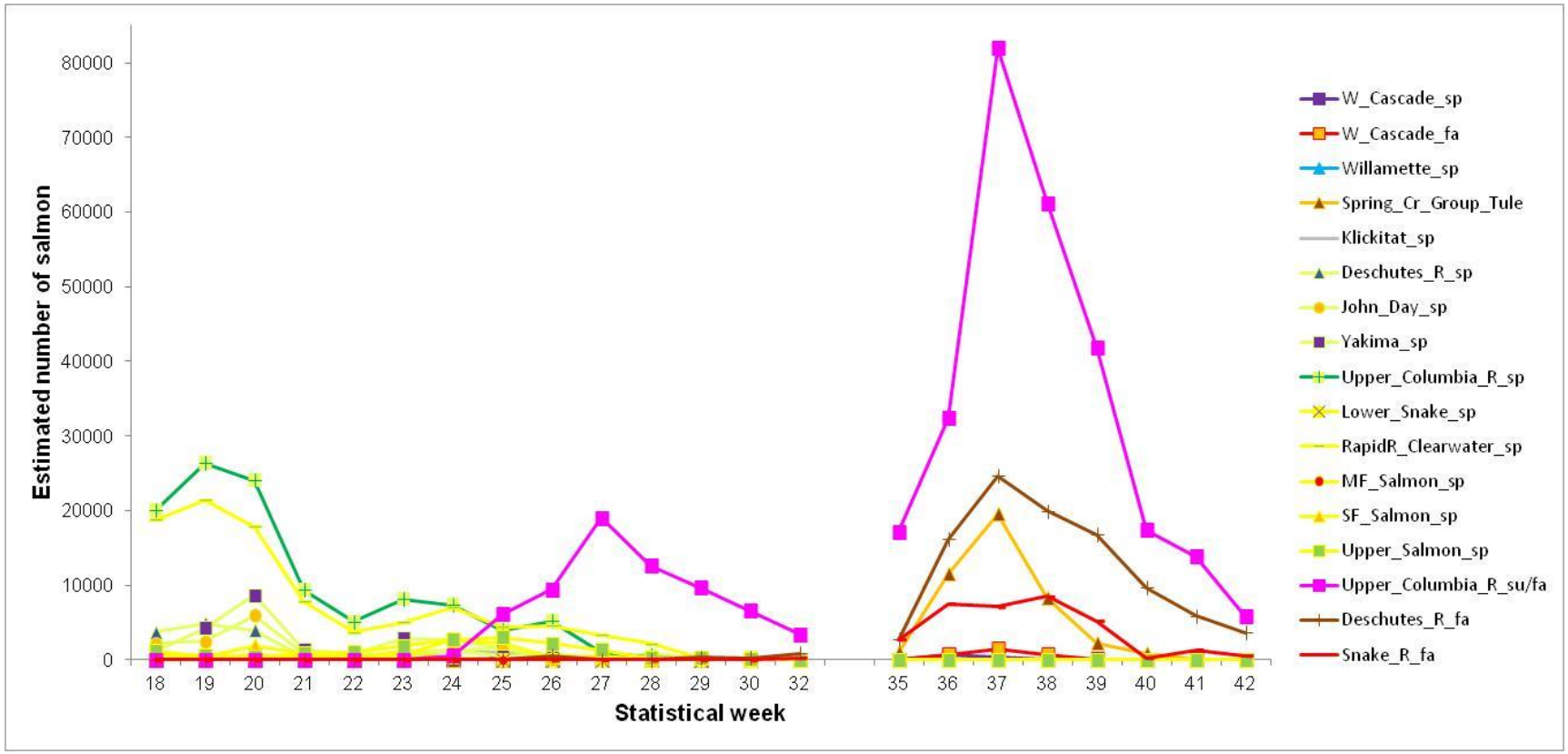


Figure 2. Estimated weekly abundance of Chinook salmon reporting groups passing Bonneville Dam in 2011. This data was generated by first estimating reporting group proportions of weekly pooled mixtures of Chinook salmon passing Bonneville Dam and then multiplying the proportions with weekly tallies of Chinook salmon at the Bonneville Dam fish counting window. Estimated abundance for weeks 31, 33, and 34 were unavailable due to insufficient sample numbers of Chinook salmon.

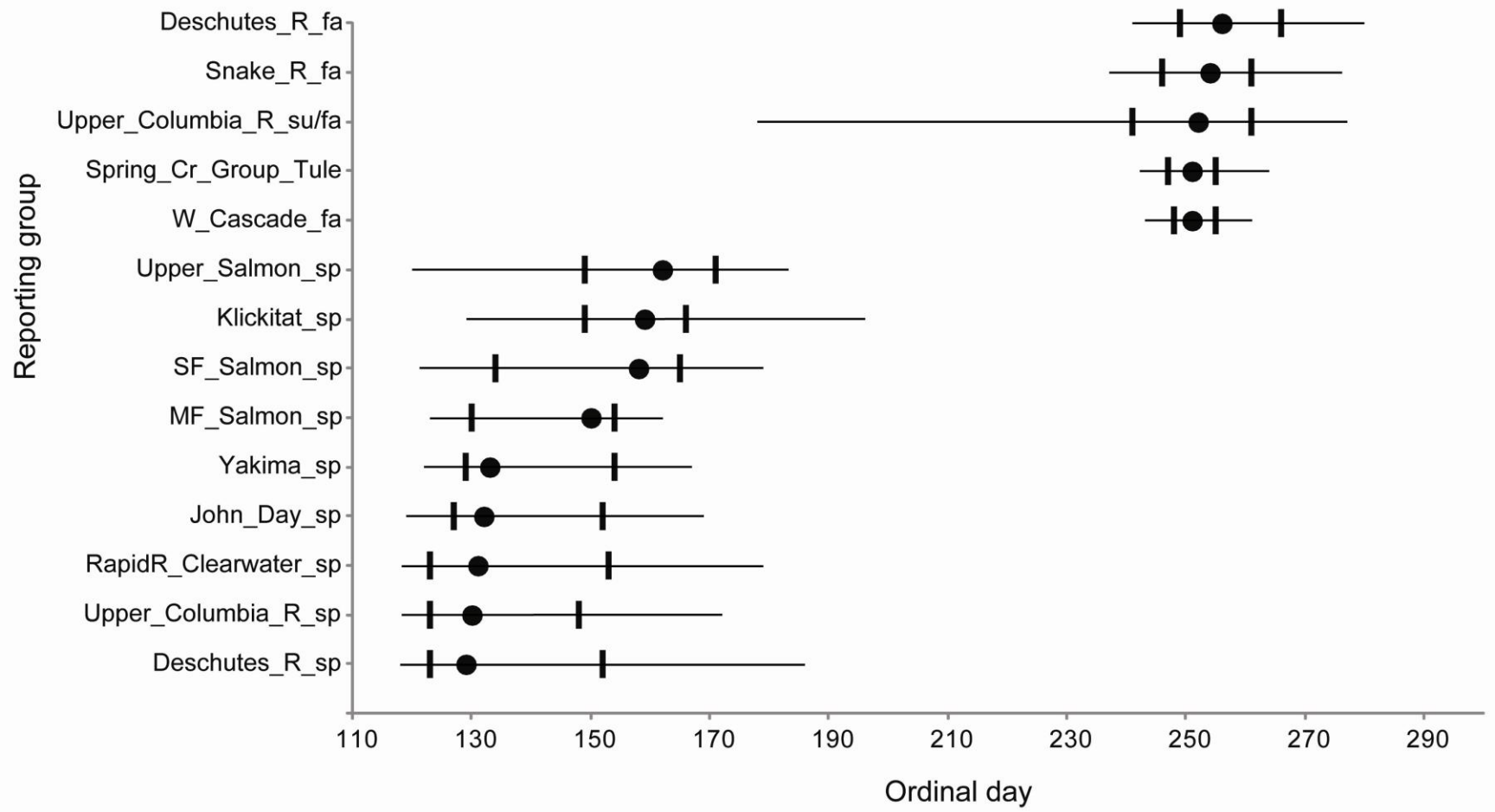


Figure 3. Columbia River Chinook salmon stock timing distributions at Bonneville Dam in 2011, including median (circles), inter-quartile (bars), and 5th and 95th percentile dates (horizontal lines). The distributions were based on the weekly estimated reporting group proportions that were applied to the total abundance of Chinook salmon tallied at the Bonneville Dam fish counting window. This method for estimating run-timing distributions minimizes bias imposed by uneven sampling.

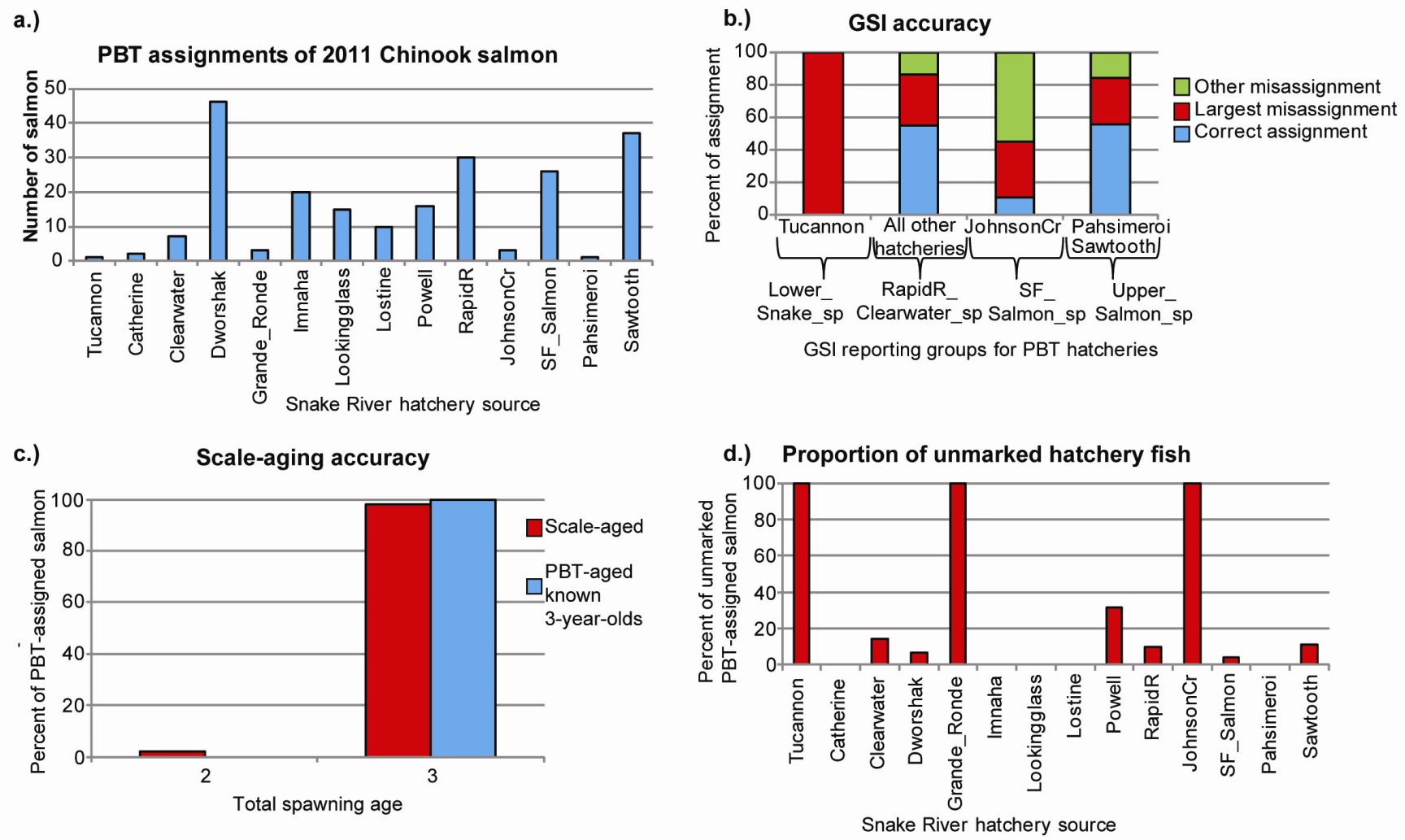


Figure 4. Parentage based tagging results for Chinook salmon passing Bonneville Dam in 2011. a.) All 3-year old salmon from spring Chinook salmon Snake River hatchery spawnings (n=217) that were sampled as returning adults at Bonneville Dam in 2011 could be assigned to source hatcheries. b.) GSI accuracy could be estimated after classifying hatcheries into reporting groups. "Largest misassignment" in most cases was due to Upper_Columbia_sp misassignments, except for the Tucannon fish (n=1) which assigned to JohnDay_sp. c.) PBT-assigned fish were known 3-year olds and tested accuracy of scale-based ages (accuracy=98%). d.) Wild (unmarked) fish assigned with PBT are due to marking error or supplementation programs.

Parentage based tagging analyses of Chinook salmon in 2011

We were able to assign 217 Chinook salmon sampled at Bonneville Dam in 2011 to the 2008 spring-run Chinook salmon broodstock from fourteen different Snake River hatcheries (Figure 4a). Using these known origin Chinook salmon, we compared the individual assignments based on GSI analysis. The fourteen Snake River hatchery sources identified by PBT were first aggregated into the appropriate GSI reporting group in order to make the assignment results comparable between methods. Tucannon hatchery was placed in the lower Snake R. reporting group, Johnson Cr. and S.F. Salmon were placed in the South Fork Salmon R. reporting group, Pahsimeroi and Sawtooth were placed in the Upper Salmon R. reporting group, and all other hatcheries were grouped into the Rapid R./Clearwater R. reporting group. GSI assignments of these PBT-assigned salmon showed concordance values below 60% for all four reporting groups (Figure 4b). In most cases, the single reporting group that could account for the largest misassignment was the Upper Columbia R. spring, however the single fish from the Tucannon hatchery was misassigned by GSI to the John Day R. Comparing the known ages of the PBT-assigned fish (3-year olds) with the "total" ages obtained by a scale-aging method (Figure 4c) demonstrated this method to be highly accurate (98%). According to the scale ages, all of the correct-aged fish spent 1 year in freshwater and 1 year in the ocean before returning to spawn. We also examined the percentage of these PBT-assigned fish that had their adipose fin intact (Figure 4d). Although marking rates varies widely across source hatcheries, for most hatcheries with sample sizes above 15 fish, the rate was near or below 10% unmarked. One exception was Powell hatchery which had a sample size above 16 fish but this sample was 31% unmarked.

Power analysis of 188-SNP steelhead baseline

The 106 steelhead reference collections were grouped into 18 steelhead reporting groups including one outgroup, Quinault Hatchery on the Washington coast (Table 3). Results from the leave-one-out analysis, a relatively conservative accuracy test, showed over half of the baseline collections (59) achieved above 75% correct individual assignment to reporting-group-of-origin. Seven of the 18 reporting groups showed an average correct individual assignment below 75%, and these reporting groups listed in descending order were Imnaha R. (70%), lower Clearwater R. (68%), Upper Columbia R. (66%), middle Columbia R. (64%), Upper Salmon R. (63%), Grande Ronde R. (56%), and lower/Little Salmon R. (53%).

Results from 100% mixture simulations for each of the baseline collections, showed that 82% of collections (89 of 106) produced estimates greater than 90% proportion of the correct reporting group (Table 3). When these proportions were averaged for each reporting group, only the following four reporting groups yielded averages below 90%: lower Clearwater (88%), Upper Salmon R. (87%), Grande Ronde R. (85%), and lower/Little Salmon R. (84%).

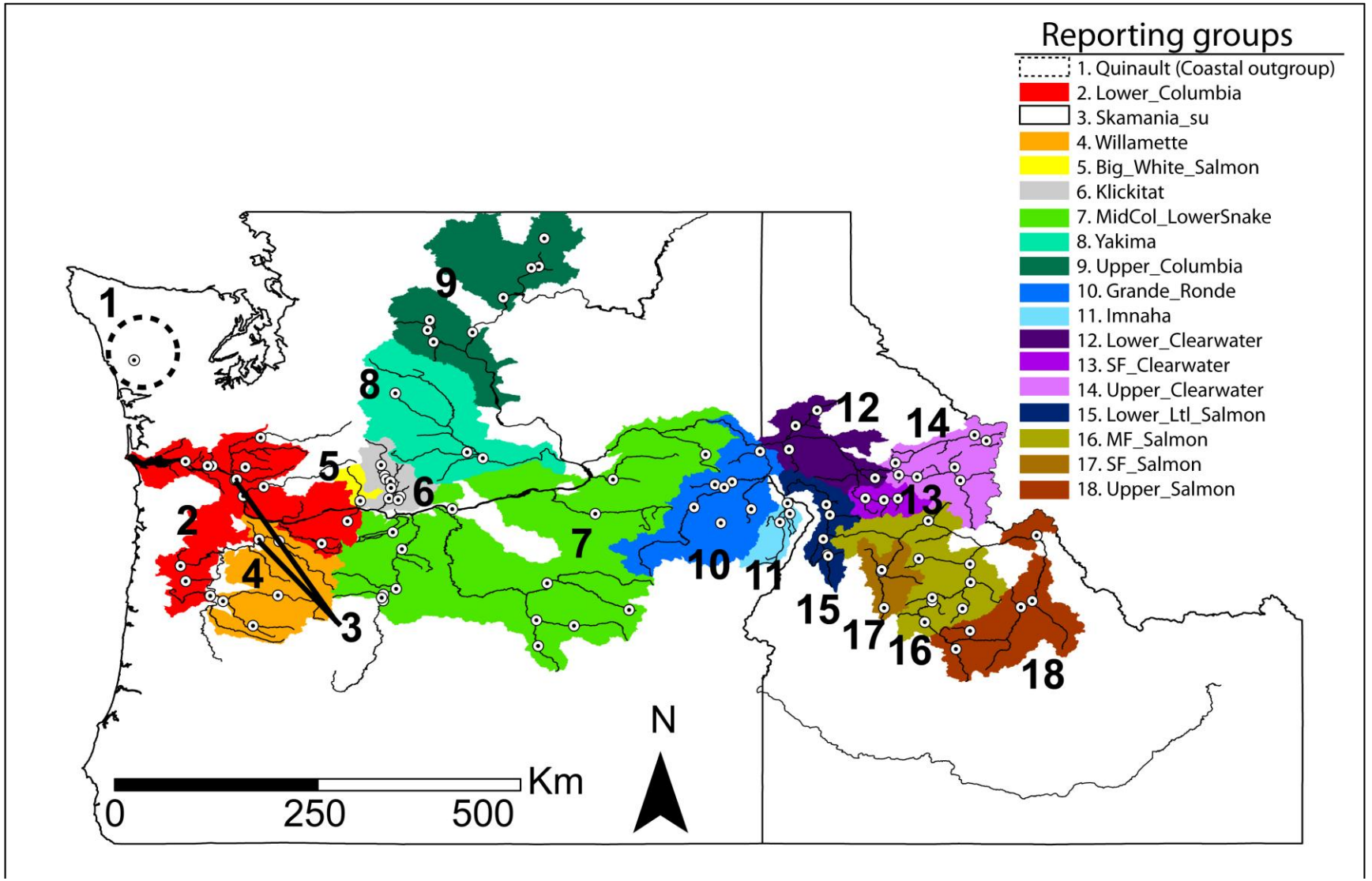


Figure 5. Map of steelhead baseline collections and reporting groups.

Table 3. Sample numbers of steelhead baseline collections and reporting group accuracy.

#	Collection	N	Reporting Group	Leave-one-out % Correct	Leave-one-out RpGr Avg.	100% simulations 100% Sims	100% simulations 95% C.I.	100% simulations RpGr Avg.
1	Quinault	92	Coast	98.1%	98.1%	100.0%	(0.9952, 1.0000)	100.0%
2	Abernathy	167	Lower_Columbia	90.9%	90.1%	99.8%	(0.9875, 1.0000)	97.9%
3	Germany	48	Lower_Columbia	90.9%		99.7%	(0.9875, 1.0000)	
4	Mill	45	Lower_Columbia	92.3%		99.3%	(0.9750, 1.0000)	
5	CowlitzBD	94	Lower_Columbia	100.0%		99.9%	(0.9927, 1.0000)	
6	KalamaW	94	Lower_Columbia	100.0%		96.3%	(0.9179, 0.9998)	
7	EFLewis	79	Lower_Columbia	74.4%		98.4%	(0.9539, 1.0000)	
8	NFLewis	94	Lower_Columbia	96.1%		99.8%	(0.9899, 1.0000)	
9	SandyStill	30	Lower_Columbia	76.9%		82.7%	(0.7458, 0.8997)	
10	HoodEM	59	Lower_Columbia	58.8%		97.1%	(0.9400, 0.9959)	
11	Coweeman	47	Lower_Columbia	90.9%		99.4%	(0.9680, 1.0000)	
12	WillametteCanyon	25	Lower_Columbia	100.0%		100.0%	(1.0000, 1.0000)	
13	WillametteLuckiamute	28	Lower_Columbia	100.0%		100.0%	(0.9962, 1.0000)	
14	WillametteWillamina	32	Lower_Columbia	100.0%		99.8%	(0.9847, 1.0000)	
15	KalamaS	94	Skamania_su	69.6%	76.9%	94.7%	(0.8955, 0.9855)	97.3%
16	ClackNFDam	60	Skamania_su	84.2%		99.9%	(0.9924, 1.0000)	
17	Clackamas	94	Willamette	90.7%	85.7%	100.0%	(0.9948, 1.0000)	99.3%
18	NFEagle	44	Willamette	76.0%		98.1%	(0.9521, 0.9979)	
19	Eagle	48	Willamette	78.9%		98.3%	(0.9572, 0.9996)	
20	LRMad	52	Willamette	92.3%		99.8%	(0.9912, 1.0000)	
21	NFSantiam	39	Willamette	93.8%		100.0%	(0.9965, 1.0000)	
22	Wiley	93	Willamette	82.6%		100.0%	(0.9931, 1.0000)	
23	MiddleColumbiaBigWhiteSalmonRiver	81	Big_White_Salmon	94.4%	94.4%	100.0%	(0.9951, 1.0000)	100.0%
24	KlickitatRiverupperTroutCreek	46	Klickitat	95.5%	76.1%	100.0%	(1.0000, 1.0000)	99.4%
25	KlickitatRiverSurveyorsCreek	39	Klickitat	100.0%		100.0%	(1.0000, 1.0000)	
26	KlickitatRiverSnyderCreek	47	Klickitat	85.7%		100.0%	(1.0000, 1.0000)	

27	KlickitatRiverlowerSummitCreek	46	Klickitat	77.8%		99.1%	(0.9764, 1.0000)	
28	KlickitatRiverlowerTroutCreek	48	Klickitat	61.9%		99.2%	(0.9736, 1.0000)	
29	KlickitatRiverLowerWhiteCreek	35	Klickitat	85.7%		99.4%	(0.9829, 1.0000)	
30	KlickitatRiverlowerLittleKlickitatRiver	46	Klickitat	63.6%		98.8%	(0.9629, 1.0000)	
31	KlickitatRiverDeadCanyonCreek	35	Klickitat	50.0%		99.0%	(0.9710, 1.0000)	
32	KlickitatRiverBowmanCreek	48	Klickitat	57.1%		99.0%	(0.9675, 1.0000)	
33	KlickitatRiverSwaleCreek	48	Klickitat	83.3%		99.6%	(0.9840, 1.0000)	
34	MiddleColumbiaFifteenmileCreek	94	Mid_Columbia	68.6%	63.7%	98.4%	(0.9615, 0.9981)	92.2%
35	Pelton	47	Mid_Columbia	35.7%		65.5%	(0.5568, 0.7361)	
36	Shitike	32	Mid_Columbia	50.0%		86.7%	(0.8013, 0.9271)	
37	DeschutesRiverBuckHollowCreek	63	Mid_Columbia	75.0%		94.0%	(0.8927, 0.9791)	
38	DeschutesRiverTroutCreek	57	Mid_Columbia	44.4%		82.2%	(0.7536, 0.8831)	
39	DeschutesRiverUpperMainstem	61	Mid_Columbia	93.8%		99.9%	(0.9930, 1.0000)	
40	Rock	266	Mid_Columbia	58.7%		96.7%	(0.9317, 0.9905)	
41	JDUpBeech	60	Mid_Columbia	64.7%		97.6%	(0.9347, 0.9962)	
42	JDLow	64	Mid_Columbia	36.4%		90.0%	(0.8361, 0.9511)	
43	MFJD	129	Mid_Columbia	86.5%		98.6%	(0.9653, 1.0000)	
44	NFJDGranite	79	Mid_Columbia	72.2%		96.8%	(0.9295, 0.9909)	
45	SFJDDeer	36	Mid_Columbia	81.8%		98.4%	(0.9533, 1.0000)	
46	Umatilla	34	Mid_Columbia	58.3%		98.0%	(0.9522, 0.9977)	
47	Touchet	89	Mid_Columbia	83.3%		98.4%	(0.9610, 0.9981)	
48	UpTucan	95	Mid_Columbia	46.5%		81.8%	(0.7289, 0.8844)	
49	YakimaNachesLittleRattlesnake	175	Yakima	78.5%	84.8%	98.9%	(0.9691, 0.9999)	98.8%
50	SatusCreek	46	Yakima	75.9%		97.8%	(0.9499, 0.9974)	
51	ToppenishCreek	46	Yakima	100.0%		99.9%	(0.9917, 1.0000)	
52	Entiat	238	Upper_Columbia	55.3%	66.2%	96.1%	(0.9224, 0.9875)	94.7%
53	Methow	93	Upper_Columbia	68.2%		93.6%	(0.8781, 0.9697)	
54	Bonaparte	100	Upper_Columbia	62.2%		96.3%	(0.9283, 0.9928)	
55	OmakCreekAdults	94	Upper_Columbia	94.6%		98.3%	(0.9595, 0.9981)	
56	OkanoganSalmonCreek	100	Upper_Columbia	62.3%		99.5%	(0.9856, 1.0000)	
57	WenatcheeChiwaukum	89	Upper_Columbia	65.2%		97.3%	(0.9481, 0.9948)	

58	LevenworthNFHIcicleCreek	43	Upper_Columbia	58.3%		88.8%	(0.8103, 0.9476)	
59	WenatcheeNason	21	Upper_Columbia	63.6%		87.8%	(0.7956, 0.9238)	
60	Asotin	49	Grande_Ronde	17.9%	55.9%	32.6%	(0.2028, 0.4337)	84.7%
61	GRCrook	98	Grande_Ronde	57.8%		91.0%	(0.8300, 0.9533)	
62	GRElk	47	Grande_Ronde	84.8%		99.1%	(0.9762, 1.0000)	
63	Lostine	45	Grande_Ronde	73.7%		94.6%	(0.9005, 0.9781)	
64	LtlMinam	48	Grande_Ronde	60.9%		96.6%	(0.9353, 0.9949)	
65	GRlower	94	Grande_Ronde	38.1%		88.5%	(0.8136, 0.9439)	
66	Wenaha	94	Grande_Ronde	58.3%		90.2%	(0.8497, 0.9509)	
67	BigSheep	71	Imnaha	73.7%	69.9%	95.0%	(0.9110, 0.9793)	90.6%
68	ImnCamp	25	Imnaha	76.9%		94.1%	(0.8805, 0.9737)	
69	ImnCow	44	Imnaha	65.4%		79.7%	(0.7043, 0.8565)	
70	ImnLight	46	Imnaha	63.6%		93.5%	(0.8774, 0.9734)	
71	EFPot	62	Lower_Clearwater	63.3%	67.8%	95.4%	(0.9200, 0.9837)	87.8%
72	BigBear	33	Lower_Clearwater	53.8%		83.8%	(0.7307, 0.9104)	
73	LtlBear	54	Lower_Clearwater	100.0%		89.1%	(0.8327, 0.9428)	
74	LapMission	51	Lower_Clearwater	54.2%		82.9%	(0.7708, 0.8844)	
75	SFCCrook	82	SF_Clearwater	87.1%	80.8%	99.2%	(0.9762, 1.0000)	96.3%
76	SFCTenmile	47	SF_Clearwater	85.2%		99.8%	(0.9903, 1.0000)	
77	SFCJohns	38	SF_Clearwater	80.0%		88.1%	(0.8084, 0.9295)	
78	MFCClear	45	SF_Clearwater	70.8%		97.9%	(0.9516, 0.9987)	
79	Storm	38	Upper_Clearwater	95.0%	89.5%	99.9%	(0.9950, 1.0000)	98.3%
80	LochsaCF	44	Upper_Clearwater	90.9%		99.3%	(0.9786, 1.0000)	
81	LochsaCanyon	46	Upper_Clearwater	76.0%		97.9%	(0.9434, 0.9999)	
82	SelBear	45	Upper_Clearwater	100.0%		100.0%	(0.9951, 1.0000)	
83	SelMoose	47	Upper_Clearwater	100.0%		99.8%	(0.9881, 1.0000)	
84	SelGedney	46	Upper_Clearwater	94.7%		99.0%	(0.9717, 1.0000)	
85	SelOhara	47	Upper_Clearwater	69.6%		92.2%	(0.8682, 0.9747)	
86	LtlBoulder	47	Lower_Ltl_Salmon	65.5%	53.1%	92.2%	(0.8661, 0.9643)	84.3%
87	LtlHazard	45	Lower_Ltl_Salmon	28.6%		69.6%	(0.6118, 0.7774)	
88	LtlRapid	47	Lower_Ltl_Salmon	65.5%		98.2%	(0.9611, 0.9976)	

89	Slate	47	Lower_Ltl_Salmon	51.9%		72.3%	(0.6209, 0.7985)	
90	WhiteBird	59	Lower_Ltl_Salmon	53.8%		89.5%	(0.8202, 0.9516)	
91	EFSF	47	SF_Salmon	76.2%	85.8%	99.6%	(0.9876, 1.0000)	99.5%
92	Secesh	45	SF_Salmon	88.0%		99.3%	(0.9792, 1.0000)	
93	Stolle	45	SF_Salmon	93.3%		99.7%	(0.9881, 1.0000)	
94	MFsalMarsh	59	MF_Salmon	95.8%	91.7%	100.0%	(0.9953, 1.0000)	99.0%
95	MFsalRapid	46	MF_Salmon	100.0%		99.9%	(0.9941, 1.0000)	
96	MFsalPistol	23	MF_Salmon	100.0%		99.9%	(0.9951, 1.0000)	
97	MFsalCamas	60	MF_Salmon	100.0%		99.5%	(0.9848, 1.0000)	
98	MFsalBig	47	MF_Salmon	96.6%		100.0%	(0.9955, 1.0000)	
99	BigMFsal	48	MF_Salmon	88.9%		99.8%	(0.9912, 1.0000)	
100	LoonMFsal	40	MF_Salmon	84.2%		99.7%	(0.9867, 1.0000)	
101	Bargamin	47	MF_Salmon	68.0%		93.3%	(0.8896, 0.9705)	
102	Sawtooth	67	Upper_Salmon	79.2%	62.8%	93.3%	(0.8764, 0.9738)	87.0%
103	WFYankee	46	Upper_Salmon	55.2%		88.4%	(0.8214, 0.9313)	
104	Morgan	45	Upper_Salmon	65.4%		96.6%	(0.9306, 0.9897)	
105	Pahsim	47	Upper_Salmon	81.0%		94.0%	(0.8903, 0.9788)	
106	NFSalm	51	Upper_Salmon	33.3%		62.5%	(0.5210, 0.7103)	

Note: Averaged values for each reporting group are shaded dark green, light green, orange, yellow, and red to indicate ranges of >90%, 80-90%, 70-80%, 60-70%, and below 60%, respectively.

Run-timing of steelhead stocks in 2009-2011

We were able to obtain sufficient sample sizes to characterize the run-timing distributions of 15 wild steelhead stocks and 10 hatchery steelhead stocks across all three years (Figure 6). These results indicate three main run-timing categories of stocks. An early run-timing category is occupied primarily by the Skamania summer-run (Median date Jul. 6th), an intermediate run-timing category includes most wild and hatchery steelhead stocks (median dates range fromJuly 22nd to Aug. 17th), and a late run-timing category includes South Fork Salmon R., South Fork Clearwater R., and upper Clearwater R. (Median dates 2nd, 7th, and 11th of Sep.). The late run-timing category is typically thought to be characteristic of B-run steelhead that return after August 25th at Bonneville Dam.

Table 4. Steelhead sampled at Bonneville Dam from 2009-2011.

	Wild			Hatchery		
Biweekly strata	2009	2010	2011	2009	2010	2011
22_23	6	10	8	45	36	11
24_25	10	30	12	32	55	20
26_27	40	55	8	68	43	15
28_29	179	154	57	235	114	35
30_31	170	178	90	270	288	66
32_33	54	86	137	132	178	237
34_35	67	24	73	240	51	151
36_37	25	33	46	125	92	115
38_39	56	31	22	244	119	98
40_41	25	21	40	263	73	111
Total	632	622	493	1654	1049	859

Note: Gray indicates sample size ≥20 and outlined boxes indicate the time interval used to estimate abundance and run-timing. These time intervals represent greater than 98% and 99% of the total numbers of fish tallied for wild and hatchery steelhead, respectively.

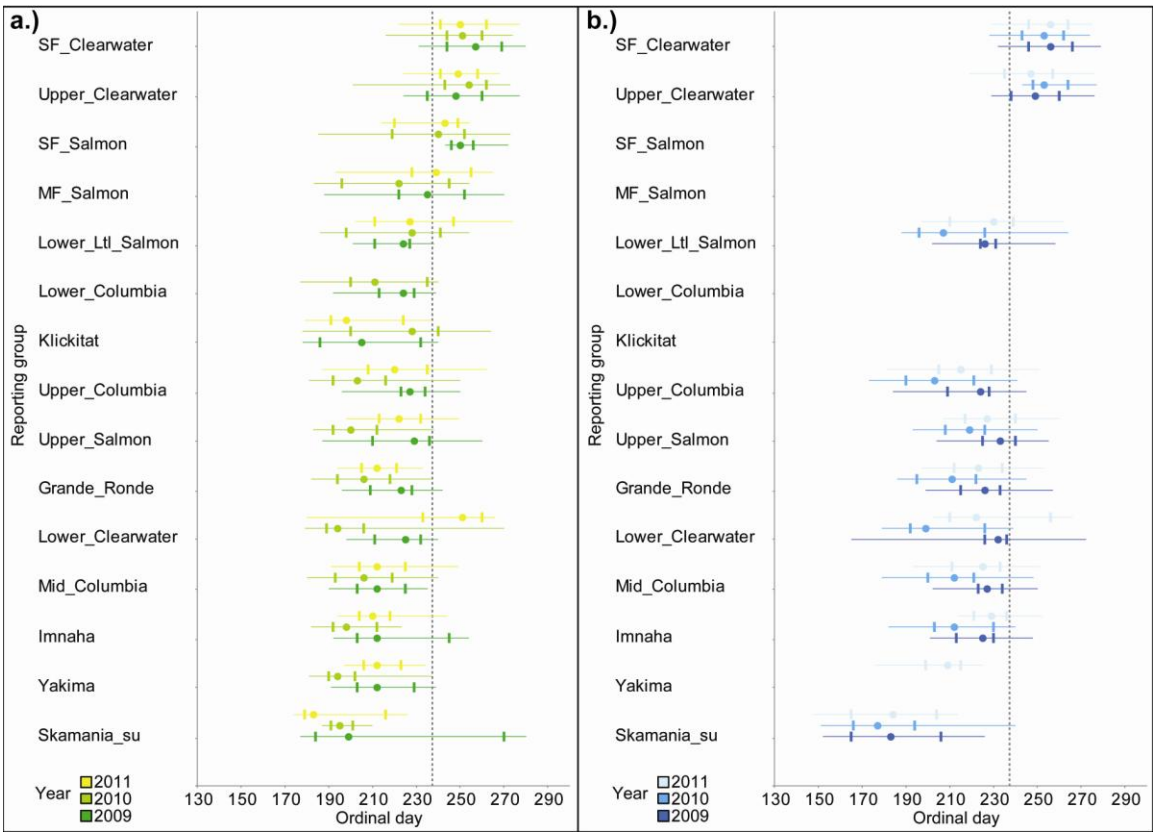


Figure 6. Run-timing distributions of (a) wild and (b) hatchery steelhead reporting groups. August 25th (dashed vertical line) classifies A-run (before the 25th) and B-run (after the 25th) steelhead. The run-timing distribution of each reporting group is indicated by median day (circle), interquartile (vertical bars), and the 5th and 95th percentile of the run (horizontal line). Reporting groups with less than 0.035% relative abundance were not included in this figure. Reporting groups were organized from top to bottom in order of average median return day across years and hatchery/wild origins.

Estimated abundance of steelhead stocks in 2009-2011

There were fourteen wild steelhead stocks passing Bonneville Dam that we estimated had an average abundance greater than 1,000 fish averaged across three years (Figure 7a). The fourteen major stocks of wild steelhead in order of magnitude were lower Columbia R. (1,000), South Fork Salmon R. (2,000), Klickitat R. (3,000), lower Clearwater R. (3,000), Middle Fork Salmon R. (5,000), Imnaha R. (5,000), lower/Little Salmon R. (5,000), Yakima R. (6,000), upper Clearwater R. (6,000), South Fork Clearwater R. (9,000), Upper Salmon R. (11,000), Grande Ronde R. (19,000), Upper Columbia R. (22,000), and middle Columbia R. (48,000). The ten major stocks of hatchery steelhead in order of magnitude were lower Clearwater R. (2,000), upper Clearwater R. (3,000), Imnaha R. (7,000), Skamania summer-run (7,000), lower/Little Salmon R. (12,000), middle Columbia R. (34,000), Grande Ronde R. (45,000), South Fork Clearwater R. (47,000), Upper Columbia R. (51,000), and Upper Salmon R. (93,000). These stock abundance estimates were based on the stock proportions that were estimated in ONCOR across biweekly strata (Table 4), and were multiplied with the total abundance of steelhead that was tallied on a daily basis at the Bonneville Dam fish counting window.

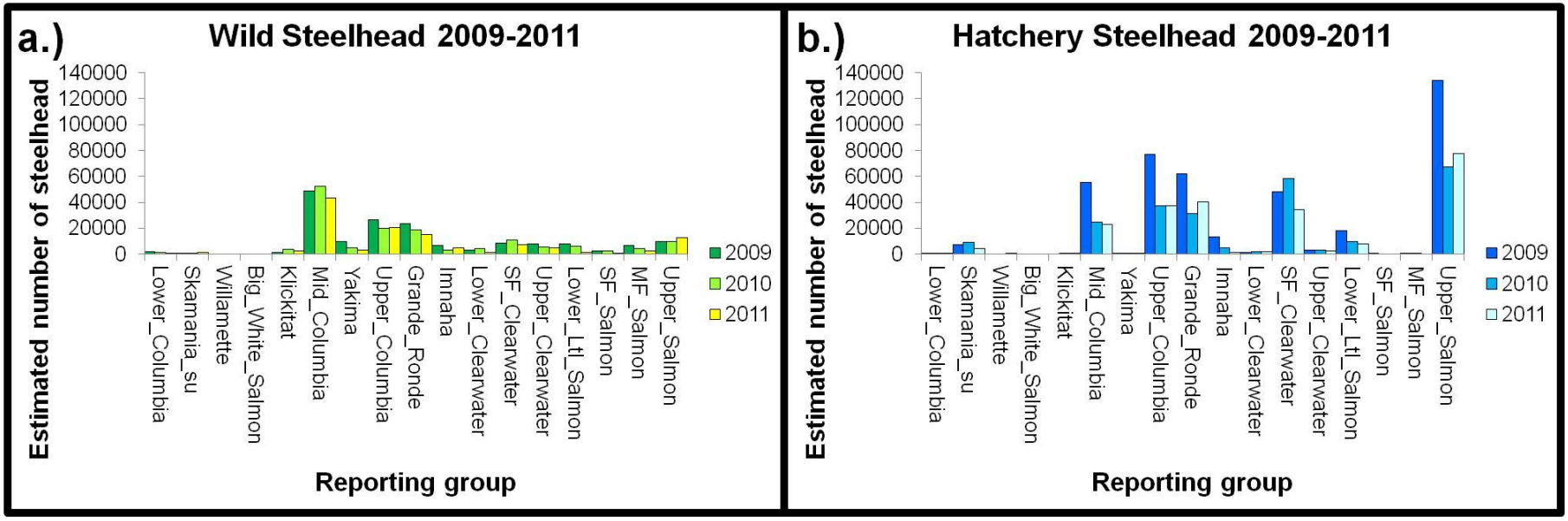


Figure 7. Estimated abundance of (a) wild and (b) hatchery steelhead. These estimates are based on the steelhead returns at Bonneville Dam during 2009-2011.

Parentage based tagging analyses of steelhead in 2011

We were able to assign 208 steelhead sampled at Bonneville Dam in 2011 to the 2008 steelhead broodstock from six different Snake River hatcheries (Figure 8a). These PBT-assigned fish were predominately from the Dworshak Hatchery (n=157, 75%). Using these known hatchery-of-origin steelhead, we compared the individual assignments based on GSI analysis. The six Snake River hatchery sources identified by PBT were first aggregated into the appropriate GSI reporting group in order to make the assignment results comparable between methods. Lyons Ferry Hatchery was placed in the Grande Ronde R. reporting group, Dworshak and Squaw Creek were placed in the South Fork Clearwater R. reporting group, and Oxbow, Pahsimeroi, and Sawtooth were all placed in the Upper Salmon R. reporting group. GSI assignments of these PBT-assigned salmon showed concordance values above 60% for all three reporting groups, and the South Fork Clearwater R. reporting group had the highest concordance (87%) and sample size (Figure 8b). In most cases, the single reporting group that could account for the largest misassignment was the middle Columbia R., however the South Fork Clearwater R. fish largely misassigned by GSI to upper Clearwater R. Comparing the known ages of the PBT-assigned fish (4-year olds) with the "total" ages obtained by a scale-aging method demonstrated this method to be have large error (Figure 8c). According to the scale ages, many of these fish were estimated to be 5-year olds and spent 1 year in freshwater and 3 years in the ocean. We also examined the percentage of these PBT-assigned fish that had their adipose fin intact (Figure 8d). All hatcheries represented by more than 10 assigned fish were found to have above 10% unmarked fish, and Sawtooth had 26% unmarked.

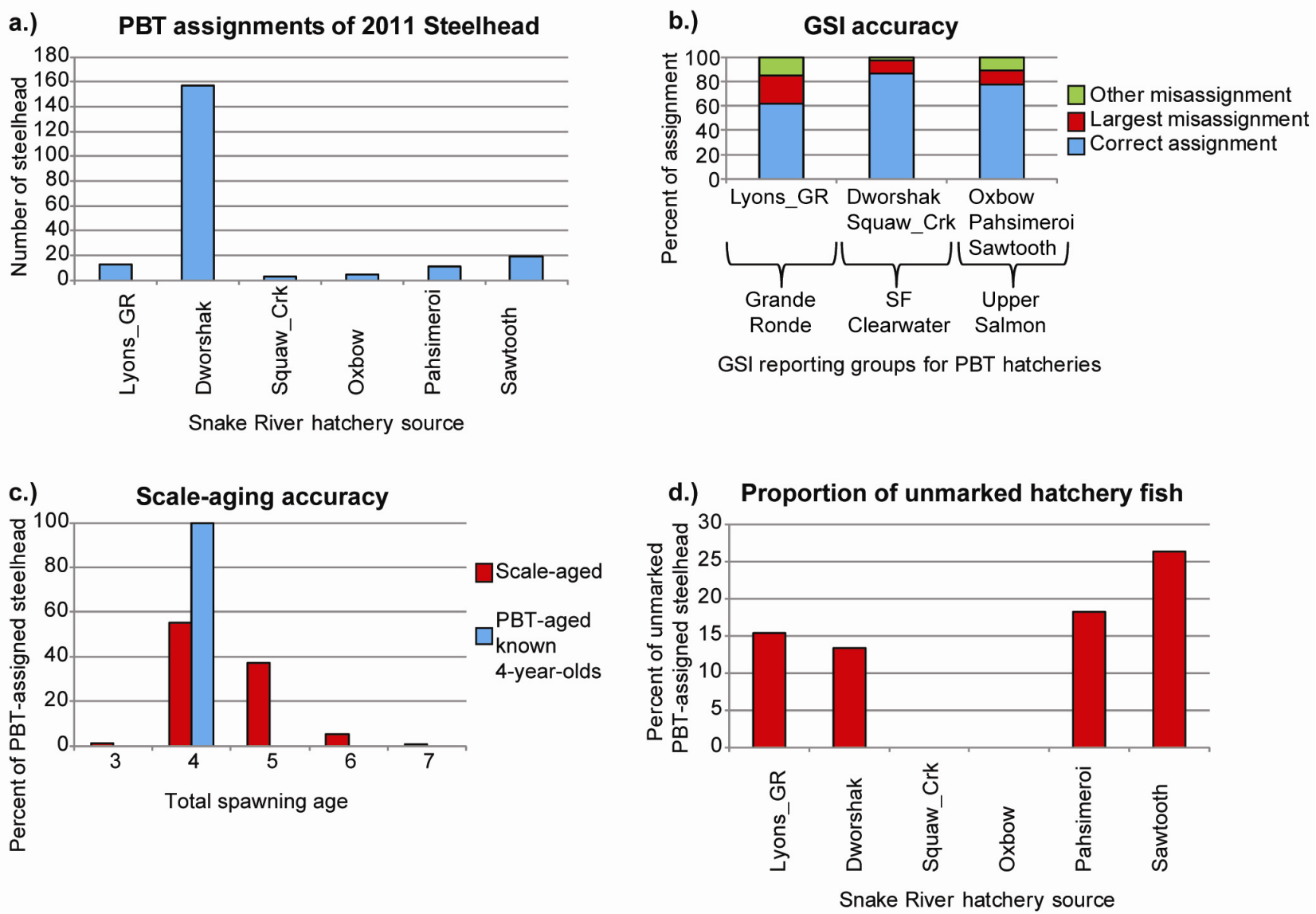


Figure 8. Parentage based tagging results for steelhead passing Bonneville Dam in 2011. a.) All 4-year old steelhead returning in 2011 were assigned to Snake River source hatcheries. b.) GSI accuracy could be estimated after classifying hatcheries into reporting groups . c.) PBT-assigned fish were known 4-year olds and tested accuracy of scale-based ages. d.) Wild (unmarked) fish assigned with PBT are due to marking error or supplementation programs.

A-run versus B-run life history analyses of steelhead

We examined how well the A-run and B-run steelhead life-history categories defined particular stocks of steelhead. We first used a fork length of 78 cm as a proxy for a B-run steelhead, and examined the run-timing of this size of fish at Bonneville (Figure 9a). August 25th did appear to delineate the run-timing distribution of fish greater than 78 cm (B-run), however, there were still a group of fish less than 78 cm that arrived after this date. There were four reporting groups that relatively high proportions (>10%) of fish larger than 78 cm assign to and these reporting groups represented tributaries typically thought to produce B-run fish, South Fork and Upper Clearwater R. and Middle Fork and South Forth Salmon R. (Figure 9b). The ocean age (based on scale-age method) of these length categories did not fit the definitions of A-run and B-run fish very well but aging for steelhead is expected to be less accurate based on the results shown in Fig. 8c. The less-than-78 cm fish (A-run) were found to be composed of both 1-year and 2-year ocean fish (Figure 9c). Two-year ocean fish are more typical of B-run life history. Finally, the distribution of ocean-ages among reporting group, did not show as strong an association of 2-year ocean fish with the South Fork and Upper Clearwater R. and Middle Fork and South Forth Salmon R. reporting groups as was seen with fork length distribution (Figure 9d).

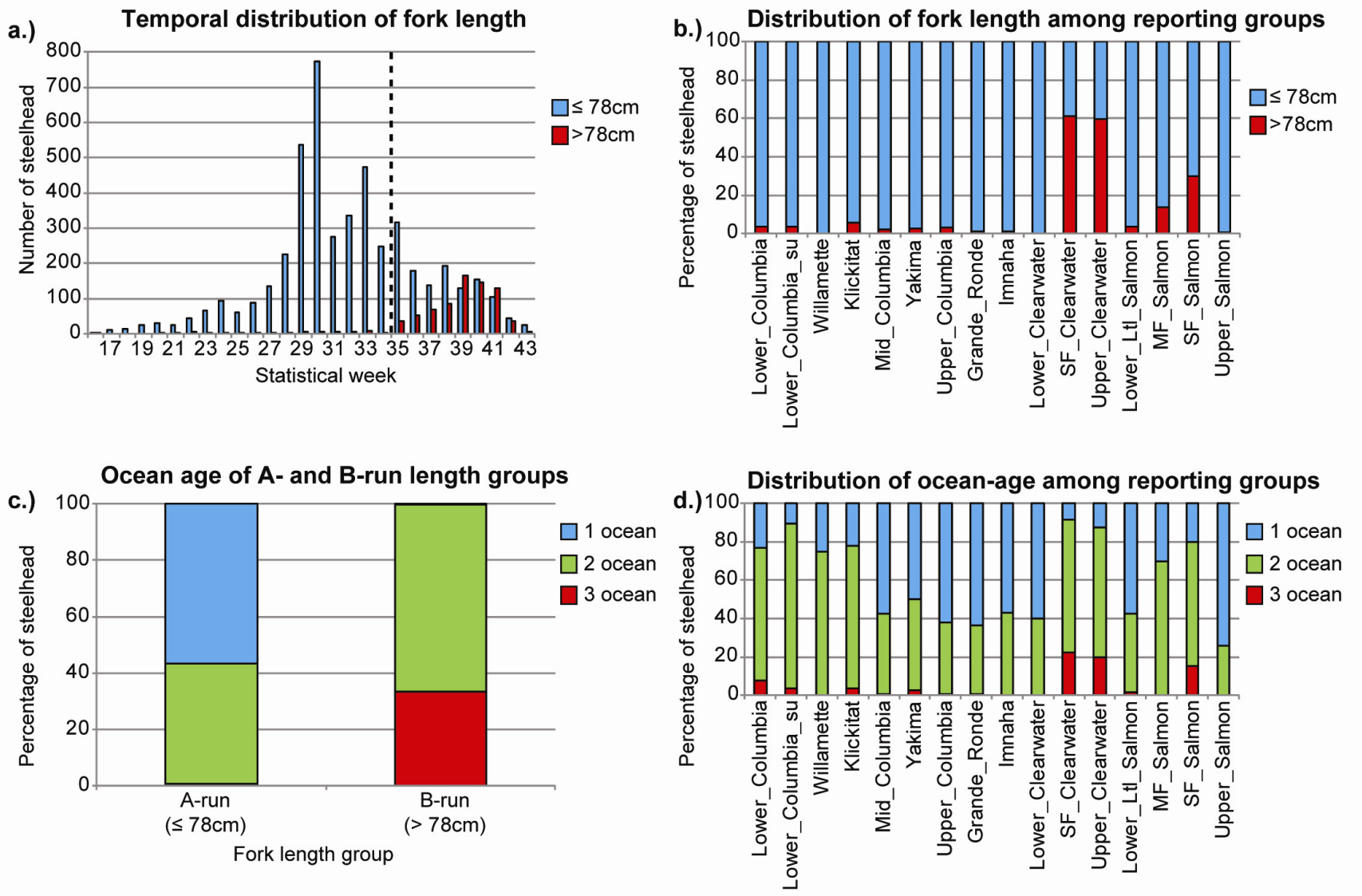


Figure 9. Steelhead A- versus B-run life-history analysis. a.) August 25th delineates run-timing of A-versus B-run fork lengths which are less and greater than 78cm, respectively. b.) GSI reporting groups show highest proportions of B-run(>78cm) in Clearwater and MF/SF Salmon Rivers. c.) Ocean-age composition of B-run consists of 2- and 3-years, but A-run also contains 2-year . d.) 2-year ocean-age is not highly predictive of B-run, however 3-ocean fish mostly found in typical B-run basins

Discussion

Management Implications

This study demonstrates great potential for the application of genetic stock identification in the management of Columbia River Chinook salmon and steelhead fisheries evidenced by the ability to estimate stock abundance, characterize run-timing distributions, and discriminate some of the major stocks by peak run-timing. Results indicate there were nine stocks of spring-run Chinook salmon and five stocks of fall-run Chinook salmon estimated to have greater than 2,000 fish pass Bonneville Dam in 2011. Further, the Middle Fork, South Fork, and Upper Salmon R. and Klickitat R stocks arrive at least two weeks later than all other spring-run stocks based on median day.

For steelhead, there were fourteen wild steelhead stocks with an estimated abundance greater than 1000 fish averaged over a three year period. Ten hatchery steelhead stocks were also estimated above this average abundance level. We described three run-timing categories which included an early Skamania summer-run, an intermediate run-timing category that contains most wild and hatchery steelhead stocks, and a late run-timing category that arrives after August 25th and includes South Fork Salmon R., South Fork Clearwater R., and upper Clearwater R. Although the fact that larger fish generally arrive after August 25th and are composed primarily of Middle Fork and South Fork Salmon R. and Clearwater R. stocks, generally support expected characteristics for typical A-run and B-run steelhead life history categories. However, there is significant overlap in all characteristics and designation of individual fish based on timing, length, and age distinctions is not clear cut. Ocean-age characteristics are also not definitive, and the use of these categories in steelhead management may require re-evaluation.

This study provided an opportunity to begin to integrate the new genetic technology of parentage based tagging (PBT) with a GSI application. The challenge imposed by long histories of exogenous stock transfers in many hatchery programs often prevents effective application of GSI in assigning hatchery fish. However, now with the prospect of expanding PBT to mass mark all hatchery fish, GSI will soon be reserved for a smaller but essential role of filling in information gaps that remain after PBT has been used to identify hatchery-origin fish. Aside from improving estimates of stock abundance and run-timing, we have also demonstrated how PBT may improve steelhead aging data and can be used to quantify the true number of hatchery fish among those that are unmarked (adipose intact). This latter ability will help to evaluate the success of supplementation programs that primarily release unmarked fish.

References

Anderson E. C. 2010. Computational algorithms and user-friendly software for parentage-based tagging of Pacific salmonids. Final report submitted to the Pacific Salmon Commission's Chinook Technical Committee (US Section). 46 p. http://swfsc.noaa.gov/textblock.aspx?Division=FED&ParentMenuId=54&id=16021.

Blankenship, S. M., M. R. Campbell, J. E. Hess, M. E. Hess, T. W. Kassler, C. C. Kozfkay, A. P. Matala, S. R. Narum, M. M. Paquin, M. P. Small, J. J. Stephenson, K. I. Warheit, and P. Moran. 2011. Major lineages and metapopulations in Columbia River *Onchorhynchus mykiss* are structured by dynamic landscape features and environments. Transactions of the American Fisheries Society 140(3):665-684.

Hess, J.E., Campbell, N.R., Matala, A.P., and Narum, S.R. 2011. 2010 Annual Report: Genetic Assessment of Columbia River Stocks. U.S. Dept. of Energy Bonneville Power Administration Report Project #2008-907-00

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. North American Journal of Fisheries Management 26(1):147-155.

Narum, S.R., Campbell, N.R., Matala, A.P., and Hess, J.E. 2010. 2009 Annual Report: Genetic Assessment of Columbia River Stocks. U.S. Dept. of Energy Bonneville Power Administration Report Project #2008-907-00

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.

Steele CA, Campbell MR, Ackerman M, McCane J, Hess MA, Campbell N, Narum SR. 2011. Parentage Based Tagging of Snake River hatchery steelhead and Chinook salmon. Bonneville Power Administration. Annual Progress Report, Project number 2010-031-00. https://research.idfg.idaho.gov/Fisheries%20Research%20Reports/Res11-111Steele2010%20Parentage%20Based%20Tagging%20Snake%20River%20Steelhead%20Salmon.pdf