Population genetic structure and recovery of chum salmon in the Lower Columbia River

Maureen P. Small[1a], Bryce Glaser[2], Todd Hillson[2], and Cherril Bowman[1]

August 16, 2011

[1] Molecular Genetics Lab, WDFW, 600 Capitol Way N, Olympia, WA

[a] corresponding author smallmps@dfw.wa.gov, 360-902-2682

[2] WDFW Fish Management – Region 5, 2108 Grand Blvd., Vancouver, WA

Abstract

The lower Columbia River formerly supported runs of over a million chum salmon. By the late 1950's, returns had decreased to as low as a few hundred fish in some years and in 1999 lower Columbia River chum salmon were listed as threatened under the Endangered Species Act. The Lower Columbia Fish Recovery Board subsequently developed recovery plans addressing ecosystem restoration down to the population level. To establish a foundation for population monitoring and assessment, we conducted a genetic analysis of chum salmon from the lower Columbia River and the Pacific Coast. Collections of chum salmon spawning aggregations from sixteen tributaries and mainstem spawning locations in the lower Columbia River and four tributaries on the Pacific Coast were genotyped at 16 microsatellite DNA loci. We calculated fundamental genetic statistics including Hardy-Weinberg equilibrium values, genetic diversity measures and effective population sizes, as baseline information to monitor genetic changes following recovery activities. Regional and local genetic relationships were assessed to identify spawning aggregates that could serve as broodstocks for reintroductions into tributaries where chum salmon populations have been extirpated or for enhancement in tributaries with extremely low population levels.

Introduction

Chum salmon (*Oncorhynchus keta*), formerly a run of over a million fish in the lower Columbia River in Washington State (WA) (NOAA 2004), were nearly extirpated in the 20th century. The National Marine Fisheries Service (NMFS) listed lower Columbia River (LCR) chum salmon as threatened under the Endangered Species Act (ESA) in 1999 (64 FR 14508, March 25, 1999). Declines resulted from overfishing coupled with changes in flow regimes and destruction of mainstem and tributary spawning habitat and rearing habitat in the Columbia River estuary (Sherwood *et al.*1990, Johnson *et al.* 1997). Historically, chum salmon spawned in 16 tributaries (Myers *et al.* 2002) and throughout the mainstem of the Columbia River up to the Walla Walla River (river kilometer (rkm) 507, Johnson *et al.* 1997). Runs were in decline by the 1930's. During the 20-year period spanning 1947-66 the catch decreased by over 90% and chum salmon had disappeared from most tributaries (Fulton 1970). By the late 1990's, the remaining run stabilized around roughly 3000 fish (WDFW 2011).

Prior to 1997, two chum salmon aggregations in the lower Columbia River were recognized as genetically distinct (Phelps et al. 1994), one sited near the mouth and the other upriver in the Columbia Gorge region, and spawners had been noted sporadically in many LCR tributaries. Using genetic analysis of samples gathered from chum salmon spawners throughout the LCR, Small et al. (2006) grouped chum salmon spawning in the mainstem and tributaries of the LCR into three ecoregional genetic aggregations: Coastal, Cascade and the Gorge. The Coastal aggregate included rain-fed tributaries near the mouth of the Columbia River that originated in the Willapa Hills (Grays and Elochoman rivers, and Skamokawa and Germany creeks); the Cascade aggregate included snowmelt-fed tributaries originating in the Cascade Range (Cowlitz and Lewis rivers); and the Gorge aggregate included tributaries mainly sustained by rain and groundwater (Duncan, Hardy and Hamilton creeks) or portions of the mainstem Columbia River near the Columbia Gorge region, as well as fish captured near Bonneville Dam destined for unknown upriver tributaries (Figure 1). The Lower Columbia/Willamette Technical Recovery Team developed three geographic strata within the Columbia River chum salmon Evolutionarily Significant Unit (ESU) that reflected this structure and incorporated Oregon chum salmon aggregates into their respective strata (Table 1). All populations within the Lower Columbia River ESU are considered at high or very high risk of extinction. Many are severely depressed and the status of several other populations is unknown (HSRG 2008).

The Lower Columbia Fish Recovery Board (LCFRB) was established in Washington State to develop and implement recovery plans for ESA listed salmon and steelhead populations (LCFRB 2004, 2010). Goals include recovering fish populations to healthy levels that allow harvest by sport, commercial, and tribal fisheries as well as enhancing and supporting the health of entire ecosystems. Recovery is to be achieved through multiple efforts including habitat restoration and protection, supportive hatcheries and harvest practices, control of non-native species, and the re-establishment of balanced predator/prey relationships (LCRFB 2004, 2010).

Chum salmon recovery in the LCR has multiple challenges. LCFRB (2010) outlines the key factors limiting chum recovery as follows: 1) tributary habitat, particularly off-channel and side channel areas critical for spawning, have been lost; preferred spawning habitat in lower segments of rivers are particularly susceptible to fine sediment accumulation, 2) estuarine habitat has been lost and suffered from loss of diversity and connectivity, 3) mainstem Columbia hydropower projects (particularly Bonneville Dam) have inundated spawning areas above the dams and variable discharge patterns continue to impact downstream spawning and incubation success. While large harvests of chum salmon occurred prior to the 1950's and likely contributed to the decline in chum salmon abundance, there are currently no directed chum salmon harvest fisheries in the LCR and incidental harvest impacts are relatively low. Negative impacts from artificial chum production programs in the LCR are also assumed to be low. Historically, there was limited hatchery supplementation using regional (56%) and out of region (44%) broodstock in tributaries near the mouth of the Columbia River in Oregon (OR) in 1929 and in WA in 1958 (Johnson *et al.* 1997). More recently, conservation hatchery supplementation programs using local broodstock and sized to available habitat have been underway since 1998 in the Grays River in the Coastal ecoregion and since 2001 in Duncan Creek in the Gorge ecoregion (Hillson 2002).

The use of additional conservation hatchery programs for reintroduction and supplementation of historic chum salmon populations is being proposed as a key element of chum salmon recovery in Washington's LCR tributaries. The Washington Department of Fish and Wildlife (WDFW) is developing a LCR chum salmon population enhancement strategy based on the framework of the Summer Chum Salmon Conservation Initiative for Hood Canal and the Strait of Juan de Fuca (WDFW and PNPTC 2000).

Monitoring and evaluating recovery efforts includes assessing temporal and spatial changes in genetic diversity. An earlier population genetic study was conducted on LCR chum salmon using samples collected through fall 2003 (Small et al. 2006) and here we compare that genetic baseline to contemporary patterns of chum salmon genetic structure. We expand the geographic extent of populations in this study to include collections from the Pacific Coast in Oregon and Washington, to broaden the examination of chum salmon population structure in and around the LCR (Table 2). We also establish criteria to identify potential donor broodstocks to support endangered populations and to repopulate and enhance tributaries where chum salmon have been extirpated or abundance is critically low. The expanded genetic baseline provides a foundation for future assessments of recovery and reintroduction efforts.

Methods

Collections

Collections included tissue or scale samples from chum salmon spawners from tributaries and spawning sites in the mainstem of lower Columbia River and from coastal rivers in Oregon and Washington (Figure 1, Table 2). Samples processed in this study were

compared to archived chum salmon collections (see Table 2) described by Small et al. (2006). The first supplementation broodstock for Grays River was constructed in 1996 by the Sea Resources hatchery from spawners collected in Willapa Bay tributaries (Steve Schroder, WDFW, pers. comm.). After the first year the broodstock was reconstructed from spawners collected from tributaries of the Grays River to promote the native genetic diversity in Grays River spawners. We include temporal samples from WDFW (archived genotypes) as well as contemporary samples from the Grays River in the analysis to examine temporal stability in a supplemented population. Several other collections included small collections from multiple years, especially for populations of low abundance. We tested for temporal stability in genetic relationships and then combined temporal collections into single collections to facilitate comparisons among tributaries.

Genotyping

Genotypes were processed for 1031 individuals from 18 collections (Table 2), at 16 microsatellite loci (Table 3). DNA was extracted with a silica membrane protocol following manufacturer's instructions (Macherey-Nagel). Microsatellite alleles were PCR-amplified using fluorescently labeled primers (see Table 3 for detailed PCR (polymerase chain reaction) information). Primers had a poly-a tail added to reverse primers (indicated by "+a" after primer name) to stabilize the reaction. PCRs were conducted in 384 well plates in 5 µl volumes employing 1 µl template with final concentrations of 1.5 mM $MgCl_2$, 200µM of each dNTP, 0.1 µl Promega GoTaq and 1X Promega PCR buffer. For all multiplexes (PCR reactions conducted for more than one locus) we used a "touch-down" cycling protocol as follows: after initial two minute denature at 94°, 3 cycles consisting of denature at 94° for 30 seconds, annealing at 60° for 30 seconds (dropping annealing temperature one degree each cycle for three cycles), extension at 72° for 60 seconds were followed 36 cycles with an annealing temperature now 52° or 50° (see Table 3 for second annealing temperature per multiplex) with the final cycle followed by a 10-minute extension at 72°. PCR products were run on an ABI-3730 automated sequencer. Archived data had been run on an ABI-3100 automated sequencer (see Small et al. 2006). A subset of samples was run on both sequencers to standardize allele mobility data generated by the two platforms such that the data were compatible. Microsatellite alleles were scored and binned using GENOTYPER and GENEMAPPER software, both from Applied Biosystems.

Statistical tests

Genotypic data per collection (Table 2 and Table 4) were examined to gather basic population information about the nature of the collections. We used the computer program FSTAT2.9.3 (Goudet 2001) to measure Hardy-Weinberg equilibrium (HWE) (genotypic proportions expected in a randomly mating population) expressed as F_{IS} values, allelic richness (number of alleles per population corrected for sample size), gene diversity (expected heterozygosity corrected for sample size – heterozygosity describes whether individuals are likely to have two different alleles at a locus), and allele frequencies. These descriptive statistics help identify sampling error, data collection error, and mixed-population samples. We assessed linkage disequilibrium (non-random genotypic associations between all possible pairs of loci) with a permutation method implemented in GENETIX (Belkhir et al. 2004) with 1000 permutations. Linkage

disequilibrium may arise if the collection includes family groups or migrants, if there has been non-random mating, or if genetic drift is large due to small population size, or if loci are physically linked. As described above, in tributary groups with small collection sizes we combined some archived genotypic data with contemporary data and calculated statistics for combined data (see Table 4).

We estimated effective population sizes using linkage disequilibrium (LD, Waples 2006) in the program LDNe (Waples and Do 2008). Since collections were composed of mixtures of year classes, the values from these measures estimated the number of breeders (N_b) giving rise to the collection, rather than the actual effective population size (N_e).

Population comparisons

We used a series of tests to explore whether collections from the same tributary differ over time and whether spawners collected in different tributaries are different from each other. We tested for significant differences in genotypic distributions between temporal collections within tributaries and among tributaries using FSTAT. We examined temporal and spatial partitioning of variance with pairwise F_{ST} tests in GENETIX and an analysis of molecular variance (AMOVA) in ARLEQUIN3.001 (Schneider et al. 2000). In pairwise F_{ST} tests we evaluated whether genetic variance was significantly different from zero with 10,000 permutations. We used AMOVA to examine the amount of genetic variance partitioned among ecoregional groups (see Table 2 for ecoregional groups: Pacific Coast, Coastal (Col. R.), Cascade, and Gorge). In a separate AMOVA including only collections from the Gorge, we estimated temporal and spatial variance with collections arranged in temporal groups and in spatial groups to determine whether temporal variance was greater than genetic variance.

We used a Bayesian analysis implemented in the program STRUCTURE 2.2 (Pritchard et al. 2000) to estimate individual and regional genetic structure. STRUCTURE sorts individuals (or portions of individuals if they appear to have mixed ancestry) into a number of hypothetical clusters (K) in order to minimize Hardy-Weinberg disequilibrium and linkage disequilibrium in the clusters. The program calculates a likelihood value for the number of clusters, given the data, with the highest likelihood value among the number of K's tested indicating the number of genetically identifiable clusters in the data set. We used the program to partition the data set into regional clusters since resolution to the population level is problematic when F_{ST} values between populations are below 0.05 (Latch et al. 2006), as found in this data set (see F_{ST} results below). Analyses from K = 1-10 were conducted in 5 independent runs per K that allowed admixture with 100,000 burn-ins and 400,000 iterations (the burn-ins move the analysis away from starting conditions to prevent them from influencing the analysis).

We used assignment tests in the program GeneClass2 (Piry et al. 2004) to examine population and regional genetic distinction. If fish assign strongly to their population of origin then genetic variance is partitioned at the population level and if fish assign weakly to their home population but assign strongly to their region or run group then genetic variance follows regional structure. The assignment test assigns fish to the collection in

which they have the highest likelihood of occurring, based on the fish's genotype and allele frequencies of all collections included as a baseline data set (the fish is removed from its own collection before calculating allele frequencies). The baseline collections here included all the contemporary and archived chum salmon collections outlined in Table 1. Assignment likelihoods to each collection were added within regional groups to calculate a relative likelihood of assignment to regional groups. The relative likelihood value was the assignment likelihood to a regional group over the sum of the assignment likelihoods. We established a threshold to evaluate assignment likelihoods: relative likelihood values over 90% were accepted as a positive assignment and fish with relative likelihood of assignment below this threshold were considered unassigned.

Genetic distance and population structure

We examined basic population genetic structure using factorial correspondence analysis (FCA) implemented in GENETIX and a dendrogram. The FCA generates composite axes from the combination of allele frequencies that describes the most genetic variation in the data and plots population centers according to their allele frequencies. Populations that are genetically similar plot near each other and populations that are genetically different plot distantly from each other. For the dendrogram, pairwise chord distances (Cavalli-Sforza and Edwards 1967) among collections were generated from allele frequencies using GENDIST in PHYLIP (Felsenstein 1993). A dendrogram illustrating genetic relationships was constructed from pairwise chord distances using the neighbor-joining (NJ) algorithm in the program NEIGHBOR in PHYLIP. Collections that are genetically similar cluster together in a dendrogram. To test the repeatability of tree branching, we made 10,000 bootstrap replicates of the pairwise chord distances using SEQBOOT, tree topologies were created for all replicates using NEIGHBOR, and a consensus tree was produced using CONSENSE in PHYLIP.

Supplementation broodstock assessment

We used genetic results and ecological information to construct a table summarizing the potential of spawning aggregates as donor broodstocks for reintroductions into streams where chum salmon have been extirpated and for enhancement of weak stocks.

Results

Population statistics

Many samples (N = 756) amplified at six or more loci and were included in the study. These samples are indicated in Table 2: where there are two numbers separated by a slash, the first is the number of samples with sufficient data and the second is the total number of samples processed in the collection. Samples that amplified at fewer than six loci (N = 275) are an indication of poor sample quality. This is a common problem for chum salmon since tissue or scale samples are often collected from carcasses of varying freshness – in older carcasses the DNA is usually degraded since the carcass is in an advanced state of decay. DNA-digesting enzymes, bacteria and funguses all break down DNA into small pieces such that PCR amplification is impossible at most loci. We attempted to maximize PCR success with these samples by using higher concentrations of

template DNA and DNA polymerase. But in many cases the DNA quality was too poor and these efforts failed to produce genotypic data.

We combined some contemporary collections (Washougal and Lacamas, Abernathy and Germany) since Lacamas is a tributary of Washougal and the mouths of Abernathy and Germany are within two rkm of each other and individual collections were too small for meaningful analysis. In tests for Hardy-Weinberg equilibrium (HWE) no locus in a contemporary collection was significant for disequilibrium after corrections for multiple tests (corrected alpha = 0.05/304) and all contemporary collections were in HWE in tests over all loci (corrected alpha = 0.05/19, Table 1). Heterozygosity, expressed as gene diversity, averaged over 80% over all collections. Heterozygosity in contemporary and archived Cowlitz collections was greater than two standard deviations (1 SD = 0.015) above the mean, and heterozygosity in the Yaquina collection was two standard deviations below the mean. Allelic richness was two standard deviations above the mean in contemporary and archived Cowlitz chum salmon collections and values were similar between contemporary and archived collections from other tributaries. Similar genetic diversity patterns in the contemporary and archived collections indicated temporal stability in genetic diversity.

Genotypic disequilibrium tests over all collections showed that genotypes at each locus were independent, but linkage disequilibrium within collections suggested possible family groups, mixture from different brood year strengths, or genetic drift due to small population sizes. There were 120 locus pairs examined per collection and in Table 1 we report the percentage of tests in which 5% of the permuted values were greater than or equal to the actual linkage value, indicating a signal of linkage. With this level, most collections had a signal of linkage since more than 5% of the locus pairs appeared to be linked. Because we were conducting multiple simultaneous tests, we relaxed the criteria to 1% of the permuted values greater than the actual value as a signal of linkage and only two archived collections had a signal of linkage at the 1% level.

We grouped contemporary and archived collections of samples from the same tributary and separated out the chum salmon from the Cowlitz River that had been identified as possible summer-run chum salmon and recalculated collection statistics (Table 4). We also combined collections from Hamilton and Hardy creeks since they were undifferentiated. Statistics were mostly similar when collections were combined, except in the Cowlitz: the combined collection was out of HWE over all loci, the fall-run collection (contemporary and archived data) was in HWE and the summer-run collection (only archived data) was out of HWE. Further, the combined collections from Ives area and Hamilton/Hardy creeks were out of HWE. Allelic richness increased when collections were combined since the minimum number of individuals with complete genotypes in a collection increased in combined collections and allelic richness is standardized to that minimum number. In the analysis with combined collections allelic richness in Cowlitz fall-run chum salmon was still two standard deviations above the mean (now 10.56). Most of the combined collections had a signal of linkage and if we relaxed the criteria to only 1% of the permuted values greater than the actual value, then the following collections had significant linkage signals: Cowlitz summer-run chum,

Hardy/Hamilton creeks, Ives area, Grays River and Skamokawa Creek. We report both levels since combining slightly differentiated temporal collections can give rise to linkage signals, as suggested by high linkage at the 5% level for combined collections.

Effective population size calculations

We calculated N_e and its 95% confidence interval (1000 bootstraps) for each collection (temporal collections combined) with the linkage disequilibrium method (LD method, Waples 2006; Waples and Do 2008) with the lowest frequency allele set at 1% to avoid bias introduced by small collections (Figure 2, Table 4). These values are to be treated cautiously since we lacked brood year information on most samples: LD calculations should be conducted on single brood year samples and these samples included representatives from multiple brood years. Further, genotypic data was incomplete for many samples, particularly in the archived data, possibly distorting true linkage values.

N_e values ranged widely. Three values were below 50 (Yaquina, Washougal and Multnomah), suggesting small effective population sizes. Yaquina Bay is near the southern end of the distribution range for chum salmon in the Pacific Northwest (Johnson et al. 1997). The low N_e concurs with expectations for population size at the end of a species distribution. Chum salmon spawning in the Washougal River and in the Multnomah area may be limited by available habitat.

We calculated N_e for Cowlitz with all collections combined and with the fall-run and summer-run chum salmon separated (Figure 2, Table 4). N_e was significantly higher (and linkage was lower) in the Cowlitz fall-run sample than in the combined fall-summer sample, supporting differentiation of the fall-run chum salmon from the summer-run chum salmon in the Cowlitz River – higher linkage in combined sample supported that there were two genetically differentiated groups in the combined sample. The smaller N_e calculated for the summer-run chum agrees with the lower abundance of the early run chum salmon in the Cowlitz River (S. Schroder, WDFW, pers. comm.).

The Sea Resources brood stock was constructed from Willapa Bay chum salmon yet their N_e value is significantly lower than the N_e for the Willapa collection (Figure 2). Hatchery broodstocks often have lower effective population sizes than their source population since the hatchery bypasses natural mate selection and in the hatchery there may be unequal sex ratios or unequal family sizes or few individuals used as hatchery broodstock (Ryman and Laikre 1991, Araki et al. 2007)

Genetic variance patterns among populations

Pairwise genotypic and F_{ST} tests indicated that most chum salmon collections from the same spawning site or tributary in different years were not genetically differentiated (data not shown). As described in Small et al. (2006), the Grays River collection from 2001 (first year of hatchery returns), were different from other Grays River collection: these were derived from Willapa broodstock and subsequent collections were derived from Grays River broodstock. Other cases of temporal variation were attributed to comparing collections composed of small samples from numerous years (eg. the contemporary Skamokawa collection included samples from five collections years, all too small to

analyze independently and combined into a single sample, also 13% of the genotypic data were missing which may impact comparisons). When temporal collections were grouped by location, pairwise genotypic and F_{ST} tests showed that collections were more similar to other collections within their ecoregional group than to collections from other regions (Table 5). Ten of the pairwise comparisons differed between the test methods (results with asterisks or exclamation marks in Table 5). In results with an asterisk (N = 8) the genotypic test indicated no differentiation and the F_{ST} test was significant and in results with an exclamation mark (N = 2), the genotypic test was significant and the F_{ST} test was not significant. These differences are likely due to missing data since GENETIX uses all genotypes in its computations while FSTAT uses only complete genotypes within collections and many collections were missing data. Also, genotypic tests tend to be very sensitive and may show significant differences when there are differences only at a single locus.

The pairwise comparison between the I-205 and Ives areas was one where the pairwise genotypic test was not significant and the pairwise F_{ST} value (0.003) was small but significantly different from zero (Table 5), suggesting weak but significant genetic variance between the collections. We examined this relationship more closely with an AMOVA with the temporal collections organized by location (I-205 and Ives areas) and with the collections organized by collection time (2000–2002, 2007-2008). There was more variance among collections when organized by location (F_{CT} = 0.0123, $p > 0.05$) than by collection time (F_{CT} = -0.017, $p > 0.05$), but neither value was significant, suggesting weak differences. We also calculated HWE values with the collections separated by location and with them together. The combined value was significant (F_{IS} = 0.031, $p = 0.0001$) due to including the 2002 Ives area collection, which was missing most of the data at two loci. Without the 2002 Ives area collection, the combined HWE value was not significant (F_{IS} = 0.011, $p = 0.033$, 95% CI = -0.00331 - 0.02168). Since a combined collection was in HWE (without the problematic collection), this supports a true lack of genetic distinction between these mainstem Columbia River spawner aggregations.

The pairwise comparisons between Ives area and Hamilton/Hardy were another case where the tests conflicted (Table 5). We examined this relationship with an AMOVA analysis and found that there was more genetic variance among collections within locations (F_{SC} = 0.0044, $p = 0.003$) than between locations (F_{CT} = 0.0001, $p > 0.05$). Some of this variation within location is likely due to high missing data (13% in the Ives area collection and 14% in the Hamilton/Hardy collection) and we concluded that these spawner groups were undifferentiated.

The Pacific Coast region had the most variance among collections within individual regions (avg. pairwise F_{ST} = 0.016, as compared to Coastal (Col. R.) (0.002), Cascade (0.012), and Gorge (0.002) - higher F_{ST} or variance indicates more genetic distinction among collections). The Pacific Coast region included collections from geographically distant tributaries that were genetically distinct from each other (Figure 1, Table 5, see dendrogram described below). The Coastal (Col. R.) region collections were mostly similar to each other. The collection from Abernathy and Germany creeks was the most

divergent of the Coastal (Col. R.) collections (Table 5) and was differentiated from all but the Elochoman collection. In the Cascade region, the Lewis collection was most similar to the Cowlitz fall-run collection and the Cowlitz summer-run collection was most similar to the Cowlitz fall-run collection, although they were significantly different. In the Gorge region, Hamilton/Hardy and I-205 collections were differentiated from other Gorge collections (except Bonneville and Washougal/Lacamas) in pairwise F_{ST} tests but not with pairwise genotypic tests. As described above, differences in test results were likely due to missing genotypic data (eg. there were 40 Washougal/Lacamas samples but only 30 with full genotypes: GENETIX used 40 samples and FSTAT used 30 samples in pairwise tests).

We used AMOVA to calculate variance among regional groups of collections. There was a significant amount of variance among ecoregional groups ($F_{ST} = 0.0166$, $p < 0.05$). In an AMOVA analysis of only the Gorge collections we organized temporal collections into three temporal groups (1990's, 2000-2002, and 2007-2008) and two spatial groups (one centered around the I-205 area and the other centered around Ives area, see Figure 1). There was less genetic variance in the temporal analysis ($F_{CT} = -0.0003$, $p = 0.75$) than in the spatial analysis ($F_{CT} = 0.0006$ $p = 0.04$), supporting weak but significant genetic structure among two spawner aggregates in the Gorge region. As described above, close examination of pairwise relationships among spawner groups revealed little differentiation but there was weak differentiation when summed over all collections in the two aggregates.

Some archived collections were missing most of the data at several loci. We removed these individuals before running the STRUCTURE analysis to minimize perturbations to the analysis from missing data. The STRUCTURE analysis supported the four ecoregional groupings (Figure 3). Results from Figure 3 are summarized in Table 6 – this presents the membership in each genetic cluster for each population averaged over all individuals in the collection. The summary is an average of the bars of color for each individual in Figure 3 and each color is associated with a single genetic cluster. STRUCTURE divides alleles in the dataset into genetic clusters or population groups, minimizing disequilibrium in the clusters, and may assign portions of an individual to different clusters if its alleles at some loci are more common in different population clusters (this is seen in Figure 3 as an individual with two or more colors in its bar of color, and the proportion of each color corresponds to the proportion of ancestry in each genetic cluster). The likelihood values for the analysis reached an asymptote around K = 4 and continued to increase slowly until K = 6, and then decreased (data not shown). At K = 3, STRUCTURE clustered the Cascade and Pacific Coast individuals together and the Coastal (Col. R.) and Gorge individuals each occupied their own cluster (data not shown). At K = 4, individuals were partitioned into four ecoregional clusters (Figure 3). While we know there is further structure in the dataset beyond K = 4 (since some collections within regions were genetically distinct), as K increased beyond 4, individuals subdivided among clusters since variance among increased number of clusters was below the resolving threshold for the program. Thus we present results only for K = 4 since this was the most likely number of genetic groups and described ecoregions.

Some interesting relationships were suggested in the STRUCTURE analysis. Although Abernathy/Germany creeks are included in the Coastal (Col. R.) ecoregional group, STRUCTURE assigned an average of 37% of their ancestry to the cluster occupied primarily by Cascade ecoregion individuals (Figure 3 and Table 6) and several individuals from Abernathy/Germany creeks were mostly assigned to the Cascade cluster (individual color bar was mostly purple, the color associated with the Cascade ecoregion). This pattern corresponds to differentiation patterns indicated in the pairwise F_{ST} tests (Table 5: there was the least genetic variance between Abernathy/Germany and Cascade collections). STRUCTURE may also have identified strays in some collections. For instance, one individual in the Elochoman collection had a red color bar and two had purple color bars, which indicated that their primary ancestry was in clusters occupied by individuals from the Gorge (red) and Cascade (purple) ecoregions, respectively. Low pairwise F_{ST} values also support gene flow among tributaries and among ecoregions through occasional strays.

GeneClass assignment test

GeneClass and STRUCTURE both assign individuals to populations or genetic clusters. However, they differ in that GeneClass calculates assignment likelihoods for the whole individual to user-defined baseline populations, whereas STRUCTURE uses no *a priori* knowledge about populations and may assign portions of an individual to different clusters or populations as described above. We use both analyses to examine relationships among populations and ecoregions from different perspectives.

GeneClass assignments further supported ecoregional distinctions. Assignments to individual populations were low (Table 7), but of the fish that did assign, over 70% in each regional group assigned correctly to their region (Table 8), indicating that genetic diversity follows ecoregional structure. Assignments to individual populations are low if populations are genetically similar such that a fish has similar likelihoods of assigning to more than one population. In that case, the relative likelihood of assignment would be below our threshold of 90% since the highest and next highest assignment likelihoods would be similar (with a 90% value, the highest assignment is at least 90 times more likely than the next most likely assignment). We used a 90% relative likelihood score as the cutoff value for a positive assignment since previous work demonstrated that this threshold is useful for striking a balance between type I and type II errors (Small, unpublished data). For the Cascade region fish with positive assignments that "mis-assigned" or assigned to another ecoregion, more of these fish misassigned to the Gorge ecoregion, suggesting that fish en route to the Gorge region stray into the Cascade tributaries. Fish with assignment values below 90% were considered unassigned. Using this criterion, roughly 60% of the fish from the Cascade region were unassigned. Unassigned fish can happen by chance if the fish has alleles that are common in more than one region, and thus similar assignment likelihoods to two regions, or it may be a signal that the fish has ancestry in more than one region from gene flow among regions. Fish may also assign weakly if the baseline population is small and has only sampled part of the genetic variation of the population.

Genetic clusters identified in FCA and dendrogram

Population centers formed roughly four basic clusters in the factorial correspondence analysis plot (Figure 4). Gorge collections clustered on the left side of the first axis and Coastal (Col. R.) collections clustered on the right side. The Pacific Coast collections plotted near the center of the first axis and Abernathy/Germany creeks, from the Coastal (Col. R.) region superimposed on the Pacific Coast cluster – Abernathy/Germany creeks was in a distinctly different space along the third axis, which is not shown in the plot. The Cascade collections plotted in the upper region of the second axis, with the Cowlitz summer-run chum collection at the top of the second axis.

The consensus dendrograms identified four major clusters of populations with high bootstrap support (Figure 5a and Figure 5b) and these clusters corresponded to the ecoregional groups. Figure 5a shows the dendrogram with uncombined collections listed in Table 2 and Figure 5b shows the dendrogram with combined collections, listed in Table 4. Bootstrap values for regional branches were lower in Figure 5a, indicating more variance with smaller collection sizes in the uncombined collections. Bootstrap values for regional branches were uniformly high in Figure 5b, indicating higher stability with larger collection sizes. Abernathy/Germany creeks joined the branch with the Coastal (Col. R.) collections near the center of the dendrogram and bootstrap support (69%) was weak for its membership in the cluster. This intermediate placement was indicated in the significant genetic variance between Germany and other Coastal (Col. R.) collections in the pairwise F_{ST} tests and the ancestry shared with the Coastal (Col. R.) and Cascade clusters in the STRUCTURE analysis. All analyses support a genetic profile intermediate to the Coastal (Col. R.) and Cascade genetic groups for Abernathy/Germany creeks. The dendrogram also suggested genetic structure within regional clusters: Washougal and it's tributary Lacamas, , form a separate sub-branch within the Gorge branch; Nehalem and Yaquina clustered together within the Pacific Coast branch and long branch lengths beyond the node and high pairwise F_{ST} values show that they are differentiated but the clustering indicates a relationship fostered by geographic proximity.

Re-introduction and Supplementation broodstock assessment

We developed a decision-making chart to identify potential donor stocks for reintroduction programs and enhancing depauperate populations (Table 9). We used the N_e calculated for spawner aggregations in this study and estimated population sizes for other tributary groups. Where N_e exceeded 200, a subset of spawners might be diverted to hatchery production for reintroduction or enhancement programs in tributaries within the same ecoregion. Where N_e was between 50 and 200, spawners might be diverted to hatchery production to supplement natural production (see discussion for protocols) within the tributary of origin. In spawner collections where N_e was larger than census size, possibly due to lack of abundance data (eg. Cowlitz), we recommended that local spawners be used only for supplementation and not as donor broodstock for reintroduction.

Discussion

Documenting patterns in genetic diversity are an important component of recovery planning for endangered species. Mapping the distribution and abundance of the extant genetic biodiversity sets the foundation for devising recovery strategies and provides tools for monitoring and assessing the success of recovery actions (Small et al. 2009). Recovery plans for endangered chum salmon in the lower Columbia River are currently underway (LCFRB 2004, 2010), and WDFW is developing a population enhancement strategy for LCR chum populations To guide and support recovery, we assessed genetic structure in LCR chum salmon throughout three ecoregions and examined temporal stability in small, endangered populations. This genetic analysis also provided guidance for developing broodstocks for reintroducing chum salmon in tributaries where they have been extirpated and developing enhancement broodstocks for chum salmon populations in tributaries where abundance is below recovery goals.

Similar to genetic structure presented by Small et al. (2006), chum salmon genetic variation showed strong ecoregional structure in the Columbia River and its tributaries, and on the Pacific Coast. Gene flow was greatest within ecoregion and minimal among ecoregions. This information will be useful for recovering imperiled or extinct chum salmon populations since it indicated that chum salmon genetic structure is at the ecoregional level. Broodstocks should thus be developed within ecoregions to recover populations within the ecoregion whenever possible. Genetic structure was weaker within ecoregions and in some areas accurate and precise assessment was complicated by missing data. But where local genetic structure exists, enhancement broodstocks should be developed from local spawners to support endangered populations and foster any local adaptation that may exist.

The recovery plan directs two uses of conservation hatchery programs: reintroduction and supplementation (LCFRB 2004, 2010). The WDFW population enhancement strategy for LCR chum salmon should consider following a modification of programs implemented in Hood Canal summer-run chum salmon recovery (WDFW and PNPTC 2007). Where the Hood Canal program restricted use of reintroduction broodstocks to no more than two tributaries where chum salmon populations had been extirpated, potential reintroduction broodstocks are more limited in the LCR and may need to be deployed in more than two tributaries. Further, the Hood Canal program specified that only local broodstock be used for supplementation, while LCR supplementation broodstocks may need to be used in multiple tributaries. Our results show little to no genetic variation among spawner aggregates within ecoregion and many aggregates have few fish returning to draw upon for broodstock. Thus, it may be more effective to use a robust supplementation broodstock from a geographically distinct, but not genetically distinct, spawner aggregate than to construct it from a few returning local spawners.

Recovery in the Cascade ecoregion is somewhat problematic; N_e was higher than might be expected since some years no chum salmon are detected returning to the tributaries in this area. Fish return in fall when flows and turbidity levels can be high. In the past, WDFW personnel often conducted only a few tributary spawning ground surveys during

the expected migration periods and only in areas where chum salmon were likely to be found spawning. Most chum salmon in this ecoregion have been found incidentally when surveying for other salmonids (mainly coho and Chinook salmon). Collecting sufficient local broodstock for supplementation will be difficult in the Cascade ecoregion. Therefore, we considered the possibility of using supplementation/reintroduction broodstock from a different ecoregion. Although salmonids are regionally adapted and the ecoregions differ in habitat characteristics, assignment tests suggest that there is some straying from other ecoregions into the Cascade tributaries. Straying was slightly higher from Gorge spawner aggregates and those populations can provide a more abundant source for a supplementation brood stock. Currently, the ongoing reintroduction program for Duncan Creek is collecting broodstock from the Gorge ecoregion, using fish from mainstem Columbia River spawning areas near Bonneville Dam (Ives, Multnomah, Horsetail and St. Cloud areas; Hillson 2009). A broodstock constructed from spawners in the I-205 area, the other large spawning population in the Gorge ecoregion, would provide the nearest genetic and geographically located source for reintroduction programs in Cascade ecoregion tributaries.

Although LCR chum salmon populations are at very low numbers, genetic drift and temporal variation were minor and contemporary genetic structure was similar to LCR chum salmon population structure presented by Small et al. (2006). There were some differences from genetic structure described in a larger study by Phelps et al. (1994). Phelps et al. (1994) showed an ambiguous genetic relationship among collections from the Satsop River and Bitter Creek in Willapa Bay: the collections clustered on the same branch in their dendrogram, but plotted distantly in the multidimensional scaling plot. Our study found a close and consistent relationship between these collections. An interesting genetic relationship between the Abernathy and Germany creeks collection and Cascade ecoregion collections, suggested by Small et al. (2006), remained consistent and was strengthened with a larger contemporary collection. Genetic analyses portrayed Abernathy/Germany creeks as genetically intermediate between the Coastal (Col. R.) and Cascade ecoregional groups. Although the tributaries originate in the Willapa Hills, the mouths of Abernathy and Germany creeks are located roughly halfway between the mouths of the Elochoman River in the Coastal (Col. R.) ecoregion and the Cowlitz River in the Cascade ecoregion. Abernathy/Germany creeks may function as a genetic stepping stone between the Coastal (Col. R.) and Cascade regions if fish destined for Cascade tributaries occasionally stray into Abernathy/Germany creeks and spawn.

In summary, we recommend following a multi-tiered population enhancement strategy for LCR chum salmon. Where artificial propagation is deemed necessary, local broodstocks should be used if they are sufficiently abundant. In tributaries where populations have been extirpated or where abundance is too low to provide for adequate broodstock collection, we suggest collecting supplementation broodstock from a geographically distinct, but not genetically distinct, abundant spawner aggregate within the ecoregion. For the Cascade ecoregion, where there is no robust local broodstock, we consider the pros and cons of using a local broodstock from a population of low abundance versus developing a robust broodstock from another ecoregion: fostering local adaptation with a local broodstock may be cancelled by enhancing genetic drift with a

small broodstock. Another option is to use a robust broodstock from a different ecoregion, specifically from the Gorge ecoregion. If a non-ecoregional broodstock is to be used for the Cascade ecoregion, we recommend the I-205 population since it is the geographically closest abundant brood source.

Acknowledgements

This project was funded by a grant from the Bonneville Power Administration and supported by Washington State General Funds. We thank all the WDFW personnel involved in sampling and thank the Oregon Department of Fish and Wildlife for providing tissue and scale samples from Oregon chum salmon. Steve Schroder (WDFW) provided life history and physical history information for chum salmon in Washington.

Literature Cited

Araki, H., Waples, R.S., Ardren, W.R., Cooper, B., and Blouin, M.S. 2007. Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. Mol. Ecol. 16(5): 953-966.

Banks, M.A., Blouin, M.S., Baldwin, B.A., Rashbrook, V.K., Fitzgerald, H.A., Blankenship, S.M., and Hedgecock, D. 1999. Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tschawytscha*). J. Hered. 90: 281-288.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2001. genetix 4.02, logiciel sous Windows TM pour la genetique des populations. Laboratoire Genone, Populations, Interactions: CNRS UMR 5000, Universite de Montpellier II, Montpellier, France.

Cairney, M., Taggart, J.B., and Hoyheim, B. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (Salmo salar L.) and cross-species amplification in other salmonids. Mol. Ecol. 9: 2175–2178.

Cavalli-Sforza, L.L. and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. Am. J. Hum. Gen. 19: 233-257.

Felsenstein, J. 1993. Phylogeny inference package (PHYLIP) Version 3.5c. University of Washington, Seattle.

Fulton LA (1970) Spawning areas and abundance of steelhead trout and coho, sockeye, and chum salmon in the Columbia river basin – past and present. NOAA Tech. Report NMFS-SSRF 618, 37 p.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 293). Available from http://www.unil.ch/dee/page6759_en.html (Accessed May 25, 2008).

Hillson, Todd D. 2002. Re-Introduction of Lower Columbia River Chum Salmon into Duncan Creek Annual Report for 2002, Report to Bonneville Power Administration, Contract No. 00007373, Project No. 2001-05300, 63 electronic pages (BPA Report DOE/BP - 00007373 -1). http://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=00007373-1

Hillson, Todd D. 2009. Re-Introduction of Lower Columbia River Chum Salmon into Duncan Creek, 2006-2007 Annual Report, Project No. 200105300, 83 electronic pages, (BPA Report DOE/BP-P112055) http://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=P112055

HSRG 2008. Summary of HSRG Findings for Chum Populations in the Lower Columbia River and Gorge. Draft Chum Population Reports and Summary of Key Findings.http://www.hatcheryreform.us/prod/Portals/_default/Documents/HSRG_Chum%20SummaryRpt_031208.pdf

Johnson, O.W., W.S. Grant, R.G. Kope, K. Neely, F.W. Waknitz, and R.S. Waples. 1997. Status review of chum salmon from Washington, Oregon, and California. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-32, 280 p.

LCFRB. Lower Columbia Fish Recovery Board. 2004. Interim Lower Columbia Salmon Recovery and Fish and Wildlife Subbasin Plan. 2004. Kelso, WA. http://www.lcfrb.gen.wa.us/December%202004%20Final%20%20Plans/lower_columbia_salmon_recovery_a.htm

LCFRB. Lower Columbia Fish Recovery Board. 2010. Washington Lower Columbia Salmon Recovery and Fish and Wildlife Subbasin Plan. 2010. Kelso, WA. http://www.lcfrb.gen.wa.us/Recovery%20Plans/March%202010%20review%20draft%20RP/RP%20Frontpage.htm

May, Christopher, David Geist. 2007. Grays River Watershed and Biological Assessment, 2006 Final Report, Project No. 200301300, 551 electronic pages, (BPA Report DOE/BP-00000652-38)

McElhany, P. M. Chilcote, J. Myers and R Beamesderfer. 2007. Viability Status of Oregon Salmon and Steelhead Populations in the Willamette and Lower Columbia Basins. http://www.nwfsc.noaa.gov/trt/wlc/trt_wlc_psr2007.cfm

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, N. Y.

Olsen, J.B., Wilson, S.L., Kretschmer, E.J., Jones, K.C., and Seeb, J.E. 2000. Characterization of 14 tetranucleotide microsatellite loci derived from sockeye salmon. Mol. Ecol. 9: 2185-2187.

Phelps, S.R., LeClair, L.L., Young, S., and Blankenship, H.L. 1994. Genetic diversity patterns of chum salmon in the Pacific Northwest. Can. J. Fish. Aquat. Sci. 51(Supplement 1): 65-83.

Piry, S., Alapetite, A., Cornuet, J.–M., Paetkau, D., Baudouin, L., and Estoup, A. 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. J. Hered. 95(6): 536-539.

Pritchard, J.K., M. Stephens, P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Raymond, M. and Rousset, F. 1995. GENEPOP (Version 3.2): population genetics software for exact tests and ecumenicism. J. Hered. 86: 248-249.

Rexroad, C.E. III, Coleman, R.L., Gustafson, A.L., Hershberger, W.K., and Killefer, J. 2002. Development of rainbow trout microsatellite markers from repeat enriched libraries. Mar. Biotechnol. 4: 12–16.

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

Ryman, N., and Laikre, L. 1991. Effects of supportive breeding on the genetically effective population size. Conserv. Biol. 5: 325-329.

Schneider, S., Roessli, D., and Excoffier, L. 2000. Arlequin version (2.001): a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. Available online at http://lgb.unige.ch/arlequin/ (Accessed July 25, 2007).

Scribner, K.T., Gust, J.R., and Fields, R.L. 1996. Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. Can. J. Fish. Aquat. Sci. 53: 833-841.

Small, M.P., Beacham, T.D., Withler, R.E., and Nelson, R.J. 1998. Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia. Mol. Ecol. 7: 141-155.

Small, Maureen P., A.E. Frye, J.F. Von Bargen, , and S.F. Young. 2006. Genetic structure of chum salmon (*Oncorhynchus keta*) populations in the lower Columbia River: are chum salmon in Cascade tributaries remnant populations? Conservation Genetics (2006) 7:65-78.

Small, M.P., K. Currens, T. H. Johnson, A. E. Frye and J. F.Von Bargen. 2009. Impacts of supplementation: Genetic diversity in supplemented and unsupplemented populations of summer chum salmon in Puget Sound. Canadian Journal of Fisheries and Aquatic Science 66(8):1216-1229.

Smith, C.T., Koop, B.F., and Nelson, R.J. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Mol. Ecol. 7: 1613-1621.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv. Gen. 7: 167-184.

Waples, R.S., and Do, C. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Mol. Ecol. Notes 8(4):753-756.

Washington Department of Fish and Wildlife (WDFW) and Point No Point Treaty Council (PNPTC) Tribes. 2000. Summer Chum Conservation Initiatives: An implementation plan to recover summer chum salmon in the Hood Canal and Strait of Juan de Fuca region. Available on the Web at: http://wdfw.wa.gov/fishing/salmon/chum/reports.html

Washington Department of Fish and Wildlife (WDFW) and Point No Point Treaty Council (PNPTC) Tribes. 2007. Summer Chum Salmon Conservation Initiative: An implementation plan to recover summer chum in the Hood Canal and Strait of Juan de Fuca Region. Supplemental report No.7, Five year review of the summer chum salmon conservation initiative. 246pp. Available at http://wdfw.wa.gov/fishing/salmon/chum/reports.html

Washington Department of Fish and Wildlife website (2011), accessed Aug. 15, 2011 (http://wdfw.wa.gov/fishing/salmon/chum/columbia_river/index.html)

Williamson, K.S., Cordes, J.F., and May, B.P. 2002. Characterization of microsatellite loci in chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. Mol. Ecol. Notes 2(1): 17-19.

Figure 1. Map of the tributaries for archived and contemporary chum salmon collections. Map was modified from two previous maps, one created by Jim Shaklee (detailed lower Columbia River tributaries map) and the inset map of major drainages in Washington State was from Shelly Snyder and Darrell Pruit at WDFW. Yaquina Bay was located below the extent of the detailed Lower Columbia River tributaries map and is indicated approximately on the inset map. Creek mouths for Germany and Abernathy creeks are roughly two rkm distant and only Germany Cr. is indicated on map. Approximate former location of Celilo Falls is indicated, just beyond extent of the detailed map. Tributaries in lower Columbia ecoregions are circled and ecoregions are identified.

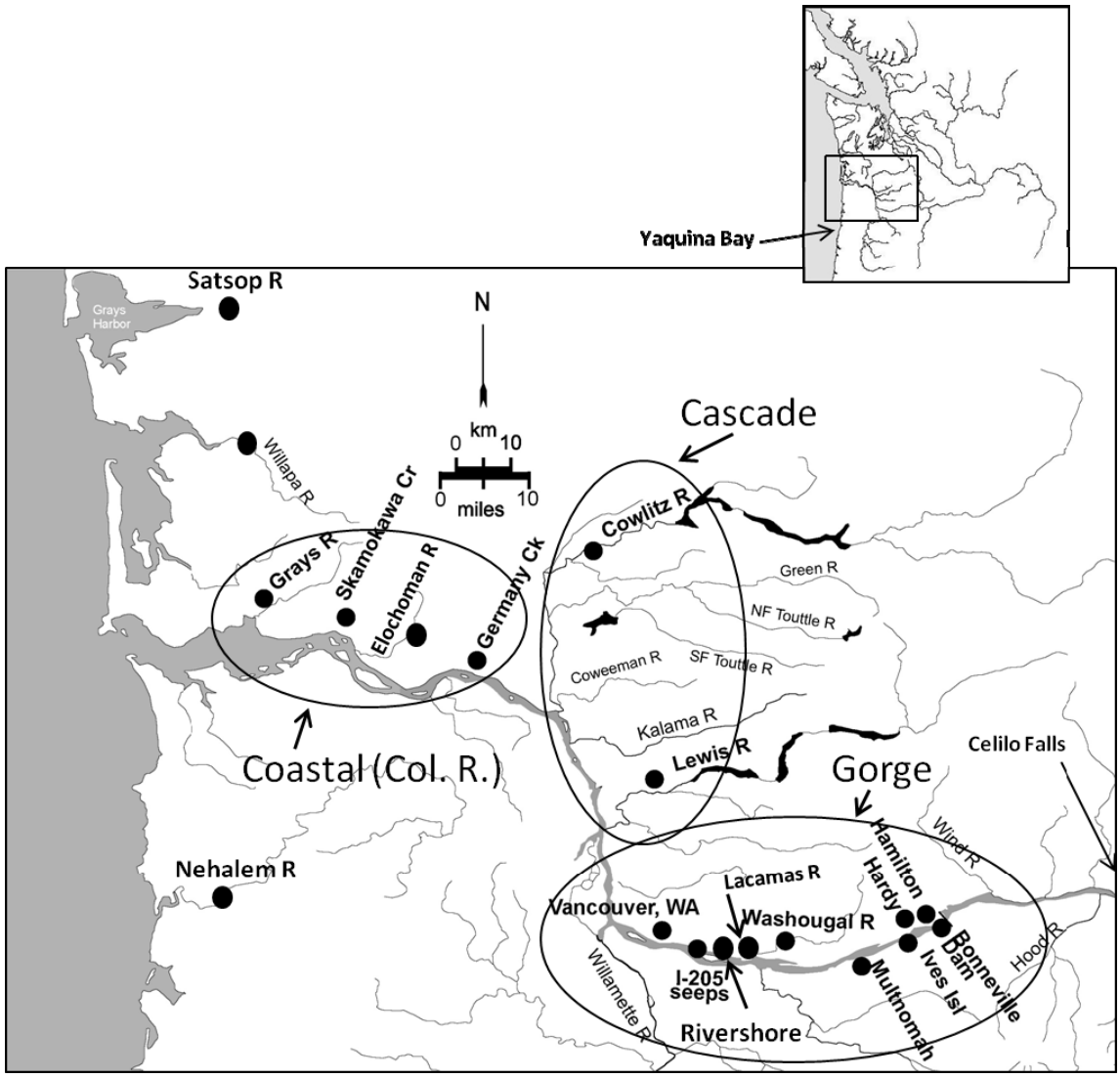

Figure 2. Graph of effective population sizes and their 95% confidence intervals calculated for collections (mixed brood years) using linkage disequilibrium (LDNe). Values were calculated for the Cowlitz collection with the summer and fall run fish combined (blue) and separated (red).

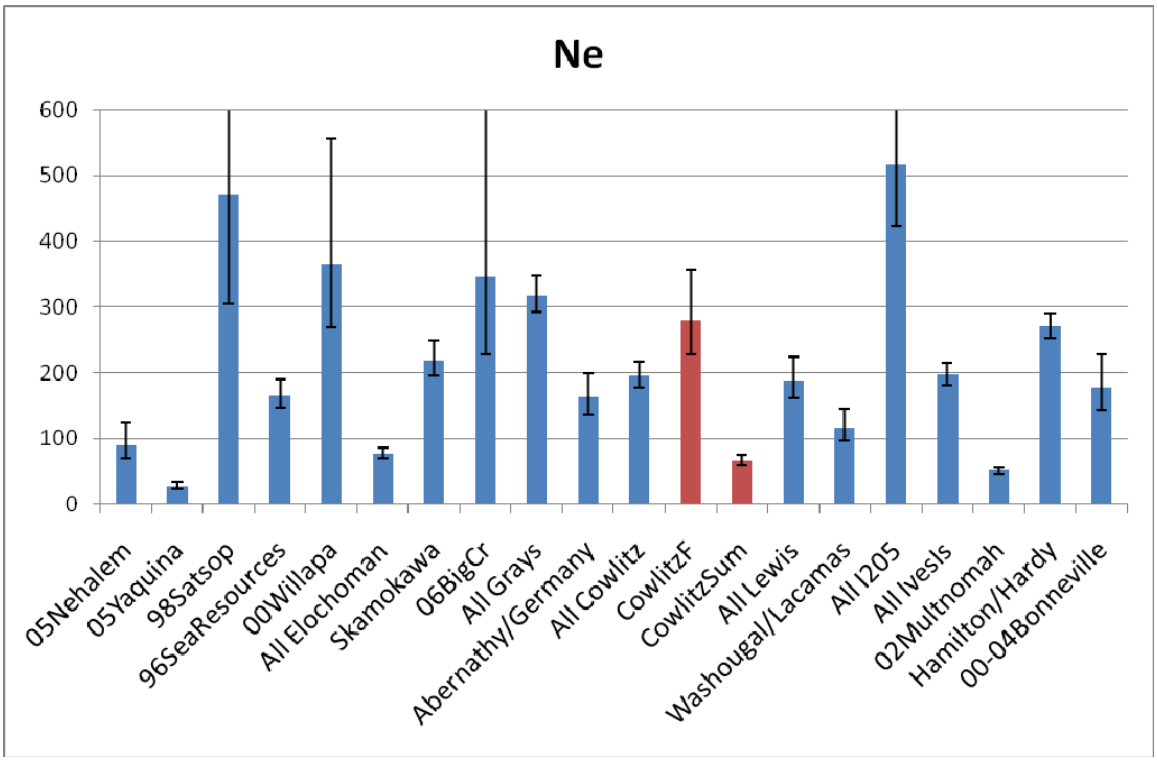

Figure 3. Individual ancestry values for combined contemporary and archived Lower Columbia River and Pacific Coast chum salmon collections from STRUCTURE analysis at K = 4. Each individual within a collection is represented by a bar of color (collections labeled and separated by black lines in far right, regions indicated at far left). The bar of color is associated with a cluster; single color suggests pure ancestry and multiple colors suggest mixed ancestry (ancestry in more than one cluster). Collections are decomposed into cluster memberships to the left of the combined color bar. Table 6 has average ancestries in each cluster over all individuals in each collection.

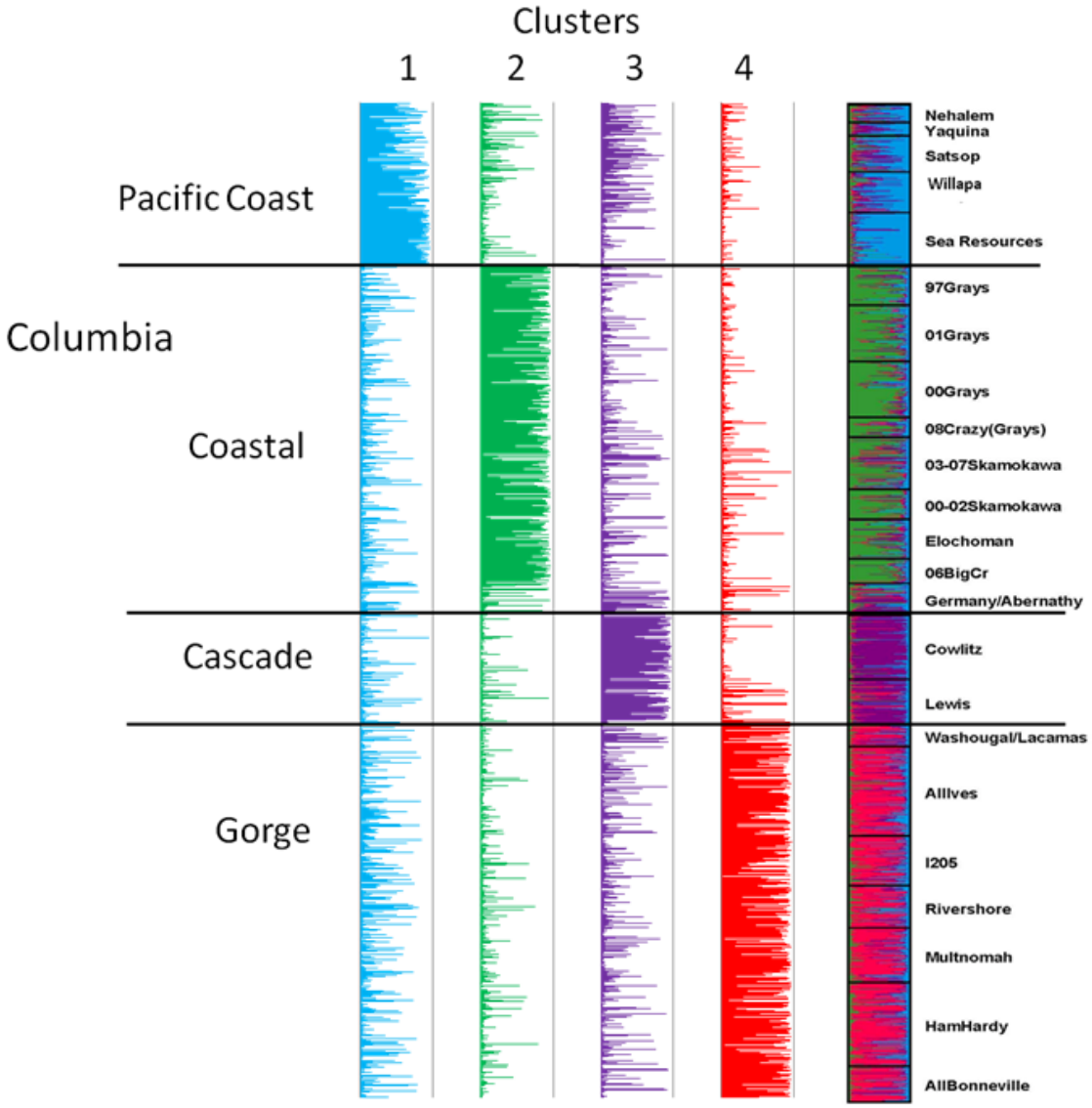

Figure 4. Factorial correspondence analysis plot of collection centers Lower Columbia chum salmon. Collection centers are identified by ecoregion except for Abernathy/Germany creeks which is identified individually as "Germany".

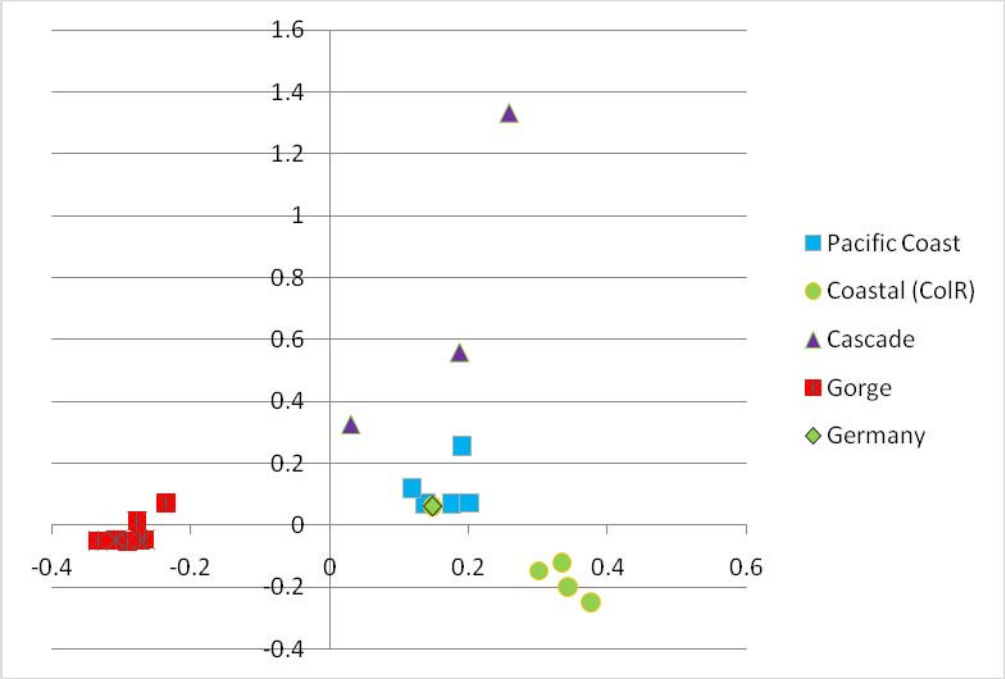

Figure 5a. Neighbor-joining dendrogram of genetic distances among chum salmon collections (see Table 2). Numbers at the nodes are bootstrap values or the percentage of trees in which the collections beyond the node occurred together in the consensus tree of 10,000 trees.

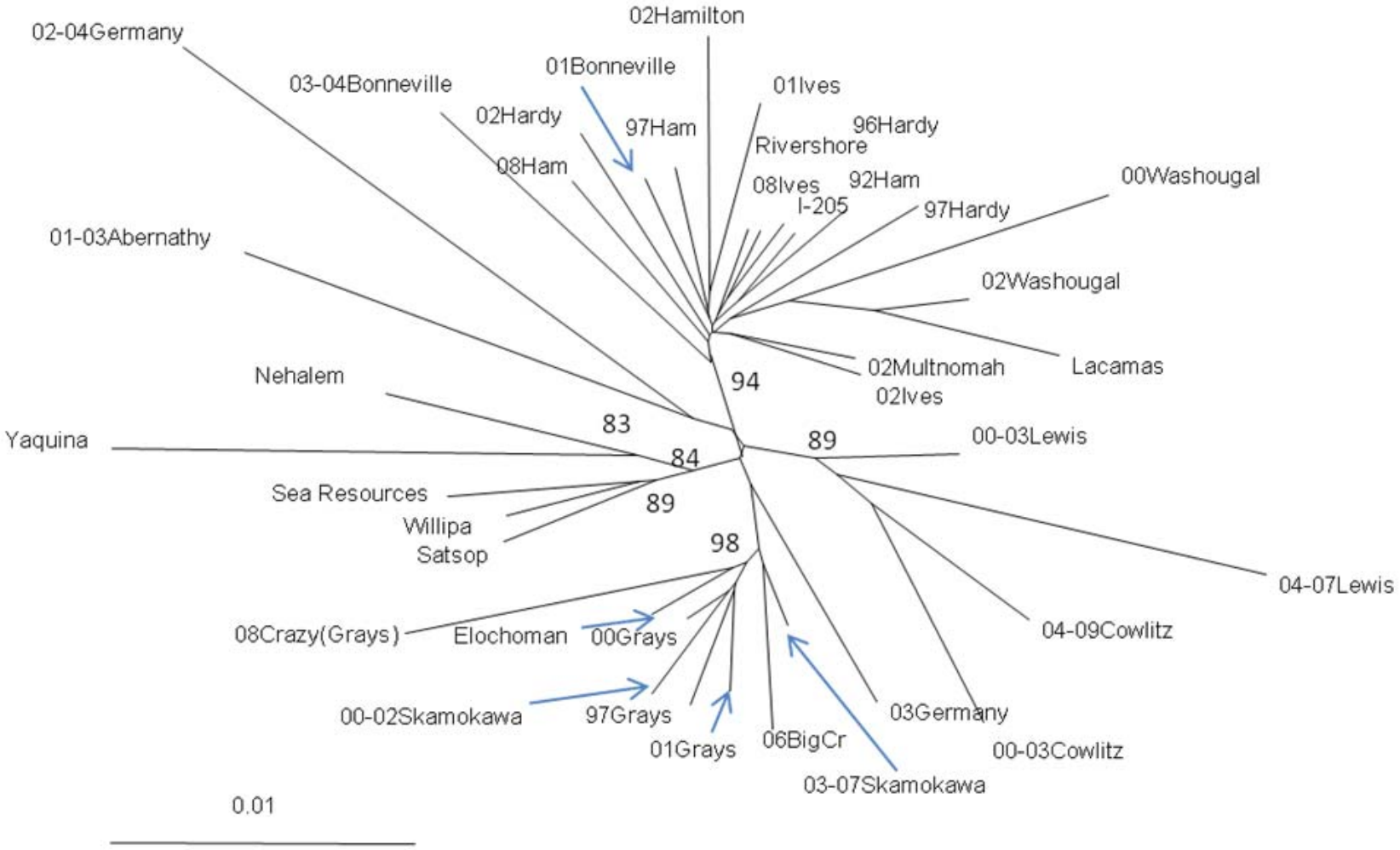

Figure 5b. Neighbor-joining dendrogram of genetic distances among combined chum salmon collections (see Table 4). Numbers at the nodes are bootstrap values or the percentage of trees in which the collections beyond the node occurred together in the consensus tree of 10,000 trees.

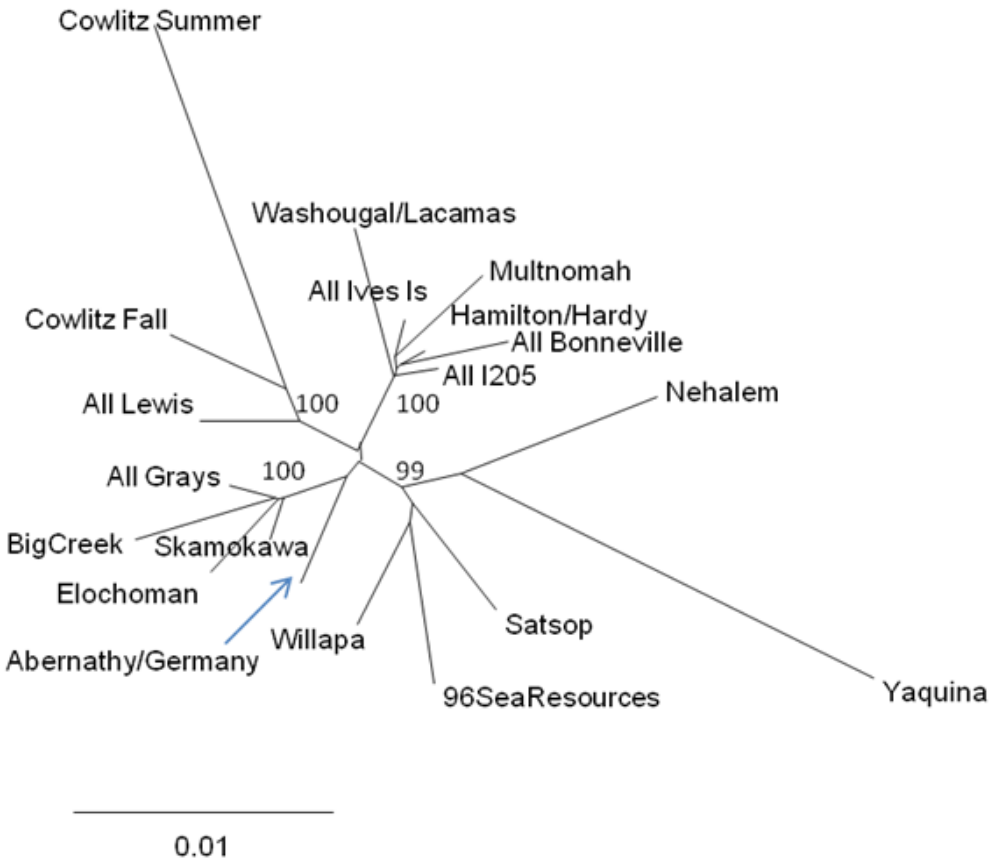

Table 1. Extinction Risk of Columbia River Chum Salmon Populations[1] as Identified by the Lower Columbia/Willamette TRT (HSRG 2008).

Populations	Extinction Risk
Coast Stratum	
Grays/Chinook (WA)	High
Elochoman (WA)	High
Mill/Abernathy/Germany (WA)	Very High
Youngs Bay Tribs. (OR)	Very High
Big Creek (OR)	Very High
Clatskanie (OR)	Very High
Scappoose (OR)	Very High
Cascade Stratum	
Cowlitz (WA)	Very High
Kalama (WA)	Very High
Lewis (WA)	Very High
Salmon (WA)	Very High
Clackamas (OR)	Very High
Sandy (OR)	Very High
Gorge Stratum	
Lower Gorge Tribs (including Washougal)	Very High/Medium
Upper Gorge Tribs	Very High/ Very High

[1] From Washington's Lower Columbia River Recovery Plan and McElhany et al. (2007) for Oregon populations

Lower Columbia chum salmon 2011, WDFW Molecular Genetics Lab

Table 2. List of chum salmon collections processed in the study and list of archived chum salmon data used in the study. Where there are two values under "N>6/N", the first value is the number of samples with six or more loci in their genotype – these were included in analyses and the second number is number of samples processed. Under "Ecoregion" is the ecoregion designation. Under "Gene Div" are gene diversity values or expected heterozygosity corrected for collection size. Under "Rich" are allelic richness values or average number of alleles per locus corrected for collection size of 5 fish (smallest number of complete genotypes in a collection). F_{IS} is the Hardy-Weinberg Equilibrium value followed by its P-value. Underlined P values were significant before corrections for multiple tests and red colored P-values were significant after corrections. Under "%Linkage" is the percentage of 120 locus pairs in linkage disequilibrium with a 5% or 1% threshold, values in red were higher than expected by chance (5% or greater).

WDFW collection codes								% Linkage	
Contemporary collections	Tributary	N>6 / N	Ecoregion	Gene Div	Rich	F_{IS}	Pval	5%	1%
98KI	Satsop	64 / 75	Pacific Coast	0.818	5.13	-0.007	0.716	3.33	0.83
00OB	Willapa	72 / 75	Pacific Coast	0.808	5.03	0.020	0.043	8.33	0.83
05OZ	Nehalem	31 / 60	Pacific Coast	0.805	5.02	-0.006	0.622	4.17	0.00
05PA	Yaquina	24 / 64	Pacific Coast	0.789	4.74	0.007	0.382	10.83	2.50
00KQ, 01KP, 01KS, 02KX, 02LD, 03IB, 03IG, 04IZ, 05HU, 05KF, 06DJ, 06FK	Elochoman	70 / 99	Coastal(ColR)	0.823	5.12	0.006	0.308	10.00	0.00
01KP, 02KX, 03IA, 02LE, 04JD	Abernathy+Germany	21	Coastal(ColR)	0.835	5.24	0.029	0.013	4.17	3.33
03IL, 04IZ, 05KS, 06FP, 07JI	Skamokawa	93 / 123	Coastal(ColR)	0.828	5.12	-0.011	0.848	10.83	4.17
08JS	CrazyCreek (Grays)	35 / 80	Coastal(ColR)	0.826	5.09	-0.021	0.824	5.00	0.83
06GA	Big Creek Hatchery	44	Coastal(ColR)	0.826	5.13	0.002	0.441	8.33	0.00
04AF, 05HU, 06DJ, 07FL, 08GD, 09GY	Cowlitz	46 / 49	Cascade	0.863	5.54	0.018	0.096	2.50	0.00
04JF, 05KR, 05KN, 06FN, 07JH	Lewis	18 / 35	Cascade	0.843	5.32	0.026	0.153	3.33	0.83
02LR, 03IM, 02OO, 03IJ, 04JJ	Washougal+Lacamas	22/41	Gorge	0.823	5.19	0.003	0.418	13.33	4.17
03IB, 04JU	Bonneville	16	Gorge	0.812	4.88	-0.006	0.622	7.50	3.33
07JQ	Rivershore (I-205)	75 / 80	Gorge	0.819	5.03	-0.011	0.821	10.00	1.67
08JJ	Hamilton	49 / 82	Gorge	0.819	5.06	-0.022	0.915	5.00	0.00
08JN	Ives area	75 / 80	Gorge	0.824	5.09	0.004	0.350	5.00	0.00
WDFW archived data									
96EC	Sea Resources	95	Pacific Coast	0.813	4.99	0.018	0.042	20.00	5.00
00KP, 01LC, 02LQ	Skamokawa	53	Coastal(ColR)	0.824	5.03	-0.007	0.672	5.00	0.00
97FT	97Grays	69	Coastal(ColR)	0.823	4.99	0.056	0.000	12.50	4.17
00NT	00Grays	99	Coastal(ColR)	0.821	5.04	0.014	0.087	7.50	4.17
01MD	01Grays	100	Coastal(ColR)	0.822	5.06	0.006	0.259	23.33	4.17
03IH	03Germany	35	Coastal(ColR)	0.838	5.30	0.043	0.006	6.67	0.83
00KX, 01LD, 01LE, 01KU, 02LC, 02LO, 03IK, 03IF	Lewis	62	Cascade	0.845	5.36	0.036	0.003	8.33	2.50
00TM, 01EH, 02GF, 02GX, 03CL	Cowlitz	68	Cascade	0.864	5.58	0.056	0.000	15.00	3.33
00IP, 01LM, 02ML	Ives Island	84	Gorge	0.809	4.89	0.139	0.000	21.67	8.33
02SO	02Multnomah	97	Gorge	0.814	4.97	0.022	0.053	8.33	0.00
00LD, 01LG, 02KY	Bonneville	46	Gorge	0.824	5.06	0.056	0.000	9.17	0.83
02LH	02Hardy	40	Gorge	0.805	4.90	-0.004	0.578	18.33	6.67
02ME	02Hamilton	29	Gorge	0.801	4.94	0.010	0.308	6.67	2.50
92HB	92Hamilton	90	Gorge	0.824	5.02	0.010	0.227	15.00	3.33
96FR	96Hardy	80	Gorge	0.815	5.01	0.032	0.004	7.50	2.50
97FR	97Hamilton	50	Gorge	0.813	4.99	0.069	0.000	8.33	2.50
97FS	97Hardy	36	Gorge	0.801	4.90	0.037	0.025	10.83	2.50
00KY, 00PT	I-205	88	Gorge	0.820	5.11	0.010	0.290	6.67	1.67

Table 3. Information for microsatellite multiplexes and loci including annealing temperature (°C) and primer concentration and color. References for primer sequences are under Citation. The +a indicated a poly-a tail on the primer. The size range of microsatellite alleles (in basepairs) and the number of alleles in this study are under "Size range" and "# alleles", respectively.

Multiplex	Locus	color	conc [µM]	Anneal T	Size range	# alleles	Citation
OkeB	One-102+a	6fam	0.140	60/50	219-330	25	Olsen et al. 2000
	One-114+a	hex	0.125		185-308	32	Olsen et al. 2000
	Ots-3M+a	ned	0.054		135-173	18	Banks et al. 1999
OkeC	Ots-1+a	6fam	0.090	60/50	152-249	37	Banks et al. 1999
	One-101+a	ned	0.070		126-281	39	Olsen et al. 2000
OkeG	One-108+a	6fam	0.086	62/52	167-396	54	Olsen et al. 2000
	Ots-103+a	vic	0.062		98-257	39	Small et al. 1998
OkeJ	One-106+a	6fam	0.080	60/50	174-326	38	Olsen et al. 2000
	Oke-3+a	hex	0.087		215-348	13	Buchholz et al. 2001
	Omy-1011+a	ned	0.063		202-275	19	Rexroad et al. 2002
OkeK	One-111+a	6fam	0.088	60/50	172-362	73	Olsen et al. 2000
	Ssa-419+a	hex	0.078		265-317	14	Cairney et al. 2000
	One-18+a	ned	0.072		169-199	11	Scribner et al. 1996
OkeL	Ots-G311+a	6fam	0.092	60/52	247-492	56	Williamson et al. 2002
	Oki-1+a	hex	0.099		194-262	16	Smith et al. 1998
	Ots-2M+a	ned	0.056		149-167	7	Banks et al. 1999

Table 4. Data for combined contemporary and archived chum salmon collections. Under "N (comb)" are the number of samples in the combined collection. Under "Gene Div" are gene diversity values or expected heterozygosity corrected for collection size. Under "Rich" are allelic richness values or average number of alleles per locus corrected for collection size of 15 individuals. F_{IS} is the Hardy-Weinberg Equilibrium value followed by its P-value. Underlined P values were significant before corrections for multiple tests and P-values in bold type were significant after corrections. Under "%Linkage" is the percentage of 120 locus pairs in linkage disequilibrium with a 5% or 1% threshold, values in red were higher than expected by chance (5% or greater). Under "N_e" is the effective population size calculated using a linkage disequilibrium method followed by the lower and upper bounds of the 95% confidence interval. The N_e values and the 95% confidence intervals are plotted in Figure 2. Statistics for the Cowlitz collection were calculated with the fall and summer chum salmon together (data in black) and separated (data in blue).

						% Linkage			N_e 95%	
	N (comb)	Gene Div	Rich	F_{IS}	Pval	5%	1%	N_e	minus	plus
Pacific Coast										
Nehalem	31	0.8049	10.21	-0.006	0.6186	4.17	0	89.7	70	122.8
Yaquina	24	0.7889	9.02	0.007	0.3819	11.67	3.33	27.1	22.4	33.7
Satsop	64	0.8179	10.65	-0.007	0.7172	3.33	0.83	469.6	305	978.9
SeaResources	95	0.8133	9.74	0.018	0.0416	18.33	3.33	165.2	146.1	189.3
Willapa	72	0.8080	10.30	0.020	0.0422	9.17	0.83	364.6	268.7	556.7
Columbia River										
Coastal										
Elochoman	70	0.8233	10.47	0.006	0.3079	9.17	0	76.4	69.1	85.1
Skamokawa	146	0.8283	10.28	-0.007	0.7914	13.33	6.67	218.8	194.6	248.8
BigCr Hatchery	44	0.8261	10.21	0.002	0.4420	6.67	0	345.6	228	688.2
Grays	303	0.8234	9.96	0.018	0.0011	24.17	9.17	318.2	292.5	347.7
Abernathy+Germany	56	0.8347	11.21	0.029	0.0137	4.17	3.33	162.4	136.7	198.5
Cascade										
Cowlitz	114	0.8646	12.45	0.040	**0.0000**	14.17	3.33	195.4	176.8	217.7
Cowlitz F	74	0.8566	12.38	0.020	0.0352	10.83	3.33	279.3	228.4	356.4
Cowlitz S	37	0.8634	12.19	0.066	**0.0000**	13.33	5.00	65.8	58.3	75.1
Lewis	80	0.8441	11.56	0.031	0.0031	8.33	3.33	187.7	161.1	223.6
Gorge										
Washougal+Lacamas	40	0.8235	10.75	0.003	0.4180	13.33	4.17	115.9	96.2	144.4
I-205	163	0.8177	10.16	0.005	0.2854	10.00	5.00	516	423.7	654.1
Ives area	159	0.8136	9.95	0.053	**0.0000**	11.67	3.33	196.6	181.1	214.3
Multnomah	97	0.8136	9.85	0.022	0.0528	9.17	1	51.2	46.9	56.1
Hamilton+Hardy	325	0.8168	10.03	0.025	**0.0000**	22.50	7.50	271.4	253.4	291.3
Bonneville	62	0.8216	10.05	0.041	0.0011	7.50	2.50	176.7	143.7	226.8

Lower Columbia chum salmon 2011, WDFW Molecular Genetics Lab

Table 5. Pairwise genotypic tests (from FSTAT) and pairwise F_{ST} values (from GENETIX) for combined chum salmon collections. The upper triangular matrix presents the pairwise F_{ST} values – all values were significant except for the values in bold and the lower triangular matrix presents the percentage of permuted values (out of 1000 permutations) that were larger than the actual value in the pairwise F_{ST} tests (pink boxed values NOT significant – more than 1% of the values were larger than the actual value). Numbers with asterisks in the lower triangular matrix indicate P-values for pairwise genotypic tests (calculated using FSTAT) where 100 % of the permuted F_{ST} values were smaller than the actual value but the pairwise genotypic test was not significant. Numbers with exclamation points in the lower triangular matrix indicate P-values for pairwise genotypic tests (calculated using FSTAT) where the pairwise genotypic test indicated a significant difference and the pairwise F_{ST} test was not significant. FSTAT uses only complete genotypes in its calculations and GENETIX uses the entire data set and results may vary when collections are missing data. See Table 4 for collection details, names were abbreviated to fit into table.

	Pacific Coast					Coastal					Columbia River: Cascade			Gorge					
	Nehalem	Yaquina	Satsop	SeaRes	Willapa	Eloch	Skamo	BigCr	All Grays	Aber/Germ	CowlitzF	CowlitzS	All Lewis	Wash/Laca	All I-205	All Ives	Multnomah	Ham/Hardy	All Bonne
Nehalem		0.0223	0.0120	0.0106	0.0121	0.0160	0.0165	0.0144	0.0166	0.0163	0.0301	0.0368	0.0275	0.0182	0.0185	0.0181	0.0220	0.0184	0.0186
Yaquina	0		0.0300	0.0303	0.0308	0.0299	0.0270	0.0280	0.0291	0.0257	0.0375	0.0362	0.0370	0.0310	0.0338	0.0295	0.0331	0.0302	0.0304
Satsop	0	0		0.0065	0.0054	0.0076	0.0124	0.0119	0.0119	0.0072	0.0210	0.0288	0.0158	0.0146	0.0119	0.0125	0.0149	0.0125	0.0126
SeaRes	0	0	0		0.0065	0.0125	0.0159	0.0158	0.0155	0.0127	0.0259	0.0345	0.0238	0.0198	0.0177	0.0189	0.0230	0.0189	0.0182
Willapa	0	0	0	0		0.0118	0.0172	0.0177	0.0162	0.0102	0.0237	0.0389	0.0220	0.0142	0.0140	0.0162	0.0197	0.0160	0.0152
Eloch	0	0	0	0	0		0.0007	0.0022	-0.0003	0.0010	0.0129	0.0277	0.0126	0.0090	0.0127	0.0123	0.0135	0.0118	0.0117
Skamo	0	0	0	0	0	19.5		0.0001	0.0008	0.0040	0.0148	0.0256	0.0137	0.0114	0.0129	0.0128	0.0139	0.0121	0.0123
BigCr	0	0	0	0	0	7	47.5		0.0019	0.0059	0.0176	0.0267	0.0157	0.0140	0.0153	0.0152	0.0187	0.0144	0.0151
All Grays	0	0	0	0	0	62.1	3!	4.1		0.0045	0.0171	0.0286	0.0155	0.0117	0.0137	0.0138	0.0146	0.0131	0.0132
Aber/Germ	0	0	0	0	0	21.1	0	0.1	0		0.0071	0.0234	0.0060	0.0045	0.0077	0.0090	0.0095	0.0091	0.0075
CowlitzF	0	0	0	0	0	0	0	0	0	0		0.0163	0.0015	0.0105	0.0177	0.0195	0.0213	0.0179	0.0159
CowlitzS	0	0	0	0	0	0	0	0	0	0	0		0.0176	0.0273	0.0288	0.0291	0.0304	0.0271	0.0259
All Lewis	0	0	0	0	0	0	0	0	0	0.1	6.4	0		0.0080	0.0126	0.0130	0.0153	0.0132	0.0107
Wash/Laca	0	0	0	0	0	0	0	0	0	1.2	0	0	0*		0.0016	0.0025	0.0017	0.0027	0.0002
All I205	0	0	0	0	0	0	0	0	0	0	0	0	0	9.7		0.0030	0.0018	0.0020	0.0006
All Ives	0	0	0	0	0	0	0	0	0	0	0	0	0	3.9	0*		0.0010	0.0013	0.0002
Multnomah	0	0	0	0	0	0	0	0	0	0*	0	0	0*	15.5	0.8*	10.1		0.0019	0.0011
Ham/Hardy	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0*	0.2*	0.4*		0.0005
All Bonne	0	0	0	0	0	0	0	0	0	0	0	0	0	41	20.3	38.4	16.8	24.5	

Table 6. Collection ancestry values for STRUCTURE analysis for K = 4 (see Figure 3 for graphic results of individual ancestry). Ancestry values for each collection in each cluster were averaged over all individuals in each collection. Order is the same as in Figure 3. Colored cells (colors match clusters in Figure 3) are values over 29%. N is the number of samples in the collection – collections with high samples numbers were subsampled (Grays, Hamilton and Hardy).

Region	Collection	K = 4 clusters				
		1	2	3	4	N
Pacific Coast	Nehalem	0.475	0.195	0.227	0.103	31
	Yaquina	0.461	0.112	0.370	0.055	24
	Satsop	0.484	0.164	0.277	0.074	64
	Willapa	0.619	0.088	0.223	0.070	72
	Sea Resources	0.814	0.070	0.086	0.031	95
Columbia River						
Coastal	All Grays	0.118	0.735	0.100	0.048	268
	Crazy Cr (Grays R)	0.133	0.653	0.133	0.081	35
	Elochoman	0.127	0.631	0.177	0.065	70
	All Skamokawa	0.103	0.668	0.139	0.091	146
	BigCr Hatchery	0.089	0.761	0.092	0.058	44
	Abernathy+Germany	0.209	0.290	0.371	0.130	56
Cascade	All Cowlitz	0.095	0.055	0.806	0.044	114
	All Lewis	0.118	0.082	0.635	0.165	80
Gorge	Washougal+Lacamas	0.145	0.060	0.284	0.512	54
	All I-205	0.177	0.083	0.114	0.627	163
	Multnomah	0.133	0.068	0.128	0.671	97
	All Ives area	0.150	0.061	0.137	0.652	159
	All Hamilton+Hardy	0.139	0.081	0.117	0.662	148
	All Bonneville	0.140	0.060	0.144	0.655	62

Lower Columbia chum salmon 2011, WDFW Molecular Genetics Lab

Table 7. Assignments to individual spawning aggregates from GeneClass self-assignment test for Columbia River and Pacific Coast chum salmon. Results tally samples with over 90% relative assignment likelihoods. Read across row for all assignments per collection. Values in bold along the diagonal are assignments back to collection or "correct" assignments. The "% correct" are correct assignments divided by the total assignments (summed across the row). The "% unassign" are the total number (N) of unassigned over the total analyzed.

						Columbia River													
	Pacific Coast					Coastal					Cascade				Gorge				
> 90%	Nehalem	Yaquina	Satsop	SeaRes	Willapa	Elocho	Skamoko	Big Cr	Grays	AberGer	CowF	CowS	Lewis	Washoug	I205	IvesIs	Multnom	HamHar	Bonnev
Nehalem	**4**	0	1	0	2	0	1	0	0	1	0	0	0	0	0	0	0	0	0
Yaquina	0	**14**	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Satsop	0	0	**9**	2	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0
SeaRes	0	0	0	**53**	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Willapa	3	0	1	0	**17**	1	0	0	0	1	0	0	1	0	0	0	0	1	0
Elocho	0	0	0	0	1	**1**	6	1	3	0	2	0	1	0	0	0	0	0	0
Skamoko	0	0	0	0	1	4	**9**	2	10	1	2	0	1	1	0	0	0	0	0
Big Cr	0	0	0	0	0	2	1	**3**	**2**	0	0	0	0	0	0	0	0	0	0
Grays	0	0	0	0	0	3	3	2	**57**	2	1	0	1	0	0	0	0	0	0
AberGer	0	0	1	0	0	2	3	1	0	**0**	3	0	1	0	0	0	0	0	0
CowF	0	0	1	1	0	0	1	0	0	2	**15**	3	4	0	0	0	0	0	0
CowS	0	0	0	0	0	0	0	0	0	0	6	**12**	2	1	0	0	0	0	0
Lewis	0	1	0	0	0	1	0	0	0	2	9	0	**15**	0	2	0	0	0	0
Washoug	0	0	0	0	1	0	0	0	0	0	2	0	1	**1**	1	0	1	1	0
I205	0	0	0	1	1	0	0	0	0	0	0	0	1	2	**9**	2	2	0	0
IvesIs	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	**17**	4	5	0
Multnom	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	5	**0**	2	0
HamHar	0	0	1	0	0	0	0	0	0	1	2	0	0	2	7	4	3	**14**	0
Bonnev	0	0	1	0	0	0	0	0	0	0	0	0	3	1	2	0	1	2	**0**
corr assign	4	14	9	53	17	1	9	5	57	0	15	12	15	1	9	17	0	14	0
Total assign	9	15	16	57	25	15	31	8	69	11	27	21	30	8	18	28	10	34	10
N unassign	22	8	48	38	42	55	115	36	234	44	50	16	50	31	143	193	87	337	51
% correct	44.44	93.33	56.25	92.98	68.00	6.67	29.03	62.50	82.61	0.00	55.56	57.14	50.00	12.50	50.00	60.71	0.00	41.18	0.00
% assign	29.03	65.22	25.00	60.00	37.31	21.43	21.23	18.18	22.77	19.64	35.06	56.76	37.50	20.00	11.04	12.61	10.31	9.09	16.13

Table 8. Summary by ecoregion of the results from GeneClass self-assignment test in Table 7. Results tally samples with over 90% relative assignment likelihoods. Read across row for all assignments per ecoregional group. Values in bold along diagonal are assignments back to ecoregion or "correct" assignments. The "% correct" are correct assignments divided by the total assignments (summed across the row). The "% assign" are the total assigned over the total analyzed.

	Pacific Coast	Coastal (Col. R.)	Cascade	Gorge
Pacific Coast	**152**	17	3	4
(Col R) Coastal	4	**422**	16	12
Cascade	3	9	**89**	16
Gorge	5	3	12	**705**
correct	152	422	89	705
Total assign	176	454	117	725
Total analyzed	286	619	194	974
% correct	86.36	92.95	76.07	97.24
% assign	61.54	73.34	60.31	74.44

Lower Columbia chum salmon 2011, WDFW Molecular Genetics Lab

Table 9. Information to guide decisions for chum salmon supplementation and re-introduction programs in lower Columbia River. Where N_e was unavailable, population size was estimated. Extirpated populations had a population size that overlapped zero and require a donor stock.

Location	Ecoregion	N_e	Recent Spawner Estimate	Potential as a Donor Stock	Potential to use local stock for enhancement and no concern for time to goal	Donor stock First Choice	Donor stock Second Choice
Youngs Bay	Coast		0-?		No (population likely too small)	Grays	Grays or Willapa?
Chinook River		162.4	0-50		No (population likely too small)	Grays	Grays
Grays Basin		318.2	> 2,500	YES	Yes	Grays	Grays
Big Creek		345.6	< 25		Yes	Big Creek	Grays
Skamokawa		218.8	< 25		Yes	Skamokawa	Grays
Elochoman/Beaver Creek		76.4	5-50		Yes	Elochoman/Beaver Creek	Grays
Mill/Abernathy/Germany			0-25		Yes	Mill/Abernathy/Germany	Grays
Clatskanie	Cascade		0-?		No (population likely too small)	Mill/Abernathy/Germany	Grays
Cowlitz (fall and sum)		195.4	0-25		Yes	Cowlitz (fall and sum)	I-205
Coweeman			0-?		No (population likely too small)	Cowlitz (fall and sum)	I-205
Kalama			0-?		No (population likely too small)	Cowlitz or Lewis	I-205
Lewis		187.7	25-75		Yes	Lewis	I-205
EF Lewis			0-?		No (population likely too small)	Lewis	I-205
I-205 (MS Col. R)	Gorge	516	> 750	YES	Yes	I-205	Any near Bon Pop
Washougal/Lacamas		115.9	0-25		Yes	I-205	Any near Bon Pop
Sandy			0-?		No (population likely too small)	I-205	Any near Bon Pop
Multnomah (MS Col. R)		51.2	50-125		Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
St Cloud (MS Col. R)			5-50		Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
Horsetail (MS Col. R)			5-50		Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
Duncan			5-50		Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
Ives Island area (MS Col. R)		196.6	> 150	YES	Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
Hardy		271.4	5-50		Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop
Hamilton (including spring channel)		271.4	> 250	YES	Yes (from any near-Bonneville population)	Any near Bon Pop	Any near Bon Pop