Final Report

Population structure and genetic assignment of bull trout (*Salvelinus confluentus*) in the Skagit River Basin

Matt J. Smith

School of Aquatic and Fishery Sciences

University of Washington

1122 NE Boat St Box 355020 Seattle, WA 98195

Submitted December, 2010

TABLE OF CONTENTS

EXECUTIVE SUMMARY

In 1999, the U.S Fish and Wildlife Service (USFWS) listed all bull trout (*Salvelinus confluentus*) populations throughout the United States as a threatened under the Endangered Species Act. The USFWS identified the possible genetic isolation of bull trout populations by Seattle City Light's hydroelectric dams as a high priority research need, and currently assumes that bull trout populations upstream of the dams are genetically distinct from those downstream of the dams. This study quantified the level of genetic diversity within and among bull trout populations in subbasins of the Skagit River both above and below Seattle City Light's Skagit River Hydroelectric Project. A genetic baseline was developed and genetic assignment tests were used to determine the composition of fluvial adult and sub-adult bull trout utilizing habitat in the Skagit River immediately downstream of the dams. Results indicated that populations above the dams were much less diverse (Heterozygosity=0.395, Allelic richness=3.12) than populations downstream of the dams (Heterozygosity=0.702, Allelic richness=6.59). Exact tests for genotypic differentiation demonstrated significant differences in genotypic distributions between most populations sampled. Pairwise F_{st} values were lower between above dam populations (Average=0.0213) than between below dam populations (Average=0.0659), indicating higher levels of geneflow occurring among above dam populations compared to that occurring among below dam populations. Above and below dam populations were highly differentiated and reproductively isolated from each other (Average pairwise F_{st} =0.2769) due to migration barriers such as the dams and historical velocity barriers which may have existed prior to dam construction. A principal coordinate analysis (PCA) and neighbor joining dendrogram supported the pooling of individual baseline collections into reporting groups for genetic assignment tests based on the subbasin in which they were collected. An analysis of molecular variance (AMOVA) demonstrated that the variation among collections within reporting groups was very low (1.49%) relative to the variation among reporting groups (15.40%). High self-assignment success rates of bull trout in the genetic baseline provided evidence for genetic differentiation and low gene flow among reporting groups, and supported the use of the baseline for determining origin of unknown fluvial bull trout from the mainstem Skagit River. Individual genetic assignment tests showed that adult and sub-adult bull trout samples of unknown origin collected from the Skagit River immediately downstream of the dams were mostly comprised of fish from Goodell (38.04%) and Cascade (35.45%), followed by smaller percentages of Illabot (12.97%), Downey (7.78%), Bacon (4.03%), and Sauk (1.73%) fish. None of the fish collected below the dams originated from above dam populations.

INTRODUCTION

Bull trout (*Salvelinus confluentus*) are iteroparous members of the Salmonidae family, and are native to the Pacific Northwest and western Canada. They may exhibit a resident, or one of three migratory life history forms, including adfluvial (migrating from tributary streams to a lake or reservoir to mature), fluvial (migrating from tributary streams to larger rivers to mature), or anadromous (migrating from freshwater to the ocean to grow and mature and returning to freshwater to spawn) behaviors (USFWS 2004).

In general, most bull trout populations appear to be consistent with the metapopulation concept, where several local populations function as one demographic unit due to occasional gene flow between them (Rieman & McIntyre 1995; Dunham & Rieman 1999). Effective management of populations exhibiting a metapopulation structure depends on the ability of management actions to maintain interconnected habitats that support diverse life histories and facilitate gene flow between populations (Rieman & Allendorf 2001). The maintenance of migratory corridors is essential not only to facilitate gene flow between populations, but to provide the potential for reestablishment of extirpated populations (Rieman & Dunham 2000). Maintaining migratory corridors may also enable the persistence of bull trout populations by allowing individuals access to unoccupied but suitable habitat, by providing a variety of foraging opportunities, and by providing refuges from disturbances (Saunders *et al.* 1991).While most bull trout populations exhibit metapopulation dynamics, some demonstrate no evidence of metapopulation structure, with geographically proximate populations being highly reproductively isolated (Kanda & Allendorf 2001). This suggests that bull trout populations in these systems have a lower probability of recolonization through dispersal from adjacent populations following population extinctions. Therefore, the long-term persistence of bull trout in these systems requires management actions that ensure that the geographically proximate populations are maintained and managed individually rather than as one demographic unit.

Differences in bull trout population structure among drainage basins highlights the importance of obtaining an accurate understanding of genetic population structure and dispersal characteristics prior to the implementation of management strategies. The physical and ecological processes influencing the genetic population structure of bull trout in one region do not necessarily accurately reflect those in another region (Whiteley *et al.* 2006). Therefore, management actions which may be appropriate for some populations could potentially be detrimental to the persistence of bull trout in other populations. A lack of information regarding the fine and broad scale breeding patterns and demography of a threatened species may result in mismanagement, either through failure to recognize its existence, or invalid assumptions about its range, trends, and its resilience (Spruell *et al.* 1999).

In 1999, the U.S Fish and Wildlife Service (USFWS) listed all bull trout populations throughout the United States as a threatened under the Endangered Species Act (USFWS 1999) The USFWS completed a recovery plan for the Puget Sound Management Unit in 2004 (USFWS 2004). The Puget Sound Management Unit consists of eight core areas, with a total of 59 local populations distributed among the core areas. A core area is defined as an area that represents a combination of suitable habitat and one or more local populations that function as one demographic unit due to occasional gene flow between them. The bull trout populations included in this study are a part of the Upper Skagit and Lower Skagit core areas, when combined include 27 local populations. Seattle City Light's (SCL) Skagit Hydroelectric Project, which consists of

three hydroelectric dams (Ross, Diablo, Gorge) was identified as the dividing line between bull trout populations in the Upper Skagit and the Lower Skagit core recovery areas (USFWS 2004). The USFWS identified the possible genetic isolation of bull trout populations by SCL's dams as a high priority research need (Ed Connor, SCL, pers. comm.). Historical records indicate that a steep and narrow bedrock gorge located at the current location of Diablo Dam probably blocked upstream migration of all migratory fish, including bull trout. Currently, bull trout in the Lower Skagit core area can migrate upstream only as far as Gorge Dam.

The Skagit River watershed possesses a high diversity of bull trout life histories, including anadromous forms that migrate into Puget Sound, fluvial forms that can range from non-migratory to highly migratory subtypes, adfluvial forms in major lakes (including Ross, Diablo, and Gorge reservoirs), and resident forms that spend most of their lives in headwater streams. It is thought that the Lower Skagit core area supports a spawning population of migratory bull trout that numbers in the thousands, likely making it the largest population in Washington (USFWS 2004). Fluvial bull trout within the Lower Skagit core area typically forage and overwinter in the larger pools of the upper portion of the mainstem Skagit River immediately downstream from SCL's Skagit Hydroelectric Project (USFWS 2004). Despite the abundance and importance of the Upper Skagit and Lower Skagit bull trout populations, the fine scale genetic structure of these spawning populations is not well understood. Maintaining the migratory life history forms has been identified as crucial to the long term persistence of bull trout populations in the Puget Sound Management Unit, yet the population composition of these migratory bull trout is uncertain. Understanding the composition of migratory bull trout that utilize habitat in the vicinity of SCL's Skagit Hydroelectric Project can help guide and prioritize management actions designed to maintain this life history form.

OBJECTIVES

The primary objective of this study was to complete an assessment of the genetic variability within and among bull trout populations of the Skagit River Basin and subbasins in the vicinity of Seattle City Light's (SCL) Skagit Hydroelectric Project. This research will be used to help identify and guide management actions that can aid in the conservation and recovery efforts of these populations. The specific objectives of the study were to:

- Identify all samples collected as bull trout, Dolly Varden, brook trout, or hybrids using a suite of diagnostic genetic markers.
- Describe the genetic diversity within all baseline collections and quantify the level of genetic differentiation among collections.
- Develop a genetic baseline for individual genetic assignment tests by pooling individual collections into appropriate reporting groups based on subbasin of collection and genetic characteristics of collections.
- Evaluate the distinctness of reporting groups and determine whether the genetic baseline was suitable for individual assignment of adult and sub-adult bull trout from the Skagit River through individual population assignment tests.

- Use individual population assignment tests to determine the composition of adult and sub-adult bull trout that utilize habitat in the Skagit River immediately downstream of Seattle City Light's Hydroelectric Project.

METHODS

Tissue Collection

Baseline collections

During 2001-2009, 595 juvenile and adult bull trout were sampled from fourteen collection sites by staff from the Washington Cooperative Fish and Wildlife Research Unit (WACFWRU) at the University of Washington and R2 Resource Consultants, Inc. (Redmond, WA). Baseline collections were from tributaries of the Skagit River in the vicinity of Seattle City Light's hydroelectric facilities and represent populations likely to contribute to the genetic diversity of the adult bull trout found in the mainstem Skagit River. Four collections were from sites located above Ross Dam, one collection was between Diablo Dam and Gorge Dam, and the remaining nine collections were from sites below the dams. Adult and juvenile samples were collected through a combination of electrofishing, snorkeling, and angling. Individuals were anesthetized with Pure Tricaine Methanesulfonate (MS-222) and a fin clip was removed and either dried on filter paper or preserved in 100% non-denatured ethanol prior to genetic analyses. Samples collected in the same location across multiple years were pooled into single collections as recommended by Waples (1990). Location of collections, sample sizes, and tissue types are presented in Table 1 and Figure 1.

Mixture collections

During 2006-2008, 435 fluvial adult and sub-adult bull trout were collected from the mainstem Skagit River from the base of Gorge Powerhouse to the confluence of the Sauk River (highlighted area in Figure 1). These samples represent a potential mixture from several spawning populations located throughout the Skagit basin. Samples were collected by staff from WACFWRU primarily by angling. Fin clips were taken from anesthetized fish and frozen prior to genetic analyses.

Laboratory Analysis

Genomic DNA from fin clips was extracted using DNeasy 96 Blood & Tissue Kits (QIAGEN). DNA from all samples was amplified at a set of sixteen microsatellite loci (Table 2) that are standardized across multiple laboratories. PCR amplification was conducted in 384-well plates in 5μL volumes containing 1.5μL genomic DNA template, forward and reverse microsatellite primers, $MgCl_2$, 0.2mM each dNTP, and 0.5 units Biolase™ Taq polymerase (see Table 2 for detailed PCR conditions). PCR products were visualized on a MegaBACE automated sequencer (Amersham Pharmacia Biotech, Inc., Piscataway, New Jersey) according to the manufacturer's protocols. Microsatellite allele size was determined with the Genetic Profiler v2.2 software package (Amersham Pharmacia Biotech, Inc., Piscataway, New Jersey).

Statistical Analysis

Evaluation of family structure

When describing population genetic structure, the sampling goal is to obtain an unbiased representation of the genetic diversity of the populations in question. Conformation to expected Hardy-Weinberg equilibrium (HWE) proportions, as measured by F_{IS} values, is used to evaluate, among other population or genetic marker characteristics, allelic biases that may be present in sample collections. Sampling juvenile fish can introduce bias if samples consist of large proportions of related individuals. Large numbers of full siblings will result in heterozygote deficiencies (positive F_{IS} values) and biased allele frequencies within collections.

We used the program COLONY (Jones & Wang 2010) to identify family structure within baseline collections that could result in a biased representation of allele frequencies. COLONY uses multilocus genotype data to infer sibship among individuals through full-pedigree likelihood methods. All baseline collections were evaluated for the presence of full sibling families. All but one full sibling was removed from further analyses when families were identified with a greater than 90% probability. Siblings with the fewest number of loci scored were removed, if the number of loci scored was equal among siblings the first sibling in the dataset was retained and all other full siblings were removed.

Identification of brook trout and Dolly Varden

Bull trout have been shown to exist in sympatry with brook trout (*Salvelinus fontinalis*) and Dolly Varden (*Salvelinus malma*). Hybridization between bull trout and brook trout and between bull trout and Dolly Varden has been documented in several watersheds where the species co-occur (Baxter *et al.* 1997; Dehaan *et al.* 2010). Several samples in the Lightning Creek and Stetattle Creek collections were identified as putative bull x brook and bull x Dolly Varden hybrids based on alleles observed and failed amplification at several diagnostic loci.

We used the Bayesian clustering method implemented in the program STRUCTURE v2.2 (Pritchard *et al.* 2000) to identify brook trout, Dolly Varden, and putative hybrids in all baseline and mixture collections. STRUCTURE assigns a proportion of each individual's multilocus genotype to each of K clusters that maximize HWE and linkage equilibrium within clusters. Individuals with brook trout or Dolly Varden alleles were expected to cluster independently from all bull trout samples. STRUCTURE runs consisted of 50,000 iterations of burn-in followed by 250,000 Markov Chain Monte Carlo steps assuming admixture for each K from 1 to 12. Five replicates for each value of K proved sufficient as likelihood values converged and the variance among runs for each K was low. Evanno's ΔK criterion was used to determine the most likely value of K, which is often found where the largest change in likelihood occurs among simulations for different K. (Evanno *et al.* 2005).

Pooling collections into reporting groups

Baseline collections from the same subbasins were pooled into seven reporting groups for genetic assignment tests. An analysis of molecular variance (AMOVA) was performed to assess the variation within and between groups in ARLEQUIN (v. 3.5). This approach hierarchically examines variance in gene frequencies due to intra-group and inter-group differences; significance of the components is determined by permuting populations within groups and populations among groups. The analysis was conducted with collections pooled into the reporting groups by the seven subbasins sampled (Figure 1): Above Dam (Upper Skagit River, Big Beaver Creek, Ruby Creek, Stetattle Creek), Goodell (Lower Goodell Creek, Upper Goodell

Creek), Bacon (Bacon Creek), Cascade (Cascade River, Marble Creek, Kindy Creek), Illabot (Illabot Creek), Sauk (South Fork Sauk River), and Downey (Downey Creek).

Within collection/reporting group diversity

All collections and reporting groups were tested for deviations from Hardy-Weinberg equilibrium and linkage equilibrium using GENEPOP v4.0 (Rousset 2008). Critical values (α=0.05) were corrected for multiple tests using a sequential Bonferroni correction (Rice 1989). Observed and unbiased expected heterozygosities were calculated in GENALEX. Allelic richness or the number of alleles corrected for sample size was estimated using a rarefaction method implemented in HP-Rare (Kalinowski 2005). Large genetic samples are expected to have more alleles than small samples. Rarefaction is a statistical technique to deal with this problem so the number of alleles in large samples can be compared with the number of alleles in small samples. Allelic richness is more sensitive to the loss of genetic variation due to small population size than expected heterozygosity, and is an important measure of the long-term evolutionary potential of populations.

We report observed and expected heterozygosities and allelic richness for: (1) individual baseline collections, (2) baseline collections pooled into reporting groups, (3) baseline collections pooled into an above dam and below dam group, and (4) mixture collections. This hierarchical structure facilitates multiple comparisons of genetic diversity among collections and groups of collections. In addition, comparisons of the amount of genetic variation in collections sampled temporally can provide evidence for loss of genetic variation because of population isolation and fragmentation due to habitat loss or other causes. Although temporal comparisons were not possible in this study due to limited numbers of temporal samples within collections, genetic diversity measures may be used by managers in the future to monitor the influence of management actions, anthropogenic effects, or natural causes on the current levels of genetic variation.

Among collection/reporting group diversity

Pairwise exact tests for genotypic differentiation were calculated for all pairs of collections and all pairs of reporting groups with Markov chain parameters (5,000 dememorizations, 1,000 batches, and 1,000 iterations/batch) using GENEPOP v4.0 (Rousset 2008). Significant differences in genotypic distributions among collections indicate that they are reproductively isolated from each other and constitute distinct populations. Weir and Cockerham's F_{st} (ARLEQUIN v. 3.5; Weir & Cockerham 1984) was calculated for all pairs of collections and reporting groups with significance based on a permutation process; a matrix plot of collection and reporting group pairwise F_{st} was created using the R (R Development Core Team. 2008) plotting function provided in ARLEQUIN (v. 3.5). Pairwise F_{st} values are a ratio of the amount of genetic variance among populations over the total variance in all baseline collections. F_{st} values that are significantly greater than zero indicate reproductive isolation among collections and that collections constitute distinct populations. Pairwise F_{st} distances were used in a Principal Coordinate Analysis (PCA option in GENALEX) to graphically represent multidimensional genetic relationships among collections and reporting groups in a two dimensional space. An unrooted neighbor-joining (NJ) dendrogram using Cavalli-Sforza and Edwards (1967) chord distance was also used to display genetic relationships using the software

POPULATIONS 1.2.14 (Langella 2001) and TREEVIEW (Page 1996). One thousand bootstrap replicates were performed to evaluate tree topology.

Baseline evaluation

Assignment tests as implemented in GeneClass2 (Piry *et al.* 2004) were used to test whether the seven subbasin reporting groups described above represent distinct reporting groups suitable for use as a genetic baseline for individual assignment of adult bull trout from the Skagit River. This test assigns fish to a baseline reporting group in which they have the highest likelihood of occurring, based on the multilocus genotype of the fish and the allele frequencies of baseline reporting groups. The test follows the leave one out method of assignment where each fish is removed from the baseline without replacement prior to calculating likelihoods of belonging to each reporting group. The program calculates a relative likelihood value, which is a ratio of the highest likelihood over the next highest likelihood value and assigns the individual to the reporting group with the highest relative likelihood value. We used a ratio threshold of 90% as positive assignment. We report home assignments (assignments to the reporting group where the individual was collected) and other assignments (assignments to reporting groups other than the group in which the individual was collected) that exceed the 90% threshold for positive assignment. Assignments below the 90% threshold are reported as unassigned individuals.

We also conducted first generation migrant tests to detect the occurrence of and examine the direction of gene flow that may be occurring among reporting groups. We used the Rannala & Mountain (1997) algorithm in GeneClass2 to calculate likelihoods that each fish originated in each of the seven reporting groups. The probability of an individual being a resident in its reporting group was computed with a Monte Carlo simulation (Paetkau *et al.* 2004). The simulation creates genotypes for 10,000 individuals from each baseline reporting group based on allele frequencies of the reporting group. The distribution of assignment likelihoods for simulated individuals is then compared to the likelihood of the individual in question. Fish were hypothesized as first generation migrants if their probability of originating in their home reporting group was represented in less than 1% of the simulated values of their home reporting group.

Assignment of mixtures

We used GeneClass2 to assign the 434 fluvial adult and sub-adult bull trout that were collected during 2006-2008 from the mainstem Skagit River to the seven subbasin reporting groups. These adult and sub-adult bull trout samples represent a potential mixture from each of the reporting groups located in the vicinity of Seattle City Light's hydroelectric facilities. We maintained the 90% likelihood ratio threshold described above as positive assignment.

An exclusion test using the Monte Carlo simulation described for the assignment of baseline samples was used to test whether the population of origin was present in the baseline reporting groups. If probabilities of an individual from the mixture coming from each baseline collection were all below the hypothesized 1%, then the individual likely originated from a population outside the baseline collections. The exclusion test differs from the relative likelihood values used for assignments. A likelihood of assignment to each reporting group always exists but likelihood values may be extremely small for all reporting groups if the population of origin

is not present in the baseline. Extremely small likelihoods may result in a deceptively high relative likelihood value and positive assignment to the most genetically similar reporting group. Identifying fish with a low probability, here 1%, of belonging to any reporting group limits the amount of false positive assignments when all potential populations are not included in the baseline. We report all fish that are likely to originate from outside the baseline and exclude them from assignments and further analyses.

We evaluate the composition of the bull trout in the mainstem Skagit River below Seattle City Light's hydroelectric facilities by reporting the percentage of fish that assign to each reporting group at or above the 90% positive assignment threshold.

There is currently a sport fishery for bull trout on the Skagit River that is managed by the Washington Department of Fish and Wildlife (WDFW). Bull trout 20 inches (508mm) or larger caught in the mainstem from the Cascade River downstream to the mouth of the Skagit River may be retained by anglers as part of the two trout daily limit for rivers. We report percentage assignments for all bull trout lengths as well as bull trout divided into groups above and below the 508mm legal length to determine whether any reporting group is contributing a disproportionate number of legal or sub-legal size fish.

Assignments of fluvial adult and sub-adult bull trout from the mainstem Skagit River were stratified by region of capture and the quarter of the year in which bull trout were captured within each region. Spatial and temporal stratification was conducted to determine the seasonal composition of bull trout present in different regions of the river immediately downstream of Seattle City Light's Hydroelectric Project. Regions were selected based on management interests (Dave Pflug, SCL, pers. comm.). The three regions include: Gorge Powerhouse downstream to Shovelspur Rapids (Above Shovelspur), Shovelspur Rapids downstream to the town of Marblemount (Shovelspur to Marblemount), and the town of Marblemount downstream to the town of Rockport (Marblemount to Rockport). There are no physical barriers such as waterfalls or dams within or between regions that would prevent bull trout from utilizing habitat in more than one region. One or more subbasins included in the genetic baseline flow into the Skagit River within each region. The Above Shovelspur region includes the confluence of Goodell Creek. The Shovelspur to Marblemount region includes the confluence of Bacon Creek. The Marblemount to Rockport region includes the confluence of the Cascade River and Illabot Creek. The Sauk River flows into the Skagit River downstream from the town of Rockport and is also included in the genetic baseline.

RESULTS

Laboratory analysis

Samples that failed to amplify at fewer than seven loci were eliminated to limit biases caused by missing data. A total of 1018 of the 1030 (98.8%) samples were successfully amplified and scored at seven or more loci (Table 1). The twelve samples that failed were rerun following a preamplification procedure designed to increase the amount of template DNA in subsequent genotyping reactions (Smith *et al.* In Review). The preamplification procedure failed to improve the data quality and these samples were eliminated from further analyses. All previously frozen tissues from the 435 adult and sub-adult bull trout that were collected during 2006-2008 from the mainstem Skagit River were preamplified prior to genotyping due to poor

amplification observed in initial screening. Following the preamplification procedure, 434 samples were successfully genotyped at seven or more loci.

Initial investigations into deviations from Hardy-Weinberg equilibrium expectations revealed a significant deficiency of heterozygotes at locus *Sco109* in a majority of the baseline collections tested. Significant heterozygote deficiencies that occur in the majority of the collections indicate either a technical problem with the amplification of the locus (i.e. large allele dropout) or the presence of a null allele in the collections. Large allele dropout can occur due to the preferential amplification of shorter DNA template during PCR, making detection of larger alleles difficult. Null alleles are alleles that fail to amplify due to base substitutions that prevent microsatellite primers from binding during PCR. We were not able to distinguish between large allele dropout and null alleles as the source of the observed heterozygote deficiencies at this locus and it was removed from further analyses.

Evaluation of family structure

A total of 161 full sibling relationships were identified through the COLONY analysis. The presence of full siblings in the juvenile bull trout baseline collections was expected. Sampling protocols were designed to achieve a sample size that would allow for the removal of full siblings while maintaining an adequate sample size for the characterization of the populations sampled. All but one full sibling was removed from further analyses to prevent a biased representation of allele frequencies caused by high proportions of related individuals in baseline collections. Siblings with the fewest number of loci scored were removed, if the number of loci scored was equal among siblings the first sibling in the dataset was retained and all other full siblings were removed. The number of full siblings removed from each collection varied from zero in the Upper Skagit River collection to 32 in the Bacon Creek collection (Table 1). Sample sizes were small in some of the baseline collections after the removal of full siblings but were not unusually small relative to other bull trout genetics studies. Pooling collections into reporting groups for genetic assignment tests helped reduce biases that can be caused by large disparities in sample sizes among reporting groups.

Identification of brook trout and Dolly Varden

Several samples from juvenile baseline collections above the dams were identified as putative Dolly Varden, brook trout, or hybrids based on alleles observed and failed amplification at several diagnostic loci. Ten Stetattle Creek samples had alleles at several loci that USFWS Abernathy Fish Technology Center (AFTC) identified in brook trout (Pat DeHaan, USFWS, pers. comm.). An additional five samples from Stetattle Creek and 24 samples from Lightning Creek had alleles at several loci that WDFW identified in Dolly Varden (Maureen Small, WDFW, pers.comm.). We included these samples in the Bayesian clustering method implemented in the program STRUCTURE to identify individuals in all baseline and mixture collections with any Dolly Varden or brook trout ancestry.

Individuals with brook trout or Dolly Varden alleles were expected to cluster independently from all bull trout samples. Individuals with partial Dolly Varden or brook trout ancestry were identified as those samples sharing some ancestry with the Dolly Varden/brook trout cluster. The results from the analysis identified K=5 as the most likely K-value using Evanno's ΔK criterion. At K=2 the samples were divided into a cluster comprised of baseline

collections from above the dams and a cluster comprised of baseline collections below the dams. At K=3 the samples identified as having Dolly Varden and brook trout alleles comprised the third cluster. At K=4 the below dam cluster split into a cluster comprised of Goodell and Bacon collections and a cluster comprised of Cascade, Marble, Kindy, Illabot, Sauk, and Downey collections. At K=5, the most likely K-value identified, the Sauk and Downey collections form a fifth cluster independent of the Cascade, Marble, Kindy, and Illabot cluster (Figure 2). The Stetattle Creek and Lightning Creek samples identified as putative Dolly Varden or brook trout had 100% ancestry in the Dolly Varden/brook trout cluster. These samples were therefore removed from subsequent analyses. After the removal of Dolly Varden juveniles, the Lightning Creek collection included only two adult bull trout samples and was therefore removed from further analyses due to insufficient sample size. Five Upper Skagit River samples, two Big Beaver Creek samples, and one Ruby Creek sample exhibited partial Dolly Varden/brook trout ancestry. However, these samples had less than 30% Dolly Varden/brook trout ancestry and were therefore included in subsequent analyses. We found no evidence of inter-specific hybrids in the baseline and mixture collections below the dams.

Within population/reporting group diversity

A total of 195 tests for Hardy-Weinberg equilibrium were performed in baseline collections and 15 of the tests were rejected at α=0.05, which was slightly higher than expected by chance (9.75 tests expected from Type I error of 0.05). The highest number of rejected tests per locus was three out of thirteen tests for locus *Omm1128* and the highest number of rejected tests per population was seven out of fifteen tests for the Bacon Creek collection. When the sequential Bonferroni correction was applied to all p-values, only two of tests were rejected and both occurred in the Bacon Creek collection (Table 3). With the exception of the Bacon Creek collection, the data indicated that there was no association between the rejected tests and a locus or population and our assumptions of Hardy-Weinberg equilibrium could be met. Pooling collections into the seven reporting groups resulted in one additional test rejected in the Above Dam reporting group (Table 3).

A total of 1,365 tests for linkage disequilibrium were performed in baseline collections and 124 of the tests were rejected at α=0.05, which was higher than expected by chance (68.25 tests expected from Type I error of 0.05). This was primarily driven by 42 of the 105 tests for linkage disequilibrium rejected in the Bacon Creek collection.

Interpreting the causes for deviations from Hardy-Weinberg equilibrium can be difficult. F_{IS} values are a measure of departure from expected Hardy-Weinberg proportions. Positive values indicate an excess of homozygotes and negative values indicate a deficit of homozygotes. The most likely biological cause for an excess of homozygotes is nonrandom mating or the presence of multiple populations in a single collection. Negative F_{IS} values may indicate small breeding populations. We report F_{IS} values for all baseline collections and pooled reporting groups in Table 3. The Bacon Creek collection deviates from expected Hardy-Weinberg proportions at a majority of the loci. F_{IS} values are primarily negative indicating small breeding populations and this is supported by the linkage disequilibrium observed in the Bacon collection. Despite the deviation of the Bacon collection from expected Hardy-Weinberg proportions, we included it in subsequent analyses as the best available characterization of the genetic diversity in Bacon Creek.

Heterozygosity and allelic richness measures indicated that baseline collections from above the dams were less diverse than collections below the dams. Unbiased expected heterozygosity ranged from 0.373 to 0.445 in above dam collections and from 0.620 to 0.696 in below dam collections. Allelic richness ranged from 2.72 to 3.93 in above dam collections and from 4.64 to 6.14 in below dam collections (Table 4).

Among population/reporting group diversity

Pairwise exact tests for genotypic differentiation demonstrated significant differences in genotypic distributions between most collections and between all reporting groups. Differences were not significant between Upper Skagit/Big Beaver, Cascade/Kindy, and Marble/Kindy collections (Table 5). Significant differences in genotypic distributions among collections indicate that they are reproductively isolated from each other and constitute distinct populations. Pairwise F_{st} values were significant for all reporting groups and collections except Upper Skagit/Big Beaver and Cascade/Kindy (Table 5). Pairwise F_{st} values were lower between above dam collections (Average=0.0213) than between below dam collections (Average=0.0659), which indicates higher levels of geneflow occurring between above dam populations compared to geneflow occurring between below dam populations. Above and below dam populations were highly differentiated and reproductively isolated from each other (Average pairwise F_{st} =0.2769) likely resulting from migration barriers such as the dams and historical velocity barriers which may have existed prior to dam construction. Graphical representations of pairwise F_{st} values between collections and between reporting groups are shown in Figure 3.

The principal coordinate analysis (PCA) showed strong genetic divergence between above dam collections and below dam collections. The first principal coordinate, which separated these collections, explained 75% of the total variation (Figure 4a). Above dam collections and below dam collections were separated from each other and a PCA was conducted independently for each group, which provided a more detailed depiction of divergence between collections (Figure 4b, c). In general, there was no strong pattern of divergence among collections observed in the independent PCA analyses. The exception was a tight clustering observed in the pooled Cascade subbasin collections. The relatively large separation between the pooled Goodell collections may be surprising but is likely due to a very small sample size in the Upper Goodell collection (n=11) which can result in artificially high pairwise F_{st} values. Despite the separation observed between these two collections, we are confident that pooling them into a single reporting group was appropriate given the conformation to expected Hardy-Weinberg proportions observed in the pooled group. The unrooted neighbor-joining (NJ) dendrogram supported the genetic divergence observed in the PCA analysis and provided additional support for the pooling of collections by subbasin for assignment tests (Figure 5). The analysis of molecular variance (AMOVA) demonstrated that the variation among collections within reporting groups was very low (1.49%) relative to the variation among reporting groups (15.40%), which further justified pooling collections (Table 6).

Baseline evaluation

The baseline assignment test results from GeneClass2 support the observed genetic distinction among reporting groups. A total of 354 fish (92.43%) assigned back to their home reporting group with a relative likelihood value that exceeded the 90% threshold for positive

assignment. Positive home assignment success rates ranged from a high of 100% in the Above Dam group to a low of 82.76% in the Bacon group (Table 7, Figure 6). Only six fish (1.57%) positively assigned back to reporting groups other than the group in which the individual was collected and 23 fish (6.01%) failed to meet the 90% threshold and were unassigned (Table 7, Figure 6). Of the six fish that positively assigned back to other reporting groups, four were in the Bacon group and two were in the Cascade group. Three out of the four Bacon fish positively assigned to the Goodell group and the fourth assigned to the Cascade group. Both of the fish from the Cascade group positively assigned to the Goodell group (data not shown).

In the first generation migrant test, three fish were identified as likely first generation migrants based on the less than 1% probability of home assignment criteria. Two fish (one from Bacon and one from Cascade) were fish that positively assigned to the Goodell group and the third fish was from the Bacon group that positively assigned to the Cascade group. The remaining three fish from the Bacon and Cascade group that positively assigned to the Goodell group had low probabilities (<5%) of home assignment but failed to meet the less than 1% criteria established.

The majority of fish positively assigning back to reporting groups other than the group in which the individual was collected and the majority of fish identified as first generation migrants assigned to the Goodell reporting group. In October 2003 floods imposed major changes on many tributaries of the Skagit River, particularly Goodell Creek where bull trout were barred from large portions of historical spawning habitats by a major landslide (Downen, WDFW, Unpublished). The results of the baseline assignment tests suggest that fish from Goodell Creek are beginning to recolonize and successfully reproduce in adjacent tributaries (Bacon Creek and Cascade River). This genetic baseline can provide an important tool that may be used to monitor the success of recolinization of bull trout that previously spawned in Goodell Creek. The high positive assignment success rates in the genetic baseline provide evidence for genetic distinction among reporting groups and support the pooling of subbasin collections into reporting groups. High assignment success rates also support the utility of the baseline for determining origin of unknown fish from the mainstem Skagit River. The low numbers of fish assigning back to reporting groups other than the group in which the individual was collected and low rates of unassigned fish suggest low levels of geneflow among reporting groups.

Assignment of mixtures

We used the genetic baseline described above to assign 434 adult and sub-adult bull trout of unknown origin that were collected during 2006-2008 in the mainstem Skagit River to their reporting group of origin. These adult and sub-adult bull trout samples represent a potential mixture from each of the reporting groups located in the vicinity of Seattle City Light's hydroelectric facilities. A total of 351 bull trout (80.9%) assigned back to a reporting group with a relative likelihood value that exceeded the 90% threshold for positive assignment. Four of these fish were identified as likely originating from populations outside of the baseline based on the 1% probability of exclusion threshold established to limit the amount of false positive assignments when all potential populations are not included in the baseline. These fish were removed from further analysis resulting in the positive assignment of 347 fish (80.0%). A total of 83 bull trout (19.1%) failed to meet the 90% threshold and were classified as unassigned. Three

of the unassigned fish were identified as likely originating from populations outside of the baseline.

The composition of the positively assigned bull trout samples collected from the mainstem Skagit River below Seattle City Light's hydroelectric facilities was mostly comprised of fish from Goodell (38.04%) and Cascade (35.45%), followed by smaller percentages of Illabot (12.97%), Downey (7.78%), Bacon (4.03%), and Sauk (1.73%) fish (Table 8, Figure 7). None of the fish collected below the dams assigned to the Above Dam reporting group.

When the samples were divided into groups of legal retention length (>508mm, n=80) and sub-legal retention length (<508mm, n=267), the percentage of Cascade fish of legal length increased to 40% of the total and the percentage of Illabott fish of legal length decreased to 10% of the total. Only slight differences were observed in the percentages of legal length versus sub-legal length fish from all other reporting groups.

Spatial stratifications by region showed that the yearly mean composition of bull trout varied considerably among regions. Bull trout collected in the Above Shovelspur region were primarily of Goodell Creek (68.6%) and Cascade River (20.1%) origin. The Shovelspur to Marblemount region included smaller percentages of Goodell (40.2%) and larger percentages of Cascade (30.7%) and Bacon Creek (12.5%) origin fish. Collections from the Marblemount to Rockport region were primarily of Cascade (41.1%) and Illabot Creek (28.7%) origin with smaller percentages of Downey Creek (13.6%) and Sauk River (7.7%) bull trout. Temporal stratification within regions demonstrates some seasonal variation in compositions occurring in each region. However, the low sample size in all regions from the July through September stratum is likely skewing the results and should be noted (Table 9, Figure 8).

DISCUSSION

Migration barriers (i.e. dams and historical velocity barriers) within the Skagit River Basin appear to be a factor affecting population structure. The results show that populations above the dams were much less genetically diverse than populations below the dams. While most collections exhibited significant genetic differentiation as measured by pairwise F_{st} values, genetic differentiation between above dam and below dam collections was considerably greater than differentiation between pairs of collections above the dams or between pairs of collections below the dams.

In general, collections from the same subbasin showed low levels of genetic differentiation suggesting high levels of geneflow within subbasins. Results from the self-assignment tests indicated that there may be some geneflow occurring between Goodell Creek and Bacon Creek and between Goodell Creek and the Cascade River. This may be due to the limited amount of spawning habitat available to bull trout from Goodell Creek following the major landslide which occurred in 2003. First generation migrants from Goodell Creek were detected in both Bacon Creek and the Cascade River suggesting that geographically proximate subbasins appear to be providing some refuge from this disturbance. The rest of the reporting groups displayed high pairwise genetic distance estimates and high rates of positive self-assignment indicating reproductive isolation.

A comprehensive genetic baseline can be a useful tool for managers. Constructing a genetic baseline is an important first step in characterizing population structure and can be used to determine the population of origin of unknown fish or mixtures of unknown fish using genetic assignment tests. Genetic assignment tests can potentially replace some types of tagging data and can be used to examine migration rates, colonization rates, and evaluate the response of bull trout to recovery efforts. The power of assignment tests are determined by the level of differentiation among sampled baseline populations, number of loci genotyped, levels of diversity at genotyped loci, and overall baseline representation of the populations potentially contributing to the mixture. If populations that have not been sampled are contributing to the mixture, fish from the mixture can be erroneously assigned to the most genetically similar population in the baseline. Therefore, it is important to include all existing populations within the baseline. While this was not feasible in this study, we attempted to control false positive assignments by identifying individuals with low probabilities of belonging to any baseline group through the use of exclusion tests. Only a few individual fish from the Skagit River adult and sub-adult were identified as likely originating from populations outside of the baseline. This suggests that the genetic baseline is fairly robust and well suited for identifying the population of origin of unknown fish from the Skagit River in the vicinity of the dams.

The observed differences in the percentage of fish among reporting groups throughout the Skagit River and the variation among regions and temporal strata could be due to a variety of biological or evolutionary factors and may also be influenced by biases introduced during sampling. The sampling protocol for adult and sub-adult bull trout was not developed specifically for this study and is not designed to isolate variables that may explain differences in the percentage contributions of reporting groups. Therefore, we make no attempts to infer which factors may be driving the observed differences. The composition estimates serve as a management tool by providing a better understanding of the population composition of the fluvial adult bull trout utilizing habitat immediately below Seattle City Light's hydroelectric facilities. This information can be used as a monitoring tool in the future to determine the influence of management actions, anthropogenic influences, or natural disturbances on the population composition. It may also be used as a tool to help guide conservation and recovery efforts by identifying habitat regions which are utilized by various bull trout populations

ACKNOWLEDGEMENTS

The funding for this project was provided by Seattle City Light. We thank Erin Lowery from the Washington Cooperative Fish and Wildlife Research Unit (WACFWRU) at the University of Washington and R2 Resource Consultants, Inc. (Redmond, WA) for collecting the samples for the project. We thank Isadora Jimenez for the laboratory processing of all samples and scoring all microsatellite loci. Maureen Small (WDFW) and Pat DeHaan (USFWS) provided expert advice regarding genetic analyses and allele standardization. The report also benefitted greatly from insight and recommendations from Ed Connor and Dave Pflug (Seattle City Light).

REFERENCES

Baxter JS, Taylor EB, Devlin RH, Hagen J, McPhail JD (1997) Evidence for natural hybridization between Dolly Varden (Salvelinus malma) and bull trout (Salvelinus confluentus) in a northcentral British Columbia watershed. *Canadian Journal of Fisheries and Aquatic Sciences* 54, 421-429.

Cavalli-Sforza L, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution* 21, 550-570.

Dehaan PW, Schwabe LT, Ardren WR (2010) Spatial patterns of hybridization between bull trout, Salvelinus confluentus, and brook trout, Salvelinus fontinalis in an Oregon stream network. *Conservation Genetics* 11, 935-949.

Dunham JB, Rieman BE (1999) Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications* 9, 642-655.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14, 2611-2620.

Jones OR, Wang JL (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* 10, 551-555.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* 5, 187-189.

Kanda N, Allendorf FW (2001) Genetic population structure of bull trout from the Flathead River Basin as shown by microsatellites and mitochondrial DNA markers. *Transactions of the American Fisheries Society* 130, 92-106.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* 13, 55-65.

Page RDM (1996) TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12, 357-358.

Piry S, Alapetite A, Cornuet JM*, et al.* (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95, 536-539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945-959.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* 94, 9197-9201.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43, 223-225.

Rieman BE, Allendorf FW (2001) Effective population size and genetic conservation criteria for bull trout. *North American Journal of Fisheries Management* 21, 756-764.

Rieman BE, Dunham JB (2000) Metapopulations and salmonids: a synthesis of life history patterns and empirical observations. *Ecology of Freshwater Fish* 9, 51-64.

Rieman BE, McIntyre JD (1995) Occurrence of bull trout in naturally fragmented habitat patches of varied size. *Transactions of the American Fisheries Society* 124, 285-296.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8, 103-106.

Saunders DA, Hobbs RJ, Margules CR (1991) Biological consequences of ecosystem fragmentation-a review. *Conservation Biology* 5, 18-32.

Service USFaW (1999) Endangered and threatened wildlife and plants; determination of threatened status for bull trout in the coterminous United States. *Fed. Reg*. [Docket 9928295, October 1990], 64(210), 58910-58933.
Service USFaW (2004) Draft Recovery Plan for the Coastal-Puget Sound Distinct Population Segment of Bull Trout (*Salvelinus confluentus*), p. 389 + xvii pp., Portland, Oregon.
Smith MJ, Pascal CE, Grauvogel Z*, et al.* (In Review) Multiplex preamplification PCR and microsatellite validation allows accurate single nucleotide polymorphism (SNP) genotyping of historical fish scales.
Spruell P, Rieman BE, Knudsen KL, Utter FM, Allendorf FW (1999) Genetic population structure within streams: microsatellite analysis of bull trout populations. *Ecology of Freshwater Fish* 8, 114-121.
Waples RS (1990) Temporal changes of allele frequency in Pacific salmon: implications for mixed-stock fishery analysis. *Canadian Journal of Fisheries and Aquatic Sciences* 47, 968-976.
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
Whiteley AR, Spruell P, Rieman BE, Allendorf FW (2006) Fine-scale genetic structure of bull trout at the southern limit of their distribution. *Transactions of the American Fisheries Society* 135, 1238-1253.

Table 1. Collection data for bull trout in the Skagit River Basin. Collections are grouped by reporting groups used in genetic assignment tests. Collection date by month and year, adult (a) or juvenile (j) sample, core recovery area (USFWS 2004) of collection, site number corresponding to map in Figure 1, number of samples collected (n), number of samples with more than 7 loci scored, full siblings removed from dataset based on COLONY analyses, Dolly Varden/brook trout removed from analyses, and final sample size used in analyses (n final).

Reporting group and collection site	Collection date	Adult/ juvenile	Core recovery area	Site number	n	n (>7 loci scored)	Full siblings removed	Dolly Varden/ brook trout	n final
Above Dam									
Upper Skagit River	Oct 01	a	U. Skagit	1	16	16	0	0	16
Lightning Creek	Aug-Sep 01,02	a,j	U. Skagit	2	32	26	1	24	n/a
Big Beaver Creek	Sep-Oct 09	a	U. Skagit	3	21	21	1	0	20
Ruby Creek	Sep-Oct 01,02,04,09	a	U. Skagit	4	43	43	3	0	40
Stetattle Creek	Jul 09-Aug 09	j	L. Skagit	5	59	56	18	15	23
Goodell									
Lower Goodell Creek	Jul 09	j	L. Skagit	6	60	59	14	0	45
Upper Goodell Creek	Jul 09	j	L. Skagit	7	19	19	8	0	11
Bacon									
Bacon Creek	Jul 09	j	L. Skagit	8	61	61	32	0	29
Cascade									
Cascade River	Jul 09	j	L. Skagit	9	39	39	9	0	30
Marble Creek	Jul 09	j	L. Skagit	10	28	28	10	0	18
Kindy Creek	Jul 09	j	L. Skagit	11	30	30	13	0	17
Illabot									
Illabot Creek	July 09	j	L. Skagit	12	70	70	26	0	44
Sauk									
South Fork Sauk River	Jul 09-Aug 09	j	L. Skagit	13	59	58	20	0	38
Downey									
Downey Creek	Aug 09	j	L. Skagit	14	58	58	6	0	52
Skagit Mixed*									
Skagit Mainstem	Aug 06-Jun 08	a	n/a		435	434	0	0	434

Table 2. PCR reaction conditions for loci amplified, including annealing temperature, primer concentration, $MgCl_2$ concentration, thermal cycling profile, and locus reference.

Locus	Anneal temperature (°C)	Primer concentration	$MgCl_2$ concentration	Thermal cycle profile	Reference
Omm1128	56	0.25uM	1.5 mM	2	Rexroad et al. (2001)
Omm1130	55	0.25uM	1.5 mM	1	Rexroad et al. (2001)
Sco102	56	0.25uM	1.5 mM	1	WDFW unpublished data
Sco105	58	0.25uM	2.0 mM	1	WDFW unpublished data
Sco106	55	0.25uM	2.0 mM	2	WDFW unpublished data
Sco107	58	0.25uM	2.0 mM	1	WDFW unpublished data
Sco109	55	0.25uM	2.0 mM	5	WDFW unpublished data
Sco200	58	0.25uM	1.5 mM	4	DeHaan and Ardren (2005)
Sco202	55	0.25uM	1.5 mM	4	DeHaan and Ardren (2005)
Sco212	55	0.25uM	2.0 mM	1	DeHaan and Ardren (2005)
Sco215	55	0.25uM	1.5 mM	4	DeHaan and Ardren (2005)
Sco216	52	0.50uM	1.5 mM	4	DeHaan and Ardren (2005)
Sco218	55	0.25uM	1.5 mM	1	DeHaan and Ardren (2005)
Sco220	60	0.25uM	2.0 mM	1	DeHaan and Ardren (2005)
Sfo18	60	0.50uM	1.5 mM	3	Angers et al. (1995)
Smm22	55	0.25uM	1.5 mM	1	Crane et al. (2004)

Cycling Profile 1: 95°C (2 min), 30 cycles of: 95°C (20 s), Tm (40 s), 72°C (1 min) followed by a final extension at 72°C (30 min)
Cycling Profile 2: 95°C (2 min), 30 cycles of: 95°C (20 s), Tm (40 s), 72°C (30 s) followed by a final extension at 72°C (10 min)
Cycling Profile 3: 95°C (2 min), 30 cycles of: 95°C (20 s), Tm (40 s), 72°C (1 min) followed by a final extension at 72°C (20 min)
Cycling Profile 4: 5 cycles of: 95°C (2 min), Tm (40 s), 72°C (1 min), 30 cycles of: 95°C (20 s), Tm (40 s), 72°C (1 min) followed by a final extension at 72°C (30 min)
Cycling Profile 5: 5 cycles of: 95°C (2 min), Tm (40 s), 72°C (1 min), 30 cycles of: 95°C (20 s), Tm (40 s), 72°C (1 min) followed by a final extension at 72°C (20 min)

Table 3. Locus-specific F_{IS} values for bull trout from the Skagit River Basin arranged by a) individual collections and b) collections pooled into reporting groups for population assignment tests; bold values indicate significant deviations from expected Hardy-Weinberg proportions, asterisks indicate significant deviations after sequential Bonferroni correction.

a)

	Locus														
Collection	*Omm1128*	*Omm1130*	*Sco102*	*Sco105*	*Sco106*	*Sco107*	*Sco200*	*Sco202*	*Sco212*	*Sco215*	*Sco216*	*Sco218*	*Sco220*	*Sfo18*	*Smm22*
U. Skagit	0.068	0.319	0.400	0.455	**0.783**	0.008	mono	-0.062	**0.443**	mono	mono	0.041	-0.079	mono	-0.005
Big Beaver	0.130	0.159	0.166	-0.108	mono	-0.036	mono	-0.027	0.157	mono	mono	0.007	0.120	mono	-0.056
Ruby	**0.024**	-0.052	0.135	0.132	0.040	-0.056	-0.026	mono	0.150	mono	mono	-0.169	-0.025	mono	0.008
Stetattle	0.051	-0.135	-0.250	0.206	0.352	-0.100	mono	mono	0.053	mono	mono	0.022	-0.177	mono	0.007
L. Goodell	-0.033	-0.013	-0.042	-0.123	0.016	-0.004	0.006	0.068	0.096	mono	-0.021	-0.074	-0.029	mono	-0.001
U. Goodell	-0.041	0.161	-0.164	-0.125	0.094	-0.257	0.007	**0.516**	-0.191	mono	-0.457	-0.111	-0.149	mono	0.126
Bacon	**-0.002***	**-0.115**	-0.414	-0.048	**0.025**	**-0.016**	**-0.117**	-0.108	**-0.100**	mono	**-0.039***	-0.103	-0.093	mono	-0.106
Cascade	-0.102	-0.076	-0.017	0.037	0.094	0.137	0.056	-0.039	0.099	mono	-0.009	0.081	0.039	mono	-0.049
Marble	-0.120	-0.060	-0.226	-0.308	-0.040	-0.059	0.064	-0.116	-0.052	mono	0.034	-0.230	-0.150	mono	0.079
Kindy	**0.065**	-0.047	0.306	-0.095	0.055	0.074	0.131	0.095	-0.095	mono	-0.321	0.176	0.101	mono	0.068
Illabot	-0.119	**0.020**	-0.117	-0.020	**0.048**	0.012	0.011	**0.279**	-0.092	mono	0.027	-0.009	-0.082	mono	0.064
S.F. Sauk	-0.040	0.102	0.066	-0.096	-0.007	-0.052	0.072	-0.199	-0.094	mono	-0.136	0.067	-0.098	mono	-0.032
Downey	0.072	0.069	0.063	-0.091	-0.062	0.001	0.029	0.029	-0.038	mono	0.027	0.017	0.038	mono	0.013

b)

	Locus														
Reporting Group	*Omm1128*	*Omm1130*	*Sco102*	*Sco105*	*Sco106*	*Sco107*	*Sco200*	*Sco202*	*Sco212*	*Sco215*	*Sco216*	*Sco218*	*Sco220*	*Sfo18*	*Smm22*
Above Dam	**0.085**	0.042	0.097	**0.146**	**0.291***	-0.051	-0.016	-0.023	**0.186**	mono	mono	-0.037	-0.025	mono	0.009
Goodell	-0.005	0.050	-0.046	-0.110	0.039	-0.033	0.011	0.210	0.044	mono	-0.112	-0.057	**-0.034**	mono	0.036
Bacon	**-0.002***	**-0.115**	-0.414	-0.048	**0.025**	**-0.016**	**-0.117**	-0.108	**-0.100**	mono	**-0.039***	-0.103	-0.093	mono	-0.106
Cascade	-0.065	-0.060	0.004	-0.065	0.043	0.075	0.093	-0.023	0.023	mono	**-0.086**	0.059	0.043	mono	0.024
Illabot	-0.119	**0.020**	-0.117	-0.020	**0.048**	0.012	0.011	**0.279**	-0.092	mono	0.027	-0.009	-0.082	mono	0.064
Sauk	-0.040	0.102	0.066	-0.096	-0.007	-0.052	0.072	-0.199	-0.094	mono	-0.136	0.067	-0.098	mono	-0.032
Downey	0.072	0.069	0.063	-0.091	-0.062	0.001	0.029	0.029	-0.038	mono	0.027	0.017	0.038	mono	0.013

Table 4. Summary of genetic statistics for bull trout from the Skagit River Basin arranged by individual collections, reporting groups, and above dam/below dam collections. Adult and sub-adult fluvial bull trout are designated as Skagit mainstem. Asterisks indicate the collections pooled into the associated reporting groups; sample size (n), observed heterozygosity (Ho), unbiased expected heterozygosity (He), and allelic richness (Ar).

	n	Ho	He	Ar
Collections				
U. Skagit*	16	0.377	0.445	3.930
Big Beaver*	20	0.356	0.378	3.150
Ruby*	40	0.382	0.385	3.110
Stetattle*	23	0.379	0.373	2.720
L. Goodell**	45	0.679	0.672	5.720
U. Goodell**	11	0.648	0.620	4.640
Bacon	29	0.717	0.661	5.160
Cascade***	30	0.667	0.681	5.700
Marble***	18	0.707	0.654	5.510
Kindy***	17	0.668	0.696	6.140
Illabot	44	0.637	0.635	4.820
S.F. Sauk	38	0.657	0.639	5.370
Downey	52	0.686	0.695	6.070
Reporting groups				
Above Dam*	99	0.375	0.395	3.260
Goodell**	56	0.674	0.675	5.710
Bacon	29	0.717	0.661	5.160
Cascade***	65	0.678	0.683	5.820
Illabot	44	0.637	0.635	4.820
Sauk	38	0.657	0.639	5.370
Downey	52	0.686	0.695	6.070
Above/Below dam				
Above dam	99	0.375	0.395	3.120
Below dam	284	0.674	0.702	6.590
Skagit mixed				
Skagit mainstem	434	0.666	0.713	6.700

Table 5. Pairwise F_{ST} values and significance of exact tests for genotypic differentiation for bull trout from the Skagit River Basin arranged by a) individual collections and b) collections pooled into reporting groups for population assignment tests; bold values indicate non-significant F_{ST} values, asterisks indicate non-significant genotypic differentiation.

a)

Collection	U. Skagit	Big Beaver	Ruby	Stetattle	L. Goodell	U. Goodell	Bacon	Cascade	Marble	Kindy	Illabot	S.F. Sauk	Downey
U. Skagit	0.0000												
Big Beaver	**0.0066***	0.0000											
Ruby	0.0242	0.0182	0.0000										
Stetattle	0.0330	0.0166	0.0147	0.0000									
L. Goodell	0.2164	0.2581	0.2771	0.2658	0.0000								
U. Goodell	0.2777	0.3290	0.3470	0.3338	0.0529	0.0000							
Bacon	0.2285	0.2785	0.3009	0.2915	0.0468	0.1032	0.0000						
Cascade	0.2229	0.2620	0.2838	0.2734	0.0307	0.0803	0.0431	0.0000					
Marble	0.2281	0.2698	0.2895	0.2844	0.0478	0.0966	0.0777	0.0236	0.0000				
Kindy	0.2369	0.2833	0.3046	0.2978	0.0275	0.0794	0.0516	**0.0004***	0.0187*	0.0000			
Illabot	0.2543	0.2972	0.3176	0.3040	0.0523	0.1217	0.0767	0.0537	0.0618	0.0464	0.0000		
S.F. Sauk	0.2515	0.2862	0.3049	0.2988	0.0718	0.1197	0.0894	0.0572	0.0747	0.0549	0.0898	0.0000	
Downey	0.2186	0.2546	0.2713	0.2689	0.0679	0.0953	0.0734	0.0485	0.0584	0.0372	0.0777	0.0497	0.0000

b)

Reporting Group	Above Dam	Goodell	Bacon	Cascade	Illabot	Sauk	Downey
Above Dam	0.0000						
Goodell	0.2840	0.0000					
Bacon	0.3148	0.0502	0.0000				
Cascade	0.2733	0.0317	0.0497	0.0000			
Illabot	0.3301	0.0578	0.0767	0.0476	0.0000		
Sauk	0.3241	0.0737	0.0894	0.0560	0.0898	0.0000	
Downey	0.2925	0.0664	0.0734	0.0440	0.0777	0.0497	0.0000

Table 6. Analysis of molecular variance (AMOVA) used to examine genetic structure among bull trout collections from the Skagit River Basin pooled into reporting groups for genetic assignment tests (*P<0.001). Variation among collections within reporting groups was very low (1.49%) relative to the variation among reporting groups (15.40%).

Source of Variation	Sum of squares	Variance components	Percentage of variation
Among groups	561.91	0.7863	15.40*
Among populations within groups	45.23	0.0759	1.49*
Within populations	3195.80	4.2441	83.12*
Total	3802.94	5.1063	

Table 7. Self-assignment to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. The test follows the leave one out method of assignment where each fish is removed from the baseline without replacement prior to calculating likelihoods of belonging to each reporting group. The number and percentage of home assignments (assignments to the reporting group where the individual was collected) and other assignments (assignments to reporting groups other than the group in which the individual was collected) that exceed the 90% threshold for positive assignment are shown. Assignments below the 90% threshold are reported as unassigned individuals.

	Baseline Assignments						
Reporting group	>90% home (n)	>90% other (n)	Unassigned (n)	Total (n)	home (%)	other (%)	unassigned (%)
Above Dam	99	0	0	99	100.00	0.00	0.00
Goodell	52	0	4	56	92.86	0.00	7.14
Bacon	24	4	1	29	82.76	13.79	3.45
Cascade	54	2	9	65	83.08	3.08	13.85
Illabot	41	0	3	44	93.18	0.00	6.82
Sauk	36	0	2	38	94.74	0.00	5.26
Downey	48	0	4	52	92.31	0.00	7.69
Total	354	6	23	383	92.43	1.57	6.01

Table 8. Mixture assignments of fluvial adult and sub-adult bull trout in the mainstem Skagit River to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. Assignments below the 90% threshold are reported as unassigned individuals. Results displayed by all samples and divided into legal (>508mm) and sub-legal (<508mm) size classes.

	Mixture Assignments					
Reporting group	all lengths (n)	>508 mm (n)	<508 mm (n)	all lengths (%)	>508 mm (%)	<508 mm (%)
Above Dam	0	0	0	0.00	0.00	0.00
Goodell	132	30	102	38.04	37.50	38.20
Bacon	14	3	11	4.03	3.75	4.12
Cascade	123	32	91	35.45	40.00	34.08
Illabot	45	8	37	12.97	10.00	13.86
Sauk	6	0	6	1.73	0.00	2.25
Downey	27	7	20	7.78	8.75	7.49
Total	347	80	267	100	100	100

Table 9. Mixture assignments of fluvial adult and sub-adult bull trout from three mainstem Skagit River regions to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. Assignments were stratified by region of capture (Above Shovelspur, Shovelspur to Marblemount, and Marblemount to Rockport) and quarter of the year within regions.

	Above Shovelspur				Shovelspur to Marblemount				Marblemount to Rockport			
Reporting Group	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
Above Dam	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Goodell	72.5%	82.5%	50.0%	69.6%	38.5%	55.6%	33.3%	33.3%	11.1%	10.9%	0.0%	5.3%
Bacon	2.5%	0.0%	0.0%	4.3%	15.4%	6.7%	16.7%	11.1%	5.6%	0.0%	0.0%	2.6%
Cascade	25.0%	17.5%	25.0%	13.0%	30.8%	31.1%	33.3%	27.8%	46.3%	50.9%	20.0%	47.4%
Illabot	0.0%	0.0%	25.0%	0.0%	7.7%	4.4%	16.7%	5.6%	16.7%	29.1%	40.0%	28.9%
Sauk	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.7%	1.8%	20.0%	5.3%
Downey	0.0%	0.0%	0.0%	13.0%	7.7%	2.2%	0.0%	22.2%	16.7%	7.3%	20.0%	10.5%
n	40	40	4	23	13	45	6	18	54	55	5	38

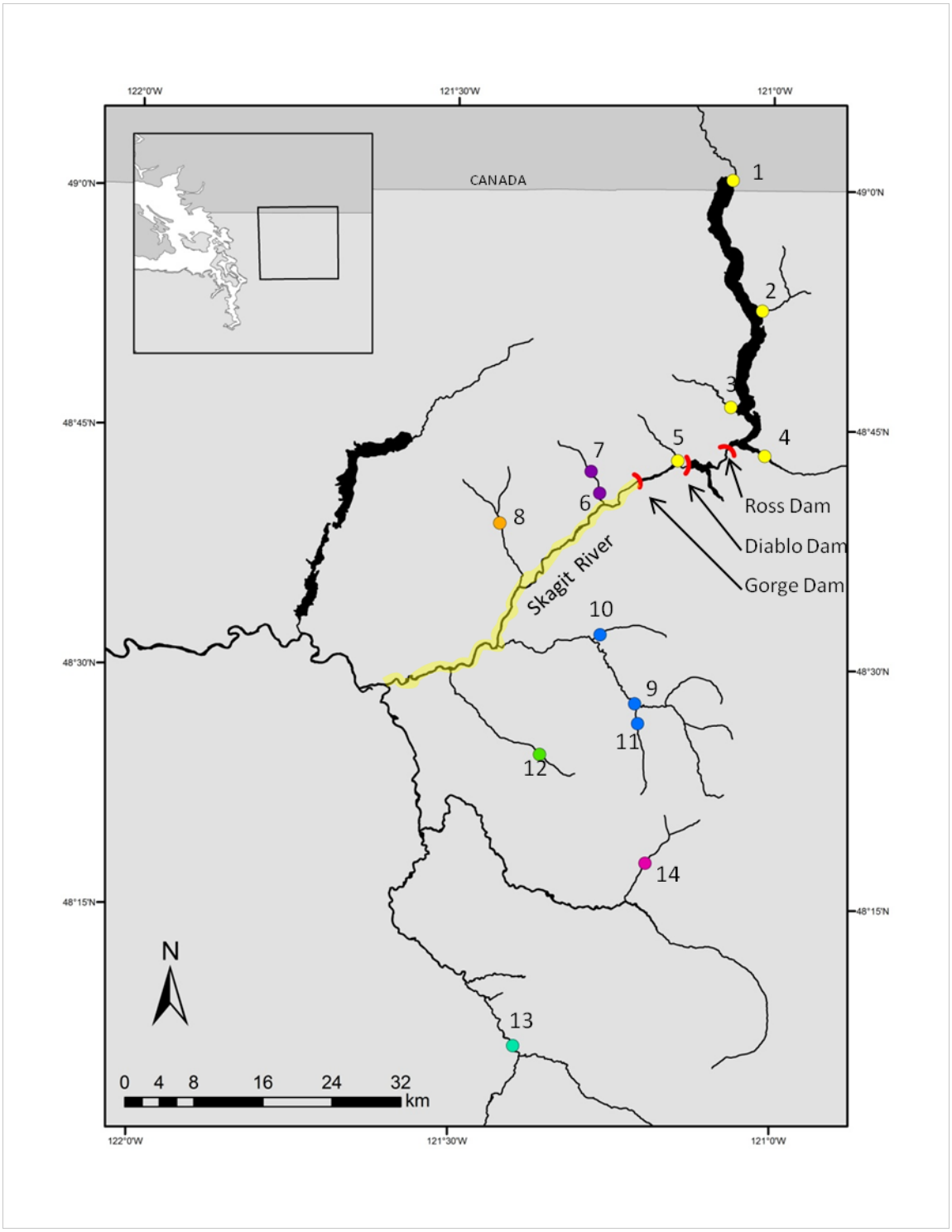

Figure 1. Map of Skagit Basin collection sites. Numbers correspond with collections shown in Table 1. Like colors represent subbasin collections pooled for genetic assignment tests. Highlighted region of the mainstem Skagit River indicates region where fluvial adult and sub-adult bull trout were sampled.

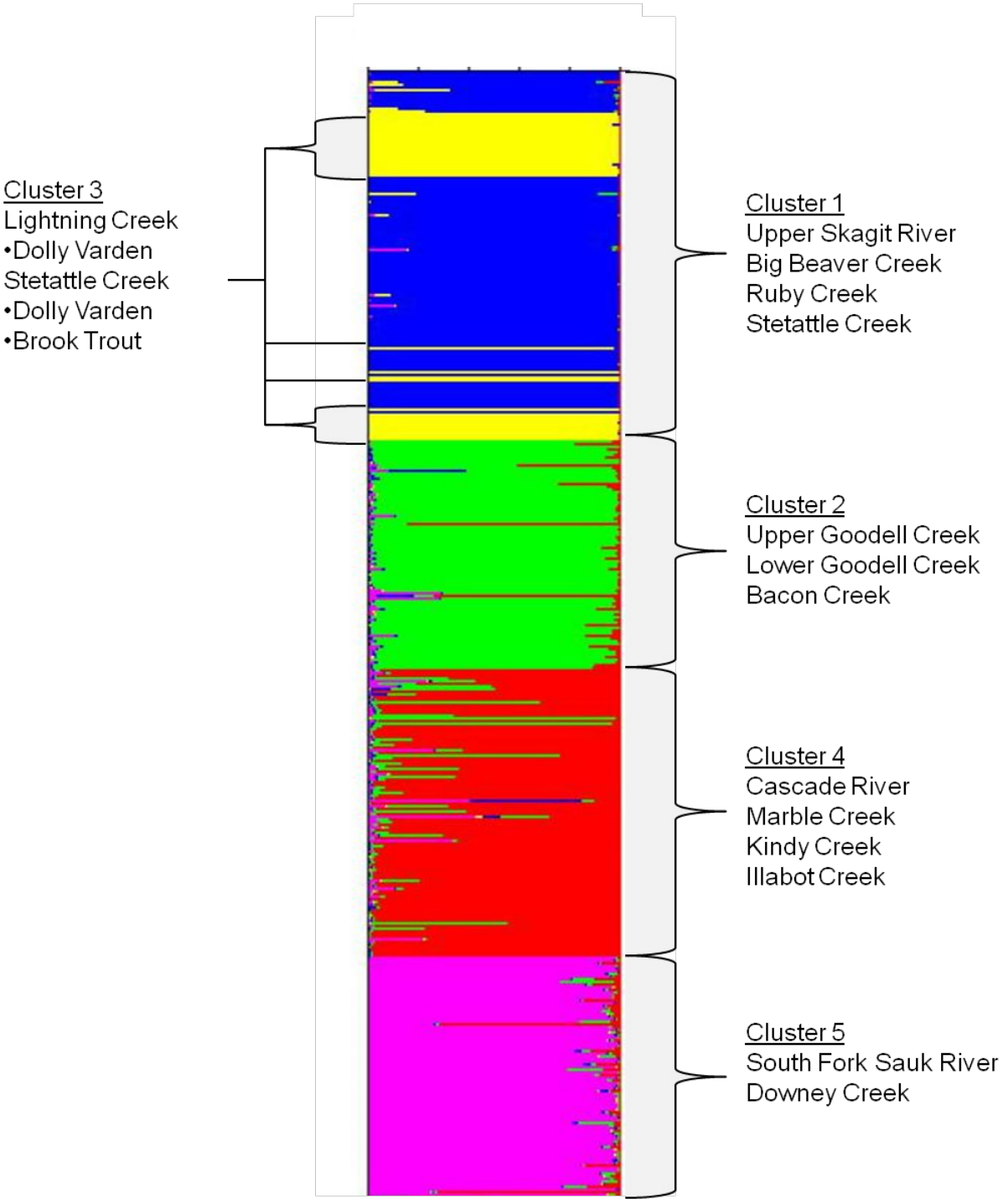

Figure 2. Clustering of ancestry values from STRUCTURE analysis for K=5. Colors represent ancestry in each of the five designated clusters. Horizontal bars represent individual bull trout ancestry; a single color indicates pure ancestry in a given cluster and multiple colors indicate mixed ancestry. Yellow bars are samples that were identified as having Dolly Varden or bull trout alleles.

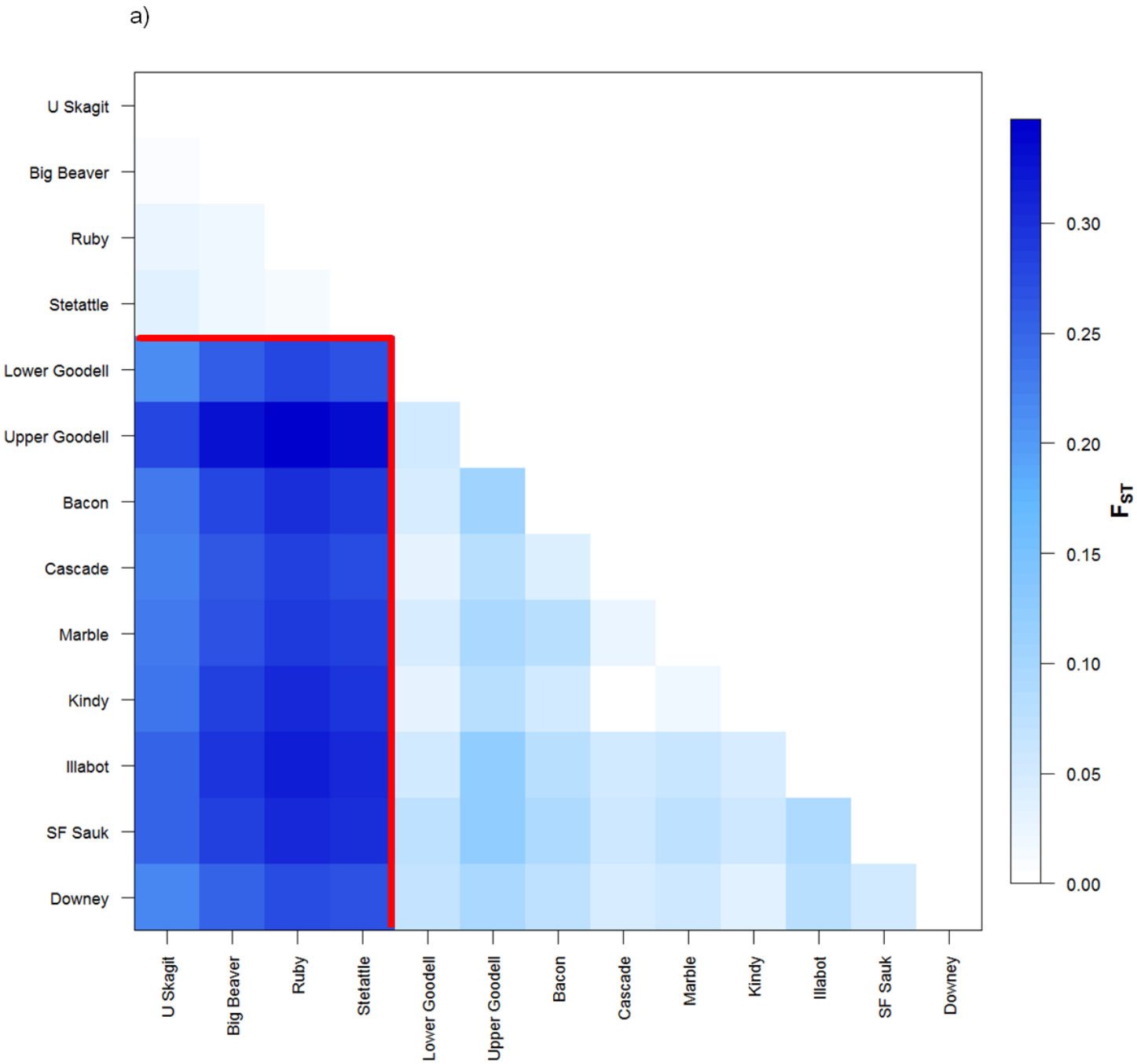

Figure 3a. Graphical depictions of pairwise F_{st} values between collections and between reporting groups for bull trout from the Skagit River Basin arranged by a) individual collections and b) collections pooled into reporting groups. Red lines indicate separation between above dam and below dam collections/reporting groups.

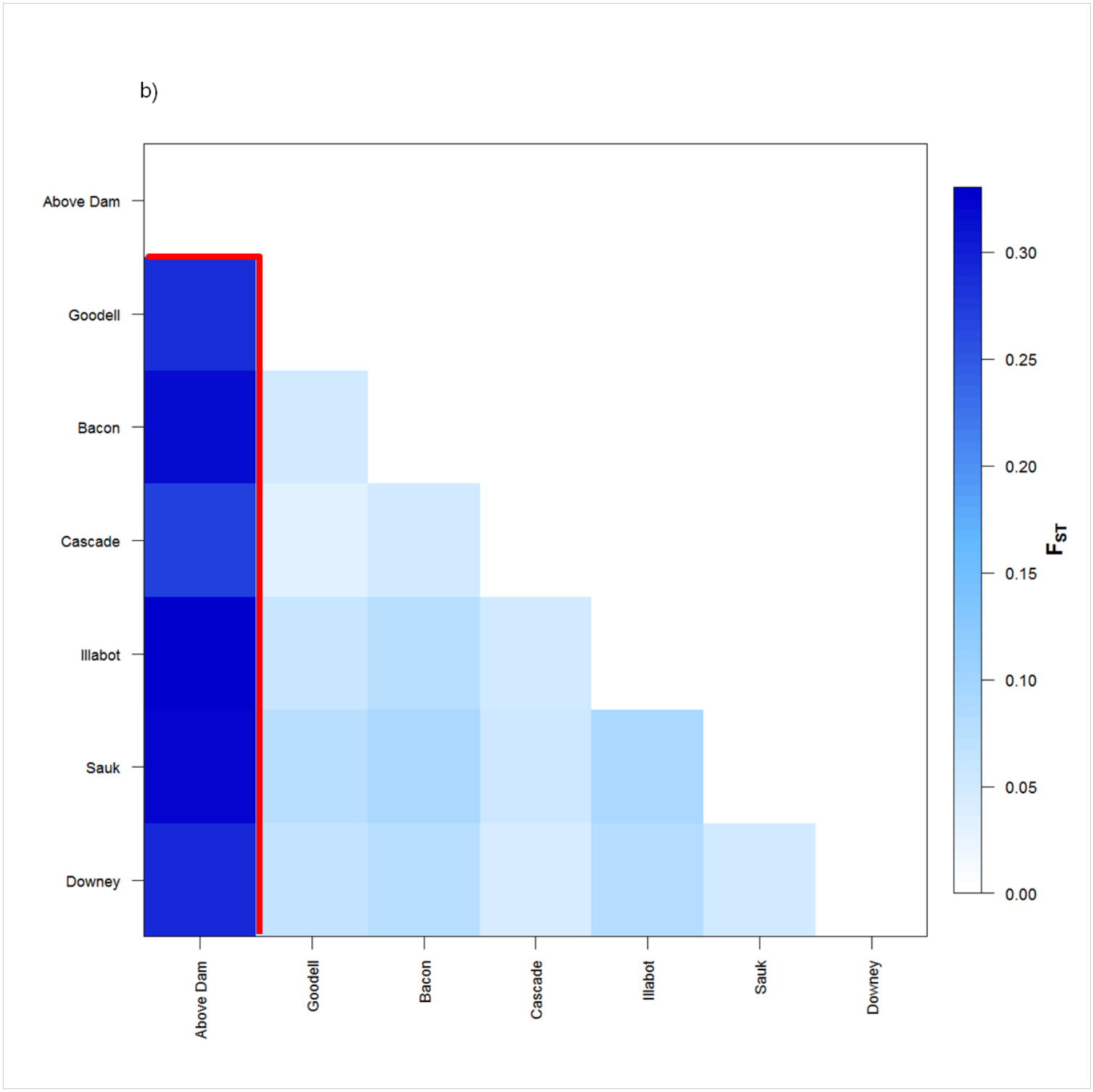

Figure 3b. Graphical depictions of pairwise F_{st} values between collections and between reporting groups for bull trout from the Skagit River Basin arranged by a) individual collections and b) collections pooled into reporting groups. Red lines indicate separation between above dam and below dam collections/reporting groups.

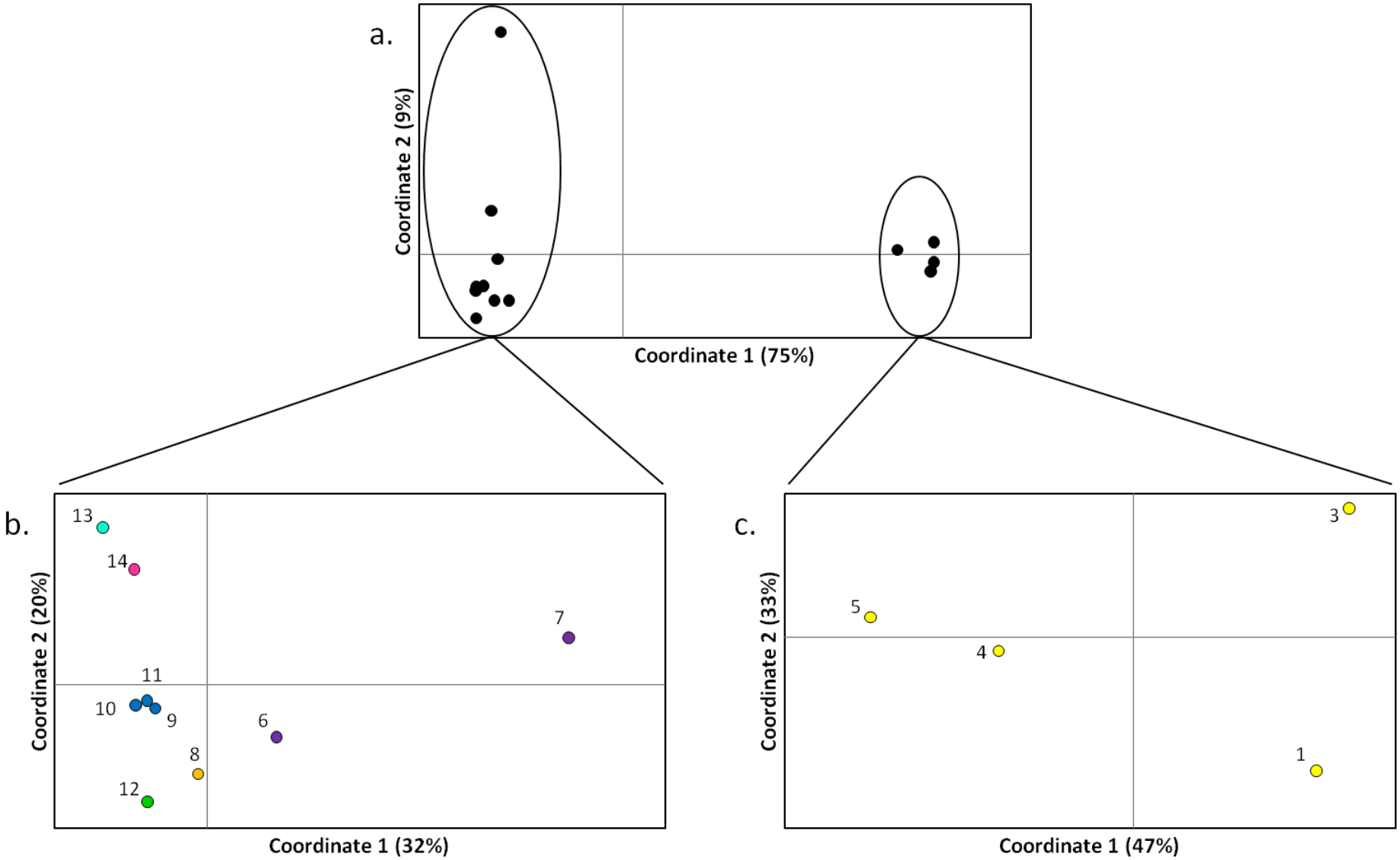

Figure 4. Principal coordinate analysis (PCA) of bull trout collections from the Skagit River Basin showing a) high genetic differentiation between above dam and below dam collections. The first principal coordinate, which separated these collections, explained 75% of the total variation. PCA was conducted independently for b) below dam collections and c) above dam collections which provided a more detailed depiction of divergence between collections; colors correspond to pooled subbasin collections in Figure 1.

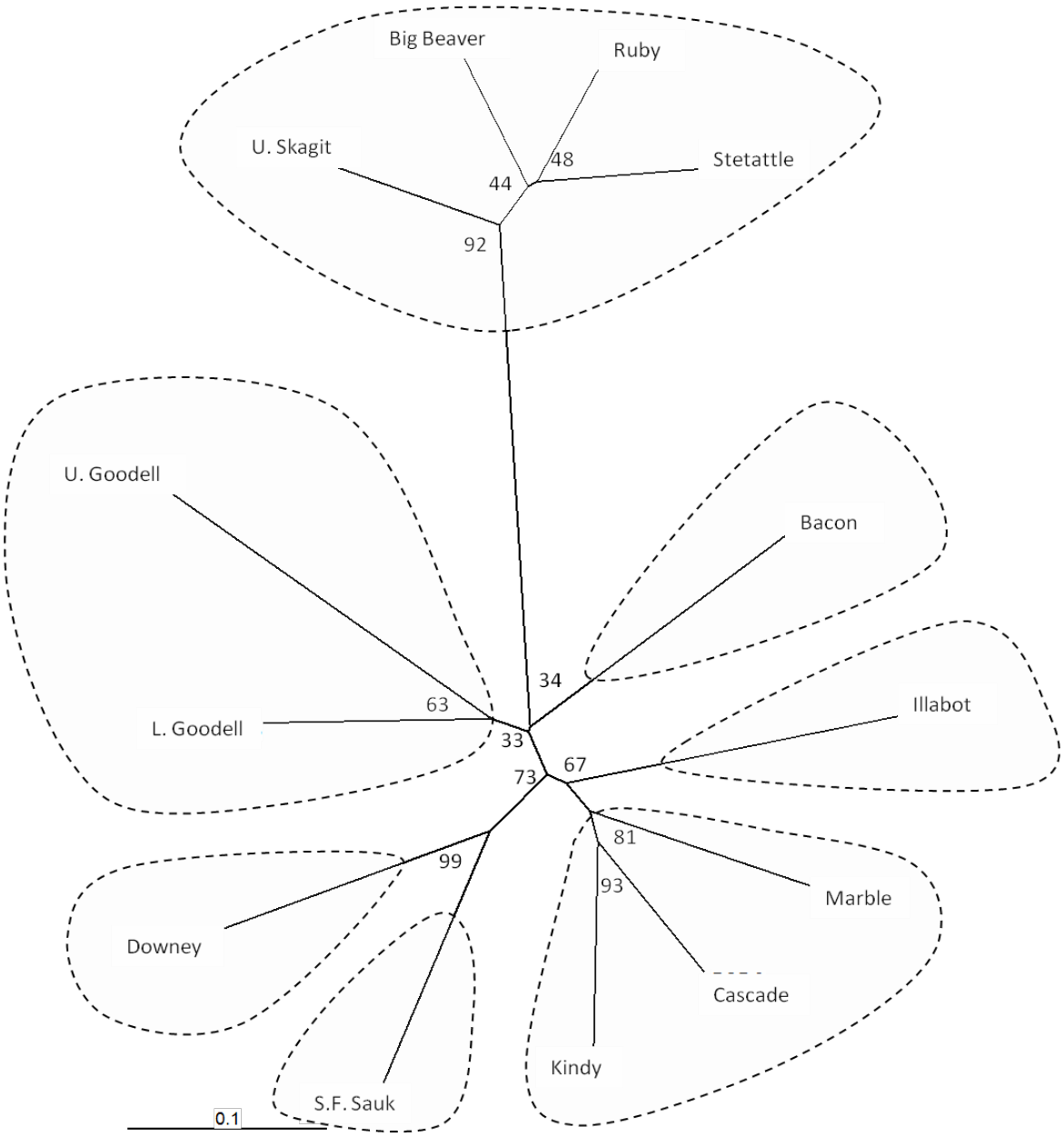

Figure 5. Unrooted neighbor-joining (NJ) dendrogram of bull trout collections from the Skagit River Basin using Cavalli-Sforza and Edwards (1967) chord distance to display genetic relationships among collections. One thousand bootstrap replicates were performed to evaluate tree topology. Percentage bootstrap support is shown. Subbasin collections pooled for genetic assignment tests are outlined with dashed lines.

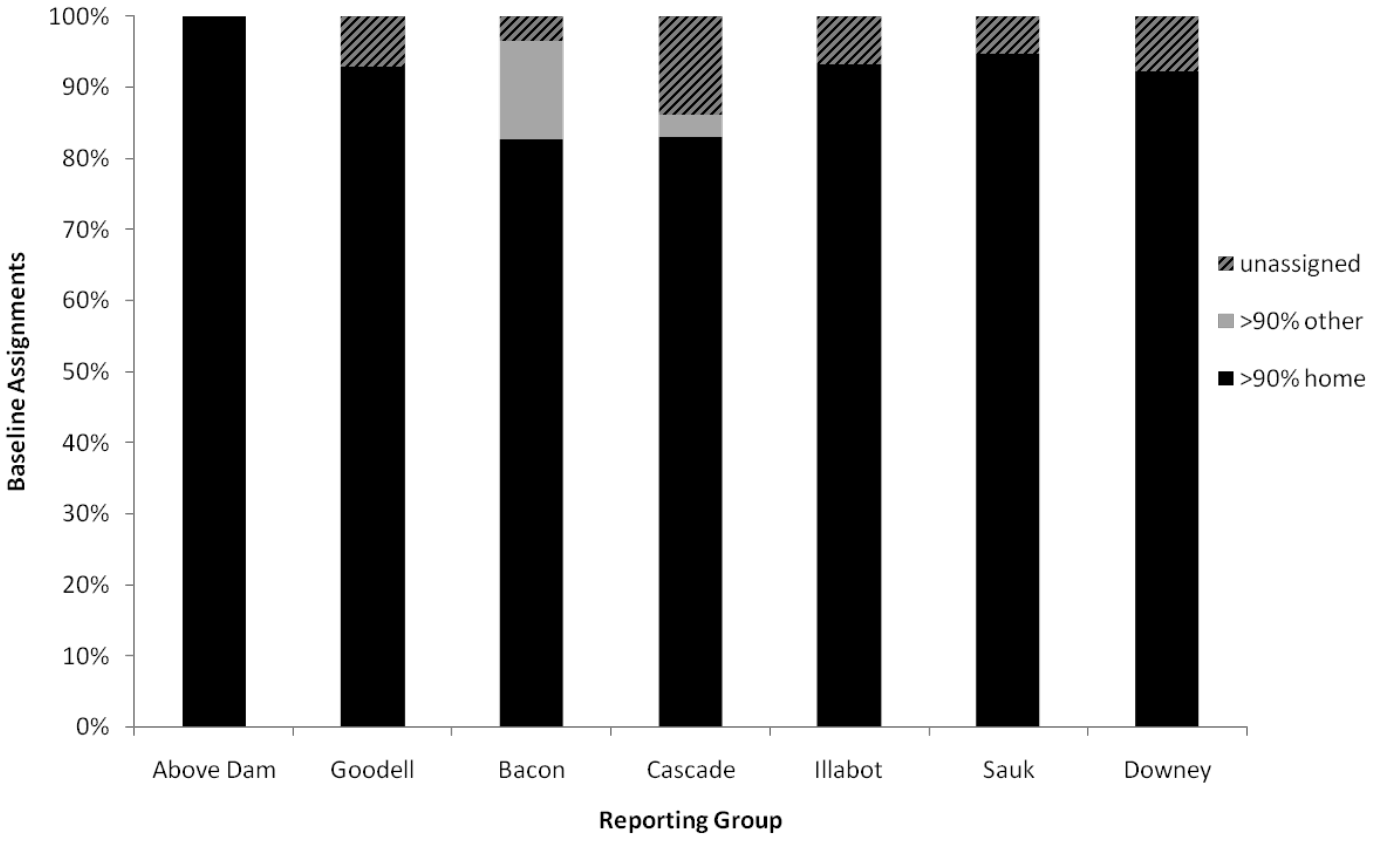

Figure 6. Self-assignment to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. The test follows the leave one out method of assignment where each fish is removed from the baseline without replacement prior to calculating likelihoods of belonging to each reporting group. The number and percentage of home assignments (assignments to the reporting group where the individual was collected) and other assignments (assignments to reporting groups other than the group in which the individual was collected) that exceed the 90% threshold for positive assignment are shown. Assignments below the 90% threshold are reported as unassigned individuals.

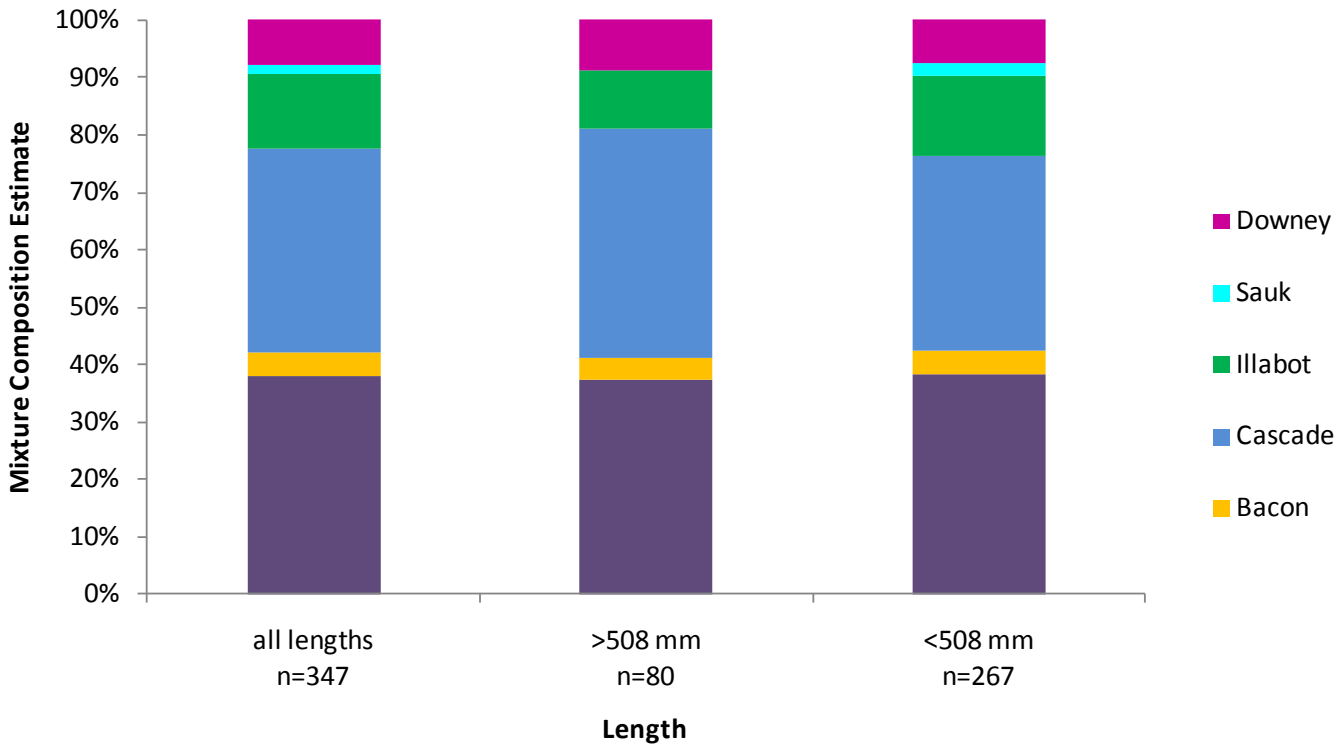

Figure 7. Mixture assignments of fluvial adult and sub-adult bull trout from the mainstem Skagit River to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. Results displayed by all samples and divided into legal (>508mm) and sub-legal (<508mm) size classes. Colors correspond to location of reporting groups in Figure1.

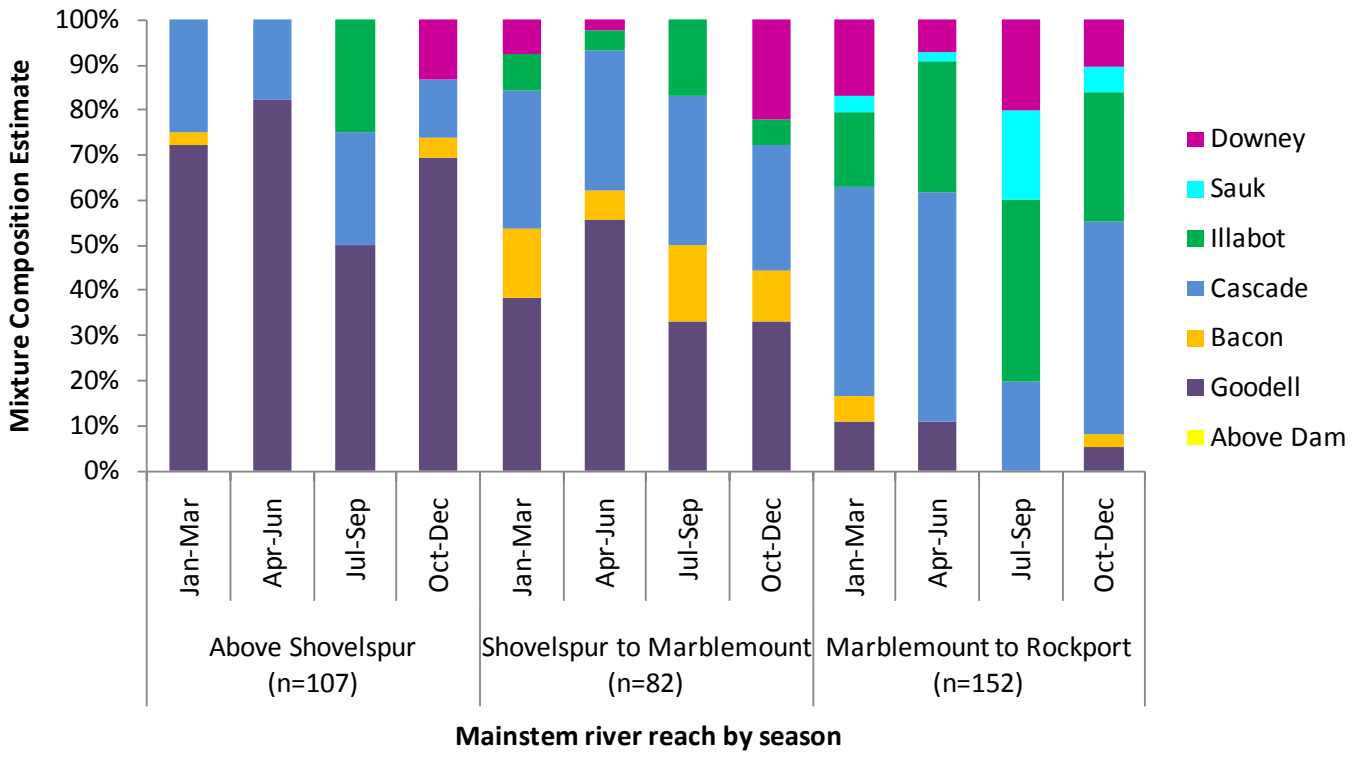

Figure 8. Mixture assignments of fluvial adult and sub-adult bull trout from three mainstem Skagit River regions to reporting group of origin using a likelihood ratio threshold of 90% as positive assignment. Assignments were stratified by region of capture (Above Shovelspur, Shovelspur to Marblemount, and Marblemount to Rockport) and quarter of the year within regions. Colors correspond to location of reporting groups in Figure1.